Universidade do Minho
Escola de Ciências

Maria Antónia Sousa Correia da Silva

Identification and functional characterization of novel plasma membrane carboxylate transporters

May 2021

Identification and functional characterization of novel plasma membrane carboxylate transporters
Maria Antónia Sousa Correia da Silva
Uminho | 2021


[DP_AEM]
DOCTORAL PROGRAM IN
APPLIED AND ENVIRONMENTAL
MICROBIOLOGY
PD+F
PROGRAMAS DE DOUTORAMENTO FCT
COMPETE
2020
NORTE2020
PROGRAMA OPERACIONAL REGIONAL DO NORTE
PORTUGAL
2020
UNIÃO EUROPEIA
Fundo Social Europeu

Universidade do Minho
Escola de Ciências

Maria Antónia Sousa Correia da Silva

Identification and functional characterization of novel plasma membrane carboxylate transporters

PhD Thesis in Biology
Specialization in Molecular Biotechnology

Work supervised by
Doctor Isabel João Soares Silva
Professor Doctor Margarida Casal

May 2021

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

ACKNOWLEDGMENTS/AGRADECIMENTOS

E no findar desta etapa, não poderia deixar de expressar o meu reconhecimento e gratidão por todos aqueles que se cruzaram comigo nestes 4 (quase 5!) anos. Muitas experiências foram partilhadas desde então, às alegrias aos momentos menos bons... uma montanha russa de momentos, rodeados de vários intervenientes dos quais merecem a minha palavra de apreço. Primeiramente, tenho de agradecer à Professora Margarida Casal, a oportunidade para voltar a Portugal e em integrar a sua equipa no âmbito do DP_AEM em 2016. Permitiu-me conhecer pessoas que verdadeiramente gostam de Ciência e trouxe-me em modo de "presente subtendido" uma cidade e pessoas que levo comigo para sempre. Por todos os ensinamentos, ideias científicas partilhadas e pelo "apertão-acreditar" quando achava que não era possível "chegar lá" passado ao longo destes tão repentinos 4 anos, expresso o meu enorme obrigada!

Em seguida, o meu maior agradecimento é dirigido à minha orientadora, à Doutora Isabel João. A João foi um presente que a vida me deu. A sua serenidade, disponibilidade ilimitada, paciência e encorajamento foram ferramentas cruciais na minha compreensão e integração total pela temática dos transportadores e por isso estou muito grata. Agradeço a sua orientação exemplar marcada por um elevado e rigoroso nível científico, interesse permanente e visão crítica em tudo que discutimos, mudando positivamente a minha perspetiva no que toca fazer Ciência. Ensinou-me também que em investigação a ordem de trabalhos não é o mais relevante, mas sim a história que podemos contar com os resultados que temos. No final, o mais importante de tudo é vibrar com o que se faz e "fazer Ciência" com verdade. Pelas infindáveis reuniões de todo o tipo, apoio, compreensão e confiança na realização deste trabalho, estou-lhe agradecida (sim, agora estou a chorar...). Obrigada por me ter acompanhado e ensinado tanto ao longo destes quatro (cinco..santinho!) anos. Além de orientadora, é uma amiga e muitas das vezes foi uma 3ªMãe. Por tudo isto e muito mais, agradeço-lhe de coração João!

Ao meu colega David Ribas, pelos ensinamentos práticos, discussões científicas e disponibilidade que teve em cada trabalho que partilhamos. TEBU, is up to you (um dia será marca registada!)!

Ao Professor Pedro Soares, Daniel, Professora Célia Pais e Ricardo pela parceria e conhecimento partilhado nas diversas áreas integradas nos projetos científicos em comum. *To Professor Robert Mans from TU Delft (NL) for the opportunity for a close collaboration with his group and Michael Sauer from BOKU (AT) for sharing ideias, and both for the avaiability from the first contact and, most of all, for the enthusiasm in sharing knowledge as great human professionals.*

Agradeço ainda aos meus colegas e amigos dos laboratórios LBM (e LBM-II) e LGM pelo espírito de partilha e boa disposição tornando os dias em laboratório mais risonhos ao longo desta etapa. Principalmente à malta fixe integrante dos n-grupos criados no fb para os almoços cantina vs grill vs raclettes-aleatórias. Obrigada ao Mário por trazer a animação e estupidez necessária, ao lab-chief Pedro pela sua amizade e companheirismo, à minha Cláu amiga de todas as horas e partner no pacotinho-açúcar, à Coralie por todo o incentivo e vai-correr-tudo-bem-amiga, ao pupilo Emanuel pela alegria e dedicação, à Catarina pela companhia nos dias longos quando a "terra gira em contramão" e ao João pela parceria durante o último ano. À minha partner-in-crime Rosie que tanto partilhamos nestes 4 anos, pelo apoio, conversas de todo tipo, mensagens para o break necessário do chocolatinho que acabava sempre numa boa conversa. Obrigada Rosie!

Não poderia esquecer as minhas miúdas partners neste doutoramento, Cátia e Barbara, pela partilha, enorme amizade, boa disposição, pausas-necessárias para uma boa playlist e por todos os abraços no momento certo.

Não esquecendo também o Departamento de Biologia, IB-S e todo o pessoal técnico integrante pelo apoio e disponibilidade prestados.

À Carol, Cris, Diana, Teixeira e Inês por todos os momentos de cumplicidade e amizade vividos durante esta fase. À Cristiane, Cris Faria e à Verónica pela vossa paciência em aturar-me, pelas caminhadas e por me fazerem desanuviar nos assuntos do dia-a-dia. A todos do POS Braga pela excelente equipa que são, concedendo-me a oportunidade de concretizar um dos meus maiores desejos, ainda durante o doutoramento, de levar a Ciência a todos e a todo lado. Às de sempre, Dani, Ana, Soraia, Renata e Susy, que mesmo estando longe fizeram-se por estar sempre presentes na minha vida. Obrigada por todo o vosso incentivo e amizade.

Especial obrigada ao core familiar Mãe, Pai, Cristina e Nuno pelo porto-seguro, apoio incondicional e compreensão nesta fase. À minha irmã Cristina agradeço o colo, mesmo quando não estás por perto e disponibilidade nos momentos críticos. E por fim, às minhas crianças Sofia e António, os meus sobrinhos, que foram o meu principal motor nestes 4 anos, também a minha motivação extra, a alegria nos dias cinzentos e deram os abraços carregados de energia durante a realização deste projeto.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

ABSTRACT

Identification and functional characterization of novel plasma membrane carboxylate transporters

Organic acids are recognized as versatile chemical compounds with a vast variety of applications in sectors ranging from food and beverages, pharmaceutical, personal care, cosmetic products, consumer goods to the chemical industry. Due to the strong demand for these compounds, alternative approaches to non-sustainable processes, e.g. chemical synthesis from petroleum derivatives, are being developed. Sustainable strategies rely on the utilization of microbial cell factories, where transporter proteins play a crucial role through the control of substrate influx and product efflux. In particular, the expression of suitable carboxylic acid exporters avoids the internal cell toxicity due to the accumulation of these compounds, while facilitating its purification from the culture broth. The biodiversity of the microbial world is an excellent pool to uncover new organic acid transporters. In this study, wild yeast species isolated from acidic environments were explored regarding their ability to utilize organic acids. The yeast *Cyberlindnera jadinii* was selected for further studies due to its ability to utilize a vast range of organic acids as sole carbon and energy source, to synthesize a variety of valuable compounds for the food and pharmaceutical sectors, its intrinsic robust characteristics, and its capacity to use inexpensive media with high productivity levels. Using an *in silico* approach, we have explored the predicted transportome of *C. jadinii* and selected putative carboxylate transporters for functional characterization by heterologous expression in *Saccharomyces cerevisiae*. A total of 16 plasma membrane carboxylate transporters, members of the AceTr, SHS, SSS, TDT, DASS, and MCT families, were characterized in terms of transport activity and specificity, structural features, and evolutionary relationships. *S. cerevisiae* was also used as expression host to deepen the knowledge of the citrate exporter CexAp from *Aspergillus niger*, a member of the DHA1 transporter family. The structural characterization of AceTr family members Ato1 and SatP led to the identification of a conserved motif essential for protein function. In addition, the plasma membrane proteins Gpr1 from *Yarrowia lipolytica* and AceP from the archaea *Methanosarcina acetivorans* were functionally characterized as acetate transporters. The role of AceTr members from *S. cerevisiae*, Ato1, Ato2, and Ato3, in the transport of monocarboxylic acids was evaluated via a combination of directed evolution, whole-genome resequencing, and reverse engineering, leading to the discovery of Ato2 and Ato3 mutants as two novel lactic acid transport proteins. Structural insights were also provided using 3D structure predictions combined with molecular docking.

KEYWORDS: biotechnology; carboxylic acids; membrane transporters; microbial cell factories

RESUMO

Identificação e caracterização funcional de novos transportadores de carboxilatos da membrana plasmática

Os ácidos orgânicos são reconhecidos como compostos versáteis com uma vasta aplicabilidade em sectores desde a área alimentar à indústria química. A forte procura por estes compostos exige abordagens alternativas à sua produção convencional, como a síntese química a partir de derivados de petróleo. As estratégias sustentáveis incluem a utilização de fábricas celulares microbianas nas quais as proteínas transportadoras desempenham um papel crucial no controlo do importe de substratos e o exporte de produtos. O exporte de ácidos carboxílicos evita a sua toxicidade intracelular, facilitando a sua purificação do meio de cultura. A biodiversidade microbiana é um excelente reservatório para a descoberta de novos transportadores de ácidos orgânicos. Neste estudo, foram estudadas espécies de leveduras provenientes de ambientes ácidos e avaliada a sua capacidade de utilização de ácidos orgânicos. A levedura *Cyberlindnera jadinii* foi selecionada para estudos posteriores dada a sua capacidade de utilização de uma gama alargada de ácidos orgânicos como fonte única de carbono e energia, para além de sintetizar vários compostos relevantes para os setores industriais. Esta levedura apresenta ainda uma grande robustez e tem capacidade para crescer em meios de cultura baratos apresentando elevados níveis de produtividade. Numa abordagem *in silico*, explorámos o transportoma inferido de *C. jadinii*, selecionando vários potenciais transportadores de carboxilatos para uma caracterização funcional por expressão heteróloga em *Saccharomyces cerevisiae*. Um total de 16 proteínas-alvo, membros das famílias AceTr, SHS, SSS, TDT, DASS e MCT, foram caracterizadas quanto à sua capacidade transporte de carboxilatos, especificidade, características estruturais e relações evolutivas. O exportador de citrato CexA de *Aspergillus niger*, membro da família DHA1, foi também caracterizado por expressão heteróloga na levedura *S. cerevisiae*. A caracterização estrutural dos membros da família AceTr, Ato1 e SatP, levou à identificação de um motivo conservado nestas proteínas transportadoras, essencial para a sua função. As proteínas Gpr1 da levedura *Yarrowia lipolytica* e AceP da archaea *Methanosarcina acetivorans* foram ainda caracterizadas como transportadores de acetato. O papel no transporte de ácidos carboxílicos de outros membros desta família, Ato1, Ato2 e Ato3 de *S. cerevisiae,* foi avaliado por uma abordagem combinada de evolução dirigida, ressequenciação do genoma e engenharia reversa, levando à descoberta dos mutantes Ato2 e Ato3 como dois novos transportadores de lactato. As características estruturais destas proteínas foram avaliadas por modelação de estruturas 3D e estudos de docking molecular.

PALAVRAS-CHAVE: ácidos carboxílicos; biotecnologia; microfábricas celulares; transportadores de membrana

Table of Contents

Acknowledgments ... III
Statement of integrity ... IV
Abstract ... V
Resumo ... VI
Table of contents ... VII
List of figures ... X
List of tables ... XIV
Acronyms ... XVII

Thesis outline ... 1
Chapter I – General introduction ... 7
1.1 Introduction ... 8
1.1.1 Transporter proteins ... 9
1.1.2 Carboxylic acids ... 11
1.1.2.1 Yeast carboxylic acid transporters ... 12
1.1.3 Transporter proteins in the biotechnological industry ... 12
1.1.4 Challenges and Perspectives ... 15
1.2 Aims of the work ... 16
1.3 References ... 17

Chapter II – Expanding the knowledge on the skillful yeast *Cyberlindnera jadinii* ... 18
2.1 Introduction ... 21
2.2 Ecology, taxonomy and evolution ... 22
2.3 Life cycle and genome organization ... 26
2.4 Morphology and physiology ... 28
2.5 Molecular tools for *C. jadinii* manipulation ... 30
2.6 Emerging biotechnological applications ... 31
2.6.1 Therapeutic applications ... 31
2.6.2 Bioproduction of valuable compounds using cost-effective carbon sources ... 32
2.6.3 Industrial applications - a patent-view ... 33
2.7 Membrane transporters characterized in *C. jadinii* ... 34
2.8 Conclusions and future perspectives ... 37
2.9 References ... 38

Chapter III – Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications ... 48
3.1 Introduction ... 51
3.2 Microbial cell factories in the production of bio-based organic acids. ... 51

3.2.1 Membrane transporters as tools for the improvement of cell factories 52
3.3 Membrane transporters in the import of renewable substrates 52
3.4 Transporter expression for the optimization of organic acid efflux 55
3.5 Engineering membrane transporters 65
3.5.1 Engineering sugar transporters 66
3.5.2 Engineering organic acid transporters 66
3.6 Future perspectives for transport engineering 67
3.7 References 69

Chapter IV – Bioprospection of non-conventional yeasts - Identification and Genomic analysis 78
4.1 Introduction 81
4.2 Materials and methods 82
4.3 Results 85
4.4 Discussion 96
4.5 Conclusions 98
4.6 References 101

Chapter V – Combined *in silico* approaches to uncover the plasma membrane carboxylate transporters in *Cyberlindnera jadinii* 107
5.1 Introduction 110
5.2 Materials and methods 111
5.3 Results 119
5.4 Discussion 143
5.5 Conclusions 148
5.6 References 150

Chapter VI – Phylogenetical, functional, and structural analysis of the CexA citrate exporter from *Aspergillus niger* 197
6.1 Introduction 200
6.2 Materials and methods 201
6.3 Results 207
6.4 Discussion 217
6.5 Conclusions 220
6.6 References 221

Chapter VII – The Acetate Uptake Transporter Family motif "NPAPLGL(M/S)" is Essential for Substrate Uptake 233
7.1 Introduction 236
7.2 Materials and methods 237

7.3 Results 242
7.4 Discussion 252
7.5 References 254

Chapter VIII – Evolutionary engineering reveals amino acid substitutions in Ato2 and Ato3 that allow improved growth of *Saccharomyces cerevisiae* on lactic Acid 257
8.1 Introduction 260
8.2 Materials and methods 261
8.3 Results 268
8.4 Discussion 274
8.5 References 276

Chapter IX – General discussion & Future perspectives 294
Discussion 295
Final remarks and future steps 298
References 300

LIST OF FIGURES

CHAPTER I

Figure 1. Membrane transport processes. 10

Figure 2. Schematic drawing of a microbial cell showing plasma membrane transporters involved in (1) substrate consumption and (2) metabolite export through the cell membrane. 13

Figure 3. SWOT analysis for biorefineries. 15

CHAPTER II

Figure 1. Evolutionary relationship of *Cyberlindnera jadinii*, a member of the Phaffomycetaceae clade..... 25

Figure 2. *C. jadinii* DSM 2361 morphological traits..... 29

Figure 3. Plasma membrane transporters functionally characterized in *Cyberlindnera jadinii*..... 36

CHAPTER III

Figure 1. Annual production of organic acids according to market sector/application in 2015 and estimated growth for 2024. 51

Figure 2. The expression of endogenous or exogenous membrane transporter genes in engineered bacteria, yeast and filamentous fungi, allows the uptake of renewable substrates, as well as the export and extracellular accumulation of specialty organic acids..... 69

CHAPTER IV

Figure 1. Growth pattern of ten wild yeast isolates..... 86

Figure 2. Colony morphology of the eight yeast species grown on solid media after 3 days of growth at 30°C. 88

Figure 3. Morphology and micromorphology of the wild yeast species. 90

Figure 4. Characterization of *Cyberlindnera jadinii* strains DSM 2361, PYCC 3092, PYCC 2578 and isolate I7. 91

Figure 5. Genetic diversity obtained for seven *Cyberlindnera jadinii* strains as disclosed by the principal component analysis (PCA) plot. 93

Figure 6. Functional annotation via EggNOG software of *Cyberlindnera jadinii* and *Saccharomyces cerevisiae* genes..... 95

Figure S1. Helmet-shaped ascospores released from deliquescent asci observed for *C. jadinii* wild isolate (I7). 99
Figures S2. Genetic diversity obtained for the seven *Cyberlindnera jadinii* strains as disclosed by the PCA plot of A) PC-2 vs PC-3 and B) PC-1 vs PC-3 components. 100

CHAPTER V

Figure 1. Initial uptake rates of labelled carboxylic acids as a function of the acid concentration of *C. jadinii* DSM 2361 cells in the presence of (A)^{14}C-acetic acid, pH 6.0; (B)^{14}C-lactic acid, pH 5.0; (C)^{14}C-succinic acid, pH 5.0 and (D)^{14}C-citric acid, pH 5.5, at 30°C. 120
Figure 2. Identification of putative membrane proteins encoded in the *C. jadinii* NRRL Y-1542 genome..... 122
Figure 3. Functional analysis of *C. jadinni* putative carboxylate transporters in *S. cerevisiae* IMX1000..... 124
Figure 4. Transport of radiolabelled carboxylates in *S. cerevisiae jen1*Δ *ato1*Δ (A,B) and IMX1000 (C) cells expressing the *C. jadinii* CjAto1-5, CjJen1-6, CjSlc16, CjSlc5, CjTDT, CjSlc13-1, CjSlc13-2 proteins and empty vector (pϕ), as negative control. 126
Figure 5. Kinetic parameters for the initial uptake rates of labelled carboxylates of *S. cerevisiae* expressing CjAto1-4 and CjJen1-6. 128
Figure 6. Citrate transport in *S. cerevisiae* IMX1000 cells expressing *C. jadinii* putative transporters..... 129
Figure 7. Maximum likelihood phylogenetic tree of AceTr family members (TCDB 2.A.96) present in ascomycetes. 131
Figure 8. Maximum likelihood phylogenetic tree of SHS family members (TCDB 2.A.1.12)..... 132
Figure 9. Maximum likelihood phylogenetic tree of SSS (SLC5 homolog) family members (TCDB 2.A.21)..... 134
Figure 10. Maximum likelihood phylogenetic tree of DASS (SLC13 homolog) family members (TC 2.A.47)..... 136
Figure 11. Predicted 3D structure of ScAto1, CjAto2, CjAto5 and CjAto6 transporters. 138
Figure 12. Predicted 3D-structure of ScJen1, CaJen2, CjJen5 and CjJen6..... 140
Figure 13. Predicted 3D structure of the *Homo sapiens* SLC5 member 8 and *C. jadinii* SLC5 homolog..... 141
Figure 14. Predicted 3D structure of the *Homo sapiens* SLC13 member 3 and CjSLC13 homolog.... 142

Figure 15. *C. jadinii* plasma membrane transporters functionally characterized within this work by heterologous expression in *S. cerevisiae* strains. 148
Figures S1. Multiple-sequence alignments..... 178
Figure S2. Topology prediction..... 182
Figures S3. Phylogeny data belonging to AceTr, SHS, SSS and DASS transporter families..... 184
Figures S4. Docking studies and pore radius trajectories..... 195

CHAPTER VI

Figure 1. Maximum likelihood phylogenetic tree of DHA1 family (TCDB 2.A.1.2) members present in eukaryotic and prokaryotic genomes..... 208
Figure 2. Growth phenotypes of *S. cerevisiae* CEN.PK 113-5D strains expressing *Aspergillus niger* CexA transporter under the control of TEF1 and TPI1 promoters..... 210
Figure 3. Uptake of ^{14}C-citric acid 10.0 mM in glucose-grown cells of *S. cerevisiae* CEN.PK 113-5D transformed with pAnCexA-TEF1..... 211
Figure 4. Citric acid transport in *S. cerevisiae* CEN.PK 113-5D cells expressing AnCexA under the control of TEF1..... 212
Figure 5. Specificity and energetics of citric acid uptake at pH 5.5, 30°C in *S. cerevisiae* CEN.PK 113-5D cells expressing AnCexA under the control of TEF1..... 212
Figure 6. Structure-function analysis of the *Aspergillus niger* CexAp..... 214
Figure 7. Structure-function study of AnCexA..... 215
Figure 8. Structural-functional analysis of AnCexA..... 216
Figure S1. Detailed AnCexA-phylogenetic tree with all homolog nominations and bootstrap values..... 226
Figure S2. Growth curve of *S. cerevisiae* CEN.PK 113-5D strain expressing pAnCexA-TEF..... 231

CHAPTER VII

Figure 1. Maximum likelihood phylogenetic tree of AceTr family (TCDB 2.A.96) present in eukaryotic and prokaryotic genomes..... 245
Figure 2. A) Initial uptake rates of radiolabeled ^{14}C-acetic acid at different concentrations by *S. cerevisiae* W303-1A *jen1*Δ *ady2*Δ cells heterologously expressing p*Gpr1*, p*Ady2* and p416GPD as a control at pH 6.0, 30°C and respective kinetic parameters..... 246

Figure 3. Acetic acid kinetics and transport energetics of *E. coli 3*Δ heterologously expressing pAceP..... 248
Figure 4. A) The percentage of 0.1 mM ^{14}C-acetic acid uptake, at pH 6.0, considering the velocity of transport found for the *S. cerevisiae ady2*Δ *jen1*Δ strains expressing pAdy2 as 100%..... 249
Figure 5. Epifluorescence microscopy localization of Ady2::GFP expressed in *S. cerevisiae ady2*Δ *jen1*Δ..... 250
Figure 6. Molecular docking of SatP with acetate and succinate..... 251

CHAPTER VIII

Figure 1. Growth rates on different carbon sources of *S. cerevisiae* reference strain CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty vector or a vector carrying the indicated organic acid transporter. 270
Figure 2. Multiple sequence alignment of *Citrobacter koseri* SatP and *Saccharomyces cerevisiae* Ady2, Ato2 and Ato3. 272
Figure 3. 3D-Models of the transporters Ato3 (dark blue) and Ato3 F95S (light blue). 273
Supplementary Figure 1. Growth of different strains on SM media with lactic acid as the sole carbon source..... 284
Supplementary Figure 2. Growth rates on SMD of *S. cerevisiae* reference strain CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty multicopy vector or a multicopy vector containing the indicated organic acid transporter gene..... 284
Supplementary Figure 3. Growth rates of *S. cerevisiae* reference strain CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty centromeric vector or a centromeric vectors containing the indicated organic acid transporter gene..... 285
Supplementary Figure 4. Growth profiles in synthetic medium (pH 5.0) with glucose as the sole carbon source of CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty multicopy vector or a multicopy vector containing the indicated organic acid transporter gene..... 285
Supplementary Figure 5. Growth profiles in synthetic medium (pH 5.0) with lactate as the sole carbon source of CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty multicopy vector or a multicopy vector containing the indicated organic acid transporter gene. 286

Supplementary Figure 6. Growth profiles in synthetic medium (pH 6.0) with acetate as the sole carbon source of CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty multicopy vector or a multicopy vector containing the indicated organic acid transporter gene. 287
Supplementary Figure 7. Growth profiles in synthetic medium (pH 5) with pyruvate as the sole carbon source of CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty multicopy vector or a multicopy vector containing the indicated organic acid transporter gene. 288
Supplementary Figure 8. 3D model of Ady2, Ady2 C755G and Ady2 C655G alleles. 289
Supplementary Figure 9. 3D model of Ato2 and Ato2 L218S allele..... 290
Supplementary figure 10. Molecular docking sites of acetate (green ligand), lactate (blue ligand) and pyruvate (orange ligand) in the predicted structure of Ady2, Ato2 and Ato3, identified using Autodock Vina..... 292

LIST OF TABLES

CHAPTER I

Table 1. Bioproduction of organic acids at the industrial level..... 14

CHAPTER II

Table 1. Main *C. jadinii* (former *C. utilis*) strains described in literature..... 24
Table 2. Genomic features of three *C. jadinii* strains and the reference *S. cerevisiae* strain S288c.....27

CHAPTER III

Table 1. Microbial organic acid transporter proteins (experimentally verified). 57

CHAPTER IV

Table 1. Identification of the wild yeast isolates by molecular typing using internal transcribed sequencing (ITS) 87
Table 2. Data characteristics and sequencing statistics from the de-novo assembly of *C. jadinii* wild and collection genomes..... 92
Table S1. Comparison for quality validation over the alignment of sequenced strains deposited in NCBI against the Cj-strains in analysis by different parameter setting in variant calling..... 99

Table S2. Predicted proteins by AUGUSTUS versus functionally annotated proteins by Cluster of Orthologous Groups (COG)..... 99

CHAPTER V

Table 1. Yeast strains used in this study. 113
Table 2. Plasmids used in this study. 114
Table 3. Oligonucleotides used for strain construction, cloning and expression..... 116
Table 4. Kinetic parameters over the ^{14}C-acetic acid (pH 6.0) transport estimated in cells of *S. cerevisiae* W303-1A *jen1*Δ *ato1*Δ expressing CjAto1-4 and CjJen1-6. 127
Table S1. Percentage of shared amino acids between A) Ato1 and B) Jen1 characterized homologs of *C. jadinii, Y. lipolytica, S. cerevisiae, C. albicans, A. nidulans, K. lactis, D. hansenii, E. coli and M. acetivorans*..... 156
Table S2. Proteome annotation of identified groups of proteins in *C. jadinii* NRRL Y-1542 genome using a bioinformatic pipeline developed for the extraction of relevant homologs of a given gene..... 157
Table S3. Docking parameters and data from the molecular docking studies..... 172

CHAPTER VI

Table 1. Yeast strains used in this study..... 203
Table 2. Plasmids used in this study..... 204
Table 3. Oligonucleotides used for strain construction, cloning and expression..... 206
Table S1. AnCexA residues predicted to present strong intramolecular interactions with citrate by molecular docking studies. 232
Table S2. Average of the binding affinity values [kcal/mol] calculated with PyRx software for the docking of AnCexA with citrate. 232

CHAPTER VII

Table 1. List of plasmids used in this study..... 238
Table 2. List of strains used in this study..... 239
Table 3. Oligonucleotides used for strain construction, cloning, expression, site-directed mutagenesis and GFP tagging..... 241

CHAPTER VIII

Table 1. Plasmids used in this study..... 263

Table 2. *Saccharomyces cerevisiae* strains used in this study 265

Table 3. Amino acid changes identified by whole-genome sequencing of single colony isolates evolved for growth in medium containing lactic acid as sole carbon source..... 269

Table 4. Average distances (in Å) between different amino acids in the constriction pore of Ady2, Ato2 Ato3 and mutated alleles, calculated using the corresponding protein model..... 273

Supplementary Table 1. Primers used in this study..... 281

Supplementary Table 2. Plasmid construction. Each plasmid was constructed via Gibson assembly of two linear DNA fragments..... 283

Supplementary Table 3. Average of the binding affinity values [kcal/mol] calculated with PyRx software for the docking of ligand in the predicted structures of wildtype and mutated Ady2, Ato2 and Ato3..... 291

Supplementary Table 4. Residues identified by molecular docking analysis as being involved in the establishment of strong interactions with the indicated ligand..... 293

ACRONYMS

3HP – 3-Hidroxylpropionic

ABC – ATP-binding cassette

amp – ampicillin

ANOVA – one-way analysis of variance

ATP – Adenosine triphosphate

BLAST – Basic Local Alignment Research Tool

bp – Base pairs

CA – Carboxylic acid

CGD – *Candida* Genome Database

CCCP – Carbonyl cyanide m-chlorophenyl hydrazine

Clustal – Cluster analysis of pairwise alignments

DNA – Deoxyribonucleic acid

GFP – green fluorescent protein

H^+ – proton

HPLC – High performance liquid chromatography

IUBMB – International Union of Biochemistry and Molecular Biology

LB – Luria-Bertani

LOMETS – Local Meta-Threading-Server

MCT – monocarboxylate transporter

MM – Minimal Media

MEGA7 – Molecular Evolutionary Genetics Analysis version 7.0

MolD – Molecular Docking

mRNA – Messenger ribonucleic acid

NCBI – National Center for Biotechnology Information

OD – Optical density

PCR – Polymerase chain reaction

PDB – Protein Database

PDR – Pleiotropic Drug Resistance

RMSD – Root-mean-square deviation

RNA – Ribonucleic acid

SGD – *Saccharomyces* Genome Database

TC – Transporter classification system

TCA – Tricarboxylic acid

TCDB – Transport Classification Database

TMHMM – Tied Mixture Hidden Markov Model

TMS – Transmembrane segments

WGD – Whole Genome Duplication

WT – Wild type

YAC – yeast artificial vectors

YNB – Yeast nitrogen base

Microorganisms abbreviations

An – Aspergillus niger

Ani – Aspergillus nidulans

Ca – Candida albicans

Cu – Candida utilis

Cg – Candida glabrata

Cj – Cyberlindnera jadinii

Dh – Debaryomyces hansenii

Ec – Escherichia coli

Sc – Saccharomyces cerevisiae

Yl – Yarrowia lipolytica

Constants and Units

Ci – Curie

Dpm – Disintegrations per minute

dry wt – Dry weight

g – Gravity acceleration

h – Hour

K_d – Diffusion constant

K_i – Inhibition constant

K_m – Michaelis Menten constant/affinity constant

L – Liter

M – Molar

min – Minute

°C – Degrees Celsius

pK_a – Acid dissociation constant

rpm – Revolutions per minute

s – Second

SD – Standard deviation

v; vol. – Volume

V_{max} – Maximum velocity

w – Weight

Transporter families

2-HCT – 2-Hydroxycarboxylate Transporter

AAEx – Aspartate:Alanine Exchanger

AceTr – Acetate Uptake Transporter

Bestrophin – Anion Channel-forming Bestrophin

CitMHS – Citrate-Mg^{2+}:H^+ (CitM) Citrate-Ca^{2+}:H^+ (CitH) Symporter

DAACS – Dicarboxylate/Amino Acid:Cation (Na^+ or H^+) Symporter

DASS – Divalent Anion:Na^+ Symporter

Dcu – C4-Dicarboxylate Uptake

DcuC – C4-dicarboxylate Uptake C

DHA1 – Drug:H^+ Antiporter-1

FNT – Formate-Nitrite Transporter

LctP – Lactate Permease

MFS – Major Facilitator Superfamily

MIP – Major Intrinsic Protein

MHS – Metabolite:H^+ Symporter

NhaC – Na^+:H^+ Antiporter

SLC – Solute carrier protein

SHS – Sialate:H^+ Symporter

SSS – Solute:Sodium Symporter

SP – Sugar porter

SulP – Sulfate Permease

TDT – Telurite-resistance/Dicarboxylate Transporter

TRAP-T – Tripartite ATP-independent Periplasmic Transporter

TTT – Tripartite Tricarboxylate Transporter

Amino acids

A	Ala	Alanine	G	Gly	Glycine
M	Met	Methionine	S	Ser	Serine
C	Cys	Cysteine	H	His	Histidine
N	Asn	Asparagine	T	Thr	Threonine
D	Asp	Aspartic acid	I	Ile	Isoleucine
P	Pro	Proline	V	Val	Valine
E	Glu	Glutamic acid	K	Lys	Lysine
Q	Gln	Glutamine	W	Trp	Tryptophan
F	Phe	Phenylalanine	L	Leu	Leucine
R	Arg	Arginine	Y	Tyr	Tyrosine

Thesis Outline

Thesis overview & Outputs

THESIS OVERVIEW

Several strategies are currently being developed for the improvement of microbial cell factories towards the production of fine chemicals by more eco-friendly and sustainable processes as an alternative to petrochemical-based ones. One of these strategies relies on the expression of cell membrane transporters, importers or exporters, and also transporters that mediate the translocation of metabolites between cell compartments. This strategy aims to improve the range and efficiency of substrate consumption, improve cell metabolic fluxes, and avoid intracellular product toxicity that consequently decreased productivities in microbial cell factories. Organic acids, such as carboxylic acids, are among the top platform chemicals due to their wide applicability. Their biosynthesis in high yields demands, aside from the engineering of the cell metabolic and enzymatic machinery, effective plasma membrane transporters able to export the acids to the extracellular media to avoid cell toxicity and further associated costs of downstream processing. In this context, increased knowledge on carboxylic acid transporters will promote the development of efficient biorefineries for the bioproduction of these acids. This work aims to expand the current knowledge on eukaryotic carboxylate transporters focusing on the identification of new carboxylate transporters, their functional characterization, and the study of structural features.
Overall, the thesis is structured into nine chapters.

Chapter I provides a general introduction to the topic of the thesis as well as the main goals of this work. The introduction describes promising microorganisms that can be used as efficient microbial cell factories, discloses the existing cell transporter systems, and addresses their role in the improvement of cell robustness. A special focus is given to carboxylic acid transporters along with their relevance in biorefineries. The status of the carboxylates market is also highlighted and the companies working on the bioproduction of organic acids at the industrial and pilot level.

Chapter II presents a first authorship published review entitled "Expanding the knowledge on the skillful yeast *Cyberlindnera jadinii*", which describes the specific features that turn this yeast attractive to the biotechnological industry. A robust phylogenetic analysis was also done in this study revealing its evolutionary position. Moreover, the relevant traits of its physiology are also described together with the solute membrane transporters so far characterized.

Chapter III includes a bibliographic revision entitled "Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications". This review focuses on microbial-based production of chemical building blocks, namely organic acids, as an alternative to chemical synthesis from petroleum derivatives. It also provides an overview of the different strategies employed for the import of renewable substrates and the optimization of organic acid efflux. Finally, the future perspectives for transporter engineering are presented, to increase the productivity and sustainability of microbial cell factories.

Chapter IV describes the bioprospection of wild yeast species for the utilization and production of carboxylic acids. The identification and characterization steps included the evaluation of morpho- and physiological traits, coupled with the evaluation of growth on organic acids as unique carbon and energy sources. The species *Cyberlindnera jadinii* were selected for further comparative genomic analysis due to its strong growth on carboxylates, presenting distinct carboxylic acids transporters whose encoding genes were still unidentified.

Chapter V presents a combination of *in silico* and genetic approaches for the study of the *Cyberlindnera jadinii* transportome. In this work, novel carboxylate transporters from *C. jadinii* were uncovered and functionally characterized by heterologous expression in the *Saccharomyces cerevisiae* IMX1000 strain (deleted in 25-genes related with carboxylate transport). Sixteen carboxylate transporters belonging to the AceTr, SHS, SSS, TDT, DASS and MCT transporter families were characterized revealing different specificities for mono-, di- and tricarboxylates. *In silico* 3D-model prediction by homology threading and docking studies indicate possible structural aspects responsible for the specificity of the distinct transporters found. In addition, phylogenetic studies revealed the evolutionary relationship between the transporters belonging to AceTr, SHS, SSS and DASS transporter families.

Chapter VI depicts the functional characterization of CexA, the recently identified citrate exporter from *Aspergillus niger*, by heterologous expression in *S. cerevisiae*. The kinetic parameters and energetics of this transporter were evaluated through the measurement of the uptake of radiolabelled substrates. AnCexA is specific for citrate and isocitrate, and the uptake of citrate is not dependent on the proton motive force. The identification of CexA homologs and subsequent phylogenetic analysis allowed the identification of conserved residues within the DHA1 transporter family. Molecular docking analysis

uncovered putative binding sites involved in substrate recognition. A site-directed mutagenesis approach was designed to uncover the structure-function relationships of these transporters.

Chapter VII presents the work published in the manuscript entitled "The Acetate Uptake Transporter family motif "NPAPLGL(M/S)" is essential for substrate uptake and binding". This study describes the functional role of the motif NPAPLGL(M/S), conserved in AceTr family members. Also, other members of the AceTr family were functionally characterized as acetate transporters in this study, namely the Gpr1 from *Yarrowia lipolytica* and the Acep from the archea *Methanosarcina acetivorans*. The phylogenetic analysis of this family suggests that these transporters play a key role in fungi since they are present in 97% of the fungal genomes analyzed.

Chapter VIII presents the work "Evolutionary engineering reveals amino acid substitutions in Ato2 and Ato3 that allow improved growth of *Saccharomyces cerevisiae* on lactic acid", where monocarboxylic acid transporters were evaluated via a combination of directed evolution, whole-genome resequencing, and reverse engineering. In this study mutated versions of Ato1, Ato2 and Ato3 transporters showed to efficiently catalyze lactic acid uptake in *S. cerevisiae*. Growth on a variety of organic acids was evaluated for cells individually overexpressing the repertoire of native and evolved lactic acid transporters. The effect of the different mutations in the Ato1, Ato2, and Ato3 structure, was predicted using 3D-structural models combined with a molecular docking analysis.

The **Chapter IX** presents the main conclusions and future perspectives of this work together with a general discussion of the results and the main outcomes of this thesis.

This thesis was developed in the Molecular and Environmental Biology Centre (CBMA), Department of Biology of the University of Minho (Portugal), supervised by Doctor Isabel João Soares-Silva and Professor Doctor Margarida Casal, under the scope of the Doctoral Program on Applied and Environmental Microbiology (Norte 2020, FCT). A scientific collaboration was established with Professor Robert Mans from the Technische Universiteit Delft (TU Delft, Netherlands), which was devoted to the work presented in chapter VIII.

OUTPUTS

The present dissertation resulted in the publication of four scientific papers, one manuscript is under review and three others are in preparation for submission, as described below:

a) Peer-reviewed scientific publications:

Ribas, D., Soares-Silva, I., Vieira, D., Sousa-Silva, M., Sá-Pessoa, J., Azevedo-Silva, J., Viegas, S., Arraiano, S., Diallinas, G., Paiva, S., Soares, P., Casal, M. (2019). The Acetate Uptake Transporter family motif "NPAPLGL(M/S)" is essential for substrate uptake and binding. *Fungal Genetics and Biology*, 122, 1-10.

Soares-Silva, I., Ribas D., Sousa-Silva, M., Azevedo-Silva, J., Rendulic, T., Casal, M. (2020). Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications. *FEMS Microbiology Letters*, *367*(15), fnaa118.

* Alves, R., Sousa-Silva, M., Vieira, D., Soares, P., Chebaro, Y, Lorenz, M.C., Casal, M., Soares-Silva, I., Paiva, S. (2020) Carboxylic acid transporters in *Candida* pathogenesis. *MBio, 11*(3).

Sousa-Silva, M., Vieira, D., Soares, P., Casal, M., Soares-Silva, I. (2021) Expanding the Knowledge on the Skillful Yeast *Cyberlindnera jadinii*. *Journal of Fungi*, *7*(1):36.

* Peer-reviewed scientific publication not included in this thesis.

b) Submitted publication:

De Valk, S., Baldi, N., Sousa-Silva, M., Casal, M., Soares-Silva, I., Mans, R. (2021). Evolutionary engineering reveals amino acid substitutions in Ato2 and Ato3 that allow improved growth of *Saccharomyces cerevisiae* on lactic acid.
Submitted to FEMS Yeast Research.

c) Articles in preparation for submission:

Combined *in silico* approaches to uncover the plasma membrane carboxylate transporters in *Cyberlindnera jadinii*
Manuscript in preparation.

Bioprospection of non-conventional yeasts – Identification and genomic analysis
Manuscript in preparation.

Phylogenetical, functional, and structural analysis of the CexA citrate exporter from *Aspergillus niger*
Manuscript in preparation.

Internship abroad:

Visiting student of Prof. Robert Mans lab at the Technische Universiteit Delft (TU Delft), Delft, Netherlands.

This internship was suspended due to the covid-19 pandemic - February to March 2020

CHAPTER I

General Introduction

1.1 INTRODUCTION

Presently there is an increasing demand for the development of more sustainable processes of fine chemical production to replace petrochemical-derived ones. Compounds such as aromatic or organic acids, which include carboxylic and sugar acids, present wide applicability in a vast range of industries [1-3]. Engineered microorganisms are emerging as attractive producers of added-value products through cutting-edge technologies and highly integrative eco-friendly approaches [1, 3-5].

To tackle the diversity of metabolites needed and reach productivity levels required to implement the production at the industrial scale, it becomes necessary to explore new expression hosts that can become more efficient in their production. Few yeasts combine the ability to produce specific metabolites with other biotechnological advantages, namely biological safety, high growth rates in simple media, simple genetic modification, and efficient secretion systems [6].

Over the years, *Saccharomyces cerevisiae* has been the preferred yeast given that it was the first eukaryotic organism having its genome fully sequenced, which in turn promoted great advances in genetic engineering. Additionally, this yeast has great tolerance to stress conditions, such as low pH values and high sugar concentrations, and natural adaptability to the harsh industrial-scale conditions with well-established fermentation [7-10]. Nonetheless, it has its drawbacks, for instance, the predisposition for the occurrence of N-hyperglycosylation of produced proteins that reduces the efficiency of protein secretion [8, 11, 12] and its metabolism being preferably fermentative, which limits biomass propagation [13].

The physiological characteristics of other non-*Saccharomyces* yeast species like *Pichia pastoris* (*Komagataella pastoris*), *Pichia kudriavzevii (Issatchenkia orientalis), Kluyveromyces marxianus*, *Yarrowia lipolytica,* or *Cyberlindnera jadinii*, make them alternative hosts for biotechnological purposes [10, 14-17]. *K. pastoris* can grow to very high cell density and is a well-established platform for the heterologous expression of proteins from a diverse range of organisms [18-20]. The species *P. kudriavzevii* exhibits tolerance to multi-stress, e.g. low pH, high temperatures, high salt concentration [21]. Several studies have also reported its potential for organic acid production (e.g. xylonic acid, lactic acid, and succinic acid) [22-24]. Over the years, *C. jadinii* has been used as a source of single-cell protein and recognized for its ability to synthesize a variety of valuable compounds for the food and pharmaceutical sectors [14]. This yeast is also capable of efficiently assimilate xylose, the main constituent of hemicellulose, an efficient and renewable energy source [25]. Contrarily to *S. cerevisiae*, these three yeast species are Crabtree-negative microorganisms being capable of utilizing substrates more efficiently under aerobic conditions, without the production of ethanol, resulting in higher biomass formation [6, 26]. However, for

some species, like *C. jadinii*, their use as expression hosts was delayed due to the lack of efficient molecular tools for genetic manipulation, such as suitable selection markers and expression cassettes [25].

In order to achieve efficient and cost-effective biobased industrial processes, it is important, among other aspects, to work with robust microbial cell factories with improved metabolic performances. Engineering the microbes' enzymatic machinery, as well as its membrane transporters, are important steps to promote the efficient utilization of specific substrates and the efflux of the produced metabolites to the extracellular medium [27]. This last step will improve the host's tolerance to the final product, e.g. organic acids, besides decreasing downstream processing costs [28]. Thus the identification of new cell membrane transporters, such as importers or exporters, or transporters that are more efficient than the ones already present so far in microbial cell factories, is a crucial step to increase cell robustness [28, 29].

1.1.1 TRANSPORTER PROTEINS

The plasma membrane controls and regulates the exchange of molecules such as amino acids, lipids, sugars, organic acids, ions among other compounds. As a permeable barrier, biological cellular membranes embrace an intrinsic selectivity function ensuring the maintenance of a constant internal environment despite the variations of the external environment [30]. Besides playing a crucial role in the maintenance of essential metabolites for cell growth, it is also involved in the secretion of waste compounds produced by the cell [31, 32]. Considering the overall cellular transmembrane transport processes, the majority is mediated by integral membrane proteins. Membrane embedded proteins possess one or more segments crossing the plasma membrane. The coupling of these proteins to the membrane occurs through residues with hydrophobic side chains that are linked to fatty acyl groups from membrane phospholipids [33]. Membrane proteins play a crucial role in many critical biological processes such as cell physiology, transmembrane signaling, trafficking, transport of substrates, adhesion, and recognition. Genome-wide scale studies revealed that MPs compose approximately 20 to 30 % of all genes present in the genome of prokaryotic and eukaryotic organisms [34].

Membrane transporters are vital for cell homeostasis promoting substrate uptake, regulation of metabolite concentration between organelles, and efflux of toxic compounds produced intracellularly [35]. The different mechanisms of transport across biological membranes are represented in Figure 1.

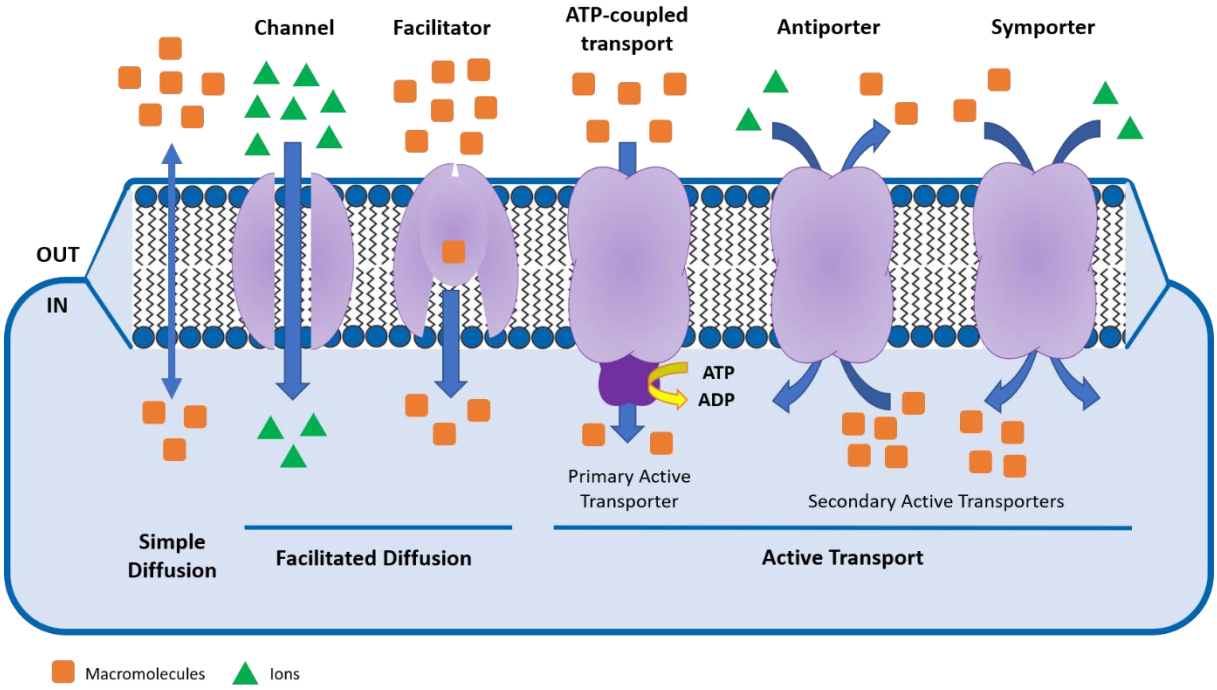


Figure 1. Membrane transport processes. Schematic representation of the several mechanisms involved in the transport of solutes across biological membranes. Initials stand for IN – intracellular space; OUT – extracellular space. Symbols depicted as triangles correspond to charge and squares to uncharged molecules.

Some compounds are translocated through a simple diffusion mechanism, crossing the hydrophobic phospholipid bilayer according to their electrochemical gradient. Such is the case of gases, like CO_2 and O_2, small neutral and polar molecules, like H_2O and ethanol, and small hydrophobic solutes such as benzene [32]. Mediated transport occurs for charged and other polar molecules that require a transporter to cross the biological membrane [36].

A facilitated diffusion mechanism occurs for molecules that cross the membrane according to concentration gradient through channels or facilitators and do not require any external source of energy [37, 38]. For charged compounds, the movement is also dependent on the membrane electric potential. Active transport mechanisms require energy. Two types of active transport systems exist, the primary active transport that depends on the recruitment of energy from metabolic processes, e.g. ATP hydrolysis or respiration events, and the secondary active transport that is coupled to the existence of electrochemical gradients of ions or protons [37]. According to the translocation of cargo molecules, secondary active transport can function as (i) an antiporter, when cargo (substrates and ions) cross the membrane in opposite directions, or (ii) a symporter, when two substrates (or substrate and ion) pass the membrane in the same direction [37, 39]. At the specificity and structural levels, it is known that

facilitators exhibit stereospecific substrate specificities and can function as monomeric proteins, whereas a great part of channels is organized in oligomeric structures [35].

Localized at the plasma membrane or in the membrane of organelles, membrane transporters are crucial components intervening in biochemical pathways [40]. The determination of its high-resolution three-dimensional (3D) structures is fundamental for a better understanding of many biological processes [41]. The characterization of these proteins is based on protein topology, α-helices or β-barrels, number of transmembrane segments (TMS), location of the N- and C-terminal ends, and subcellular location [42, 43]. Several techniques are available for structure characterization (i) crystallographic, such as X-ray crystallography, electron microscopy, and nuclear magnetic resonance spectroscopy methods, (ii) homology-based for comparative protein modelling, homology modelling or protein threading methods, and (iii) the study of structure-function relationships of protein transporters [44-46]. The latter approach uses functional analysis of mutated templates of proteins. This has become a valuable tool for examining transporter structures in the cellular environment as well as a tool devoted to *in vitro* analysis [47].

1.1.2 CARBOXYLIC ACIDS

Carboxylic acids (CAs) are ubiquitous compounds that can be used by microorganisms as sole carbon and energy sources or final products or by-products of fermentative processes. Additionally, they are platform chemicals constituting primary building blocks commonly used in food and chemical industries [37, 48, 49]. CAs are described as organic acids with a carboxyl group associated (formula R-COOH), being R the monovalent functional group. They can be classified in mono-(1-COOH), di- (2–COOH), tri- (3–COOH) as well tetra-carboxylic acids (4-COOH) depending on the number of carboxyl groups associated with the short-chain fatty acids compound. Following deprotonation of the carboxyl group, the conjugate base turns into a carboxylate anion [36, 48]. As weak acids, CAs can dissociate partially into protons and carboxylates in solution. This level of dissociation is strongly dependent on the pK_a and pH of the aqueous solution. At a pH value below the pK_a of the acid, CAs are present predominantly on the undissociated form (R-COOH) as lipid-soluble substances, being able to cross the biological membrane by simple diffusion. At a pH value above the pK_a of the acid, the charged anionic form predominates, requiring a transporter to cross the biological membrane [36].

1.1.2.1 YEAST CARBOXYLIC ACID TRANSPORTERS

The first carboxylic acid transporter functionally characterized in yeast was the dicarboxylate transporter from *K. lactis* [50], later identified as being encoded by the Kl*Jen2* [51]. The dicarboxylate transporter

Mae1 from *Schizosaccharomyces pombe* was the first carboxylate transporter coding gene cloned in yeasts [52, 53]. The Mae1 transporter has 10 TMSs and belongs to the Tellurite-resistance/Dicarboxylate Transporter (TDT) Family (TC 2.A.16), transports malate and maleic, malonic, oxaloacetic, and succinic acids (C4 di-CAs) by a proton-symporter mechanism, being expressed in cells grown on glucose. Mae1 behaves both as an importer or an exporter, dependent on substrate and proton gradient over the plasma membrane [37, 52]. The first CA transporter system characterized in *S. cerevisiae* was a lactate-proton symporter shared by acetic, propionic, and pyruvic acids, repressed by glucose and inducible by the substrate [54, 55]. Since the permeability of the plasma membrane for lactic acid increases with the pH, the diffusion constant associated with the acid becomes higher following an increase of the extracellular pH [54]. The identification of *JEN1*, the gene encoding the lactate-proton symporter in *S. cerevisiae* was carried out by Casal's group [55]. Jen1 has 12 TMSs and is a member of the major facilitator superfamily (MFS) and the sialate:H^+ Symporter (SHS) family (TC 2.A.1.12.2), with specificity for lactate, pyruvate, acetate, propionate, and selenite [55, 56]. Later on, another carboxylate transporter was described by the same group, the Ady2 (TC 2.A.96.1.4), recently renamed as ATO1 (from acetate transporter ortholog) [57], that belongs to the Acetate Uptake Transporter (AceTr) family [58]. Ato1 has 283 amino acids arranged in 6 TMSs and is involved in the transport of acetate, formate, propionate, and lactate being also hypothesized its role on ammonia export [58, 59]. Two Ato1 homologs can be found in *S. cerevisiae*, Ato2 (former Fun 34), and Ato3, but their function is not solved [60, 61].

Several homologs of the Jen1 and Ato1 were functionally characterized as membrane transporters in other organisms, presenting a broad range of specificities, including the transport of mono-, di- and tricarboxylic acids [27, 62, 63].

1.1.3 TRANSPORTER PROTEINS IN THE BIOTECHNOLOGICAL INDUSTRY

Transporters play an important role in the biotechnology field. The expression of adequate transporters can improve substrate consumption rates and increase the range of substrates utilized by microbial cell factories [29, 64]. On the other hand, the export of newly synthesized CAs can decrease product toxicity and increase productivity levels. These are two relevant strategies for the improvement of biotechnological processes (Figure 2).

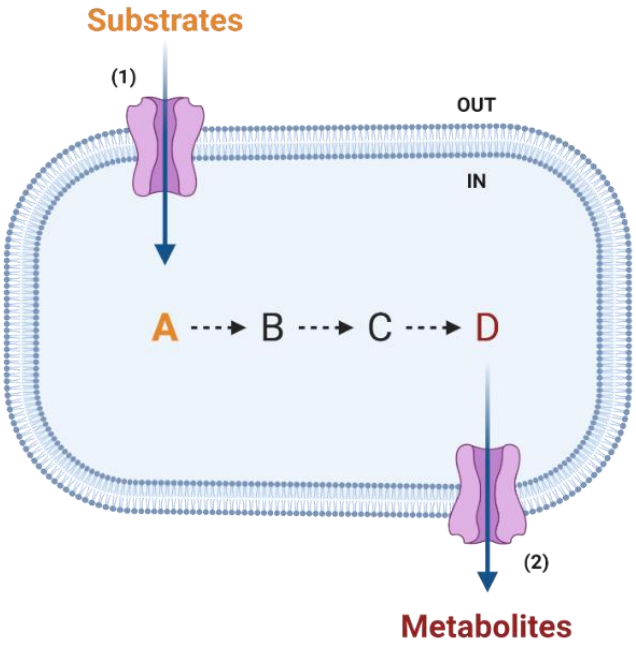


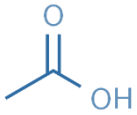
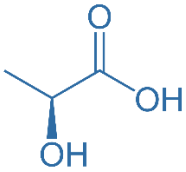
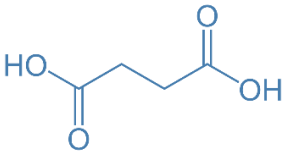
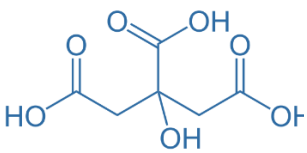
Figure 2. Schematic drawing of a microbial cell showing plasma membrane transporters involved in (1) substrate consumption and (2) metabolite export through the cell membrane. A to D represents the metabolic pathway, having A as a substrate D as a product. Initials stand for IN - intracellular space; OUT - extracellular space.

The status of the CAs market was reviewed by López-Garzón and Straathof team (2014) and recently by De Jong and colleagues (2020) [65-67]. Organic acids can be divided into short-chain fatty acids (acetic, pyruvic, propionic, and butyric), hydroxy acids (functional group with an alcohol and carboxyl group; lactic acid and 3-hydroxypropionic), di-CAs (fumaric, malic, succinic, muconic, and glutaric), tri-CAs (acrylic, citric, isocitric, and gluconic) [48, 64]. At the industrial level, some organic acids are being produced as bio-based products by some major corporations. Such is the case of lactic acid, produced by Cargill and NatureWorks, 3-hydroxypropionic acid by Cargill/Novozymes and OPX/Dow Chemical, succinic acid by Myriant, DSM and BASF, and 1,4-butanediol, synthesized from succinic acid by Genomatica [64, 66, 68]. Fumaric, malic, and succinic acids are some of the dicarboxylic acids naturally produced by microorganisms. These CAs can be generated from abundant renewable biomass having a value-added as they are used as building-block chemicals in industry [64]. Along with dicarboxylic acids, there is also an increasing interest in tricarboxylic acids given its worldwide market demand. For instance, the global market for citric acid was 2.39 million tons in 2020 being used as a precursor for the most varied applications [64, 69, 70]. This market is expected to grow in the next years, with a compound annual growth rate (CAGR) of 4.0% in the forecast period of 2021-2026 reaching a volume of 2.91 million tons by 2026 [70]. In the long-term, the ultimate goal of the biotechnical industry is to efficiently produce these compounds via industrial fermentations. Table 1 outlines the current state of bioproduction of carboxylic

acids, the currently commercialized, ones under industrial development together with ones on a pilot scale.

All the organic acids mentioned in table 1 correspond to a Technology Readiness Level (TRL) of 8 to 9 which means that part of their production is in a stage of a first-kind commercial system with their manufacturing issues solved (system qualified) or in a fully commercial application with the product already available for consumers [71]. The manipulation of metabolic processes, fermentation conditions such as pH, substrate, and other medium components, can increase the product yield as well as its purity [64]. As previously mentioned, CAs exporters are crucial elements for cell factory optimization and there is an increasing demand for specific carboxylic acid membrane transporters. Also, an increasing knowledge of the mechanisms underlying CAs transport is essential to achieve an efficient biological production of these carboxylates, suitable for industrial production levels [29].

Table 1. Bioproduction of organic acids at the industrial level. Overview of final products and microorganisms used as a platform for their production. Adapted from the EU commission report, 2015 [71], Chen and Nielsen, 2016 [68], and De Jong *et al.*, 2020 [66].

Carboxylic acids	Structural formula	Microorganism	Companies
Acetic acid [Mono-CA]		*Gluconobacter frateurii* *Acetobacter ocidentalis*	Wacker; Sekag; Godovari Biorefineries Ltd; Zeachem
Lactic acid [Mono-CA]		*Lactobacillus sp.*	Purac; Galactic; NatureWorks
		Saccharomyces cerevisiae *Pichia kudriavzevii*	VTT; Cargill; BASF; Myriant; Corbion; NatureWorks; Galactic, Henan Jindan; BBCA
Succinic acid [Di-CA]		*Saccharomyces cerevisiae* *Basfia succiniciproducens* *Escherichia coli* *Corynebacterium glutamicum*	DSM; Genomatica; BioAmber; Myriant; Succinity (joint venture from BASF/Corbion); Reverdia (Roquette)
Citric acid [Tri-CA]		*Aspergillus niger*	Cargill; DSM; BBCA; Ensign; TTCA; RZBC

Biorefineries promote the development of a more sustainable economy, withal there are still some drawbacks to its full implementation. A SWOT analysis is outlined in figure 3 showing the strengths, weaknesses, opportunities, and threats within this topic.

Figure 3. SWOT analysis for biorefineries. Combination of strengths, weaknesses, opportunities, and threats of using biobased production, adapted from De Jong *et al.*, 2020 [66] and Spekreijse *et al.*, 2019 [72].

1.1.4 Challenges and Perspectives

One of the major challenges of this work is to identify novel CAs exporters, as most of the transporters so far identified are CAs importers. As the export of these acids is a major limiting step for the production of CAs in the biorefineries, the identification of novel CAs exporters will certainly contribute to the development of this field. By using metabolic engineering tools allied to the expression of suitable

membrane transporters, new and more efficient strains can be developed for the production of bio-based carboxylates [64, 73]. Increased knowledge in carboxylic acid transport is thus necessary to develop novel tools for cell factory optimization and increased production of these compounds [29]. The ability of yeasts to use and degrade a broad diversity of carbon sources and their natural ability to produce organic acids turns them into ideal microorganisms for the construction of future biorefineries. Over the last two decades, as aforementioned, M. Casal's group has worked extensively on the transport mechanisms of CA and the identification of CA transporters. Two genes encoding CA permeases in yeast were identified in our group, namely Jen1 and Ato1. Other functional characterizations were also carried out at the level of intracellular signaling, endocytic trafficking, and protein turnover of these transporters [55, 74]. Besides these, other CA transporter homologs were characterized in *Candida albicans* (CaJen1 and CaJen2), *K. lactis* (KlJen1 and KlJen2), *Debaryomyces hansenii* (DH27, DH17, DH18, and DH24), *Aspergillus nidullans* (AcpA), and *E. coli* (SatP) and *Methanosarcina acetivorans* (AceP) [60, 75-79]. In this work, we will focus on the identification of novel transporters that belong to these and other transporter families, aiming at the identification of relevant transporters for the improvement of microbial cell factories.

1.2 AIMS OF THE WORK

This thesis aims at identifying new carboxylate transporters in yeasts and explore their biotechnological application for CA production. The work was focused on the structural and functional features of CA transporters for a better understanding of the mechanisms of transport that can be relevant to improve industrial microbe cell factories.

The main objectives of this work are:

- To screen novel plasma membrane transporter proteins for mono-, di- and tricarboxylic acids, such as acetic, lactic, fumaric, malic, succinic, and citric acids.
- To characterize the structure of the target transporters using bioinformatic tools.
- To study the kinetics, specificity, and energetics of the target transporters.
- To unveil transporters that behave as exporters since they are key players in increasing the tolerance and productivity of organic acids in microbial cell factories.

The final goal of this work is to contribute to the improvement of metabolic engineered strains to be used for industrial exploitation, aiming at the large-scale production of organic acids. Finally, the outcomes of this project are expected to provide novel insights for the biotechnology industry, through the utilization of membrane transporters in the microbial production of organic acids.

1.3 REFERENCES

1. Hara, K., Araki, M., Okai, N., Wakai, S., Hasunuma, T.,Kondo, A., *Development of bio-based fine chemical production through synthetic bioengineering.* Microbial cell factories, **2014**. 13(1): p. 1-19
2. Coban, H., *Organic acids as antimicrobial food agents: applications and microbial productions.* Bioprocess and biosystems engineering, **2020**. 43(4): p. 569-591
3. Panda, S., Sahu, L., Behera, S.,Ray, R., *Research and production of organic acids and industrial potential.* Bioprocessing for Biomolecules Production, **2019**: p. 195-209
4. Branduardi, P., Smeraldi, C.,Porro, D., *Metabolically engineered yeasts:'potential'industrial applications.* Journal of Molecular Microbiology and Biotechnology, **2008**. 15(1): p. 31-40
5. Jang, Y., Kim, B., Shin, J., Choi, Y., Choi, S., Song, C., Lee, J., Park, H.,Lee, S., *Bio-based production of C2–C6 platform chemicals.* Biotechnology and bioengineering, **2012**. 109(10): p. 2437-2459
6. Mattanovich, D., Branduardi, P., Dato, L., Gasser, B., Sauer, M.,Porro, D., *Recombinant protein production in yeasts.* Recombinant gene expression, **2012**: p. 329-358
7. Goffeau, A., Barrell, B.G., Bussey, H., Davis, R., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J., Jacq, C.,Johnston, M., *Life with 6000 genes.* Science, **1996**. 274(5287): p. 546-567
8. Mattanovich, D., Sauer, M.,Gasser, B., *Yeast biotechnology: teaching the old dog new tricks.* Microbial cell factories, **2014**. 13(1): p. 1-5
9. Kavšček, M., Stražar, M., Curk, T., Natter, K.,Petrovič, U., *Yeast as a cell factory: current state and perspectives.* Microbial cell factories, **2015**. 14(1): p. 1-10
10. Gomes, A., Carmo, T., Carvalho, L., Bahia, F.,Parachin, N., *Comparison of yeasts as hosts for recombinant protein production.* Microorganisms, **2018**. 6(2): p. 38
11. Kim, H., Yoo, S.J.,Kang, H.A., *Yeast synthetic biology for the production of recombinant therapeutic proteins.* FEMS Yeast Res, **2014**. 15: p. 1-16
12. Zhang, N.,An, Z., *Heterologous protein expression in yeasts and filamentous fungi.* Manual of Industrial Microbiology and Biotechnology, **2010**: p. 145-156
13. Nevoigt, E., *Progress in metabolic engineering of Saccharomyces cerevisiae.* Microbiol. Mol. Biol. Rev., **2008**. 72(3): p. 379-412
14. Sousa-Silva, M., Vieira, D., Soares, P., Casal, M.,Soares-Silva, I., *Expanding the Knowledge on the Skillful Yeast Cyberlindnera jadinii.* Journal of Fungi, **2021**. 7(1): p. 36
15. Nombela, C., Gil, C.,Chaffin, W., *Non-conventional protein secretionin yeast.* Trends in microbiology, **2006**. 14(1): p. 15-21
16. Radecka, D., Mukherjee, V., Mateo, R., Stojiljkovic, M., Foulquié-Moreno, M.,Thevelein, J., *Looking beyond Saccharomyces: the potential of non-conventional yeast species for desirable traits in bioethanol fermentation.* FEMS Yeast Res, **2015**. 15(6)
17. Fonseca, G., Heinzle, E., Wittmann, C.,Gombert, A., *The yeast Kluyveromyces marxianus and its biotechnological potential.* Applied microbiology and biotechnology, **2008**. 79(3): p. 339-354
18. Schwarzhans, J., Luttermann, T., Geier, M., Kalinowski, J.,Friehs, K., *Towards systems metabolic engineering in Pichia pastoris.* Biotechnology advances, **2017**. 35(6): p. 681-710

19. Öberg, F., Sjöhamn, J., Conner, M., Bill, R.,Hedfalk, K., *Improving recombinant eukaryotic membrane protein yields in Pichia pastoris: the importance of codon optimization and clone selection.* Molecular membrane biology, **2011**. 28(6): p. 398-411
20. Bill, R., *Playing catch-up with Escherichia coli: using yeast to increase success rates in recombinant protein production experiments.* Frontiers in microbiology, **2014**. 5: p. 85
21. Sun, W., Vila-Santa, A., Liu, N., Prozorov, T., Xie, D., Faria, N., Ferreira, F., Mira, N.,Shao, Z., *Metabolic engineering of an acid-tolerant yeast strain Pichia kudriavzevii for itaconic acid production.* Metabolic engineering communications, **2020**: p. e00124
22. Park, H., Bae, J., Ko, H., Lee, S., Sung, B., Han, J.,Sohn, J., *Low-pH production of d-lactic acid using newly isolated acid tolerant yeast Pichia kudriavzevii NG7.* Biotechnology and bioengineering, **2018**. 115(9): p. 2232-2242
23. Toivari, M., Vehkomäki, M., Nygård, Y., Penttilä, M., Ruohonen, L.,Wiebe, M., *Low pH D-xylonate production with Pichia kudriavzevii.* Bioresource technology, **2013**. 133: p. 555-562
24. Xiao, H., Shao, Z., Jiang, Y., Dole, S.,Zhao, H., *Exploiting Issatchenkia orientalis SD108 for succinic acid production.* Microbial cell factories, **2014**. 13(1): p. 121, 10.1186/s12934-014-0121-4
25. Buerth, C., Tielker, D.,Ernst, J.F., *Candida utilis and Cyberlindnera (Pichia) jadinii: yeast relatives with expanding applications.* Applied microbiology and biotechnology, **2016**. 100(16): p. 6981-6990
26. Dashko, S., Zhou, N., Compagno, C.,Piškur, J., *Why, when, and how did yeast evolve alcoholic fermentation?* FEMS Yeast Res, **2014**. 14(6): p. 826-832
27. Soares-Silva, I., Ribas, D., Sousa-Silva, M., Azevedo-Silva, J., Rendulić, T.,Casal, M., *Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications.* FEMS microbiology letters, **2020**
28. Van der Hoek, S.,Borodina, I., *Transporter engineering in microbial cell factories: the ins, the outs, and the in-betweens.* Current opinion in biotechnology, **2020**. 66: p. 186
29. Kell, D.B., Swainston, N., Pir, P.,Oliver, S.G., *Membrane transporter engineering in industrial biotechnology and whole cell biocatalysis.* Trends in biotechnology, **2015**. 33(4): p. 237-246
30. de la Serna, J.B., Schütz, G.J., Eggeling, C.,Cebecauer, M., *There Is No Simple Model of the Plasma Membrane Organization.* Frontiers in cell and developmental biology, **2016**. 4
31. Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D.,Darnell, J., *Molecular cell biology 4th edition.* National Center for Biotechnology InformationÕs Bookshelf, **2000**
32. Cooper, G.M.,Ganem, D., *The Cell: A Molecular Approach.* Nature Medicine, **1997**. 3(9): p. 1042-1042
33. Cournia, Z., Allen, T.W., Andricioaei, I., Antonny, B., Baum, D., Brannigan, G., Buchete, N.-V., Deckman, J.T., Delemotte, L.,Del Val, C., *Membrane protein structure, function, and dynamics: a perspective from experiments and theory.* The Journal of membrane biology, **2015**. 248(4): p. 611-640
34. Wallin, E.,Heijne, G.V., *Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms.* Protein Science, **1998**. 7(4): p. 1029-1038
35. Saier, M.H., *A functional-phylogenetic classification system for transmembrane solute transporters.* Microbiology and Molecular Biology Reviews, **2000**. 64(2): p. 354-411

36. Casal, M., Queirós, O., Talaia, G., Ribas, D.,Paiva, S., *Carboxylic acids plasma membrane transporters in Saccharomyces cerevisiae*, in *Yeast Membrane Transport***2016**, Springer. p. 229-251
37. Casal, M., Paiva, S., Queirós, O.,Soares-Silva, I., *Transport of carboxylic acids in yeasts.* FEMS microbiology reviews, **2008**. 32(6): p. 974-994
38. Conde, A., Diallinas, G., Chaumont, F., Chaves, M.,Gerós, H., *Transporters, channels, or simple diffusion? Dogmas, atypical roles and complexity in transport systems.* The international journal of biochemistry & cell biology, **2010**. 42(6): p. 857-868
39. Henderson, P.J.F., *Membrane Proteins for Secondary Active Transport and their Molecular Mechanisms.* In Comprehensive Biophysics, Bioenergetics., **2012**(Academic Press): p. 265-288
40. Ubarretxena-Belandia, I.,Stokes, D.L., *Present and future of membrane protein structure determination by electron crystallography.* Advances in protein chemistry and structural biology, **2010**. 81: p. 33
41. McCusker, E.C., Bagnéris, C., Naylor, C.E., Cole, A.R., D'avanzo, N., Nichols, C.G.,Wallace, B.A., *Structure of a bacterial voltage-gated sodium channel pore reveals mechanisms of opening and closing.* Nature communications, **2012**. 3: p. 1102
42. Punta, M., Forrest, L.R., Bigelow, H., Kernytsky, A., Liu, J.,Rost, B., *Membrane protein prediction methods.* Methods, **2007**. 41(4): p. 460-474
43. Hennerdal, A.,Elofsson, A., *Rapid membrane protein topology prediction.* Bioinformatics, **2011**. 27(9): p. 1322-1323
44. Carpenter, E.P., Beis, K., Cameron, A.D.,Iwata, S., *Overcoming the challenges of membrane protein crystallography.* Current opinion in structural biology, **2008**. 18(5): p. 581-586
45. Ravna, A.W.,Sylte, I., *Homology modeling of transporter proteins (carriers and ion channels).* Homology Modeling: Methods and Protocols, **2012**: p. 281-299
46. Diallinas, G., *Dissection of transporter function: from genetics to structure.* Trends in Genetics, **2016**. 32(9): p. 576-590
47. Colas, C., Ung, P.M.-U.,Schlessinger, A., *SLC transporters: structure, function, and drug discovery.* MedChemComm, **2016**. 7(6): p. 1069-1081
48. Murali, N., Srinivas, K.,Ahring, B.K., *Biochemical Production and Separation of Carboxylic Acids for Biorefinery Applications.* Fermentation, **2017**. 3(2): p. 22
49. Hamid, O., Puzanov, I., Dummer, R., Schachter, J., Daud, A., Schadendorf, D., Blank, C., Cranmer, L.D., Robert, C., Pavlick, A.C., Gonzalez, R., Hodi, F.S., Ascierto, P.A., Salama, A.K.S., Margolin, K.A., Gangadhar, T.C., Wei, Z., Ebbinghaus, S., Ibrahim, N.,Ribas, A., *Final analysis of a randomised trial comparing pembrolizumab versus investigator-choice chemotherapy for ipilimumab-refractory advanced melanoma.* Eur J Cancer, **2017**. 86: p. 37-45, 10.1016/j.ejca.2017.07.022
50. Zmijewski Jr, M.J.,MacQuillan, A.M., *Dual effects of glucose on dicarboxylic acid transport in Kluyveromyces lactis.* Canadian journal of microbiology, **1975**. 21(4): p. 473-480
51. Lodi, T., Fontanesi, F., Ferrero, I.,Donnini, C., *Carboxylic acids permeases in yeast: two genes in Kluyveromyces lactis.* Gene, **2004**. 339: p. 111-119, https://doi.org/10.1016/j.gene.2004.06.019

52. Osothsilp, C.,Subden, R., *Malate transport in Schizosaccharomyces pombe.* Journal of bacteriology, **1986**. 168(3): p. 1439-1443
53. Grobler, J., Bauer, F., Subden, R.E.,Van Vuuren, H.J., *The mae1 gene of Schizosaccharomyces pombe encodes a permease for malate and other C4 dicarboxylic acids.* Yeast, **1995**. 11(15): p. 1485-1491
54. Cássio, F., Leao, C.,Van Uden, N., *Transport of lactate and other short-chain monocarboxylates in the yeast Saccharomyces cerevisiae.* Applied and Environmental Microbiology, **1987**. 53(3): p. 509-513
55. Casal, M., Paiva, S., Andrade, R.P., Gancedo, C.,Leão, C., *The lactate-proton symport of Saccharomyces cerevisiae is encoded by JEN1.* Journal of bacteriology, **1999**. 181(8): p. 2620-2623
56. McDermott, J.R., Rosen, B.P.,Liu, Z., *Jen1p: a high affinity selenite transporter in yeast.* Molecular biology of the cell, **2010**. 21(22): p. 3934-3941
57. Alves, R., Sousa-Silva, M., Vieira, D., Soares, P., Chebaro, Y., Lorenz, M., Casal, M., Soares-Silva, I.,Paiva, S., *Carboxylic Acid Transporters in Candida Pathogenesis.* MBio, **2020**. 11(3): p. e00156-20, 10.1128/mBio.00156-20
58. Paiva, S., Devaux, F., Barbosa, S., Jacq, C.,Casal, M., *Ady2p is essential for the acetate permease activity in the yeast Saccharomyces cerevisiae.* Yeast, **2004**. 21(3): p. 201-210
59. Palková, Z., Devaux, F., Ricicová, M., Mináriková, L., Le Crom, S.,Jacq, C., *Ammonia pulses and metabolic oscillations guide yeast colony development.* Molecular biology of the cell, **2002**. 13(11): p. 3901-3914
60. Ribas, D., Soares-Silva, I., Vieira, D., Sousa-Silva, M., Sá-Pessoa, J., Azevedo-Silva, J., Viegas, S., Arraiano, C., Diallinas, G., Paiva, S., Soares, P.,Casal, M., *The acetate uptake transporter family motif "NPAPLGL (M/S)" is essential for substrate uptake.* Fungal Genetics and Biology, **2019**. 122: p. 1-10
61. Guaragnella, N.,Butow, R., *ATO3 encoding a putative outward ammonium transporter is an RTG-independent retrograde responsive gene regulated by GCN4 and the Ssy1-Ptr3-Ssy5 amino acid sensor system.* Journal of Biological Chemistry, **2003**. 278(46): p. 45882-45887
62. Xi, Y., Zhan, T., Xu, H., Chen, J., Bi, C., Fan, F.,Zhang, X., *Characterization of JEN family carboxylate transporters from the acid-tolerant yeast Pichia kudriavzevii and their applications in succinic acid production.* Microbial biotechnology, **2021**
63. Sá-Pessoa, J., Paiva, S., Ribas, D., Silva, I., Viegas, S., Arraiano, C.,Casal, M., *SATP (YaaH), a succinate–acetate transporter protein in Escherichia coli.* Biochemical journal, **2013**. 454(3): p. 585-595
64. Wang, J., Lin, M., Xu, M.,Yang, S.-T., *Anaerobic fermentation for production of carboxylic acids as bulk chemicals from renewable biomass*, in *Anaerobes in Biotechnology***2016**, Springer. p. 323-361
65. López-Garzón, C.S.,Straathof, A.J., *Recovery of carboxylic acids produced by fermentation.* Biotechnology advances, **2014**. 32(5): p. 873-904
66. De Jong, E., Stichnothe, H., Bell, G.,Jørgensen, H., *Bio-Based Chemicals: a 2020 Update.* Paris: IEA Bioenergy, **2020**

67. Straathof, A.J., *Transformation of biomass into commodity chemicals using enzymes or cells.* Chemical reviews, **2013**. 114(3): p. 1871-1908
68. Chen, Y.,Nielsen, J., *Biobased organic acids production by metabolically engineered microorganisms.* Current opinion in biotechnology, **2016**. 37: p. 165-172
69. Roehr, M.,Kubicek, C.P., *Citric acid and Further organic acids.* Biotechnology Set, Second Edition, **1996**: p. 363-379
70. Research, E.-E.M., *Global Citric Acid Market Report and Forecast 2021-2026*, **2020**,
71. Taylor, R., Nattrass, L., Alberts, G., Robson, P., Chudziak, C., Bauen, A., Libelli, I., Lotti, G., Prussi, M.,Nistri, R., *From the sugar platform to biofuels and biochemicals: final report for the European Commission Directorate-General Energy*, **2015**, E4tech/Re-CORD/Wageningen UR,
72. Spekreijse, J., Lammens, T., Parisi, C., Ronzon, T.,Vis, M., *Insights into the European market of bio-based chemicals.* Publications Office of the European Union, Luxembourg, **2019**
73. Yang, S.-T., Liu, X.,Zhang, Y., *Metabolic engineering–applications, methods, and challenges.* Bioprocessing for Value-Added Products from Renewable Resources: New Technologies and Applications, **2011**. 73
74. Andrade, R.,Casal, M., *Expression of the lactate permease gene JEN1 from the yeast Saccharomyces cerevisiae.* Fungal Genetics and Biology, **2001**. 32(2): p. 105-111
75. Soares-Silva, I., Schuller, D., Andrade, R.P., Baltazar, F., Cássio, F.,Casal, M., *Functional expression of the lactate permease Jen1p of Saccharomyces cerevisiae in Pichia pastoris.* Biochemical Journal, **2003**. 376(3): p. 781-787
76. Queirós, O., Pereira, L., Paiva, S., Moradas-Ferreira, P.,Casal, M., *Functional analysis of Kluyveromyces lactis carboxylic acids permeases: heterologous expression of KlJEN1 and KlJEN2 genes.* Current genetics, **2007**. 51(3): p. 161-169
77. Soares-Silva, I., Ribas, D., Foskolou, I.P., Barata, B., Bessa, D., Paiva, S., Queirós, O.,Casal, M., *The Debaryomyces hansenii carboxylate transporters Jen1 homologues are functional in Saccharomyces cerevisiae.* FEMS yeast research, **2015**. 15(8)
78. Sá-Pessoa, J., Paiva, S., Ribas, D., Silva, I.J., Viegas, S.C., Arraiano, C.M.,Casal, M., *SATP (YaaH), a succinate–acetate transporter protein in Escherichia coli.* Biochemical journal, **2013**. 454(3): p. 585-595
79. Sá-Pessoa, J., Amillis, S., Casal, M.,Diallinas, G., *Expression and specificity profile of the major acetate transporter AcpA in Aspergillus nidulans.* Fungal Genetics and Biology, **2015**. 76: p. 93-103

CHAPTER II

Expanding the Knowledge on the Skillful Yeast *Cyberlindnera jadinii*

Adapted from:

Sousa-Silva, M.[1,2], Vieira, D.[1,2], Soares. P.[1,2], Casal, M.[1,2], Soares-Silva, I.[1,2] (2021). Expanding the Knowledge on the Skillful Yeast *Cyberlindnera jadinii*. *Journal of Fungi*, *7*(1):36. doi.org/10.3390/jof7010036

[1] Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, Portugal

[2] Centre of Molecular and Environmental Biology (CBMA), Department of Biology, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal

CHAPTER II

Expanding the Knowledge on the Skillful Yeast *Cyberlindnera jadinii*

ABSTRACT

Cyberlindnera jadinii is widely used as a source of single-cell protein and is known for its ability to synthesize a great variety of valuable compounds for the food and pharmaceutical industries. Its capacity to produce compounds such as food additives, supplements, and organic acids, among other fine chemicals, has turned it into an attractive microorganism in the biotechnology field. In this review, we performed a robust phylogenetic analysis using the core proteome of *C. jadinii* and other fungal species, from Asco- to Basidiomycota, to elucidate the evolutionary roots of this species. In addition, we report the evolution of this species nomenclature over-time and the existence of a teleomorph (*C. jadinii*) and anamorph state (*Candida utilis*) and summarize the current nomenclature of most common strains. Finally, we highlight relevant traits of its physiology, the solute membrane transporters so far characterized, as well as the molecular tools currently available for its genomic manipulation. The emerging applications of this yeast reinforce its potential in the white biotechnology sector. Nonetheless, it is necessary to expand the knowledge on its metabolism, regulatory networks, and transport mechanisms, as well as to develop more robust genetic manipulation systems and synthetic biology tools to promote the full exploitation of *C. jadinii*.

KEYWORDS: *Cyberlindnera jadinii*; phylogeny; life cycle; genome; physiology; biotechnology applications; membrane transporter systems

2.1 INTRODUCTION

The future of our society challenges researchers to find novel technologies to address global environmental problems, mitigate ecosystems' damage, and biodiversity losses, as the current model of development based on natural resources exploitation is unsustainable. Exploring microorganisms for the production of platform chemicals constitutes an alternative approach to avoid the use of nonrenewable petrochemical-based derivatives. Developing applications for the industrial sphere using biological systems instead of classical chemical catalysts is the main focus of white biotechnology [1]. In microbial-based industrial processes, several features have to be addressed to obtain robust cell factories capable of achieving superior metabolic performances, such as the optimization of metabolic fluxes, membrane, and transporter engineering, and in-creased tolerance against harsh industrial conditions and toxic compounds [2]. In addition, specifications like the cost of feedstock, product yield and productivity, and downstream processing have to be taken in account to develop successful industrial approaches [1]. Withal the fact that *Saccharomyces cerevisiae* is by far the most relevant industrial yeast species, *Cyberlindnera jadinii* is an example of the so-called non-Saccharomyces yeasts [3] claiming for a place as a relevant contributor to the industrial biotechnological sector. The yeast *C. jadinii* is able to produce valuable bioproducts being an attractive source of biomass enriched in protein and vitamins. The richness of protein content, around 50% of dry cell weight, and amino acid diversity turn its biomass ideal as a source of protein supplement for animal feed and human consumption [4]. The high degree of tolerance to environmental changes occurring during fermentation turn *C. jadinii* an alternative to other established cell factory systems [5]. As a Crabtree-negative yeast, it has one of the highest respiratory capacity among characterized yeast species, being considered ideal for continuous cell cultures [6]. The Food and Drug Administration (FDA) attributed the "General Regarded As Safe" (GRAS) status to this yeast, recognizing it as safe and suitable for supplying food additives and dietary supplements for humans [5, 7-9]. The ability to produce relevant compounds, to grow in a wide range of temperatures, to use inexpensive media with high productivity levels turns it an industrially relevant microorganism [8, 10-12]. Recent efforts have developed *C. jadinii* molecular tools for metabolic engineering processes and the overexpression of proteins. In the past, the uncertainty of this yeast polyploidy, together with the lack of suitable selection markers and expression cassettes, [10] delayed its widely use as cell factory. With this review, we intend to compile the existing knowledge on this yeast, that will allow the development of future strategies to strengthen the role of *C. jadinii* in the biotechnology sector. We start by reviewing the nomenclature of this species, altered several times over time. An update on the current nomenclature of the most relevant

strains is also presented. We establish the evolutionary relationship of *C. jadinii* within other fungi with a complete genome available. The most relevant morphological and physiological traits are also here described, together with the genetic manipulation tools and expression systems currently available. Moreover, we present a summary of all the plasma membrane transporter systems so far characterized in this yeast, as they are key-players for cell factory optimization. Finally, we will focus on the biotechnological potential of this yeast and highlight the future challenges to achieve the full exploitation of this industrially valuable microorganism.

2.2 ECOLOGY, TAXONOMY AND EVOLUTION

The natural environment of *Cyberlindnera jadinii* is still an open question. It is thought that it may be associated with the decomposition of plant material in nature, as it is able to assimilate pentoses and tolerate lignocellulosic by-products [3], displays great fermentative ability, and is copiotrophic [13]. The current laboratory strains were isolated from distinct environments, namely, from the pus of a woman abscess (CBS 1600/NRRL Y-1542), a cow with mastitis (CBS 4885/NRRL Y-6756), a yeast deposit from a distillery (CBS 567), yeast cell factories (CBS 621), and flowers (CBS 2160) [5-9, 12]. The extensive nomenclature revisions of this species are well described in "The tortuous history of *Candida utilis*" by Barnet [14]. In 1926 this yeast was isolated from several German yeast factories, which had been cultivated without a systematic name during the time of World War I for food and fodder [14]. It was first named "*Torula utilis"*, being later referred to as *Torulopsis utilis* (1934). The "food yeast" was also designated as *Saccharomyces jadinii* (1932), *Hansenula jadinii* (1951), *Candida utilis* (1952), *Pichia jadinii* (1984), and *Lindnera jadinii* (2008) [12, 14-16]. From the aforementioned, *C. utilis* was the nomenclature most commonly used, having almost 1000 published papers in PubMed (results available at: https://pubmed.ncbi.nlm.nih.gov/?term=%22candida+utilis%22; Accessed October 16, 2020). The *Candida* genus comprised species that form pseudohyphae or true hyphae with blastoconidia, among other standard characteristics [17-19]. In the classification system implemented in 1952 by Lodder and Kreger van Rij, the *Candida* genus included yeasts that produce only simple pseudohyphae [14]. At that time, the majority of the isolates was renamed as *C. utilis* [18]. Later, the *C. utilis* was established as an asexual state of a known ascosporogenous yeast, *Hansenula jadinii*, as it was found to share some similarities between phenotypic traits [20]. In 1984, even though concurring with a publication of an extensive chapter of the genus *Hansenula*, Kurtzman moved most of the *Hansenula* species to the *Pichia* genus, due to their "deoxyribonucleic acid relatedness." Thus, *C. utilis* was renamed *Pichia jadinii*. A

quarter of a century later, this yeast species was again renamed as *Lindnera jadinii* based on analyses of nucleotide sequence divergence in the genes coding for large and small-subunit rRNAs [12]. Species integrated into the *Lindnera* differ considerably in ascospore morphology ranging from spherical to hat-shaped or Saturn-shaped spores. In addition, this clade includes both hetero- and homothallic species and physiological features as fermenting glucose and assimilating a variety of sugars, polyols, and other carbon sources are defining characteristics of the *Lindnera* genus [12]. Finally, 1 year later, the genus *Lindnera* was replaced by *Cyberlindnera*, as the later homonym defined a validly published plant genus [15]. This substitution occurred in 21 new species, including *Cyberlindnera jadinii* [15]. In summary, any of the aforementioned nomenclatures reported in the literature may refer to the same organism since *C. utilis* is the anamorph state and *C. jadinii* the teleomorph state [14]. The anamorph represents the asexual stage of a fungus contrasting with the teleomorph form that defines the sexual stage of the same fungus [21]. The primary name of a species relies on the sexual state or teleomorph, but a second valid name may rely on the asexual state or anamorph [22]. However, this should only happen when teleomorphs have not been found for a specific species or it is not clear if a particular teleomorph is the same species as a particular anamorph. Accordingly, since 2013, the International Botanical Congress states that the system for allowing separate names for the anamorph state should end [23]. The new International Code of Nomenclature for algae, fungi, and plants, the Melbourne Code, supports the directive that fungal species or higher taxon should be assigned with a single valid name. Accordingly, anamorph yeast genera like *Candida* should be revised to turn the genus consistent with phylogenetic affinities [19]. Notwithstanding, the reclassification of several *C. utilis* as *C. jadinii* in several culture type collections is still confusing, reaching a point where the same strain is designated as *C. utilis* and *C jadinii* in different culture type collections. This aspect, together with the previous nomenclatures used in research papers, leads to unnecessary misunderstandings. To clarify this, Table 1 presents the alternative designations of the main laboratory strains.

Table 1. Main *C. jadinii* (former *C. utilis*) strains described in literature

Nomenclature in Literature	Current Nomenclature	Isolation Source	Reference
Candida utilis NBRC 0988	*C. jadinii* ATCC 9950; CBS 5609; DSM 2361; NBRC 0988; NCYC 707; NRRL Y-900	Yeast factory in Germany	[24]
C. utilis ATCC 9256 [a]	*C. jadinii* NRRL Y-1084; CBS 841; CCRC 20334; DSM 70167; NCYC 359; VKM Y-768; VTT C-79091	Unknown	[25, 26]
C. utilis ATCC 9226 [a] NBRC 1086	*C. jadinii* VTT C-71015; FMJ 4026; NBRC 1086	Unknown	[25, 27, 28]
C. utilis IGC 3092	*C. jadinii* PYCC 3092; CBS 890; VKM Y-33	Unknown	[29-31]
C. utilis CCY 39-38-18	*C. jadinii* CCY 029-38-18 [b]	Unknown	[32]
C. utilis NCYC 708	*C. jadinii* NCYC 708; ATCC 42181; CBS 5947; VTT C-84157	Unknown	[33]
C. utilis CBS 4885 NRRL Y-6756	*C. jadinii* CBS 4885; NRRL Y-6756; NBRC 10708	Cow with mastitis	[34]
C. utilis CBS 567 NRRL Y-1509	*C. jadinii* CBS 567; NRRL Y-1509	Yeast deposit in distillery	[34]
C. utilis CBS 2160	*C. jadinii* CBS 2160	Flower of *Taraxacum* sp.	[34]
C. utilis CBS 621	*C. jadinii* CBS 621; NRRL Y-7586; ATCC 22023; PYCC 4182	Yeast factories	[35]
C. utilis CBS 1600	*C. jadinii* CBS 1600; NRRL Y-1542; ATCC 18201	Pus of a woman abscess	[16]

[a] This strain has been discontinued in ATCC.
[b] The strain number reported in the literature is not available in the Culture Collection of Yeasts (CCY), all *C. jadinii* strains are registered as 029-38-XX, including *C. jadinii* 029-38-18, the likely match to CCY 39-38-18.

C. jadinii belongs to the phylum Ascomycota, subphylum Saccharomycotina. The members of this subphylum constitute a monophyletic group of ascomycetes that are well defined by ultrastructural and DNA characteristics [13]. These include lower amounts of chitin in overall polysaccharide composition at cell walls, being unable to stain with diazonium blue, low content of guanine and cytosine (G + C < 50%) at nuclear DNA, and presence of continuous holoblastic bud formation with wall layers. *C. jadinii* belongs to the Saccharomycetes class, Saccharomycetidae family, Saccharomycetales order, and the *Cyberlindnera* genus. However, a comprehensive phylogenetic analysis and evolutionary relationship are still missing for this species [36-38]. Aiming at filling this gap, we performed a robust phylogeny reconstruction [36, 39-41]. As can be depicted in Figure 1, this yeast localizes in the Phaffomycetaceae clade together with *Cyberlindnera fabianii* and *Wickerhamomyces ciferri*. The nearest neighbors belong to the Saccharomycetaceae family, which includes a clade with *S. cerevisiae/Torulaspora delbrueckii* species and another clade with *Eremothecium gossypii* (former *Ashbya gossypii*), *Kluyveromyces lactis/marxianus,* and *Lachancea* species.

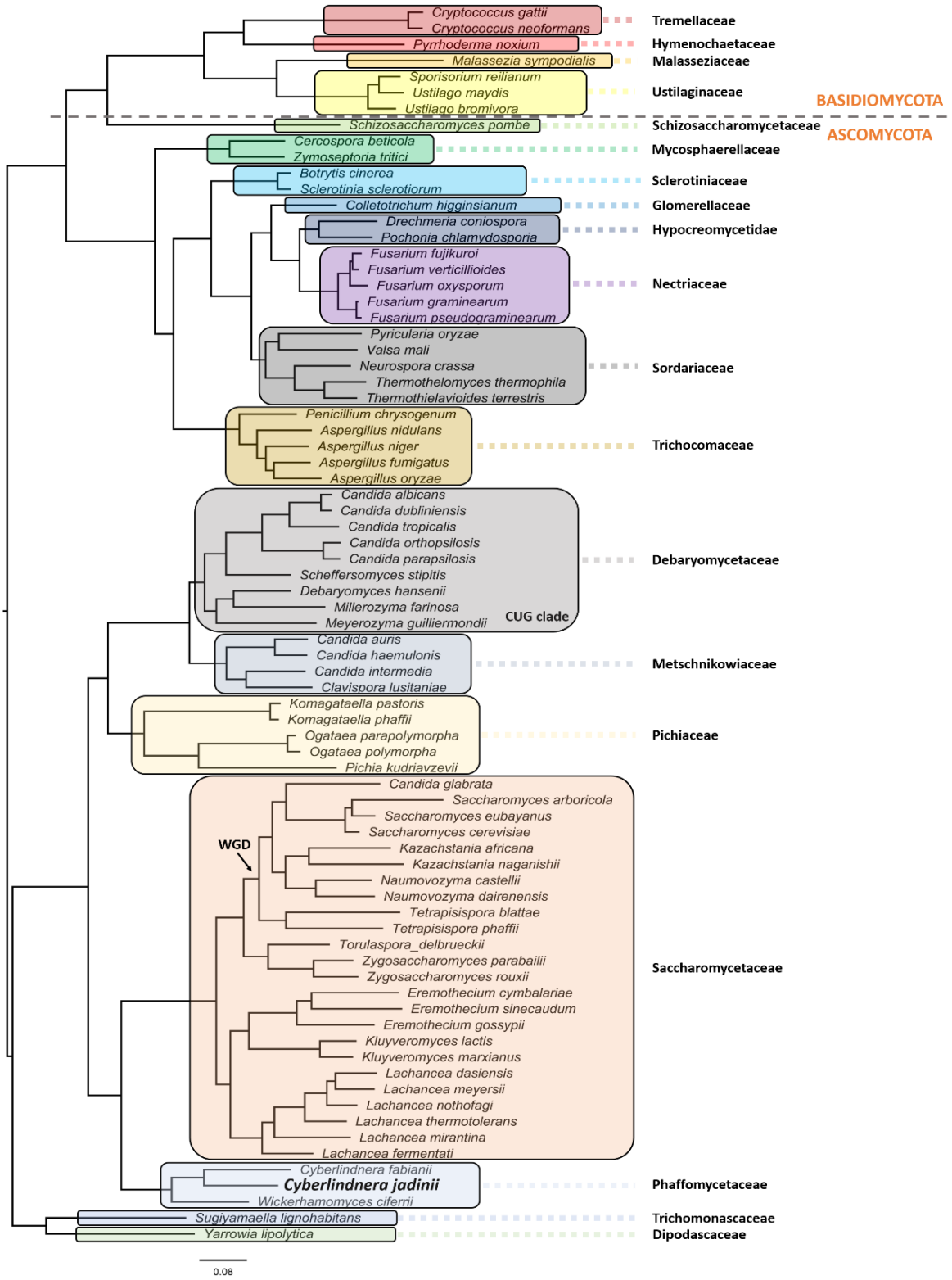


Figure 1. Evolutionary relationship of *Cyberlindnera jadinii*, a member of the Phaffomycetaceae clade. The phylogenetic reconstruction was obtained using the following parameters: maximum likelihood in IQ-TREE (http://www.iqtree.org), the model of amino acid evolution JTT (Jones-Taylor-Thornton), and four gamma-distributed rates. Homologues were detected for 1567 proteins across the proteome of 77 selected fungal species from NCBI. The 1567 set of proteins were aligned and then concatenated in order to use in the phylogenetic analysis. These proteins offer a clear high-resolution evolutionary view of the different species, as they are essential proteins beyond the specific

biology of the different yeasts. Bootstrapping provided values of 100% for all the nodes. Yeast and fungi families are highlighted with different colors and shades. The phylogenetic relationships reflect evolutionary ancestries, independently of adaptations and overall gene contents within the various species. All families with more than one representative species in the analyses formed monophyletic groups.

Despite the previous genus nomenclature adopted for *C. jadinii* (Candida and Pichia), it is phylogenetically distant from the Debaryomycetaceae and Pichiaceae families that include the *Candida* species, except *C. glabrata*, the *Pichia kudriavzevii,* and *Ogataea* species. *Komagataella phaffi* is as expected included in the Pichiaceae clade, together with *Komagataella pastoris* [36, 40]. The Trichomonascaceae/Dipodascaceae clade, formerly known as the *Yarrowia* clade, includes now the *Sugiyamaella lignohabitans* species together with *Yarrowia lipolytica*, and is the most distant yeast clade, except for the Schizosaccharomycetaceae that clusters with all the Basiodiomycota. The filamentous fungi *Neurospora crassa* and *Fusarium graminearum* are in different clades as members of the Sordariaceae and Nectriaceae clades, respectively [40]. In addition, in the Sordariaceae clade, the phylogenetic position of *Thermothelomyces thermophila* species (*Myceliophthora thermophila*) was uncovered [39, 42].

2.3 LIFE CYCLE AND GENOME ORGANIZATION

Kurtzman and colleagues proposed *C. jadinii* as the teleomorphic parental species of *C. utilis*, due to the 85% reassociation rate obtained between genomic DNA of the two yeast species and to the high similarities of ribosomal RNA sequences [20, 34]. The formation of ascospores allied with genomic sequencing data confirmed the diploidy of *C. jadinii* NRRL Y-1542 strain and the identification of *MAT*a and *MAT*α genes allelic locations [20, 37]. Ikushima et al. studied the polyploid of several *C. utilis* strains (NBRC0396, 0619, 0626, 0639, 0988, 1086, and 10707) detecting an overall ploidy switching between 2n and 5n [43]. Later, Kondo and colleagues inferred, through the analysis of *C. utilis* ATCC 9950 transformants the presence of a diploid state, although some years later, the sequential disruption of the *URA3* and *PDC1* locus suggested the tetraploidy of this strain [25, 43, 44]. A fluorescence-activated cell sorting analysis pointed out a ploidy of 3n to 5n in this latter *C. utilis* strain, following the aforementioned data by Ikushima *et al.*[16, 43]. Furthermore, a single nucleotide polymorphism analysis suggested that the *C. utilis* NBRC0988 genome was triploid Overall, these results suggested that *C. utilis* has derived from the parental yeast *C. jadinii* through triploidization pursuing an unexplained sequence of genetic events [16].

Recently, Krahulec and colleagues analyzed the *C. utilis* CCY 39-38-18 genome ploidy through the evaluation of the maltase gene copy number in deleted mutants, pointing out the tetraploidy of this strain [32]. Despite the existing ploidy variation, the diploid state of *C. jadinii* impelled its genetic manipulation and subsequent utilization in the biotechnological industry [16, 37]. Table 2 summarizes the genetic features of the *C. jadinii* strains sequenced so far [24]. Although the GC-content is quite similar, the genome size is different among the distinct strains evaluated (Table 2). When compared to *S. cerevisiae, C. jadinii* has an increased genome size and higher GC-content. The genomic features of *C. jadinii* strains are different from two closed phylogenetic species *Wickerhamomyces ciferrii and Cyberlindnera fabianii* that, respectively, contain a G-C ratio of 30.4% and 44.4%, genome sizes of 15.9 and 12.3 Mb, and a total of 6702 and 5944 CDS [45, 46].

Table 2. Genomic features of three *C. jadinii* strains and the reference *S. cerevisiae* strain S288c.

Components	***Cyberlindnera jadinii* Strains**			***Saccharomyces cerevisiae* S288c**
	NBRC 0988	**CBS 1600**	**NRRL Y-1542**	
NCBI assembly reference	GCA_000328385.1	GCA_001245095.1	GCA_001661405.1	GCA_000146045.2
Assembly level	Chromosomes	Scaffold	Scaffold	Complete genome
Genome size	14.3 Mb	12.7 Mb	13.0 Mb	12.2 Mb
Genes [a]	8864	5566	6184	6002
No of scaffolds [b]	1002	7	76	17
Scaffold N50 [b]	189,765	2,123,196	700,888	924,431
No. of contigs [b]	1163	91	392	17
Contig N50 (bp) [b]	158,681	287,918	111,555	924,431
No. of chromosomes	13	–	–	16
GC-content (%)	44.7	44.5	44.6	38.3
Total of CDS [a]	8646	5057	6032	5771
Gene annotation	[24]	[16]	[37]	[47, 48]

[a] Total number of predicted genes and protein-coding genes (CDS) are taken from original publications or subsequent annotations. [b] Data retrieved by Joint Genome Institute (JGI)—Integrated Microbial Genomes & Microbiomes system (https://img.jgi.doe.gov/).

Among the genomic indicators presented here, the genome size and predicted/protein-coding genes seem to be strain dependent, whereas the GC-content is species independent, which is in accordance with the previously reported ploidy variations. Some differences were also detected among *C. jadinii* strains considering their specific genetic features, namely, the NBRC 0988 strain has 6417 predicted open reading frames (ORFs) comprising 16 unique ORFs [5, 10, 24], whereas the CBS1600 strain has

5689 ORFs, including 64 unique ORFs [10]. In 2015, Rupp and colleagues revealed a close haploid consensus sequences sizes, 12.7 Mbp for *C. jadinii* and 12.8 Mbp for *C. utilis* with an overall sequence identity of 98% [16].

2.4 Morphology and Physiology

The *C. jadinii* microscopic view provided by Kurtzman et al. (2011) has shown the diversity of cell shapes and sizes [20, 34] after 10–30 days at 25 °C in 5% Malt Extract Agar media. The cell patterns of *C. jadinii* CBS 1600 varied from ellipsoidal to elongated occurring in single cells or in pairs. Some pseudohyphae forms were also detected with diameter balanced between (2.5–8.0) and (4.1–11.2) µm. Figure 2 shows *C. jadinii* DSM 2361 strain cultivated on YPD or Malt Extract Agar media for 3 (A and C) and 12 days (B and D) at 30 °C.

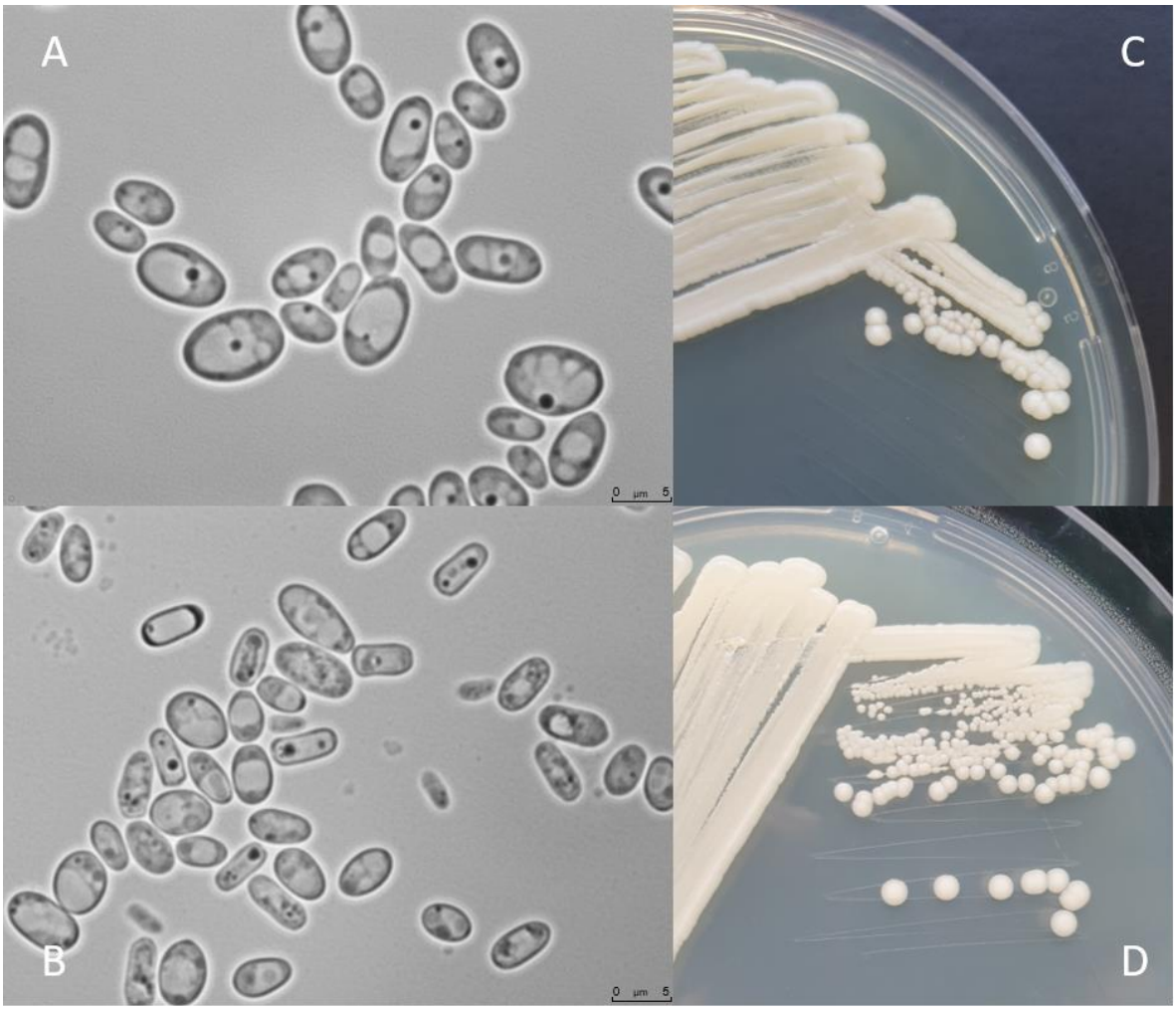


Figure 2. *C. jadinii* DSM 2361 morphological traits. Microscopic photographs of *C. jadinii* cells after 3 days (**A**) and 12 days (**B**) of growth at 30 °C on yeast extract-peptone-dextrose media. Scale bars are 5.0 µm. Macro-morphological features of *C. jadinii* after 3 days of growth in YPD (**C**) and Malt Extract Agar (**D**) media, at 30 °C.

The colonies are white, round, with a smooth texture, an entire margin, and a convex elevation trait (C and D). Cells present an ellipsoidal to elongated form, with a diameter between 5 and 7.5 µm (A and B). Yeast cell morphology can be tightly influenced by the environment. These modifications can affect the fermentation performance by inducing rheological changes that can influence mass and heat transfer alterations in the bioreactor [49]. However, in a study performed by Pinheiro et al. (2014), the CBS 621 strain cultured in a pressurized-environment triggered with 12 bar air pressure presented no significant differences in cell size and shape [35]. *C. jadinii* is a homothallic species and forming hat-shaped ascospores that can be present in a number of one to four in unconjugated deliquescent asci [34]. *Cyberlindnera* species can assimilate several compounds, namely, sugars and organic acids. The robust fermentation characteristics of *C. jadinii* allow growth in a diversity of substrates from biomass-derived

wastes, including hardwood hydrolysates from the pulp industry, being able to assimilate glucose, arabinose, sucrose, raffinose, and D-xylose [8-10, 50]. As previously mentioned, *C. jadinii* is a Crabtree-negative yeast, reaching higher cell yields under aerobic conditions [51-53] than Crabtree-positive species. The Crabtree-negative effect favors the respiration process over fermentation, enabling the development of phenotypes relevant for protein production [54]. This species has a high tolerance to elevated temperatures, being able to grow in a broad spectrum of temperatures from 19 to 37°C [37] and to tolerate long-term mild acid pHs (~3.5) [55]. Another relevant property is the ability to release proteins to the extracellular medium [56]. Significant lipase and protease content were achieved using a wild *C. jadinii* strain isolated from spoiled soybean oil, using solid-state fermentation [57, 58]. *C. jadinii* assimilates alcohols, acetaldehyde, organic acids, namely, monocarboxylates (DL-lactate), dicarboxylates (succinate), and tricarboxylates (citrate), sugar acids (D-gluconate), and various nitrogen sources comprising nitrate, ammonium hydroxide, as well as amino acids [8, 10, 12, 37]. A set of metabolic advantages, as the high metabolic flux in TCA cycle, the great amino acid synthesis ability, and strong protein secretion turns *C. jadinii* a yet underexplored host for bioprocesses. An incomplete understanding of genetics, metabolism, and cellular physiology combined with a lack of advanced molecular tools for genome edition and metabolic engineering manipulation of *C. jadinii* hampered its development for cell factory utilization.

2.5 MOLECULAR TOOLS FOR *C. JADINII* MANIPULATION

The genetic manipulation of *C. jadinii* has enabled the heterologous expression of various genes, resulting in the improvement of metabolic traits targeted for the optimization of exogenous product formation [8, 10, 44, 56]. The establishment of genetic transformation methods allowed the efficient production of enzymes, carotenoids compounds, and organic acids such as L-lactic acid [44, 56, 59, 60]. The development of an integrative transformation vector for the *C. jadinii* ATCC 9226 and *C. utilis* ATCC 9256 strains relied on a gene encoding a mutated ribosomal protein L41, conferring cycloheximide resistance as a dominant selection marker [25]. Dominant markers were used for cell transformation of *aph, hph, nat,* and *ble* genes, conferring resistance to G418, hygromycin B, nourseothricin, or zeocin, respectively, and the endogenous gene *YAP1*, conferring cycloheximide resistance [56, 60-62]. A multiple gene

disruption method based on the Cre-loxP system allowing the reuse of selection markers was developed for *C. jadinii* NBRC0988 [43]. Auxotrophic *ura3* strains were transformed with the expression vectors further integrated in the rDNA locus or in other chromosomal loci (e.g., *TDH3*). Single-plasmid integrations were reported for the *TDH3* and *HIS3* loci [16, 25, 56] as well as multiple plasmid integrations for the rDNA and *URA3* loci for the expression of heterologous genes in high copy number (up to 90 copies) [63]. High plasmid stability was observed mainly for integrations into the *URA3* and *HIS3* loci, in contrast to the integrations in the rDNA, while integrations at the *TDH3* locus were reported to be both stable and unstable [56, 64, 65]. Two chromosomal autonomously replicating sequences (ARS) were uncovered in *C. jadinii* ATCC 9226 and *C. utilis* ATCC 9256. Six plasmids harboring these ARS were obtained using a G418-resistance marker. The low copy number plasmids pCARS6 (CuARS1 region) and pCARS7 (CuARS2 region) presented the highest transformation efficiency [26]. A set of promoters were also explored to promote an efficient expression in *C. jadinii*, the *TDH3*, homolog of *TDH3* from *S. cerevisiae*, encoding glycerol-3-phosphate dehydrogenase, as well *PGK1* and *PDC1* promoters from *C. jadinii* NBRC0988 encoding the phosphoglycerate kinase and pyruvate decarboxylase, respectively [44, 64]. Furthermore, Kunigo *et al.* identified the highly xylose-inducible, glucose-repressed promoters of *XDH1* and *GXS1* genes, encoding a NAD-xylitol dehydrogenase and glucose/xylose symporter, respectively [56]. Promoters of the genes encoding the plasma membrane ATPase Pma1, Rpl29/Rpl31 ribosomal proteins, and Rpl41, as well as P14/P57 promoters from unknown chromosomal loci were used for the production of valuable products for the food industry, namely, for the secretion of heterologous proteins [25, 66-68]. CRISPR-Cas9 has quickly become the preferred targeted genome-editing technology for the genetic manipulation of yeasts [69], being extensively used in *S. cerevisiae*. The adaptation of the type II CRISPR/Cas system has been successfully used for the genetic manipulation of non-*Saccharomyces* species, such as *Y. lipolytica, K. pastoris, K. lactis, Schizosaccharomyces pombe,* and some pathogenic yeast species, such as *Candida albicans* and *Cryptococcus neoformans* [70, 71]. Recently, a patent application reported the development of a CRISPR/Cas9 system that was applied to knock out and insert exogenous genes in *C. jadinii* ATCC 22023 [72] . This strategy also uncovered the triploidy of this strain.

2.6 EMERGING BIOTECHNOLOGICAL APPLICATIONS

2.6.1 THERAPEUTIC APPLICATIONS

Being an edible yeast, *C. jadinii* has the potential to target the gastrointestinal tract in humans and animals [10, 55]. *C. jadinii*'s robust growth characteristics including insensitivity to low pH and temperatures up to 40 °C, allow its transit in the gastrointestinal tract without losing viability [55, 73]. Additionally, like other food constituents, intact/partially degraded *C. jadinii* cells may adhere to M cells in the small intestine. Upon this, they are translocated to antigen-presenting cells of Peyer's plaques or to other lymphoid tissue connected with the gastrointestinal tract [74, 75]. The ingestion of engineered *C. jadinii* cells, carrying a myelin oligodendrocyte glycoprotein antigen on its surface, promoted tolerance to self-antigens in a mouse model of the autoimmune multiple sclerosis (MS) disease [76]. *C. jadinii* cells expressing the immunodominant MOG_{35-55} epitope of a myelin protein on their surface, fused with the native fungal Gas1 cell wall protein, prevented the typical MS symptoms in this animal model [76]. The cell surface display of antigens by *C. jadinii* seems to modulate immune responses, either by suppression to combat autoimmune disease or through immune stimulation, enabling the creation of edible vaccines [73, 76, 77]. *C. jadinii* cells were also applied as probiotic agents against fungal infections (particularly oral candidiasis) as an antagonist to the relevant human fungal pathogens *Candida albicans* (strains SC5312, 10341, and GDH2346), *Aspergillus* sp., and *Fusarium* sp. [73]. The unidentified toxins secreted by *C. jadinii* act as antagonistic compounds conditioning the growth, systemic invasion, and disease caused by these fungal pathogens. Buerth et al. [10] reported the role of *C. jadinii* and also *W. farinosa* as antagonistic agents against *C. albicans* by inhibiting its growth and morphogenesis, proposing their exploitation for the formulation of new prebiotic compounds and strategies to tackle candidiasis. *C. jadinii* was also described to secrete heterologous proteins into the growth medium, including lipase B from *Candida antarctica* (CalB) [56]. The signal sequence of the enzyme invertase, one of the most predominant proteins of the *C. jadinii* secretome, allowed high secretion levels of recombinant CalB [5]. Lipases have a great potential application in substitution therapies, where metabolic deficiency is overcome by external administration of these enzymes in diseased conditions [78]. Furthermore, lipase activity can go from activation of tumor necrosis factor, having a relevant role over the treatment of malignant tumors, to the treatment of gastrointestinal disturbances, digestive allergies, or dyspepsias [78]. In addition, CalB is involved in enzymatic resolutions, desymmetrization, and aminolysis events with application in not only pharmaceutical but also biotech industry, having a role in polymer production [78,

79]. *C. jadinii* is also able to synthesize (R)-phenylacetylcarbinol (L-PAC), the pharmaceutical precursor for L-ephedrine and pseudoephedrine, relevant compounds used in the treatment of nasal congestion [80-83]. The function of 30 ATP-binding cassette transporters (ABC) transporters was studied by the amplification of predicted *C. jadinii* gDNA ORFs. The function of putative multidrug efflux pumps was evaluated by heterologous expression in *S. cerevisiae* ADΔ, a strain disrupted in seven of its major multidrug efflux pumps: Pdr5p, Pdr10p, Pdr15p, Snq2p, Pdr11p, Ycf1p, and Yor1p [84]. This strategy uncovered the mechanism of action of CjCdr1*, C. jadinii*'s closest homolog of the multidrug efflux pump *C. albicans* Cdr1. The characterization of *C. jadinii* multidrug efflux systems can turn *C. jadinii* into an appealing host for the development of novel antimicrobial agents [85], as it is imperative to understand the structure, function, and expression of multidrug efflux pumps in order to develop optimal novel antimicrobial agents.

2.6.2 BIOPRODUCTION OF VALUABLE COMPOUNDS USING COST-EFFECTIVE CARBON SOURCES

Recombinant *C. jadinii* strains were developed for the production of a variety of compounds from food supplements such as vitamins (biotin) [86], carotenoids (lycopene, β-carotene, and astaxanthin) [60, 68], to proteins (α-amylase, monellin) [64], antioxidant glutathione [87], polysaccharides (glucomannan) [88, 89], organic acids (L-lactic acid) [44, 59], and ribonucleic acids [5, 24]. Additionally, the production of secreted enzymes such as invertase and phospholipase B (NBRC 1086 strain) [27, 28] was also explored. Cells of *C. jadinii* DSM 2361 were successfully engineered for the secretion of *Penicillium simplicissimum* xylanase (PsXynA) to the culture medium [65] allowing cells to grow on xylan as the sole carbon source. Cells expressing the xylose reductase from *Candida shehatae* and the native xylitol dehydrogenase, in combination with further multiple site-directed mutations in coenzyme binding sites, resulted in the highest titer of 17.4 g/L of ethanol from 50 g/L of xylose in 20 h [90]. Organic acids present in industrial waste streams (e.g., acetic acid, propionic, or butyric acid) have been demonstrated to be suitable substrates for biomass production, reaching biomass yields varying from 30% to 40% in batch cultures, while in continuous cultures, an average of 44–55% was achieved [91, 92]. Despite its already important role as an industrial microorganism, further developments are still necessary to fully explore the biotechnological potential of this yeast.

2.6.3 INDUSTRIAL APPLICATIONS - A PATENT-VIEW

In recent years, the applications of *C. jadinii* were extended to cosmetic and health care products and to the chemical-process industry for the production of chiral chemicals, as well as for agriculture and wine making. In this last application, *C. jadinii* yeast was used in the production of loquat wine, being introduced after *S. cerevisiae* fermentation to reduce acid content and enhance the aroma [93, 94]. It is also used in another fermentation process, for the production of an alcohol-free fruit wine rich in lovastatin [95]. In the cosmetic industry, a β-D-glucan polysaccharide produced by *C. jadinii* was applied in formulations of several products, i.e., body lotions [96], sunscreen cream [97], facial cleanser [98], toner [99], eye cream [100], shampoo [101], body wash [102] and hand-care cream [103].This bioproduct is mainly added for its properties as a moisturizing agent and for conferring oxidation and radiation resistance. *C. jadinii* was also used for the efficient production of a recombinant uricase, active in humans and with greater stability and/or activity than naturally occurring enzymes [104]. This enzyme can be used for the treatment of hyperuricemia-related diseases or other human pathological symptoms. Considering the chemical industry, the bioproduction of methyl fluorophenyl methyl propionate was achieved with a developed reduction method using *C. jadinii* as a biocatalyst [105]. The obtained chemical, (2S,3S)-3-(4-fluorophenyl)-3-hydroxy-2-methyl methyl propionate, is reported to be produced in high yield and with a high level of the enantiomeric excess rate. The wide applicability of this chiral building-block chemical can go from the synthesis of chiral drugs, fine chemicals to pesticides [105]. In agriculture, a consortium of strains that include *C. jadinii* was incorporated in an organomineral granular fertilizer containing among other components fulvic acid, and a natural mineral component (activated natural siliceous zeolite-containing rock). Its properties allow the reduction in the amount of fertilizer introduced in the soil with a prolonging action improved [106]. As *C. jadinii* is capable of efficiently converting a cadmium form from contaminated soil, it is now being proposed for soil bioremediation [107]. In addition, a microbial soil conditioner for lithified soil was also developed involving a *C. jadinii* strain along with *Bacillus megaterium*, *Bacillus subtilis*, *Rhodopseudomonas palustris,* and *Azotobacter chroococcum* species. The full interaction among the aforementioned strains was claimed to improve the soil ecological environment and alter the soil lithiation event, thereby contributing for the purpose of turning the soil suitable for farming [108].

2.7 MEMBRANE TRANSPORTERS CHARACTERIZED IN *C. JADINII*

The production of biocompounds in high yields requires the optimization of several processes, including membrane transport of solutes to improve the entrance of substrates in the cell, exchange of products within cell organelles, and the efflux of metabolites to the extracellular medium, increasing the cell's tolerance to toxic final products, and decreasing downstream processing costs [109]. In *C. jadinii*, several plasma transporters were physiologically or genetically characterized (Figure 3).

In this yeast, copper (Cu^{2+}) transport is biphasic, energy-dependent, and relatively specific. Uptake is inhibited completely by 2,4-dinitrophenol (DNP), but carbonyl cyanide m-chlorophenylhydrazone (CCCP) had relatively little effect (Figure 3A) [33, 110]. The uptake follows a Michaelis-Menten kinetic with mean values for K_t = 3.1 μM and V_{max}, 0.5 nmol min^{-1} per mg (dry wt.) and has an optimal pH between 5 and 5.5. No exchange of K^+ for Cu^{2+} could be detected during Cu^{2+} uptake, and Cu^{2+} efflux from preloaded cells was not observed [33]. A high-affinity energy- and pH-dependent manganese (Mn^{2+}) importer was reported in *C. jadinii* (Figure 3B) [111]. With an apparent half-saturation constant K_t of 16.4 nM and a V_{max} of 1.01 nmol min^{-1} mg^{-1} dry wt., this transporter was shown to be highly specific for Mn^{2+} uptake. Efflux studies demonstrated that the metabolic exchange of labeled ^{54}Mn occurred to a small extent, being unaffected by a 100-fold molar excess of Mg^{2+}, Zn^{2+}, Ca^{2+}, Co^{2+}, Ni^{2+}, and Cu^{2+}, but inhibited 30–40% by a 1000-fold molar excess of Mg^{2+}, Zn^{2+}, Ca^{2+}, Co^{2+}, Ni^{2+} [110, 111]. The zinc uptake-system described in *C. jadinii* is energy-dependent and apparently unidirectional as no exchange occurs between intracellular accumulated ^{65}Zn and cold external Zn^{2+} (Figure 3C) [112]. This transporter exhibits a high-affinity for Zn^{2+} (K_m = 0.36 μM) with a V_{max} of 2.2 nmol min^{-1} per mg dry wt. of cells. The regulation of zinc homeostasis occurs either by altering the levels of a cytoplasmic zinc-sequestering macromolecule or by inhibition of zinc efflux through a membrane carrier [112-114]. The saturable and unidirectional sulfate transporter (Figure 3D) is pH-, temperature-, and energy-dependent with a K_m of 1.43 mM, being competitively inhibited by molybdate, selenate, thiosulfate, chromate, and sulfite [115, 116]. The activity of this sulfate transporter is controlled by the pool of external sulfur compounds as well as by the mitochondrial metabolism [116]. The proton-symporter for nitrate (Figure 3E) is repressed by ammonium [117-119]. Ali and Hipkin [118] reported that the addition of 1, 2, or 10 mM ammonium resulted in an inhibition of nitrate uptake close to 30%. Studies using 3,3′-dipropylthiadicarbocyanine, a fluorogenic probe used to detect and measure alterations in transmembrane potential, indicated that the proton-linked uptake of nitrate, amino acids, or glucose during energy metabolism tended to depolarize the plasma membrane of *C. jadinii* cells [117].

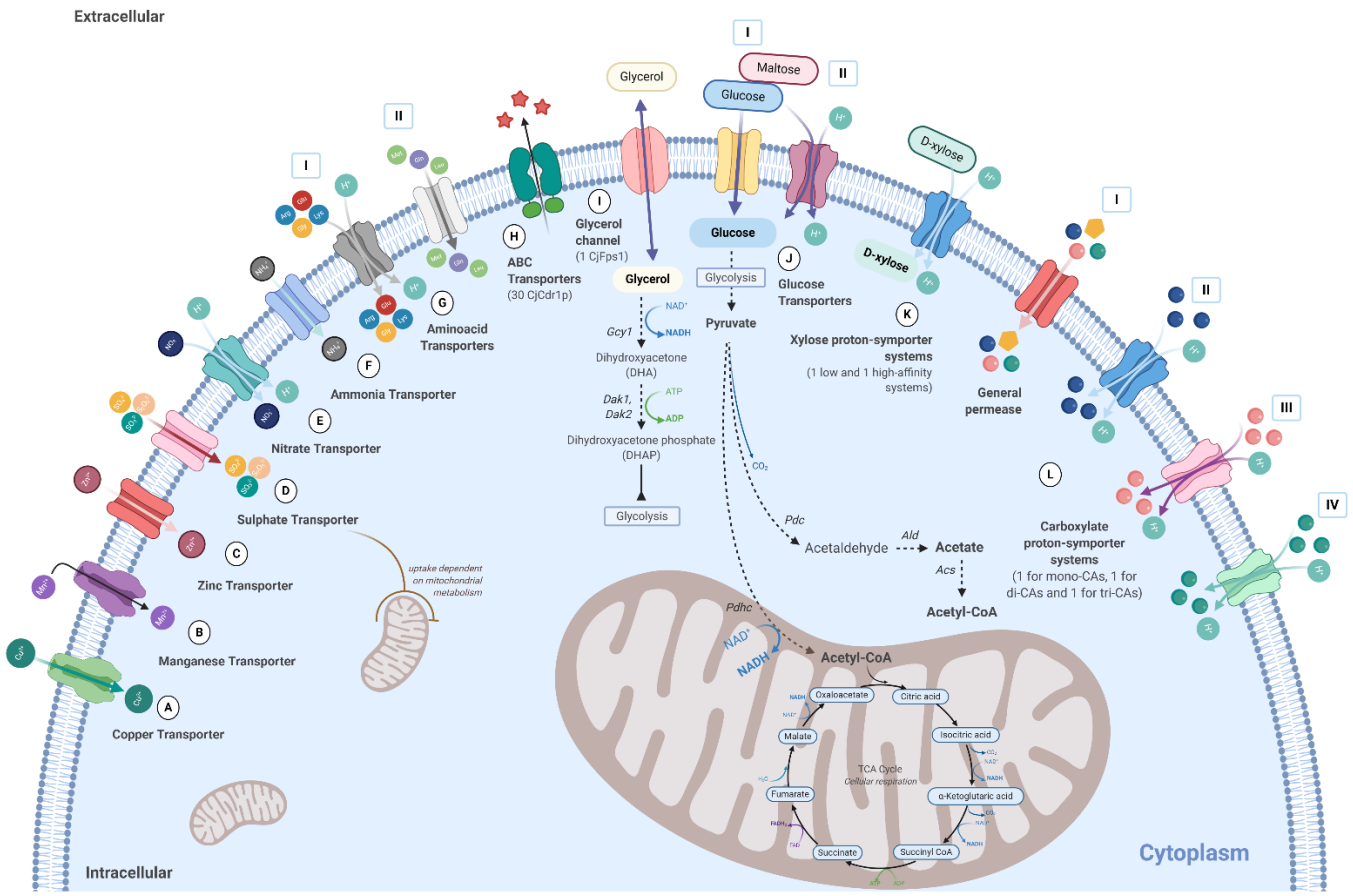


Figure 3. Plasma membrane transporters functionally characterized in *Cyberlindnera jadinii* (A–L). The main metabolic pathways of yeast general metabolism are also presented. Symbols represent the specificities uncovered for each of the protein-system: drugs—pink star, amino acids—yellow pentagon, monocarboxylic acids—blue filled circle, dicarboxylic acids—rose filled circle, and tricarboxylic acids—green filled circle. I-IV corresponds to different transporter systems with the same type of substrate G-I: amino acid proton symporter; G-II: facilitated diffusion of L-methionine, L-glutamine, and L-leucine; J-I glucose proton-symporter; J-II glucose facilitated diffusion; L-I facilitated diffusion of the undissociated form of carboxylic acids (general permease); L-II monocarboxylate proton symporter; L-III dicarboxylate-proton symporter; L-IV tricarboxylate-proton symporter. Initials stand for Mg^{2+}, magnesium; Mn^{2+}, manganese; SO_4^{2-}, sulfate; SO_3^{2-}, sulfite; $S_2O_3^{2-}$, thiosulfate; NH_4^+, ammonia; NO_3^-, nitrate; Cu^{2+}, copper; Zn^{2+}, zinc; H^+, proton; Gln, L-glutamine; Met, L-methionine; Leu, L-leucine; Arg, arginine; Lys, lysine; Gly, glycine; Glu, glutamate; *Gcy1*, glycerol dehydrogenase; *Dak1, Dak2* dihydroxyacetone kinases; *Pdhc,* pyruvate dehydrogenase complex; *Pdc,* pyruvate decarboxylase; *Ald*, aldolase; *Acs*, acetyl-CoA synthetase; *Acetyl-CoA*, acetyl coenzyme A; TCA cycle, tricarboxylic acid cycle; *NAD*, reduced form from nicotinamide adenine dinucleotide; *NADH,* nicotinamide adenine dinucleotide; *FADH,* flavin adenine dinucleotide; *FAD*, reduced form from flavin adenine dinucleotide; *ATP,* adenosine triphosphate; *ADP*, adenosine diphosphate; CO_2, carbon dioxide; H_2O, water.

An ammonia carrier (Figure 3F), revealed by spectrophotometry, presents a K_m of 1.0 μM of NH_4^+ [120]. Several amino acid proton-symporters (Figure 3G, I and II), namely, for arginine, lysine, glycine, and glutamate were uncovered in *C. jadinii* [121, 122]. A permease for L-glutamine (K_m = 410 μM), a high (K_m = 23 μM) and low-affinity (K_m = 495 μM) transporter for L-methionine and a high (K_m = 5.6 μM) and a low affinity (K_m = 530 μM) for L-leucine were also described [123].

Thirty ABC transporters (Figure 3H) from different transporter subfamilies were found by homology search on the genome of the strain NBRC0988 [85]. The expression of the *C. albicans* Pdr1 homolog CjCdr1 in the *S. cerevisiae* ADΔ strain conferred resistance to geneticin (75 μg), micafungin (40 μg), and nystatin (500 μg). These two proteins present similar substrate specificity, although CjCdr1 is more resistant to Rhodamine 6G [85]. The *C. jadinii* aquaglyceroporin CjFps1 (Figure 3I) shares 38% of identity with the *S. cerevisiae* Fps1. The heterologous expression of *CjFPS1* in a glycerol-consuming *S. cerevisiae* wild-type strain (CBS 6412-13A) promoted cellular growth improvement on glycerol as the sole carbon source [124].

Two different systems were found for glucose transport in *C. jadinii*, a proton-symporter and a facilitated diffusion (Figure 3J) [125-127], however, the respective genes remain unidentified. The high-affinity proton-symporter presents a K_m of 15 μM glucose, displays a stoichiometry of 1:1, and is partially constitutive, appearing in cells grown on gluconeogenic substrates such as lactate, ethanol, and glycerol. This transporter is repressed by high glucose concentration but is induced by glucose up to 0.7 mM [125], a behavior also found for the maltose-uptake system, indicating that both systems share a common glucose control pathway [125, 128]. Barnett and Sims (1976) proposed that the differences in glucose *vs.* maltose affinity can be due to allosteric mechanisms associated with a multimeric transporter or to the hysteretic behavior of a monomeric transporter [126], conditions that have not yet been solved as the correspondent genes(s) remain unidentified. The facilitated diffusion mechanism (Figure 3J-II) is found in cells growing on glucose at concentrations higher than 10 mM and presents complex kinetics of glucose transport whose K_m oscillates between 2 and 70 mM [125, 128]. It was also reported that uric acid enters *C. jadinii* cells by the glucose-dependent active transport [129].

The low- and the high-affinity systems for D-xylose transport (Figure 3K) with K_m values of 67.6 ± 3.2 and 1.9 ± 1.2 mM, respectively, act as proton-symporters with distinct modes of regulation. The starvation of glucose-grown cells decreases the K_m value of the low-affinity system (K_m = 10.5 ± 2.6 mM) [130]. The high-affinity system found during starvation requires protein synthesis and is inactive when cells are exposed to glucose, through a process independent of protein synthesis. Glucose and acetic acid inhibited both the high- and low-affinity xylose transport systems [130].

Four mediated transport systems for organic acids are described in *C. jadinii* (Figure 3L) [29-31]. The monocarboxylate proton symport, shared by the L- and D-lactate (K_m 0.06 mM), pyruvate (K_m 0.03 mM), propionate (K_m 0.05 mM), and acetate (K_m 0.1 mM) (Figure 3L-II) is active over a pH range of 3.0–6.0, with an optimum of activity at pH 5.0 [29]. The dicarboxylate-proton symporter (Figure 3L-III) is shared by L-malate (K_m 4.0 ± 0.5 μM), succinate (K_m 0.03 ± 0.01 mM), fumarate, oxaloacetate, and α-ketoglutarate [30, 31, 131]. The tricarboxylate-proton symporter (Figure 3L-IV) presents a K_m of 0.056 mM for citrate and is competitively inhibited by isocitric acid, while aconitic, tricarballylic, trimesic, and hemimellitic acids did not affect citrate uptake [30]. All these carboxylic acid transporters are inducible by the respective substrates, being subjected to glucose repression as well as by acid concentrations higher than 3% (w/v) [31].]. The facilitated diffusion for the undissociated form of the acids (Figure 3L-I), which is likely to operate as a general organic permease, is active at pH below 5.0. It accepts mono-, di-, and tricarboxylates as well as glycine and glutamic acid [31, 132]. The following kinetic parameters were obtained at pH 3.0: (a) V_{max} 0.516 nmol of malic acid s^{-1} per mg (dry wt. of cells) and K_m 1.529 ± 0.024 mM malic acid, (b) V_{max} 0.585 nmol of succinic acid s^{-1} per mg (dry wt. of cells) and K_m 1.789 ± 0.089 succinic acid, and (c) V_{max} of 1.14 nmol of s^{-1} per mg (dry wt. of cells) and K_m of 0.59 mM citric acid [29-31]. Despite being functionally characterized for several years, the genes encoding these proteins remain to be identified [133]. Recently, our group identified six genes homologous to Sc*ATO1* and six genes homologous to Sc*JEN1* (unpublished results). In *S. cerevisiae*, these two genes encode distinct monocarboxylate transporters [134, 135] and its homologs, present in bacteria, archaea and eukaryotes, are able to transport mono- and dicarboxylates [136-142]. The expression of membrane transporters allied to metabolic engineering tools is crucial to develop new and more efficient strains to produce bio-based compounds. Thus, increased knowledge of transporter proteins will enable the development of improved cell factories [109, 143, 144].

2.8 Conclusions and future perspectives

As biotechnological applications expand, it becomes necessary to explore novel expression hosts as more efficient and robust microbial cell factories are demanded. Over the years, *Cyberlindnera jadinii* has been widely explored as a source of single-cell protein, having the ability to produce vitamins (e.g., biotin), organic acids (e.g., glutamate), and proteins (e.g., enzymes). The capacity to utilize and degrade a great variety of carbon sources and its natural ability to produce significant compounds make *C. jadinii* an

attractive microorganism for industrial applications [5, 7, 8, 10]. Additionally, several features turn this yeast an ideal platform for biotechnological processes, like the higher level of tolerance to changes occurring during growth and multiplication conditions [5, 55]. However, *C. jadinii* is still lagging behind when compared to other non-*Saccharomyces* yeasts, mainly due to the inexistence of extensive knowledge on its metabolism, regulatory networks, and transport mechanisms. The genomic characterization of several strains is also necessary to reveal the genetic features underlying the existing interspecies variability, particularly between its teleomorph and anamorph state. The potential of this yeast as a therapeutic agent is due to its already known antagonistic effects on human pathogens and utilization as a probiotic agent. Its protein secretion system is another attractive feature for the heterologous expression of soluble proteins. Nonetheless, only with the improvement of advanced genetic manipulation systems and the development of synthetic biology tools will the full exploitation of *C. jadinii* biotechnological potential be achieved. These and other advances will certainly allow *C. jadinii* to become a robust microbial cell factory in an expanding era of metabolite bioproduction.

2.9 REFERENCES

1. Heux, S., Meynial-Salles, I., O'Donohue, M.,Dumon, C., *White biotechnology: state of the art strategies for the development of biocatalysts for biorefining.* Biotechnology advances, **2015**. 33(8): p. 1653-1670
2. Gong, Z., Nielsen, J.,Zhou, Y.J., *Engineering Robustness of Microbial Cell Factories.* Biotechnol J, **2017**. 12(10), 10.1002/biot.201700014
3. Sibirny, A., *Non-conventional Yeasts: from Basic Research to Application.***2019**: Springer
4. Martínez, E., Santos, J., Araujo, G., Souza, S., Rodrigues, R.,Canettieri, E., *Production of Single Cell Protein (SCP) from Vinasse.*, in *Fungal Biorefineries***2018**, Springer. p. 215-238, https://doi.org/10.1007/978-3-319-90379-8_10
5. Buerth, C., Heilmann, C., Klis, F., de Koster, C., Ernst, J.,Tielker, D., *Growth-dependent secretome of Candida utilis.* Microbiology, **2011**. 157(9): p. 2493-2503, 10.1099/mic.0.049320-0
6. Barford, J., *The technology of aerobic yeast growth.* Yeast biotechnology, **1987**: p. 200-230, https://doi.org/10.1007/978-94-009-3119-0_7
7. Miura, Y., Kettoku, M., Kato, M., Kobayashi, K.,Kondo, K., *High level production of thermostable alpha-amylase from Sulfolobus solfataricus in high-cell density culture of the food yeast Candida utilis.* Journal of molecular microbiology and biotechnology, **1999**. 1(1): p. 129-134
8. Bekatorou, A., Psarianos, C.,Koutinas, A., *Production of food grade yeasts.* Food Technology and Biotechnology, **2006**. 44(3): p. 407-415

9. Boze, H., Moulin, G.,Galzy, P., *Production of food and fodder yeasts.* Critical reviews in biotechnology, **1992**. 12(1-2): p. 65-86, https://doi.org/10.3109/07388559209069188
10. Buerth, C., Tielker, D.,Ernst, J., *Candida utilis and Cyberlindnera (Pichia) jadinii: yeast relatives with expanding applications.* Applied microbiology and biotechnology, **2016**. 100(16): p. 6981-6990, https://doi.org/10.1007/s00253-016-7700-8
11. Klein, R.,Faureau, M., *Chapter 8: The Candida species: biochemistry, molecular biology, and industrial applications*, in *Food Biotechnology: Microorganisms*, Y.H.H.a.G.G. Khachatourians, Editor **1995**. p. 297
12. Kurtzman, C., Robnett, C.,Basehoar-Powers, E., *Phylogenetic relationships among species of Pichia, Issatchenkia and Williopsis determined from multigene sequence analysis, and the proposal of Barnettozyma gen. nov., Lindnera gen. nov. and Wickerhamomyces gen. nov.* FEMS yeast research, **2008**. 8(6): p. 939-954, 10.1111/j.1567-1364.2008.00419.x
13. Suh, S., Blackwell, M., Kurtzman, C.,Lachance, M., *Phylogenetics of Saccharomycetales, the ascomycete yeasts.* Mycologia, **2006**. 98(6): p. 1006-1017
14. Barnett, J., *A history of research on yeasts 8: taxonomy.* Yeast, **2004**. 21(14): p. 1141-1193, 10.1002/yea.1154.
15. Minter, D., *Cyberlindnera, a replacement name for Lindnera Kurtzman et al., nom. illegit.* Mycotaxon, **2009**. 110: p. 473-476
16. Rupp, O., Brinkrolf, K., Buerth, C., Kunigo, M., Schneider, J., Jaenicke, S., Goesmann, A., Pühler, A., Jaeger, K.,Ernst, J., *The structure of the Cyberlindnera jadinii genome and its relation to Candida utilis analyzed by the occurrence of single nucleotide polymorphisms.* Journal of biotechnology, **2015**. 211: p. 20-30
17. Murray, P., Baron, E., Pfaller, M., Tenover, F.,Yolken, R., *Manual of microbiology*, **1995**, Washington,
18. Lodder, J.,Kreger-Van Rij, J.W., *The yeasts, a taxonomic study*, **1952**, North Holland Publishing Company: Amesterdam
19. Daniel, H., Lachance, M.,Kurtzman, C., *On the reclassification of species assigned to Candida and other anamorphic ascomycetous yeast genera based on phylogenetic circumscription.* Antonie van Leeuwenhoek, **2014**. 106(1): p. 67-84
20. Kurtzman, C., Johnson, C.,Smiley, M., *Determination of conspecificity of Candida utilis and Hansenula jadinii through DNA reassociation.* Mycologia, **1979**. 71(4): p. 844-847
21. Deak, T., *Characteristics and properties of Foodborne yeasts*, in *Handbook of food spoilage yeasts*, T. Deak, Editor **2007**, CRC press
22. Kurtzman, C., Fell, J.,Boekhout, T., *Definition, classification and nomenclature of the yeasts*, in *The yeasts: A Taxonomic Study***2011**, Elsevier. p. 3-5
23. Hawksworth, D., *Managing and coping with names of pleomorphic fungi in a period of transition.* IMA fungus, **2012**. 3(1): p. 15-24
24. Tomita, Y., Ikeo, K., Tamakawa, H., Gojobori, T.,Ikushima, S., *Genome and transcriptome analysis of the food-yeast Candida utilis.* PLoS One, **2012**. 7(5): p. e37226
25. Kondo, K., Saito, T., Kajiwara, S., Takagi, M.,Misawa, N., *A transformation system for the yeast Candida utilis: use of a modified endogenous ribosomal protein gene as a drug-resistant marker*

and ribosomal DNA as an integration target for vector DNA. Journal of bacteriology, **1995**. 177(24): p. 7171-7177

26. Ikushima, S., Minato, T.,Kondo, K., *Identification and application of novel autonomously replicating sequences (ARSs) for promoter-cloning and co-transformation in Candida utilis.* Bioscience, biotechnology, and biochemistry, **2009**. 73(1): p. 152-159
27. Belcarz, A., Ginalska, G., Lobarzewski, J.,Penel, C., *The novel non-glycosylated invertase from Candida utilis (the properties and the conditions of production and purification).* Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology, **2002**. 1594(1): p. 40-53
28. Fujino, S., Akiyama, D., Akaboshi, S., Fujita, T., Watanabe, Y.,Tamai, Y., *Purification and characterization of phospholipase B from Candida utilis.* Bioscience, biotechnology, and biochemistry, **2006**. 70(2): p. 377-386
29. Leão, C.,Van Uden, N., *Transport of lactate and other short-chain monocarboxylates in the yeast Candida utilis.* Applied microbiology and biotechnology, **1986**. 23(5): p. 389-393
30. Cássio, F.,Leão, C., *Low-and high-affinity transport systems for citric acid in the yeast Candida utilis.* Applied and Environmental Microbiology, **1991**. 57(12): p. 3623-3628
31. Cássio, F.,Leão, C., *A comparative study on the transport of L (-) malic acid and other short-chain carboxylic acids in the yeast Candida utilis: Evidence for a general organic acid permease.* Yeast, **1993**. 9(7): p. 743-752
32. Krahulec, J., Lišková, V., Boňková, H., Lichvariková, A., Šafranek, M.,Turňa, J., *The ploidy determination of the biotechnologically important yeast Candida utilis.* Journal of Applied Genetics, **2020**: p. 1-12
33. Ross, I.,Parkin, M., *Uptake of copper by Candida utilis.* Mycological Research, **1989**. 93(1): p. 33-37
34. Kurtzman, C., *Chapter 42 - Lindnera Kurtzman, Robnett & Basehoar-Powers (2008)* in *The yeasts: a taxonomic study*, C. Kurtzman, J. Fell, and T. Boekhout, Editors. **2011**, Elsevier. p. 521-543
35. Pinheiro, R., Lopes, M., Belo, I.,Mota, M., *Candida utilis metabolism and morphology under increased air pressure up to 12 bar.* Process Biochemistry, **2014**. 49(3): p. 374-379
36. Hittinger, C., Rokas, A., Bai, F., Boekhout, T., Gonçalves, P., Jeffries, T., Kominek, J., Lachance, M., Libkind, D.,Rosa, C., *Genomics and the making of yeast biodiversity.* Current opinion in genetics & development, **2015**. 35("Genomes and Evolution" Special Issue): p. 100
37. Riley, R., Haridas, S., Wolfe, K., Lopes, M., Hittinger, C., Göker, M., Salamov, A., Wisecaver, J., Long, T.,Calvey, C., *Comparative genomics of biotechnologically important yeasts.* Proceedings of the National Academy of Sciences, **2016**. 113(35): p. 9882-9887
38. Shen, X., Opulente, D., Kominek, J., Zhou, X., Steenwyk, J., Buh, K., Haase, M., Wisecaver, J., Wang, M.,Doering, D., *Tempo and mode of genome evolution in the budding yeast subphylum.* Cell, **2018**. 175(6): p. 1533-1545. e20
39. Marcet-Houben, M.,Gabaldón, T., *Evolutionary and functional patterns of shared gene neighbourhood in fungi.* Nature microbiology, **2019**. 4(12): p. 2383-2392
40. Shen, X., Zhou, X., Kominek, J., Kurtzman, C., Hittinger, C.,Rokas, A., *Reconstructing the backbone of the Saccharomycotina yeast phylogeny using genome-scale data.* G3: Genes, Genomes, Genetics, **2016**. 6(12): p. 3927-3939

41. Alves, R., Sousa-Silva, M., Vieira, D., Soares, P., Chebaro, Y., Lorenz, M., Casal, M., Soares-Silva, I.,Paiva, S., *Carboxylic Acid Transporters in Candida Pathogenesis.* MBio, **2020**. 11(3): p. e00156-20, 10.1128/mBio.00156-20
42. Thanh, V., Thuy, N., Huong, H., Hien, D., Hang, D., Anh, D., Hüttner, S., Larsbrink, J.,Olsson, L., *Surveying of acid-tolerant thermophilic lignocellulolytic fungi in Vietnam reveals surprisingly high genetic diversity.* Scientific reports, **2019**. 9(1): p. 3674, 10.1038/s41598-019-40213-5
43. Ikushima, S., Fujii, T.,Kobayashi, O., *Efficient gene disruption in the high-ploidy yeast Candida utilis using the Cre-loxP system.* Bioscience, biotechnology, and biochemistry, **2009**. 73(4): p. 879-884
44. Ikushima, S., Fujii, T., Kobayashi, O., Yoshida, S.,Yoshida, A., *Genetic engineering of Candida utilis yeast for efficient production of L-lactic acid.* Bioscience, biotechnology, and biochemistry, **2009**. 73(8): p. 1818-1824
45. Freel, K., Sarilar, V., Neuvéglise, C., Devillers, H., Friedrich, A.,Schacherer, J., *Genome Sequence of the Yeast Cyberlindnera fabianii (Hansenula fabianii).* Genome Announc, **2014**. 2(4): p. e00638-14, 10.1128/genomeA.00638-14
46. Schneider, J., Andrea, H., Blom, J., Jaenicke, S., Rückert, C., Schorsch, C., Szczepanowski, R., Farwick, M., Goesmann, A.,Pühler, A., *Draft genome sequence of Wickerhamomyces ciferrii NRRL Y-1031 F-60-10.* **2012**. 11(12): p. 1582–1583
47. Foury, F., Roganti, T., Lecrenier, N.,Purnelle, B., *The complete sequence of the mitochondrial genome of Saccharomyces cerevisiae.* FEBS letters, **1998**. 440(3): p. 325-331
48. Goffeau, A., Barrell, B., Bussey, H., Davis, R., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J., Jacq, C.,Johnston, M., *Life with 6000 genes.* Science, **1996**. 274(5287): p. 546-567
49. Coelho, M., Amaral, P.,Belo, I., *Yarrowia lipolytica: an industrial workhorse*, in *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology*, A. Méndez-Vilas, Editor **2010**, Formatex: Badajoz. p. 930-940
50. Andreeva, E., Pozmogova, I.,Rabotnova, I., *Principle growth indices of a chemostat Candida utilis culture resistant to acid pH values.* Mikrobiologiia, **1979**. 48(3): p. 481
51. Verduyn, C., *Physiology of yeasts in relation to biomass yields.*, in *Quantitative Aspects of Growth and Metabolism of Microorganisms***1992**, Springer. p. 325-353
52. Ordaz, L., Lopez, R., Melchy, O.,Torre, M., *Effect of high-cell-density fermentation of Candida utilis on kinetic parameters and the shift to respiro-fermentative metabolism.* Applied microbiology and biotechnology, **2001**. 57(3): p. 374-378
53. Dashko, S., Zhou, N., Compagno, C.,Piškur, J., *Why, when, and how did yeast evolve alcoholic fermentation?* FEMS Yeast Research, **2014**. 14(6): p. 826-832
54. Gomes, A., Carmo, T., Carvalho, L., Bahia, F.,Parachin, N., *Comparison of yeasts as hosts for recombinant protein production.* Microorganisms, **2018**. 6(2): p. 38
55. Wang, D., Zhang, J., Dong, Y., Wei, G.,Qi, B., *Glutathione is involved in physiological response of Candida utilis to acid stress.* Applied microbiology and biotechnology, **2015**. 99(24): p. 10669-10679
56. Kunigo, M., Buerth, C., Tielker, D.,Ernst, J., *Heterologous protein secretion by Candida utilis.* Applied microbiology and biotechnology, **2013**. 97(16): p. 7357-7368

57. Moftah, O., Grbavčić, S., Žuža, M., Luković, N., Bezbradica, D.,Knežević-Jugović, Z., *Adding value to the oil cake as a waste from oil processing industry: production of lipase and protease by Candida utilis in solid state fermentation.* Applied biochemistry and biotechnology, **2012**. 166(2): p. 348-364
58. Grbavčić, S., Dimitrijević-Branković, S., Bezbradica, D., Šiler-Marinković, S.,Knežević, Z., *Effect of fermentation conditions on lipase production by Candida utilis.* Journal of the Serbian Chemical Society, **2007**. 72
59. Tamakawa, H., Ikushima, S.,Yoshida, S., *Efficient production of L-lactic acid from xylose by a recombinant Candida utilis strain.* Journal of bioscience and bioengineering, **2012**. 113(1): p. 73-75
60. Shimada, H., Kondo, K., Fraser, P., Miura, Y., Saito, T.,Misawa, N., *Increased Carotenoid Production by the Food Yeast Candida utilis through Metabolic Engineering of the Isoprenoid Pathway.* Applied and Environmental Microbiology, **1998**. 64(7): p. 2676-2680
61. Boňková, H., Osadská, M., Krahulec, J., Lišková, V., Stuchlík, S.,Turňa, J., *Upstream regulatory regions controlling the expression of the Candida utilis maltase gene.* Journal of biotechnology, **2014**. 189: p. 136-142
62. Iwakiri, R., Noda, Y., Adachi, H.,Yoda, K., *Isolation of the YAP1 homologue of Candida utilis and its use as an efficient selection marker.* Yeast, **2005**. 22(13): p. 1079-1087
63. Tamakawa, H., Mita, T., Yokoyama, A., Ikushima, S.,Yoshida, S., *Metabolic engineering of Candida utilis for isopropanol production.* Applied microbiology and biotechnology, **2013**. 97(14): p. 6231-6239
64. Kondo, K., Miura, Y., Sone, H.,Kobayashi, K., *High-level expression of a sweet protein, monellin, in the food yeast Candida utilis.* Nature biotechnology, **1997**. 15(5): p. 453
65. Kunigo, M., Buerth, C.,Ernst, J., *Secreted xylanase XynA mediates utilization of xylan as sole carbon source in Candida utilis.* Applied microbiology and biotechnology, **2015**. 99(19): p. 8055-8064
66. Miura, Y., Kondo, K., Shimada, H., Saito, T., Nakamura, K.,Misawa, N., *Production of lycopene by the food yeast, Candida utilis that does not naturally synthesize carotenoid.* Biotechnology and bioengineering, **1998**. 58(2-3): p. 306-308
67. Iwakiri, R., Noda, Y., Adachi, H.,Yoda, K., *Isolation and characterization of promoters suitable for a multidrug-resistant marker CuYAP1 in the yeast Candida utilis.* Yeast, **2006**. 23(1): p. 23-34
68. Miura, Y., Kondo, K., Saito, T., Shimada, H., Fraser, P.,Misawa, N., *Production of the carotenoids lycopene, β-carotene, and astaxanthin in the food yeast Candida utilis.* Applied and Environmental Microbiology, **1998**. 64(4): p. 1226-1229
69. Sander, J.,Joung, J., *CRISPR-Cas systems for editing, regulating and targeting genomes.* Nature biotechnology, **2014**. 32(4): p. 347
70. Cai, P., Gao, J.,Zhou, Y., *CRISPR-mediated genome editing in non-conventional yeasts for biotechnological applications.* Microbial cell factories, **2019**. 18(1): p. 63
71. Stovicek, V., Holkenbrink, C.,Borodina, I., *CRISPR/Cas system for yeast genome engineering: advances and applications.* FEMS Yeast Research, **2017**. 17(5)

72. Zehuan, L., Jianghai, L.,Yu, D., *Method for verifying feasibility of CRISPR-Cas9 system for knocking out Candida utilis target gene*, **2019**: China,
73. Mukherjee, P., Chandra, J., Retuerto, M., Sikaroodi, M., Brown, R., Jurevic, R., Salata, R., Lederman, M., Gillevet, P.,Ghannoum, M., *Oral mycobiome analysis of HIV-infected patients: identification of Pichia as an antagonist of opportunistic fungi.* PLoS pathogens, **2014**. 10(3): p. e1003996
74. De Jesus, M., Rodriguez, A., Yagita, H., Ostroff, G.,Mantis, N., *Sampling of Candida albicans and Candida tropicalis by langerin-positive dendritic cells in mouse Peyer's patches.* Immunology letters, **2015**. 168(1): p. 64-72
75. Thévenot, J., Cordonnier, C., Rougeron, A., Le Goff, O., Nguyen, H., Denis, S., Alric, M., Livrelli, V.,Blanquet-Diot, S., *Enterohemorrhagic Escherichia coli infection has donor-dependent effect on human gut microbiota and may be antagonized by probiotic yeast during interaction with Peyer's patches.* Applied microbiology and biotechnology, **2015**. 99(21): p. 9097-9110
76. Buerth, C., Mausberg, A., Heininger, M., Hartung, H., Kieseier, B.,Ernst, J., *Oral tolerance induction in experimental autoimmune encephalomyelitis with Candida utilis expressing the immunogenic MOG35-55 peptide.* PLoS One, **2016**. 11(5): p. e0155082
77. Feng, T.,Elson, C., *Adaptive immunity in the host–microbiota dialog.* Mucosal immunology, **2011**. 4(1): p. 15
78. Singh, R., Singh, T.,Singh, A., *Chapter 9 - Enzymes as Diagnostic Tools*, in *Advances in Enzyme Technology*, R.S. Singh, et al., Editors. **2019**, Elsevier. p. 225-271, https://doi.org/10.1016/B978-0-444-64114-4.00009-1
79. Żymańczyk-Duda, E., Brzezińska-Rodak, M., Klimek-Ochab, M., Duda, M.,Zerka, A., *Yeast as a Versatile Tool in Biotechnology.* Yeast: Industrial Applications, **2017**: p. 1
80. Khan, T.,Daugulis, A., *Application of solid–liquid TPPBs to the production of L-phenylacetylcarbinol from benzaldehyde using Candida utilis.* Biotechnology and bioengineering, **2010**. 107(4): p. 633-641
81. Tripathi, C., Agarwal, S.,Basu, S., *Production of L-phenylacetylcarbinol by fermentation.* Journal of fermentation and bioengineering, **1997**. 84(6): p. 487-492
82. Shin, H.,Rogers, P., *Biotransformation of benzeldehyde to L-phenylacetylcarbinol, an intermediate in L-ephedrine production, by immobilized Candida utilis.* Applied microbiology and biotechnology, **1995**. 44(1-2): p. 7-14
83. Rogers, P., Shin, H.,Wang, B., *Biotransformation for L-ephedrine production.*, in *Biotreatment, Downstream Processing and Modelling***1997**, Springer. p. 33-59
84. Lamping, E., Monk, B., Niimi, K., Holmes, A., Tsao, S., Tanabe, K., Niimi, M., Uehara, Y.,Cannon, R., *Characterization of Three Classes of Membrane Proteins Involved in Fungal Azole Resistance by Functional Hyperexpression in Saccharomyces cerevisiae.* Eukaryotic Cell, **2007**. 6(7): p. 1150-1165, 10.1128/ec.00091-07
85. Watanasrisin, W., Iwatani, S., Oura, T., Tomita, Y., Ikushima, S., Chindamporn, A., Niimi, M., Niimi, K., Lamping, E.,Cannon, R., *Identification and characterization of Candida utilis multidrug efflux transporter CuCdr1p.* FEMS yeast research, **2016**. 16(4): p. fow042
86. Hong, Y., Chen, Y., Farh, L., Yang, W., Liao, C.,Shiuan, D., *Recombinant Candida utilis for the production of biotin.* Applied microbiology and biotechnology, **2006**. 71(2): p. 211

87. Li, Y., Wei, G.,Chen, J., *Glutathione: a review on biotechnological production.* Applied microbiology and biotechnology, **2004**. 66(3): p. 233-242
88. Kogan, G., Šandula, J.,Šimkovicová, V., *Glucomannan from Candida utilis.* Folia microbiologica, **1993**. 38(3): p. 219-224
89. Ruszova, E., Pavek, S., Hajkova, V., Jandova, S., Velebny, V., Papezikova, I.,Kubala, L., *Photoprotective effects of glucomannan isolated from Candida utilis.* Carbohydrate research, **2008**. 343(3): p. 501-511
90. Tamakawa, H., Ikushima, S.,Yoshida, S., *Ethanol production from xylose by a recombinant Candida utilis strain expressing protein-engineered xylose reductase and xylitol dehydrogenase.* Bioscience, biotechnology, and biochemistry, **2011**. 75(10): p. 1994-2000
91. Hui, Y.H.,Khachatourians, G., *Food biotechnology: microorganisms*, ed. Y.H. Hui and G. Khachatourians**1994**: Wiley-Interscience
92. Maugeri-Filho, F.,Goma, G., *SCP production from organic acids with Candida utilis.* Revista de Microbiologia, **1988**. 19(4): p. 446-452
93. Luqiang, H., Yi, Z., Liping, Z., Zebin, G., Shaoxiao, Z., Yuting, T.,Baodong, Z., *Application of Candida utilis to loquat wine making and making method for loquat wine*, in *China***2012**: China,
94. Bangxiang, X., Bin, H., Jing, H., Qianwei, Y., Qingyin, H.,Yi, Z., *Dry loquat wine and making method thereof*, in *China***2015**,
95. Yingchun, C., Xiaoyuan, J., Yaokun, C., Qichang, C., Yinglong, C.,Yingbing, C., *Preparation method of alcohol-free fruit wine being rich in lovastatin*, **2019**: China,
96. Xia, M.,Lantian, L., *Body lotion containing Candida utilis beta-D-glucan*, **2018**: China,
97. Xia, M.,Lantian, L., *Sunscreen cream with Candida utilis beta-D-glucan*, **2018**: China,
98. Xia, M.,Lantian, L., *Facial cleanser containing Candida utilis beta-D-glucan*, **2018**: China,
99. Xia, M.,Lantian, L., *Toner containing beta-D-glucan of Candida utilis*, **2018**: China,
100. Xia, M.,Lantian, L., *Eye cream containing beta-D-glucan of Candida utilis*, **2018**: China,
101. Xia, M.,Lantian, L., *Shampoo containing Candida utilis beta-D-glucan*, **2018**: China,
102. Xia, M.,Lantian, L., *Body wash with Candida utilis beta-D-glucan*, **2018**: China,
103. Xia, M., Lantian, L., Li, S.,Yan, H., *Hand care cream with Candida utilis beta-D-glucan*, **2018**: China,
104. Shentao, L., Hai, P.,Shuang, L., *Preparation method of recombinant uricase*, **2015**: China,
105. Junhong, L., Xin, G., Fengxia, S., Pengcheng, L.,Meihong, Z., *Candida utilis reduction method for producing methyl fluorophenyl methyl propionate*, **2015**: China,
106. Firgatovich, B.F., Gennadevich, G.Y., Yurevna, G.E.,Mikhajlovna, K.E., *Organomineral granular fertilizer*, R. Federation, Editor **2020**: Russia,
107. Lianyang, B., Xueduan, L., Huidan, J., Hongwei, L.,Yan, D., *Application of microorganism bacterium agent capable of efficiently converting form of cadmium in contaminated soil biologically*, **2020**: China,
108. Guochen, C., Yanli, G., Changna, C.,Yue, Z., *Microbial soil conditioner for lithified soil*, **2015**: China,

109. Soares-Silva, I., Ribas, D., Sousa-Silva, M., Azevedo-Silva, J., Rendulić, T.,Casal, M., *Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications.* FEMS microbiology letters, **2020**
110. Parkin, M.,Ross, I., *Uptake of copper and manganese by the yeast, Candida utilis.* Microbios letters, **1985**. 29(115-116): p. 115-120
111. Parkin, M.,Ross, I., *The specific uptake of manganese in the yeast Candida utilis.* Microbiology, **1986**. 132(8): p. 2155-2160
112. Lawford, H., Pik, J., Lawford, G., Williams, T.,Kligerman, A., *Hyperaccumulation of zinc by zinc-depleted Candida utilis grown in chemostat culture.* Canadian Journal of Microbiology, **1980**. 26(1): p. 71-76
113. Failla, M., Benedict, C.,Weinberg, E., *Accumulation and storage of Zn^{2+} by Candida utilis.* Microbiology, **1976**. 94(1): p. 23-36
114. Failla, M.,Weinberg, E., *Cyclic accumulation of zinc by Candida utilis during growth in batch culture.* Microbiology, **1977**. 99(1): p. 85-97
115. García, M., Benitez, J., Delgado, J.,Kotyk, A., *Isolation of sulphate transport defective mutants of Candida utilis: Further evidence for a common transport system for sulphate, sulphite and thiosulphate.* Folia microbiologica, **1983**. 28(1): p. 1-5
116. Benitez, J., Alonso, A., Delgado, J.,Kotyk, A., *Sulphate transport in Candida utilis.* Folia microbiologica, **1983**. 28(1): p. 6-11
117. Eddy, A.,Hopkins, P., *The putative electrogenic nitrate-proton symport of the yeast Candida utilis. Comparison with the systems absorbing glucose or lactate.* Biochemical journal, **1985**. 231(2): p. 291-297
118. Ali, A.,Hipkin, C., *Nitrate assimilation in Candida nitratophila and other yeasts.* Archives of Microbiology, **1986**. 144(3): p. 263-267
119. Barnett, J., *A history of research on yeasts 13. Active transport and the uptake of various metabolites.* Yeast, **2008**. 25(10): p. 689-731
120. De Koning, W., Zwart, K.,Harder, W., *A spectrophotometric method for the determination of the kinetic parameters of NH_4^+ uptake by yeast cells.* FEMS microbiology letters, **1984**. 23(2-3): p. 257-262
121. Eddy, A., Philo, R., Earnshaw, P.,Brocklehurst, R., *Some common aspects of active solute transport in yeast and mouse ascites tumour cells.*, in *Biochemistry of Membrane Transport***1977**, Springer. p. 250-260
122. Eddy, A., Seaston, A., Gardner, D.,Hacking, C., *The thermodynamic efficiency of cotransport mechanisms with special reference to proton and anion transport in yeast.* Annals of the New York Academy of Sciences, **1980**. 341(1): p. 494-509
123. Robins, R.,Davies, D., *The role of glutathione in amino-acid absorption. Lack of correlation between glutathione turnover and amino-acid absorption by the yeast Candida utilis.* Biochemical journal, **1981**. 194(1): p. 63-70
124. Klein, M., Islam, Z., Knudsen, P., Carrillo, M., Swinnen, S., Workman, M.,Nevoigt, E., *The expression of glycerol facilitators from various yeast species improves growth on glycerol of Saccharomyces cerevisiae.* Metabolic engineering communications, **2016**. 3: p. 252-257

125. Peinado, J., Cameira-Dos-Santos, P.,Loureiro-Dias, M., *Regulation of glucose transport in Candida utilis.* Microbiology, **1989**. 135(1): p. 195-201
126. Barnett, J.,Sims, A., *A note on the kinetics of uptake of d-glucose by the food yeast, Candida utilis.* Archives of Microbiology, **1976**. 111(1-2): p. 193-194
127. Flores, C., Rodríguez, C., Petit, T.,Gancedo, C., *Carbohydrate and energy-yielding metabolism in non-conventional yeasts.* FEMS microbiology reviews, **2000**. 24(4): p. 507-529
128. Broek, P., Gompel, A., Luttik, M., Pronk, J.,Leeuwen, C., *Mechanism of glucose and maltose transport in plasma-membrane vesicles from the yeast Candida utilis.* Biochemical journal, **1997**. 321(2): p. 487-495
129. Roush, A.,Domnas, A., *Induced biosynthesis of uricase in yeast.* Science, **1956**. 124(3212): p. 125-126
130. Kilian, S., Prior, B.,Du Preez, J., *The kinetics and regulation of M-xylose transport in Candida utilis.* World Journal of Microbiology and Biotechnology, **1993**. 9(3): p. 357-360
131. Saayman, M., Van Vuuren, H., Van Zyl, W.,Viljoen-Bloom, M., *Differential uptake of fumarate by Candida utilis and Schizosaccharomyces pombe.* Applied microbiology and biotechnology, **2000**. 54(6): p. 792-798
132. Roush, A., Questiaux, L.,Domnas, A., *The active transport and metabolism of purines in the yeast, Candida utilis.* Journal of Cellular and Comparative Physiology, **1959**. 54(3): p. 275-286
133. Casal, M., Paiva, S., Queirós, O.,Soares-Silva, I., *Transport of carboxylic acids in yeasts.* FEMS Microbiology Reviews, **2008**. 32(6): p. 974-94, 10.1111/j.1574-6976.2008.00128.x
134. Paiva, S., Devaux, F., Barbosa, S., Jacq, C.,Casal, M., *Ady2p is essential for the acetate permease activity in the yeast Saccharomyces cerevisiae.* Yeast, **2004**. 21(3): p. 201-210
135. Casal, M., Paiva, S., Andrade, R., Gancedo, C.,Leão, C., *The lactate-proton symport of Saccharomyces cerevisiae is encoded by JEN1.* Journal of bacteriology, **1999**. 181(8): p. 2620-2623
136. Soares-Silva, I., Schuller, D., Andrade, R., Baltazar, F., Cássio, F.,Casal, M., *Functional expression of the lactate permease Jen1p of Saccharomyces cerevisiae in Pichia pastoris.* Biochemical Journal, **2003**. 376(3): p. 781-787
137. Queirós, O., Pereira, L., Paiva, S., Moradas-Ferreira, P.,Casal, M., *Functional analysis of Kluyveromyces lactis carboxylic acids permeases: heterologous expression of KlJEN1 and KlJEN2 genes.* Current genetics, **2007**. 51(3): p. 161-169
138. Soares-Silva, I., Ribas, D., Foskolou, I., Barata, B., Bessa, D., Paiva, S., Queirós, O.,Casal, M., *The Debaryomyces hansenii carboxylate transporters Jen1 homologues are functional in Saccharomyces cerevisiae.* FEMS yeast research, **2015**. 15(8)
139. Sá-Pessoa, J., Paiva, S., Ribas, D., Silva, I., Viegas, S., Arraiano, C.,Casal, M., *SATP (YaaH), a succinate–acetate transporter protein in Escherichia coli.* Biochemical journal, **2013**. 454(3): p. 585-595
140. Sá-Pessoa, J., Amillis, S., Casal, M.,Diallinas, G., *Expression and specificity profile of the major acetate transporter AcpA in Aspergillus nidulans.* Fungal Genetics and Biology, **2015**. 76: p. 93-103
141. Ribas, D., Soares-Silva, I., Vieira, D., Sousa-Silva, M., Sá-Pessoa, J., Azevedo-Silva, J., Viegas, S., Arraiano, C., Diallinas, G., Paiva, S., Soares, P.,Casal, M., *The acetate uptake transporter family*

motif "NPAPLGL (M/S)" is essential for substrate uptake. Fungal Genetics and Biology, **2019**. 122: p. 1-10

142. Rossi Goncalves, I., Dantas, R.C.C., Ferreira, M.L., Batistao, D., Gontijo-Filho, P.P.,Ribas, R.M., *Carbapenem-resistant Pseudomonas aeruginosa: association with virulence genes and biofilm formation.* Braz J Microbiol, **2017**. 48(2): p. 211-217, 10.1016/j.bjm.2016.11.004
143. Wang, J., Lin, M., Xu, M.,Yang, S., *Anaerobic fermentation for production of carboxylic acids as bulk chemicals from renewable biomass*, in *Anaerobes in Biotechnology***2016**, Springer. p. 323-361
144. Kell, D., Swainston, N., Pir, P.,Oliver, S., *Membrane transporter engineering in industrial biotechnology and whole cell biocatalysis.* Trends in biotechnology, **2015**. 33(4): p. 237-246

CHAPTER III

Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications

Adapted from:

Soares-Silva, I.[1,2], Ribas, D.[1,2], Sousa-Silva, M.[1,2], Azevedo-Silva, J.[1,2], Rendulic, T.[1,2], Casal, M.[1,2] (2020). Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications. *FEMS Microbiology Letters*, *367*(15), fnaa118. doi.org/10.1093/femsle/fnaa118

[1] Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, Portugal
[2] Centre of Molecular and Environmental Biology (CBMA), Department of Biology, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal

Personal contribution: I was involved in the manuscript writing of section 3.4 Transporter expression for the optimization of organic acid efflux.

Note: Part of this chapter was previously reported in the Doctoral thesis of David Ribas, which is entitled "Characterization of carboxylate transporters: functional and structural studies for the improvement of cell factories" and is available in the "Repositorium" of the University of Minho.

CHAPTER III

Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications

ABSTRACT

Organic acids such as monocarboxylic acids, dicarboxylic acids or even more complex molecules such as sugar acids, have displayed great applicability in the industry as these compounds are used as platform chemicals for polymer, food, agricultural and pharmaceutical sectors. Chemical synthesis of these compounds from petroleum derivatives is currently their major source of production. However, increasing environmental concerns have prompted the production of organic acids by microorganisms. The current trend is the exploitation of industrial biowastes to sustain microbial cell growth and valorize biomass conversion into organic acids. One of the major bottlenecks for the efficient and cost-effective bioproduction is the export of organic acids through the microbial plasma membrane. Membrane transporter proteins are crucial elements for the optimization of substrate import and final product export. Several transporters have been expressed in organic acid-producing species, resulting in increased final product titers in the extracellular medium and higher productivity levels. In this review, the state of the art of plasma membrane transport of organic acids is presented, along with the implications for industrial biotechnology.

KEYWORDS: Industrial biotechnology; Cell factories; Carboxylic acids; Transporter proteins; Permease

3.1 INTRODUCTION

Organic acids are an essential group of platform chemicals produced by microbes. Most of the organic acids produced industrially are used in the food industry. Currently, the major source of production of these compounds is the chemical synthesis from petroleum derivatives. Nonetheless, several organic acids are already industrially generated *via* microbial cell-factories, including succinic, lactic, citric, gluconic and acetic acid [1]. Microbial production of organic acids comprises several membrane transport processes, mostly controlled by membrane proteins, namely substrate import, transport of metabolites between organelles and product export. These processes, critical for the bioproduction of organic acids, are the major topic of this review.

3.2 MICROBIAL CELL FACTORIES IN THE PRODUCTION OF BIO-BASED ORGANIC ACIDS

The global organic acids market was valued at 17 billion euros in 2016. The forecasts predict an annual growth of 8.3%, which should reach 30 billion euros by the year 2023 [2] with an impact in a broad range of industrial sectors (Figure 1).

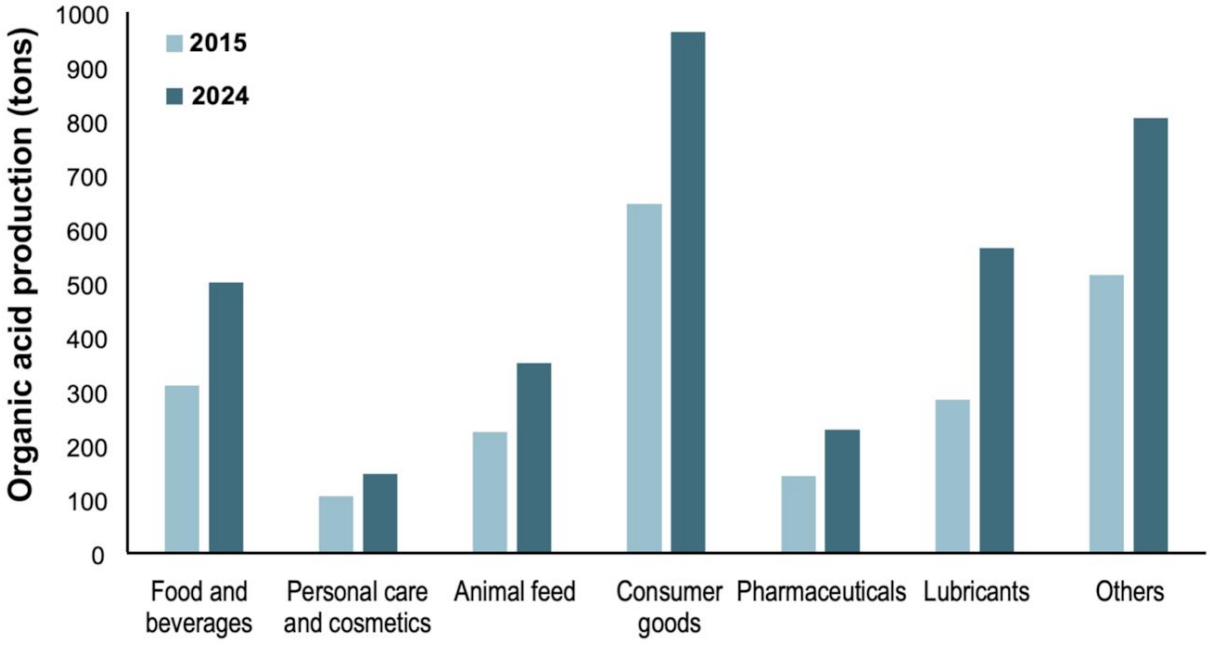


Figure 1. Annual production of organic acids according to market sector/application in 2015 and estimated growth for 2024 (adapted from https://www.alliedmarketresearch.com/organic-acids-market).

The most significant contributions to this growth are the use of renewable resources, the rising market, and the growing demand from developing countries for bio-based organic acids. The industry of microbial organic acid production is under continuous development to increase cell factory productivity, yields and range of products. Along with classical strain engineering approaches and adaptive laboratory evolution (ALE) strategies, the development of recombinant DNA technology together with synthetic biology has allowed the rational engineering of organic acid-producing microbes. Meanwhile, beyond the classical industrial microbes, such as *Escherichia coli, Saccharomyces cerevisiae* and *Corynebacterium glutamicum*, we have witnessed the appearance of other species isolated from natural sources, displaying a high capacity to generate organic acids [3-5].

3.2.1 MEMBRANE TRANSPORTERS AS TOOLS FOR THE IMPROVEMENT OF CELL FACTORIES

Nowadays, most industrial microorganisms are metabolically engineered to produce specific products and/or to metabolize specific substrates. For decades, the transport mechanisms and energetics of these compounds were underestimated, and most attention was given to the engineering of metabolic pathways. Recently, the scientific community and biotech companies focused their efforts on transporter engineering, envisaging the development and improvement of microbial cell factories [6-8].

The microbial fermentation industry faces two major bottlenecks in the production line: the first relates to product accumulation and toxicity inside the cell and low product titers in the extracellular medium; the second is associated with cell factory capacity to assimilate carbon and energy sources for product biosynthesis. These two obstacles in microbial fermentation can be overcome by transport activity of endogenous or exogenous membrane transporters, importers and/or exporters, as well as their genetic manipulation regarding expression levels and generation of mutant alleles with increased transport capacity or altered specificity [6, 7, 9].

3.3 MEMBRANE TRANSPORTERS IN THE IMPORT OF RENEWABLE SUBSTRATES

The most used renewable feedstocks in the bio-based industrial production of carboxylic acids are cheese whey, lignocellulosic biomass, glycerol and pectin-rich wastes [1]. Lactose is the most abundant sugar in cheese whey; xylose along with arabinose are abundant sugars in lignocellulosic hydrolysates; pectin-rich wastes, such as citrus and beet pulp, are rich in galacturonic acid [1, 10]. Frequently, these sugars are

hardly assimilated and metabolized by microbial cell factories. This bottleneck is associated with the lack of membrane proteins or extracellular enzymes capable of respectively taking up or converting these substrates into assimilated forms. Therefore, membrane transporters are engineered in microbial cell factories to increase the efficiency of substrate influx, by altering transporter specificity, affinity and/or capacity, ultimately leading to improved production yields [6, 10, 11].

Extensive efforts were devoted to the use of genetically engineered *E. coli* as a sustainable platform for the production of industrially important compounds, including organic acids (for a review see Chen *et al.* 2013; Yang *et al.* 2020) [12, 13]. One of the few examples reporting the engineering of membrane transporters for the uptake of substrates to improve organic acid production in *E. coli* is the study by Wu, *et al.* (2018). Here, the increased production of catechol (gluconic acid precursor) was achieved after the co-expression of the catechol biosynthetic pathway and the transporter CouP, which enabled the uptake of aromatic compounds present in lignin. In a more recent work, Khunnonkwao *et al.* (2018) described the improvement of succinic acid production upon re-engineering of xylose transporters in *E. coli*. The filamentous fungus *Aspergillus niger* is the oldest industrial workhorse due to its great robustness to extreme acid environments and better fitness for industrial fermentation [16]. Genome design and metabolic engineering approaches to optimize the *A. niger* cell factory for industrial citric acid production can be found in Tong *et al.* (2019). However, few transporter engineering strategies for substrate influx were described in this species. The endogenous low-affinity glucose transporter Lgt1 was expressed in the citrate-producing *A. niger* H915-1 strain, under the control of the low-pH-inducible promoter Pgas, leading to enhanced glucose absorption during the acid producing period and enhanced citrate production [17]. Yeasts are considered one of the most promising groups of industrial microorganisms to produce organic acids and ethanol. Thus hereafter, emphasis will be given on the functional expression of xylose, arabinose, lactose, glycerol and galacturonic acid transporters in attempts to improve organic acid biorefinery applications in yeasts.

XYLOSE

The heterologous expression of xylose transporters in *S. cerevisiae* for the conversion of this lignocellulosic sugar into ethanol was extensively reported in the literature. More than 80 heterologous xylose transporters or putative xylose transporters have already been expressed in *S. cerevisiae* (for a review see Moysés *et al.* 2016) [18]. Significant examples include *SUT1*, *SUT2*, *XUT1*, *XUT3* (Xyp33), *XUT4*, Xyp29 (*STL12*), *SUT3* (Xyp37) from *Scheffersomyces stipitis*, *GXF1* from *Candida intermedia*, At5g59250 from *Arabidopsis thaliana*, An29-2 and An25 from *Neurospora crassa*, *xtrD* from

A. nidulans, MgT05196 from *Meyerozyma guilliermondii* and *Xylh* from *Debaryomyces hansenii.* More than 80% of these putative transporters or annotated sugar transporters were not functional in *S. cerevisiae*, probably due to misfolding or improper localization [18]. The ones that were properly expressed in the HXT-null *S. cerevisiae* strain displayed activity for xylose transport, but the majority showed a preference for glucose over xylose.

ARABINOSE

Along with xylose, arabinose is the second most abundant pentose sugar present in lignocellulosic hydrolysates. *S. cerevisiae* possesses endogenous arabinose transporters (Hxt9 and Hxt10, Gal2), Gal2 being the most prominent [19]. However, the consumption of this sugar pentose is inefficient and inhibited by glucose, since Gal2 exhibits a much lower affinity for arabinose than for glucose or galactose [20, 21]. To improve the kinetics of arabinose uptake, transporter genes from other organisms have been functionally expressed in engineered arabinose-metabolizing *S. cerevisiae* strains. Two characterized L-arabinose transporters, LAT-1 from *N. crassa* and MtLAT-1 from *Myceliophthora thermophila* were expressed in this yeast [22]. The expression of both transporters in a *S. cerevisiae* strain containing a L-arabinose metabolic pathway resulted in a much faster L-arabinose utilization, greater biomass accumulation, and higher ethanol production than the control strain. Expression of the *Pc*AraT arabinose transporter from *Penicillium chrysogenum* enabled growth on arabinose in the presence of glucose in a *S. cerevisiae* strain deficient in hexose phosphorylation and able to metabolize arabinose [23]. This transporter showed significantly higher affinity for arabinose compared to the endogenous Gal2 and had far less pronounced inhibition of arabinose uptake in the presence of glucose or xylose.

LACTOSE

A recombinant *S. cerevisiae* flocculent strain heterologously expressing the β-galactosidase *LAC4* and lactose permease *LAC12* genes from *Kluyveromyces lactis* was used for ethanol production from lactose in a continuous culture operation [24]. This approach resulted in an ethanol production yield of 0.51 g/g of lactose. In continuous fermentation conditions, this engineered *S. cerevisiae* strain reached an ethanol productivity of 11 g/L/h, which represented a 7-fold rise compared with values of ethanol productivity from lactose previously mentioned in the literature [24]. The expression of the lactose transporter CDT-1, the intracellular β-galactosidase GH1–1 from *N. crassa* and the lactate dehydrogenase ldhA from *Rhizopus oryzae* in *S. cerevisiae* allowed the production of lactic acid from lactose, cow's milk, or

whey [25]. A lactic acid yield of 0.358 g/g lactose was achieved from a Yeast extract-Peptone medium containing about 80 g/L whey.

GLYCEROL

Heterologous expression of glycerol facilitators (Fps1 homologs) from non-*Saccharomyces* yeast species that show superior growth on glycerol, e.g. *Cyberlindnera jadinii*, *Komagataella pastoris*, *Pachysolen tannophilus* and *Yarrowia lipolytica*, improved the maximum specific growth rates of the *S. cerevisiae* CBS 6412–13A strain by 30–40% in synthetic glycerol medium [26]. Conversely, no improvement was visible after the overexpression of the endogenous *S. cerevisiae FPS1* gene. Deletion of the endogenous glycerol/H^+ symporter *STL1* did not impair the superior growth of these strains. A significant increase in ethanol production (from none to 8.5 g/L) was obtained upon the expression of the heterologous aquaglyceroporin *CjFPS1* from *C. jadinii* in the strain CBS DHA, which catabolizes glycerol via the dihydroxyacetone (DHA) pathway [27]. Further optimizations, including the reduction of oxygen availability in the shake flask cultures, increased the ethanol titer up to 15.7 g/L.

GALACTURONIC ACID

The introduction of the galacturonic acid transporter GAT1 from *N. crassa*, along with a fungal reductive pathway for galacturonic acid catabolism (*gaaA*, *gaaC* and *gad* from *A. niger* and *lgd1* from *Trichoderma reesei*), allowed the engineered *S. cerevisiae* strain to metabolize galacturonic acid [28]. This strain was only able to catabolize galacturonic acid when a co-substrate was added (fructose). Tracing experiments with ^{13}C-galacturonic acid revealed its conversion into glycerol [28]. Recently, the expression of another galacturonic acid transporter, GatA from *A. niger*, allowed a more rapid consumption of this acid [29]. The involvement of endogenous yeast hexose transporters Hxt1–7 and Gal2 in the glucose-inhibited uptake of undissociated galacturonic acid in acidic conditions was also uncovered. Expression of glucose-insensitive GatA coupled with uronate dehydrogenase allowed the engineered *S. cerevisiae* strain to produce 8 g/L of meso-galactaric acid from citrus peel waste supplemented with additional glucose [29].

3.4 TRANSPORTER EXPRESSION FOR THE OPTIMIZATION OF ORGANIC ACID EFFLUX

Regarding the microbial production of organic acids, several reports pointed to the effectiveness and contribution of membrane transporters for product efflux. The organic acid transporters that thus far have been functionally characterized in yeast, fungi and bacteria are listed in Table1.

Table 1. Microbial organic acid transporter proteins (experimentally verified). Table includes the transporter family, the species, the Transport Classification Database (TC number), the number of transmembrane segments (TMS), description of the transporter activity and references.

Family	Transporter protein	Species	TC number	TMS*	Description	References
2-HCT	CimH (YxkJ)	*Bacillus subtilis*	2.A.24.2.4	10	Electroneutral L-Malate/Citrate:H^+ symporter; Citrate (K_m 10 μM), L-Malate (K_m 1.5 mM)	[75]
	MaeN (YufR)		2.A.24.2.3	11	Malate:Na^+ symporter	[76]
	CitS	*Klebsiella pneumoniae*	2.A.24.1.1	12	Sodium:Citrate symporter	[77]
	MleP	*Lactococcus lactis*	2.A.24	13	Sodium:Citrate symporter; Malate (K_m 0.46 mM), Lactate (K_m 4.6 mM)	[78-80]
	CitP (CitN)		2.A.24.3.1	12	Electrogenic Citrate:L-Lactate exchanger Citrate (K_m 56 μM), Malate (K_m 0.1 mM), Lactate (K_m 26 mM)	[32]
	CitP	*Leuconostoc mesenteroides*	2.A.24.3.2	13	Citrate:Lactate antiporter Citrate, Citramalate, Malate, 2-Hydroxyisobutyrate and Lactate	[31, 81]
AAEx	SucE	*Corynebacterium glutamicum*	2.A.81.1.3	9	Succinate exporter	[82]
AceTr	SatP	*Escherichia coli*	2.A.96.1.1	6 in hexameric channels	Acetate, Lactate and Succinate transporter; Acetate (K_m 1.24 mM) and Succinate (K_m 1.18 mM)	[83]
	AceP	*Methanosarcina acetivorans*	2.A.96.1	6	Acetate transporter; Acetate (K_m 0.49 mM)	[59]
	Ato1 (Ady2)	*Saccharomyces cerevisiae*	2.A.96.1.4	6	Acetate permease; Acetate (K_m 0.84 mM), Lactate, Propionate and Formate	[58, 59]
	Gpr1	*Yarrowia lipolytica*	2.A.96.1.2	6	Acetate transporter; Acetate (K_m 0.95 mM)	[58, 59, 84]
Bestrophin	Best1 (AN2251)	*Aspergillus nidulans*	1.A.46.2.1	4	Ca^{2+}-activated anion-selective channel; Citrate, Propionate, Benzoate and Sorbate	[85, 86]

CitMHS	CitM	*Bacillus subtilis*	2.A.11.1.1	9	Citrate or D-Isocitrate divalent metal:H^+ symporter (K_m 35-63 μM), Metal (in order of preference): Mg^{2+}, Mn^{2+}, Ni^{2+}, Zn^{2+} and Co^{2+}	[87, 88]
	CitH (CitN)		2.A.11.1.2	11	Citrate divalent metal:H^+ symporter (K_m 35-63 μM); Metal (in order of preference): Ca^{2+}, Ba^{2+} and Sr^{2+}	[88]
	YRAO		2.A.11.1.5	13	Citrate:H^+ symporter	[89]
	CitH	*Corynebacterium glutamicum*	2.A.11.1.6	10	Divalent cation:citrate; Citrate transport in complex with Ca^{2+} or Sr^{2+}	[90]
DAACS	Dct	*Aspergillus carbonarius*	2.A.23.1.7	9	Fumarate, L-Aspartate: symporter	[91]
		Actinobacillus succinogenes	2.A.23	12	Malate and Citrate exporter	[41]
	DctA	*Bacillus subtilis*	2.A.23.1.6	8	Dicarboxylate:H^+ symporter: Succinate (K_m 2.6 μM), Fumarate	[92, 93]
		Corynebacterium glutamicum	2.A.23	7	Dicarboxylate:H^+ symporter; L-Malate (K_m 736 μM), Fumarate (K_m 232 μM), Succinate (K_m 218 μM), Oxaloacetate and Glyoxylate	[94]
		Escherichia coli	2.A.23.1.7	8	Aerobic dicarboxylate transporter: Succinate (K_m 25 μM), Orotate, Fumarate and L- and D-Malate	[95-97]
DASS	DccT (DcsT)	*Corynebacterium glutamicum*	2.A.47.1.12	14	Aerobic sodium dicarboxylate transporter Succinate (K_m 30 μM), Fumarate (K_m 79 μM), Malate (K_m 360 μM) and Oxaloacetate	[94, 98, 99]
	TtdT (YgjE)	*Escherichia coli*	2.A.47.3.3	12	L-Tartrate:Succinate antiporter L-Tartrate (K_m 700 μM) uptake, Succinate (K_m 400 μM) efflux	[100]
	CitT		2.A.47.3.2	13	Citrate:Succinate antiporter Citrate uptake and efflux of Succinate, Fumarate and Tartrate.	[101]
	SLC13		2.A.47	14	Citrate exporter	[41]

	SdcA	*Actinobacillus succinogenes*	2.A.47.5.3	13	Dicarboxylate: Na^+ transporter Fumarate (K_m 536 μM) and Succinate (K_m 389 μM) uptake	[102]
	SdcS	*Staphylococcus aureus*	2.A.47.1.11	14	Dicarboxylate: Na^+ symporter Succinate (K_m 7 mM), Malate (K_m 8 mM), Fumarate (K_m 15 mM), Aspartate and α-Ketoglutarate transporter	[103, 104]
Dcu	DcuA	*Escherichia coli*	2.A.13.1.1	11	Anaerobic antiporter of Aspartate, Malate, Fumarate and Succinate Uptake and efflux of Fumarate	[45, 105]
	DcuB		2.A.13.1.2	11	Anaerobic antiporter of Aspartate, Malate, Fumarate and Succinate; Uptake and efflux of Fumarate and Citrate exporter	[41, 45, 100, 105]
DcuC	DcuC	*Escherichia coli*	2.A.61.1.1	12	Anaerobic electroneutral C_4-dicarboxylate exchanger; Dicarboxylate-proton symporter; Citrate exporter	[41, 45-47]
DHA1	CexA	*Aspergillus niger*	2.A.1.2	12	Citrate exporter	[62]
FNT	FocA	*Escherichia coli*	1.A.16.1.1	6 in pentameric channels (PDB 3KCU)	Exporter of Acetate (K_m 23.9 mM), Lactate (K_m 96 mM) and Pyruvate (K_m 11.6 mM); Uptake/efflux of Formate (K_m 11.7 mM)	[106, 107]
	PfFNT	*Plasmodium falciparum*	1.A.16.2.7	6	Lactate:H^+ symporter D-Lactate, Pyruvate, Acetate and Formate	[108, 109]
LctP	LldP	*Escherichia coli*	2.A.14.1.1	12	Lactate permease L-Lactate, D-Lactate and Glycolate	[110]
	GlcA (YghK)		2.A.14.1.2	13	Glycolate permease L-Lactate, D-Lactate and Glycolate	[110, 111]
	LutP	*Bacillus subtilis*	2.A.14.1.3	14	Lactate permease	[112]
		Bacillus coagulans	2.A.14.1	14	Lactate permease	[113]
MFS	MfsA	*Aspergillus terreus*	2.A.1	12	Dicarboxylate transporter; Itaconate exporter	[52, 53]

	Itp1	*Ustilago maydis*	2.A.1	12	Itaconate exporter	[114]
MHS	Dehp2	*Burkholderia caribensis* MBA4	2.A.1.6.11	12	Acetate/haloacid transporter Acetate, Chloroacetate, Bromoacetate, 2-Chloropropionate; low affinity to Glycolate, Lactate and Pyruvate	[115, 116]
	Deh4p		2.A.1.6.13	11	Acetate/Monochloroacetate (haloacid) permease Acetate (K_m 5.5 μM) and Monochloroacetate (K_m 9 μM)	[115, 116]
NhaC	MleN (YqkI)	*Bacillus subtilis*	2.A.35.1.2	10	Malate:Lactate antiporter coupled with proton uptake and sodium efflux $Malic^{2-}$-$2H^+$: Na^+-$Lactate^-$	[117]
SHS	Jen1	*Saccharomyces cerevisiae*	2.A.1.12.2	12	Lactate/Pyruvate:H^+ symporter Acetate (K_m 4.8 mM), Lactate (K_m 0.2 mM), Propionate, Pyruvate (K_m 0.7 mM), Selenite	[56, 57, 118, 119]
	CaJen1	*Candida albicans*	2.A.1.12	10	Monocarboxylate permease Lactate (K_m 0.33 mM) Pyruvate and Propionate	[120]
	CaJen2		2.A.1.12	10	Dicarboxylate permease Succinate (K_m 0.49 mM), Malate (K_m 0.12 mM) Affinity for the sugar acids Gluconate, Xylarate and Mucate	[35, 121]
	DH17	*Debaryomyces hansenii*	2.A.1.12	12	Malate permease (K_m 0.27 mM)	[122]
	DH18		2.A.1.12	12	Succinate permease (K_m 0.31 mM)	
	DH24		2.A.1.12	12	Succinate permease (K_m 0.16 mM)	
	DH27		2.A.1.12	12	Acetate permease (K_m 0.94 mM)	
	KlJen1	*Kluyveromyces lactis*	2.A.1.12	12	Monocarboxylate permease Lactate (K_m 2.08 mM), Pyruvate	[123, 124]
	KlJen2		2.A.1.12	11	Dicarboxylate permease Malate (K_m 0.15 mM), Succinate (K_m 0.11 mM), Fumarate; Affinity for the sugar acids Gluconate and Saccharate	[59, 123, 124]

SSS	MctC	*Corynebacterium glutamicum*	2.A.21.7.3	13	Acetate/Propionate:H^+ symporter Pyruvate (K_m 250 μM), Acetate (Km 31 μM), Propionate (K_m 9 μM)	[125]
	ActP (YjcG)	*Escherichia coli*	2.A.21.7.2	13	Acetate (K_m 5.4 μM) and Glyoxylate coupling ion: proton transporter, with affinity for Tellurite	[126, 127]
	ActP1	*Rhodobacter capsulatus*	2.A.21.7.4	14	Acetate permease Acetate (K_m 1.89 mM), Pyruvate, Lactate, Tellurite (K_m 163 μM)	[128-130]
ST	KgtP (WitA)	*Escherichia coli*	2.A.1.6.2	12	α -Ketoglutarate (Oxoglutarate):symporter Arabinose exporter	[131]
SulP	DauA (YchM)	*Escherichia coli*	2.A.53.3.11	11	Aerobic Succinate transporter Succinate (K_m 0.56 mM), Aspartate and Fumarate	[97]
TDT	Mae1	*Schizosaccharomyces pombe*	2.A.16.2.1	10	Malate:H^+ symporter Oxaloacetate, Malonate, Succinate, Fumarate and Thio-malate; Exporter of Fumarate, Succinate and Malate	[41, 132-134]
	Ssu1 Ssu2	*Ustilago trichophora*	2.A.16.	9	Dicarboxylate transporter Malate and α-Ketoglutarate	[39]
TRAP-T	DctPQM	*Rhodobacter capsulatus*	2.A.56.1.1	12 (DctM) + 4 (DctQ) + receptor	Tripartite dicarboxylate: H^+ symporter Malate (K_d 8.4 μM) competitively inhibited by Fumarate (K_i 2 μM) and Succinate (K_i 8 μM)	[135]
TTT	TctABC	*Corynebacterium glutamicum*	2.A.80.1.4	12(TctA) + 4(TctB) + 1(TctC)	Citrate transport in complex with Ca^{2+} or Mg^{2+}	[90]

* Number of TMS predicted with the TMHMM software (http://www.cbs.dtu.dk/services/TMHMM/) or verified

TCDB Families: 2-HCT – 2-Hydroxycarboxylate Transporter; AAEx – Aspartate:Alanine Exchanger; AceTr – Acetate Uptake Transporter; Bestrophin – Anion Channel-forming Bestrophin; CitMHS – Citrate-Mg^{2+}:H^+ (CitM) Citrate-Ca^{2+}:H^+ (CitH) Symporter; DAACS – Dicarboxylate/Amino Acid:Cation (Na^+ or H^+) Symporter; DASS –

Divalent Anion:Na^+ Symporter; Dcu – C_4-Dicarboxylate Uptake; DcuC – C_4-dicarboxylate Uptake C; DHA1 – Drug:H^+ Antiporter-1; FNT – Formate-Nitrite Transporter; LctP – Lactate Permease; MFS – Major Facilitator Superfamily; MHS – Metabolite:H^+ Symporter; NhaC – Na^+:H^+ Antiporter; SHS – Sialate:H^+ Symporter; SSS – Solute:Sodium Symporter; ST – Sugar transporter; SulP – Sulfate Permease; TDT – Telurite-resistance/Dicarboxylate Transporter; TRAP-T – Tripartite ATP-independent Periplasmic Transporter; TTT – Tripartite Tricarboxylate Transporter. na – not annotated at TC Database.

Among these transporters, a great majority belong to the 2-hydroxycarboxylate transporter (2-HCT) (TC 2.A.24), Divalent Anion:Na^+ Symporter (DASS) (TC 2.A.47) and Sialate:H^+ symporter (SHS) (TC 2.A.1.12) families. Members of 2-HCT are involved in the transport of di- and tricarboxylate substrates (malate/citrate uptake) with either Na^+ or H^+ as the co-substrate and precursor/product exchangers [30]. Some members mediate the transport of monocarboxylate substrates, 2-hydroxyisobutyrate and D-lactate [31, 32]. The integral membrane proteins of the DASS family are conserved from bacteria to humans. DASS proteins typically mediate the coupled uptake of Na^+ ions and dicarboxylate, tricarboxylate, or sulfate (for a review see Lu 2019). A total of six members of DASS present a broad range of substrates from mono-, di- to tricarboxylates. The SHS transporter family, despite only having two distinct family members, the sialic acid transporter NanT, and the lactate/pyruvate:H^+ symporter orthologues, is the one with most members characterized in yeast, accepting mainly mono- and dicarboxylates as well as sugar acids [34, 35]. Next, we will highlight the transporters that had an impact on the improvement of cell factories.

GLUTAMIC ACID

The bacterium *Corynebacterium glutamicum* is used in microbial biotechnology for the production of glutamic acid. Glutamate efflux, triggered by increased mechanic tension, was associated with the activation of the channel NCgl1221 (MscCG) [36, 37], belonging to the MscS Family (TCDB 1.A.23 The Small Conductance Mechanosensitive Ion Channel). However, the activation mechanism of *C. glutamicum* mechanosensitive channels is not fully understood (for a review see Nakayama *et al.* 2019). Several channels of this family are described to play a critical role in product efflux of other amino acids, namely lysine, isoleucine, threonine, methionine, and others [6, 11].

MALIC ACID

In the natural malic acid producer *Ustilago trichophora* RK089, the overexpression of two endogenous malate transporter genes improved the production yields by 54% [39]. The overexpression of pyruvate carboxylase (*pyc*) together with two malate dehydrogenases (*mdh1*, *mdh2*), and two malate transporters (*ssu1, ssu2*) was carried out in a laboratory-evolved *U. trichophora* strain that reached an extracellular malate titer of 120 g/L. Wild-type *S. cerevisiae* strains produce low levels of malate. High yield production of malic acid required the elimination of alcoholic fermentation, which in this yeast occurs under fully aerobic conditions when high concentrations of sugar are present [40]. The metabolic engineering of a *S. cerevisiae* strain allowed an increase up to 10-fold of malic acid titer relative to the control strain [40].This was achieved through the engineering of a glucose-tolerant, C_2-independent pyruvate decarboxylase-negative strain, together with: i)

the overexpression of the endogenous pyruvate carboxylase encoded by *PYC2*, ii) the overexpression of an allele of the peroxisomal malate dehydrogenase *MDH3* gene targeted to the cytoplasm, and iii) the functional expression of the *S. pombe* malate transporter *Sp*Mae1. These modifications *per se* improved malate production, and the combination of all genetic modifications reached a malate titer of approximately 59 g/L [40]. Recently, seven dicarboxylic acid transporters were expressed in a *S. cerevisiae* strain engineered for dicarboxylic acid production [41]. In this work, the expression of the *Sp*Mae1 homologous gene from *Aspergillus carbonarius, Ac*Dct, increased malate titer up to 12-fold. Upon *Sp*Mae1 expression, the following titers were obtained for malate, (8 fold-4.3 g/L) succinate (3 fold-2.6 g/L) and fumarate (5 fold-0.33 g/L).

FUMARIC ACID

The overexpression of the *S. cerevisiae* mitochondrial succinate-fumarate carrier *SFC1* gene enhanced fumarate export and production by 47.6 % in this yeast [42]. A *S. cerevisiae* strain engineered for the production of fumarate, deleted in the fumarase *FUM1* gene, and expressing the *RoPYC* pyruvate carboxylase gene of *R. oryzae* and the endogenous *SFC1* gene, resulted in a titer of 1.7 g/L of fumarate in batch culture.

Using a different approach, fumarate production in *Candida glabrata* was improved by overexpressing the Sfc1 mitochondrial carrier in combination with the heterologous expression of *Sp*Mae1[43]. This work established the metabolic engineering of the tricarboxylic acid cycle in *C. glabrata* to construct the oxidative pathway for fumarate production. Thus, a set of genetic modifications to manipulate the oxidative pathway was applied in the α-ketoglutarate dehydrogenase complex, succinyl-CoA synthetase, and succinate dehydrogenase. As a result, the *C. glabrata* producer strain reached a fumarate titer of 8.24 g/L. Overexpression of the argininosuccinate lyase gene led to a fumarate increase up to 9.96 g/L. The additional expression of two dicarboxylic acid transporters, Sfc1 and *Sp*Mae1, allowed an improvement of fumarate production (15.76 g/L) [43]. In *E. coli* a set of C_4-dicarboxylate transporters from different organisms were cloned in a fumaric acid-producing strain deleted in the genes *fumABC*, *frdABCD*, *iclR* and *arcA*, to evaluate their impact on the production of this acid [44]. It was the overexpression of the endogenous transporters *DcuB*, an anaerobic fumarate–succinate antiporter [45], and *DcuC*, a C_4-dicarboxylate carrier that promotes succinate efflux during glucose fermentation [46], that displayed the highest impact on the production of fumaric acid. These lead to an increase of fumaric acid yield by 48.5% and 53.1%, respectively. In fed-batch fermentation culture, the fumaric acid producer strain overexpressing the *dcuB* gene reached 9.42 g/L of fumaric acid after 50 hours [44].

SUCCINIC ACID

Modulation of the simultaneous expression of *E. coli* transporter genes *dcuB* and *dcuC* led to a 34% increase of succinic acid titer in an engineered *E. coli* strain [47]. In this work, four *E. coli* Dcu C_4-dicarboxylate transporters were exploited for succinate export. Single deletion of *dcuA* or *dcuD* did not affect the export of this organic acid, while *dcuB* and *dcuC* deletion led to 15% and 11% decrease of succinate extracellular titer, respectively. The combined deletion of *dcuB* and *dcuC* genes resulted in a 90% decrease of succinate titer. As a result, a ribosome binding site library was investigated to modulate and increase the co-expression of *dcuB* and *dcuC*, which led to a 34% increase of succinate titer produced by *E. coli* [47]. In *C. glutamicum*, the overexpression of the endogenous succinate exporter, SucE, increased succinate yield in an engineered strain [48]. A dual-route for anaerobic succinate production was devised, involving the reconstruction of the glyoxylate pathway by overexpressing isocitrate lyase, malate synthase, and citrate synthase. This succinate producer strain reached a succinate yield of 1.34 mol/mol of glucose. The additional overexpression of the endogenous succinate exporter, SucE, increased succinate yield to 1.43 mol/mol of glucose. In anaerobic fed-batch fermentation, the *C. glutamicum* succinate producer strain overexpressing SucE led to a titer of 109 g/L succinate.

ITACONIC ACID

Expression of two *Aspergillus terreus* genes encoding organic acid transporters, *mttA* and *mfsA*, increased itaconic acid production in an *A. niger* strain expressing the cis-aconitate decarboxylase [49, 50]. MttA is a mitochondrial tricarboxylic acid transporter that preferentially transports cis-aconitate instead of citrate [51]. MfsA is an itaconate plasma membrane exporter [52, 53]. The strains expressing *mttA* or *mfsA* displayed an increased itaconic acid (1.5 g/L) production when compared with an *A. niger* strain expressing only cis-aconitate decarboxylase (0.8 g/L) [54]. Interestingly, the production did not increase further when both transporters were co-expressed (0.9 g/L). In a previous study in *A. terreus* the overexpression of a bacterial hemoglobin (vgb) led to an increased dissolved oxygen level, having a strong effect on itaconic acid production [55]. Additional optimization was achieved by overexpression of the fungal hemoglobin domain *hbd1* and deletion of the oxaloacetate acetylhydrolase *oahA* gene, in combination with controlled batch fermentation conditions, resulting in the increase of the production level from 0.8 to 2.5 g/L of itaconic acid [54]. In a subsequent study, a titer of 26.2 g/L and a maximum production rate of 0.35 g/L/h were reached by overexpressing the cytosolic citrate synthase *citB* [53].

LACTIC ACID

The *S. cerevisiae* genome encodes at least two plasma membrane monocarboxylate transporters, Jen1 [56, 57] and Ady2 [58, 59] with distinct specificities, mode of action and regulation mechanisms [34, 60]. In a *S. cerevisiae* strain engineered for lactate production, the constitutive expression of these two transporters resulted in a higher accumulation of lactic acid in the extracellular medium [61]. In this study, the authors expressed the lactate-dehydrogenase *LDH* gene from *L.casei* in the *S. cerevisiae* W303-1A parental strain and in the three isogenic strains *jen1*Δ, *ady2*Δ and *jen1*Δ*ady2*Δ to allow lactate production. All the deleted strains expressing *LDH* were able to produce higher titers of lactic acid compared with the parental isogenic strain. Moreover, the constitutive expression of *JEN1* or *ADY2* genes, along with *LDH*, resulted in the higher external accumulation of lactic acid in the presence of glucose. Upon glucose depletion, lactate consumption was also more pronounced in cells expressing Jen1 and/or Ady2, suggesting the involvement of these transporters in both the import and export of lactic acid [61].

CITRIC ACID

In a recent work, Steiger *et al.* (2019) identified CexA, the longtime sought citrate exporter from *A. niger*. The constitutive and inducible overexpression of CexA in the native citric acid-producing species *A. niger*, resulted in significant increases in secreted citric acid [62]. The inducible system reached 109 g/L citric acid, five times higher than the parental wild-type strain and three times higher than the constitutive expression system.

3.5 ENGINEERING MEMBRANE TRANSPORTERS

Membrane transporters, like any protein, can display substrate promiscuity, altered conformation, distinct affinity and capacity depending on physiological conditions, as well as alterations in folding and stability. Finding a membrane transporter, either for import or export, might not be enough to achieve the levels of cell factory productivity needed to obtain a cost-effective and sustainable bioproduction process. The engineering of membrane proteins can diminish these constraints by tuning the activity towards specific conditions and substrates. This approach is frequently achieved by ALE experiments, mutagenesis or recombination involving methods of synthetic biology [6, 11, 18]. ALE of host organisms combined with the identification of responsible genetic changes and subsequent reverse engineering, is a powerful approach to obtain novel or improved substrate specificity of membrane transporters.

3.5.1 ENGINEERING SUGAR TRANSPORTERS

Whereas it is common to look for exogenous transporters to be cloned into producer strains, the endogenous transportome can be used as a pool of transporters for cell factory optimization. One such example is the complex landscape of the *S. cerevisiae* genome that includes 20 transporter proteins belonging to the Hexose Transporter (HXT) Family, with great potential to be exploited in cell factories for the uptake of renewable sugars from lignocellulosic wastes [18, 63]. By using molecular modeling and docking studies, the endogenous *S. cerevisiae* Gal2 transporter was engineered to improve L-arabinose transport capacity [64]. In this study, nine residues were found to interact with L-arabinose. Rational protein design by directed mutagenesis allowed an increase of transporter capacity for L-arabinose. Besides the gain of function associated with arabinose transport capacity, the F85S mutation specifically improved xylose transport [65]. In another study, the combination of computer-assisted modeling, site-directed mutagenesis, error-prone PCR approaches and selective growth conditions, resulted in the identification of residues in both Hxt7 and Gal2 that yielded glucose-insensitive xylose transporters [66]. The mutant Gal2-N376F had the highest affinity for D-xylose, along with a moderate transport rate for this pentose sugar, and completely lost the ability to transport hexoses [66].

To obtain a transporter able to sufficiently import arabinose in the presence of glucose and xylose, a strain deficient in glucose phosphorylation and able to metabolize arabinose was created [67]. Subsequently, the engineered strain was grown in medium with these three sugars. This way, conditions were met where arabinose became the only metabolizable sugar within the media, while glucose and xylose were exerting selective pressure towards the evolution of an arabinose transporter uninhibited by glucose and xylose. Consequently, mutations within hexose transporter Gal2 in residues T89 and N376 were found to significantly increase the K_m value of Gal2 for glucose, and decrease the K_m value for arabinose, enabling superior growth of the engineered strain in a medium containing the three sugars [67].

3.5.2 ENGINEERING ORGANIC ACID TRANSPORTERS

In an attempt to evolve an efficient fumarate exporter in *S. cerevisiae*, a knock-out strategy was implemented in which fumaric acid was turned into the energetically more favorable catabolic product, by deletion of the fumarase (*FUM1*) and glucose 6-phosphate dehydrogenase (*ZWF1*) genes [68]. The malate and succinate transporter DCT_02 from *A. niger* was used as a template expected to evolve into an efficient fumarate exporter. However, the evolution experiment did not yield the desired results, since

only malate and succinate were secreted to the extracellular medium, and further strategy refinement is required [68].

Through a rationally designed site-directed mutagenesis strategy, the substrate specificity of the yeast Jen1 monocarboxylate transporter was altered to acquire the ability to transport the dicarboxylic acids succinate (F270A and F270G) [69] and saccharate (S271Q) [35].

In two independent evolution experiments, *S. cerevisiae* strains deficient in Jen1 were evolved for growth on lactate as sole carbon and energy source [70]. Whole-genome resequencing of evolved strains uncovered the presence of single nucleotide changes in the acetate transporter gene *ADY2* (C755G/L219V and C655G/A252G). These Ady2 mutated alleles encode efficient lactate transporters.

Presently, new relevant roles of protein transporters are being uncovered, namely at the level of improving industrial strain's tolerance to by-products. An example of the complexity of the roles of transporters in regulatory networks is reported by Zang *et al*. (2017) who, in the presence of 3.6 g/L acetic acid pH 3.7, observed an increment of 14.7% in the final ethanol concentration for the *S. cerevisiae* strain lacking the *ADY2* gene [71]. By impairing acetate uptake from the extracellular space, the accumulation of intracellular acetate was reduced, and as consequence cells acquired increased tolerance towards this organic acid.

3.6 FUTURE PERSPECTIVES FOR TRANSPORT ENGINEERING

It is expected that in the near future, biorefineries increase the production of platform chemicals from renewable resources [72]. The exploitation of industrial biowastes to sustain microbial cell growth and valorize biomass conversion into organic acids is one of these current trends. Achieving optimal processes requires industrially robust strains. One of the major bottlenecks for the efficient and cost-effective bioproduction of organic acids is their export through the microbial plasma membrane. Membrane transporter proteins are thus crucial elements for the optimization of this process. In this review, we presented examples of the most relevant and emerging cell factories for the production of organic acids, as well as the engineering strategies used to turn them into efficient producers of this family of compounds.

In recent years, a great effort was dedicated to transporter engineering, envisaging the development and improvement of microbial cell factories. Examples of transporters engineered in producer strains, especially in the yeast *S. cerevisiae*, are summarized in Figure 2.

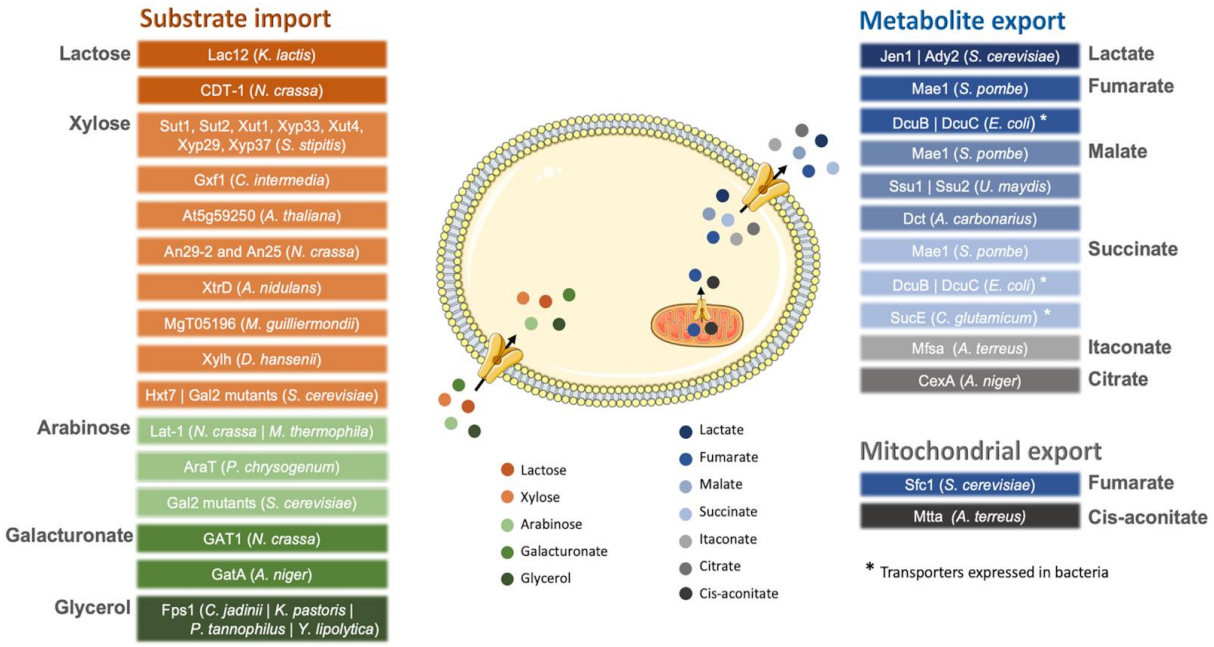


Figure 2. The expression of endogenous or exogenous membrane transporter genes in engineered bacteria, yeast and filamentous fungi, allows the uptake of renewable substrates, as well as the export and extracellular accumulation of specialty organic acids. Transporters on the left were expressed in the plasma membrane, to promote the import of substrates. The transporters on the right were expressed either in the inner mitochondrial membrane or in the plasma membrane, to promote the export of organic acids. The black arrows indicate the direction of the transport, either to the cytoplasm, out of the mitochondria or to the extracellular medium. Transporters expressed in bacteria are marked with *. The figure was produced using the vector image bank of Servier Medical Art (http://smart.servier.com/).

Despite these advances, as the functional and structural characterization of membrane proteins is still a cumbersome process, the redesigning and engineering of optimized cell membrane transporters for industrial organic acid production is still at an early stage [6, 7]. Different strategies can be followed to obtain improved transporters, namely with higher activity, altered substrate specificity and product selectivity. ALE can be a suitable approach when the desired transport process is directly linked to a selective advantage, such as the import of a sole carbon source necessary for growth. Nevertheless, its employment to generate improved transporters for the efflux of solutes can represent a demanding challenge. The complexity of biorefineries relies on many factors, including the optimization of several transporters, with complementary kinetic and regulatory properties [67]. Structure-based or computer simulation-based protein engineering is a powerful approach. However, these methods are hampered by the low number of robust three-dimensional structural models of transporter proteins. According to the Protein Data Bank (https://www.rcsb.org), transporter 3D structures account for less than 10% of the total database entries, showing that membrane proteins remain until now mostly uncharacterized, which evidences the need to increase the existing knowledge on this field. The recent identification of the gene

encoding the long-sought citrate exporter from *A. niger* [62, 73], is an example of the effort that must be carried out towards the identification of new transporters.

The biodiversity of the microbial world is an excellent pool to uncover relevant transporters for organic acid production. Achieving the efficient heterologous expression of transporters is crucial to improve the robustness of microbial cell factories. For instance, the proper expression of bacterial transporters in fungi is constrained due to membrane incompatibility, low expression levels, and folding difficulties [74], limiting the options for prokaryotic transporter expression in eukaryotic cells. Still, the versatility and plasticity of membrane transporters suggest a promising future towards the optimization and implementation of platform chemicals bioproduction at the industrial scale.

3.7 REFERENCES

1. Alonso, S., M. Rendueles, and M. Díaz, *Microbial production of specialty organic acids from renewable and waste materials.* Critical Reviews in Biotechnology, 2015. **35**(4): p. 497-513.
2. Sahu, Y., *Organic Acids Market by Type (Acetic Acid, Citric Acid, Formic Acid, Lactic Acid, Itaconic Acid, Succinic Acid, Gluconic Acid, Ascorbic Acid, Fumaric Acid, and Propionic Acid), Source (Biomass, Molasses, Starch, Chemical Synthesis, Agro-Industrial Residue), End-User (Food & Beverage, Animal Feed, Chemicals & Industrial, Pharmaceuticals, Personal Care, Agriculture) - Global Opportunity Analysis and Industry Forecast, 2017-2023*, 2017. p. 260. Available in https://www.alliedmarketresearch.com/organic-acids-market.
3. Becker, J. and C. Wittmann, *Advanced biotechnology: metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products.* Angewandte Chemie (International Edition), 2015. **54**(11): p. 3328-3350.
4. Buschke, N., et al., *Metabolic engineering of industrial platform microorganisms for biorefinery applications – Optimization of substrate spectrum and process robustness by rational and evolutive strategies.* Bioresource Technology, 2013. **135**: p. 544-554.
5. Na, D., et al., *Systems metabolic engineering of Escherichia coli for chemicals, materials, biofuels, and pharmaceuticals*, in *Systems Metabolic Engineering*, C. Wittmann and S.Y. Lee, Editors. 2012, Springer Netherlands: Dordrecht. p. 117-149.
6. Kell, D.B., et al., *Membrane transporter engineering in industrial biotechnology and whole cell biocatalysis.* Trends in Biotechnology, 2015. **33**(4): p. 237-246.
7. Boyarskiy, S. and D. Tullman-Ercek, *Getting pumped: membrane efflux transporters for enhanced biomolecule production.* Current Opinion in Chemical Biology, 2015. **28**: p. 15-19.
8. Sauer, M., et al., *Microbial production of organic acids: expanding the markets.* Trends in Biotechnology, 2008. **26**(2): p. 100-108.
9. Sauer, M., et al., *16 years research on lactic acid production with yeast - ready for the market?* Biotechnol Genet Eng Rev, 2010. **27**: p. 229-56.

10. Deng, W., Y. Wang, and N. Yan, *Production of organic acids from biomass resources.* Curr Opin Green Sustain Chem, 2016. **2**: p. 54-58.
11. Van Dyk, T.K., *Bacterial efflux transport in Biotechnology* Advances in Applied Microbiology, 2008. **63**(07): p. 231-247.
12. Chen, X., et al., *Metabolic engineering of Escherichia coli: a sustainable industrial platform for bio-based chemical production.* Biotechnol Adv, 2013. **31**(8): p. 1200-23.
13. Yang, D., et al., *Metabolic engineering of Escherichia coli for natural product biosynthesis.* Trends Biotechnol, 2020.
14. Wu, W., F. Liu, and S. Singh, *Toward engineering E. coli with an autoregulatory system for lignin valorization.* Proc Natl Acad Sci U S A, 2018. **115**(12): p. 2970-2975.
15. Khunnonkwao, P., et al., *Re-engineering Escherichia coli KJ122 to enhance the utilization of xylose and xylose/glucose mixture for efficient succinate production in mineral salt medium.* Appl Microbiol Biotechnol, 2018. **102**(1): p. 127-141.
16. Tong, Z., et al., *Systems metabolic engineering for citric acid production by Aspergillus niger in the post-genomic era.* Microb Cell Fact, 2019. **18**(1): p. 28.
17. Liu, L., et al., *Method for reconstructing Aspergillus niger to increase citrate production* U.S.P. Application, Editor 2018: China.
18. Moysés, D.N., et al., *Xylose fermentation by Saccharomyces cerevisiae: challenges and prospects.* International Journal of Molecular Sciences, 2016. **17**(3): p. 207.
19. Wang, C., et al., *Improvement of L-arabinose fermentation by modifying the metabolic pathway and transport in Saccharomyces cerevisiae.* Biomed Res Int, 2013. **2013**: p. 461204.
20. Becker, J. and E. Boles, *A modified Saccharomyces cerevisiae strain that consumes L-Arabinose and produces ethanol.* Appl Environ Microbiol, 2003. **69**(7): p. 4144-50.
21. Subtil, T. and E. Boles, *Improving L-arabinose utilization of pentose fermenting Saccharomyces cerevisiae cells by heterologous expression of L-arabinose transporting sugar transporters.* Biotechnol Biofuels, 2011. **4**: p. 38.
22. Li, J., et al., *Functional analysis of two l-arabinose transporters from filamentous fungi reveals promising characteristics for improved pentose utilization in Saccharomyces cerevisiae.* Appl Environ Microbiol, 2015. **81**(12): p. 4062-4070.
23. Bracher, J.M., et al., *The Penicillium chrysogenum transporter PcAraT enables high-affinity, glucose-insensitive l-arabinose transport in Saccharomyces cerevisiae.* Biotechnol Biofuels, 2018. **11**: p. 63.
24. Domingues, L., et al., *Continuous ethanol fermentation of lactose by a recombinant flocculating Saccharomyces cerevisiae strain.* Biotechnol Bioeng, 1999. **64**(6): p. 692-697.
25. Turner, T.L., et al., *Short communication: Conversion of lactose and whey into lactic acid by engineered yeast.* J Dairy Sci, 2017. **100**(1): p. 124-128.
26. Klein, M., et al., *The expression of glycerol facilitators from various yeast species improves growth on glycerol of Saccharomyces cerevisiae.* Metab Eng Commun, 2016. **3**: p. 252-257.
27. Asskamp, M.R., M. Klein, and E. Nevoigt, *Saccharomyces cerevisiae exhibiting a modified route for uptake and catabolism of glycerol forms significant amounts of ethanol from this carbon source considered as 'non-fermentable'.* Biotechnol Biofuels, 2019. **12**: p. 257.

28. Biz, A., et al., *The introduction of the fungal D-galacturonate pathway enables the consumption of D-galacturonic acid by Saccharomyces cerevisiae.* Microb Cell Fact, 2016. **15**(1): p. 144.
29. Protzko, R.J., et al., *Engineering Saccharomyces cerevisiae for co-utilization of d-galacturonic acid and d-glucose from citrus peel waste.* Nat Commun, 2018. **9**(1): p. 5059.
30. Sobczak, I. and J.S. Lolkema, *The 2-hydroxycarboxylate transporter family: physiology, structure, and mechanism.* Microbiol. Mol. Biol. Rev., 2005. **69**(4): p. 665-695.
31. Bandell, M., et al., *Mechanism of the Citrate Transporters in Carbohydrate and Citrate Cometabolism in Lactococcus and Leuconostoc Species.* Appl. Environ. Microbiol., 1998. **64**(5): p. 1594-1600.
32. Pudlik, A.M. and J.S. Lolkema, *Substrate specificity of the citrate transporter CitP of Lactococcus lactis.* Journal of bacteriology, 2012. **194**(14): p. 3627-3635.
33. Lu, M., *Structure and mechanism of the divalent anion/Na^+ symporter.* Int J Mol Sci, 2019. **20**(2).
34. Casal, M., et al., *Transport of carboxylic acids in yeasts.* FEMS Microbiol Rev, 2008. **32**(6): p. 974-94.
35. Donida, B., et al., *Oxidative damage and redox in Lysosomal Storage Disorders: Biochemical markers.* Clin Chim Acta, 2017. **466**: p. 46-53.
36. Nakamura, J., et al., *Mutations of the Corynebacterium glutamicum NCgl1221 gene, encoding a mechanosensitive channel homolog, induce L-glutamic acid production.* Appl Environ Microbiol, 2007. **73**(14): p. 4491-8.
37. Becker, M., et al., *Glutamate efflux mediated by Corynebacterium glutamicum MscCG, Escherichia coli MscS, and their derivatives.* Biochim Biophys Acta, 2013. **1828**(4): p. 1230-40.
38. Nakayama, Y., et al., *"Force-From-Lipids" mechanosensation in Corynebacterium glutamicum.* Biophys Rev, 2019. **11**(3): p. 327-333.
39. Zambanini, T., et al., *Metabolic engineering of Ustilago trichophora TZ1 for improved malic acid production.* Metab Eng Commun, 2017. **4**: p. 12-21.
40. Zelle, R.M., et al., *Malic acid production by Saccharomyces cerevisiae: engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export.* Appl Environ Microbiol, 2008. **74**(9): p. 2766-2777.
41. Darbani, B., et al., *Engineering energetically efficient transport of dicarboxylic acids in yeast Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 2019.
42. Xu, G., et al., *Fumaric acid production in Saccharomyces cerevisiae by in silico aided metabolic engineering.* PLoS One, 2012. **7**(12): p. e52086.
43. Chen, X., et al., *Mitochondrial engineering of the TCA cycle for fumarate production.* Metab Eng, 2015. **31**: p. 62-73.
44. Zhang, T., et al., *Regulating C_4-dicarboxylate transporters for improving fumaric acid production.* RSC Advances, 2017. **7**(5): p. 2897-2904.
45. Zientz, E., S. Six, and G. Unden, *Identification of a third secondary carrier (DcuC) for anaerobic C4-dicarboxylate transport in Escherichia coli: roles of the three Dcu carriers in uptake and exchange.* Journal of bacteriology, 1996. **178**(24): p. 7241-7247.

46. Zientz, E., et al., *Functioning of DcuC as the C4-Dicarboxylate Carrier during Glucose Fermentation by Escherichia coli.* Journal of bacteriology, 1999. **181**(12): p. 3716-3720.
47. Chen, J., et al., *Activating C_4-dicarboxylate transporters DcuB and DcuC for improving succinate production.* Appl Microbiol Biotechnol, 2014. **98**(5): p. 2197-2205.
48. Zhu, N., et al., *Improved succinate production in Corynebacterium glutamicum by engineering glyoxylate pathway and succinate export system.* Biotechnology Letters, 2014. **36**(3): p. 553-560.
49. Li, A., et al., *Enhanced itaconic acid production in Aspergillus niger using genetic modification and medium optimization.* BMC Biotechnology, 2012. **12**: p. 57-57.
50. Li, A., et al., *A clone-based transcriptomics approach for the identification of genes relevant for itaconic acid production in Aspergillus.* Fungal Genet Biol, 2011. **48**(6): p. 602-611.
51. Steiger, M.G., et al., *Characterizing MttA as a mitochondrial cis-aconitic acid transporter by metabolic engineering.* Metab Eng, 2016. **35**: p. 95-104.
52. Huang, X., et al., *Improving itaconic acid production through genetic engineering of an industrial Aspergillus terreus strain.* Microb Cell Fact, 2014. **13**(1): p. 119.
53. Hossain, A.H., et al., *Rewiring a secondary metabolite pathway towards itaconic acid production in Aspergillus niger.* Microb Cell Fact, 2016. **15**(1): p. 130.
54. Li, A., et al., *Reduced by-product formation and modified oxygen availability improve itaconic acid production in Aspergillus niger.* Appl Microbiol Biotechnol, 2013. **97**(9): p. 3901-3911.
55. Lin, Y.H., et al., *Intracellular expression of Vitreoscilla hemoglobin in Aspergillus terreus to alleviate the effect of a short break in aeration during culture.* Biotechnol Lett, 2004. **26**(13): p. 1067-72.
56. Casal, M., et al., *The lactate-proton symport of Saccharomyces cerevisiae is encoded by JEN1.* J Bacteriol, 1999. **181**(8): p. 2620-3.
57. Soares-Silva, I., et al., *Functional expression of the lactate permease Jen1p of Saccharomyces cerevisiae in Pichia pastoris.* Biochem J, 2003. **376**(Pt 3): p. 781-7.
58. Paiva, S., et al., *Ady2p is essential for the acetate permease activity in the yeast Saccharomyces cerevisiae.* Yeast, 2004. **21**(3): p. 201-210.
59. Ribas, D., et al., *The acetate uptake transporter family motif "NPAPLGL(M/S)" is essential for substrate uptake.* Fungal Genet Biol, 2019. **122**: p. 1-10.
60. Casal, M., et al., *Carboxylic acids plasma membrane transporters in Saccharomyces cerevisiae.* Adv Exp Med Biol, 2016. **892**: p. 229-251.
61. Pacheco, A., et al., *Lactic acid production in Saccharomyces cerevisiae is modulated by expression of the monocarboxylate transporters Jen1 and Ady2.* FEMS Yeast Res, 2012. **12**(3): p. 375-81.
62. Steiger, M.G., et al., *Engineering of the citrate exporter protein enables high citric acid production in Aspergillus niger.* Metab Eng, 2019. **52**: p. 224-231.
63. Kruckeberg, A.L., *The hexose transporter family of Saccharomyces cerevisiae.* Archives of Microbiology, 1996. **166**(5): p. 283-292.
64. Subtil, T. and E. Boles, *Improving L-arabinose utilization of pentose fermenting Saccharomyces cerevisiae cells by heterologous expression of L-arabinose transporting sugar transporters.* Biotechnol Biofuels, 2011. **4**: p. 38-38.

65. Rovira, M., et al., *A Cu(I)/Cu(III) prototypical organometallic mechanism for the deactivation of an active pincer-like Cu(I) catalyst in Ullmann-type couplings.* Chem Commun (Camb), 2017. **53**(62): p. 8786-8789.
66. Farwick, A., et al., *Engineering of yeast hexose transporters to transport d-xylose without inhibition by d-glucose.* Proc Natl Acad Sci U S A, 2014. **111**(14): p. 5159-5164.
67. Verhoeven, M.D., et al., *Laboratory evolution of a glucose-phosphorylation-deficient, arabinose-fermenting S. cerevisiae strain reveals mutations in GAL2 that enable glucose-insensitive l-arabinose uptake.* FEMS Yeast Res, 2018. **18**(6).
68. Shah, M., *Dicarboxylic acids transport, metabolism and production in aerobic Saccharomyces cerevisiae.* , in *Department of Biotechnology, Faculty of Applied Sciences*2016, Delft University of Technology.
69. Soares-Silva, I., et al., *A substrate translocation trajectory in a cytoplasm-facing topological model of the monocarboxylate/H+ symporter Jen1p.* Mol Microbiol, 2011. **81**(3): p. 805-17.
70. de Kok, S., et al., *Laboratory evolution of new lactate transporter genes in a jen1 delta mutant of Saccharomyces cerevisiae and their identification as ADY2 alleles by whole-genome resequencing and transcriptome analysis.* FEMS Yeast Res, 2012. **12**(3): p. 359-374.
71. Zhang, M., et al., *Deletion of acetate transporter gene ADY2 improved tolerance of Saccharomyces cerevisiae against multiple stresses and enhanced ethanol production in the presence of acetic acid.* Bioresource Technology, 2017. **245**: p. 1461-1468.
72. Takkellapati, S., T. Li, and M.A. Gonzalez, *An Overview of Biorefinery Derived Platform Chemicals from a Cellulose and Hemicellulose Biorefinery.* Clean Technol Environ Policy, 2018. **20**(7): p. 1615-1630.
73. Odoni, D.I., et al., *Aspergillus niger citrate exporter revealed by comparison of two alternative citrate producing conditions.* FEMS Microbiol Lett, 2019. **366**(7).
74. Young, E., et al., *Functional survey for heterologous sugar transport proteins, using Saccharomyces cerevisiae as a host.* Appl Environ Microbiol, 2011. **77**(10): p. 3311-9.
75. Krom, B.P., R. Aardema, and J.S. Lolkema, *Bacillus subtilis YxkJ is a secondary transporter of the 2-hydroxycarboxylate transporter family that transports L-malate and citrate.* Journal of bacteriology, 2001. **183**(20): p. 5862-5869.
76. Tanaka, K., K. Kobayashi, and N. Ogasawara, *The Bacillus subtilis YufLM two-component system regulates the expression of the malate transporters MaeN (YufR) and YflS, and is essential for utilization of malate in minimal medium.* Microbiology, 2003. **149**(9): p. 2317-2329.
77. Kebbel, F., et al., *Structure and substrate-induced conformational changes of the secondary citrate/sodium symporter CitS revealed by electron crystallography.* Structure, 2013. **21**(7): p. 1243-1250.
78. Bandell, M., et al., *Membrane potential-generating malate (MleP) and citrate (CitP) transporters of lactic acid bacteria re Homologous Proteins Substrate specificity of the 2-hydroxycarboxylate transporter family.* Journal of Biological Chemistry, 1997. **272**(29): p. 18140-18146.
79. Pudlik, A.M. and J.S. Lolkema, *Citrate uptake in exchange with intermediates in the citrate metabolic pathway in Lactococcus lactis IL1403.* Journal of bacteriology, 2011. **193**(3): p. 706-714.

80. Poolman, B., et al., *Malolactic fermentation: electrogenic malate uptake and malate/lactate antiport generate metabolic energy.* Journal of bacteriology, 1991. **173**(19): p. 6030-6037.
81. Marty-Teysset, C., et al., *The citrate metabolic pathway in Leuconostoc mesenteroides: expression, amino acid synthesis, and alpha-ketocarboxylate transport.* Journal of bacteriology, 1996. **178**(21): p. 6209-6215.
82. Fukui, K., et al., *Identification of succinate exporter in Corynebacterium glutamicum and its physiological roles under anaerobic conditions.* Journal of biotechnology, 2011. **154**(1): p. 25-34.
83. Sá-Pessoa, J., et al., *SATP (YaaH), a succinate–acetate transporter protein in Escherichia coli.* Biochemical journal, 2013. **454**(3): p. 585-595.
84. Augstein, A., et al., *Characterization, localization and functional analysis of Gpr1p, a protein affecting sensitivity to acetic acid in the yeast Yarrowia lipolytica.* Microbiology, 2003. **149**(3): p. 589-600.
85. Galagan, J.E., et al., *Sequencing of Aspergillus nidulans and comparative analysis with A. fumigatus and A. oryzae.* Nature, 2005. **438**(7071): p. 1105-1115.
86. Roberts, S.K., J. Milnes, and M. Caddick, *Characterisation of AnBEST1, a functional anion channel in the plasma membrane of the filamentous fungus, Aspergillus nidulans.* Fungal Genet Biol, 2011. **48**(9): p. 928-938.
87. Li, H. and A.J. Pajor, *Functional characterization of CitM, the Mg^{2+}-citrate transporter.* J Membr Biol, 2002. **185**(1): p. 9.
88. Krom, B.P., et al., *Complementary Metal Ion Specificity of the Metal-Citrate Transporters CitM and CitH of Bacillus subtilis.* Journal of bacteriology, 2000. **182**(22): p. 6374-6381.
89. Watanabe, S., et al., *Complete sequence of the first chimera genome constructed by cloning the whole genome of Synechocystis Strain PCC6803 into the Bacillus subtilis 168 genome.* Journal of bacteriology, 2012. **194**(24): p. 7007-7007.
90. Brocker, M., et al., *Citrate utilization by Corynebacterium glutamicum is controlled by the CitAB two-component system through positive regulation of the citrate transport genes citH and tctCBA.* Journal of bacteriology, 2009. **191**(12): p. 3869-3880.
91. Yang, L., et al., *Overexpression of a C_4-dicarboxylate transporter is the key for rerouting citric acid to C_4-dicarboxylic acid production in Aspergillus carbonarius.* Microb Cell Fact, 2017. **16**(1): p. 43.
92. Asai, K., et al., *Regulation of the transport system for C4-dicarboxylic acids in Bacillus subtilis.* Microbiology, 2000. **146**(2): p. 263-271.
93. Groeneveld, M., et al., *Biochemical characterization of the C_4-dicarboxylate transporter DctA from Bacillus subtilis.* Journal of bacteriology, 2010. **192**(11): p. 2900-2907.
94. Youn, J.-W., et al., *Identification and characterization of the dicarboxylate uptake system DccT in Corynebacterium glutamicum.* Journal of bacteriology, 2008. **190**(19): p. 6458-6466.
95. Kay, W.W. and H.L. Kornberg, *The Uptake of C4-Dicarboxylic Acids by Escherichia coli.* European Journal of Biochemistry, 1971. **18**(2): p. 274-281.
96. Baker, K.E., et al., *Utilization of orotate as a pyrimidine source by Salmonella typhimurium and Escherichia coli requires the dicarboxylate transport protein encoded by dctA.* Journal of bacteriology, 1996. **178**(24): p. 7099-7105.

97. Karinou, E., et al., *The Escherichia coli SLC26 homologue YchM (DauA) is a C4-dicarboxylic acid transporter.* Molecular Microbiology, 2013. **87**(3): p. 623-640.
98. Ebbighausen, H., B. Weil, and R. Krämer, *Carrier-mediated acetate uptake in Corynebacterium glutamicum.* Archives of Microbiology, 1991. **155**(5): p. 505-510.
99. Teramoto, H., et al., *Identification of a gene encoding a transporter essential for utilization of C_4 dicarboxylates in Corynebacterium glutamicum.* Appl Environ Microbiol, 2008. **74**(17): p. 5290-5296.
100. Kim, O.B. and G. Unden, *The L-Tartrate/Succinate Antiporter TtdT (YgjE) of L-Tartrate Fermentation in Escherichia coli.* Journal of bacteriology, 2007. **189**(5): p. 1597-1603.
101. Pos, K.M., P. Dimroth, and M. Bott, *The Escherichia coli Citrate Carrier CitT: a Member of a Novel Eubacterial Transporter Family Related to the 2-Oxoglutarate/Malate Translocator from Spinach Chloroplasts.* Journal of bacteriology, 1998. **180**(16): p. 4160-4165.
102. Rhie, M.N., et al., *A Na^+-coupled C_4-dicarboxylate transporter (Asuc_0304) and aerobic growth of Actinobacillus succinogenes on C_4-dicarboxylates.* Microbiology, 2014. **160**(Pt 7): p. 1533-1544.
103. Hall, J.A. and A.M. Pajor, *Functional characterization of a Na^+-coupled dicarboxylate carrier protein from Staphylococcus aureus.* J Bacteriol, 2005. **187**(15): p. 5189-94.
104. Hall, J.A. and A.M. Pajor, *Functional reconstitution of SdcS, a Na^+-coupled dicarboxylate carrier protein from Staphylococcus aureus.* J Bacteriol, 2007. **189**(3): p. 880-5.
105. Six, S., et al., *Escherichia coli possesses two homologous anaerobic C4-dicarboxylate membrane transporters (DcuA and DcuB) distinct from the aerobic dicarboxylate transport system (Dct).* Journal of bacteriology, 1994. **176**(21): p. 6470-6478.
106. Lü, W., et al., *The formate channel FocA exports the products of mixed-acid fermentation.* Proc Natl Acad Sci U S A, 2012. **109**(33): p. 13254-13259.
107. Wang, Y., et al., *Structure of the formate transporter FocA reveals a pentameric aquaporin-like channel.* Nature, 2009. **462**: p. 467.
108. Marchetti, R.V., et al., *A lactate and formate transporter in the intraerythrocytic malaria parasite, Plasmodium falciparum.* Nat Commun, 2015. **6**: p. 6721.
109. Wu, B., et al., *Identity of a Plasmodium lactate/H+ symporter structurally unrelated to human transporters.* Nat Commun, 2015. **6**: p. 6284.
110. Núñez, M.F., et al., *The gene yghK linked to the glc operon of Escherichia coli encodes a permease for glycolate that is structurally and functionally similar to L-lactate permease.* Microbiology, 2001. **147**(4): p. 1069-1077.
111. Núñez, M.a.F., et al., *Transport of l-Lactate, d-Lactate, and Glycolate by the LldP and GlcA Membrane Carriers of Escherichia coli.* Biochem Biophys Res Commun, 2002. **290**(2): p. 824-829.
112. Chai, Y., R. Kolter, and R. Losick, *A widely conserved gene cluster required for lactate utilization in Bacillus subtilis and its involvement in biofilm formation.* J Bacteriol, 2009. **191**(8): p. 2423-30.
113. Wang, Y., et al., *Elucidating the role and regulation of a lactate permease as lactate transporter in Bacillus coagulans DSM1.* Appl Environ Microbiol, 2019. **85**(14).
114. Geiser, E., et al., *Ustilago maydis produces itaconic acid via the unusual intermediate trans-aconitate.* Microb Biotechnol, 2016. **9**(1): p. 116-26.

115. Su, X., et al., *Transport of haloacids across biological membranes.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 2016. **1858**(12): p. 3061-3070.
116. Su, X. and J.S.H. Tsang, *Existence of a robust haloacid transport system in a Burkholderia species bacterium.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 2013. **1828**(2): p. 187-192.
117. Wei, Y., et al., *Bacillus subtilis YqkI Is a Novel Malic/Na+-Lactate Antiporter That Enhances Growth on Malate at Low Protonmotive Force.* Journal of Biological Chemistry, 2000. **275**(39): p. 30287-30292.
118. McDermott, J.R., B.P. Rosen, and Z. Liu, *Jen1p: a high affinity selenite transporter in yeast.* Mol Biol Cell, 2010. **21**(22): p. 3934-41.
119. Soares-Silva, I., et al., *The conserved sequence NXX[S/T]HX[S/T]QDXXXT of the lactate/pyruvate:H+ symporter subfamily defines the function of the substrate translocation pathway.* Mol Membr Biol, 2007. **24**(5-6): p. 464-474.
120. Soares-Silva, I., et al., *The disruption of JEN1 from Candida albicans impairs the transport of lactate.* Molecular membrane biology, 2004. **21**(6): p. 403-411.
121. Vieira, N., et al., *Functional specialization and differential regulation of short-chain carboxylic acid transporters in the pathogen Candida albicans.* Molecular Microbiology, 2010. **75**(6): p. 1337-1354.
122. Soares-Silva, I., et al., *The Debaryomyces hansenii carboxylate transporters Jen1 homologues are functional in Saccharomyces cerevisiae.* FEMS Yeast Res, 2015. **15**(8).
123. Lodi, T., et al., *Carboxylic acids permeases in yeast: two genes in Kluyveromyces lactis.* Gene, 2004. **339**: p. 111-9.
124. Queirós, O., et al., *Functional analysis of Kluyveromyces lactis carboxylic acids permeases: heterologous expression of KlJEN1 and KlJEN2 genes.* Current Genetics, 2007. **51**(3): p. 161-169.
125. Jolkver, E., et al., *Identification and characterization of a bacterial transport system for the uptake of pyruvate, propionate, and acetate in Corynebacterium glutamicum.* Journal of bacteriology, 2009. **191**(3): p. 940-948.
126. Gimenez, R., et al., *The Gene yjcG, Cotranscribed with the Gene acs, Encodes an Acetate Permease in Escherichia coli.* Journal of bacteriology, 2003. **185**(21): p. 6448-6455.
127. Elías, A., et al., *The ActP acetate transporter acts prior to the PitA phosphate carrier in tellurite uptake by Escherichia coli.* Microbiological Research, 2015. **177**: p. 15-21.
128. Borghese, R., et al., *On the role of a specific insert in acetate permeases (ActP) for tellurite uptake in bacteria: Functional and structural studies.* Journal of Inorganic Biochemistry, 2016. **163**: p. 103-109.
129. Borghese, R., S. Cicerano, and D. Zannoni, *Fructose increases the resistance of Rhodobacter capsulatus to the toxic oxyanion tellurite through repression of acetate permease (ActP).* Antonie van Leeuwenhoek, 2011. **100**(4): p. 655.
130. Borghese, R. and D. Zannoni, *Acetate permease (ActP) Is responsible for tellurite (TeO32-) uptake and resistance in cells of the facultative phototroph Rhodobacter capsulatus.* Appl Environ Microbiol, 2010. **76**(3): p. 942-944.
131. Koita, K. and C.V. Rao, *Identification and Analysis of the Putative Pentose Sugar Efflux Transporters in Escherichia coli.* PLoS One, 2012. **7**(8): p. e43700.

132. Camarasa, C., et al., *Characterization of Schizosaccharomyces pombe malate permease by expression in Saccharomyces cerevisiae.* Appl Environ Microbiol 2001. **67**(9): p. 4144-4151.
133. Grobler, J., et al., *The mae1 gene of Schizosaccharomyces pombe encodes a permease for malate and other C_4 dicarboxylic acids.* Yeast, 1995. **11**(15): p. 1485-91.
134. Osawa, H. and H. Matsumoto, *Cytotoxic thio-malate is transported by both an aluminum-responsive malate efflux pathway in wheat and the MAE1 malate permease in Schizosaccharomyces pombe.* Planta, 2006. **224**(2): p. 462.
135. Forward, J.A., et al., *TRAP transporters: a new family of periplasmic solute transport systems encoded by the dctPQM genes of Rhodobacter capsulatus and by homologs in diverse gram-negative bacteria.* Journal of bacteriology, 1997. **179**(17): p. 5482-549

CHAPTER IV

Bioprospection of non-conventional yeasts – Identification and genomic analysis

Note:

The work presented in this chapter is in preparation for publication.

CHAPTER IV

Bioprospection of non-conventional yeasts – Identification and genomic analysis

ABSTRACT

The exploitation of biodiversity has captured the interest of the food, pharmaceutical, and even fuel industries. In particular, microorganisms isolated from natural environments present properties that can be relevant for industrial processes, e.g. tolerance to extreme pH values or temperatures, and the ability to utilize and produce specific compounds. Yeasts are recognized as model organisms for basic research but are also commonly used microorganisms for the production of added-value compounds, such as chemical building blocks and biofuels. In this work, a group of wild yeasts isolated from acidic environments, developed by fruit and vegetable wastes, was analyzed for their ability to grow on carboxylic acids as carbon sources. The isolates here identified belonged to the Pichiaceae, Debaryomycetaceae and Metschnikowiaceae families. The yeast *Cyberlindnera jadinii* behaved as a promising cell factory due to its ability to use carboxylic acids as sole carbon and energy sources. To further explore the distinct capacities to assimilate organic acids presented by *C. jadinii*, we have explored the genomic features of four *C. jadinii* strains, including wild isolates and laboratorial strains.

KEYWORDS: Non-*Saccharomyces* yeasts; Physiological characterization; Carboxylic acids*; Cyberlindnera jadinii;* Whole Genome resequencing

4.1 INTRODUCTION

As the production rates of global food waste are predicted to increase in the near future, the development of strategies for sustainable waste management has become more urgent than ever [1-3]. The food industry generates great amounts of biogenic residues. Strategies for treating these waste materials are ancient and contribute to environmental degradation and reduction of the economic profit. Activities such as composting, landfilling, anaerobic digestion, incineration, animal feeding, among other actions, represent some of the current procedures for biowaste treatment [4-6].

Considering the high amount of by-products and wastes created by food processing industries, novel solutions for its valorization are required. Such promotion of sustainable uses of the available natural resources and subsequent reduction of environmental footprint constitute the main pillars implemented in any industrial sector aiming at achieving circular bioeconomy [7, 8]. Therefore, several sustainable approaches have been suggested for converting solid wastes into added-value products: (i) direct use of biological material, where biowastes are used with or without modifications for feed applications, (ii) energy recovery, where the step from burning biomass arise from anaerobic digestion allowing the recovery of its energy contents, (iii) material recovery by biochemical or chemical extraction and/or conversion of the biomass into other valuable products such as platform chemicals, fertilizers and solvents with potential commercial applications [5, 9]. In this latter solution, studies have shown that certain biowaste-based products are capable of competing with and even outperforming conventional methods [5]. In the fruit and vegetable transforming industry, the by-products can be used for instance as chemical feedstock for fermentations as these are rich in carbohydrates, proteins, lipids, phosphorus, water, along with aromatic and aliphatic compounds [6, 10-13]. Another viable alternative for biowaste valorization is the vast and robust microbial diversity present in such ecological niches. The acidic environments developed by fruit and vegetable wastes can promote the growth of microorganisms with interesting properties. These features are triggered mostly by the microhabitat present in foods in which microbes become adapted by a phenomena designated as "domestication" [14, 15].

Different microbes can be unveiled in foods depending on the pH present [15]. Products derived from fruits, fruit juices, fermented fruits, vegetables, dairy products and salad dressings are considered as low-acid foods, meaning that the pH present is generally above 5.2 [15, 16]. In some fruits, where pH value is around 4.0, microbial colonization by molds and yeast occurs, being associated with food spoilage. In case of vegetable wastes with higher pH values (~pH 6.0), these are more susceptible to bacterial spoilage [15]. The proliferation of bacteria is associated with a tight pH range when compared to yeasts and molds. For instance, lactic and acetic acid bacteria grow well at pH values ranging from 5.0

to 6.0. Considering yeasts and molds, these species have higher tolerance to acidic pH values. Yeast grow better in pH values ranging from 4.0 to 6.0, molds have a better growth in pHs from 3.5 to 5.0 [15]. In this context, microorganisms tolerant to acidic environments are good candidates for the production of valuable chemicals, such as organic acids [16-18]. The great tolerance of *S. cerevisiae* to low pH values and to high sugar concentration [19, 20], together with its genetic manipulation, prompted its use in the biotechnology industry. Nevertheless, *S. cerevisiae* has its drawbacks, as the absence of an efficient system for product secretion [21]. Synthetic biology and metabolic engineering are important tools for the design of robust yeast cell factories [22-24]. Some microorganisms naturally produce chemicals, such as carboxylic acids [25-27]. Carboxylic acids can have multiple industrial applications as food additives, pharmaceutical and cosmetic excipients [28]. These compounds can be used particularly as chemical intermediates for the production of several products as biodegradable polymers, bulk-chemicals, potentially replacing petroleum-based and synthetic ones [29-31]. Carboxylic acids can be used by microorganisms as sole carbon and energy sources, or can be final products or by-products of the fermentative processes [25]. Some non-*Saccharomyces* yeast species like *Kluyveromyces lactis, Debaryomyces hansenii, Pichia kudriavzevii* (former *Issatchenkia orientalis*), *Komagataella pastoris* (former *Pichia pastoris*), *Candida utilis* (teleomorph state as *Cyberlindnera jadinii*) revealed promising physiological characteristics that turn them as alternative hosts for biotechnological purposes [14, 18, 27, 32, 33]. In this work we exploited the microbial biodiversity present in agro-food by-products to identify novel robust yeast strains adapted to such environments. The genetic and phenotypic variability from these natural isolates was evaluated, along with their morphological traits. The indigenous species described in this work might display a great potential to become promising biosustainable platforms, given their ability to utilize carboxylic acids present in industrial biowastes as sole carbon and energy sources.

4.2 Materials and Methods

4.2.1 Microorganisms

The yeast isolates used in this work belong to the Microbial collection of the Department of Biology of the University of Minho and derived from the TransBio Project (FP7 KBBE – N°289603). Microorganisms from this collection were isolated from agro-food by-products specifically mix salad, from a vegetable processing industry. In this study, ten yeast isolates were selected for further analysis based on preliminary screenings targeted for the isolation of organic acid producing yeasts.

4.2.2 CULTURE CONDITIONS

Cell cultures were maintained on YPD (yeast extract 1% w/v, bacto-peptone 1% w/v, dextrose 2% w/v, agar 2% w/v). Yeast cells were grown in Malt extract agar 5% w/v (MEA; *Merck*) and Sporulation media (SM; dextrose 0.05% w/v, bacto-peptone 0.1% w/v, potassium acetate 1% w/v (*Difco*), yeast extract 0.1% w/v and agar 2% w/v) for morphological characterization. The physiological traits of isolates were evaluated through growth in yeast nitrogen base 0.67% w/v (YNB; *Difco*™, USA) supplemented with different carbon and energy sources: glucose 2% w/v, acetic acid 0.5% v/v (pH 6.0), lactic acid 0.5% v/v, fumaric acid 1% w/v, malic acid 1% w/v, succinic acid 1% w/v and citric acid 1% w/v (pH 5.5), at 30°C, pH 5.0. All compounds used, from glucose to organic acids, were produced by *Sigma Aldrich/Merck* (Germany). Cells were grown in liquid YPD media at 30°C, 200 rpm. Overnight cultures were adjusted to an OD_{640nm} of 0.1. Cells were serially diluted (1:10) 3 times and 3 µL drops of each dilution were spotted in plates with desired media and incubated at 30°C for 2 days. The phenotypic analysis was carried out in triplicate, and results were converted to a classification system between 0 and 3 (0: without growth on solid media; 1: weak growth; 2: moderate growth; 3: strong growth on solid media) [34]. Phenotypic variability data was analyzed using the Orange software [35].

4.2.3 MOLECULAR TYPING AND SEQUENCE ANALYSIS

The genomic DNA was extracted using the protocol described by Lõoke *et al.,* (2011) [36]. The universal primer pair ITS1 (5'-TCCGTAGGTGAACCTGCGG-3') and ITS4 (5'-TCCTCCGCTTATTGATATGC-3') was used to amplify the intervening hypervariable ribosomal ITS1-5.8S-ITS2 region of each yeast isolate [37, 38]. PCR was carried out using the following conditions: initial denaturation at 95°C for 6 mins followed by 35 cycles of denaturation (95°C for 20 seconds), annealing (53°C for 20 seconds) and extension (72°C for 1 minute), followed by a final extension (5 minutes at 72°C). A negative control was included in each run by replacing the template DNA with sterile water in the PCR mixture. DNA was visualized by electrophoresis in a 1% (w/v) agarose gel, quantified with the NanoDrop ND-1000 (Thermo Scientific, USA) and sequenced by Eurofins Genomics (Germany). The collected nucleotide sequences were aligned using ClustalW (www.genome.jp/tools-bin/clustalw) to obtain the complete sequence. For the identification of yeast species, the sequences were compared to the reference data available in NCBI (BLASTn algorithm – sequence analysis tool) by default and MYCObank (pairwise sequence alignment, www.mycobank.org/page/Pairwise_alignment) databases. Identification was based on the probability of yeast matching and similarity between standard sequences.

4.2.4 MORPHOLOGICAL CHARACTERIZATION OF YEAST SPECIES

The morphological characterization of yeast isolates included macro- and micromorphology classification according to the guidelines provided by Kurtzman *et al.* (2011) and Wickerham (1951) [39, 40]. Macromorphology analysis evaluated colonies grown in the culture media YPD, MEA and SM. Several features were analyzed: (i) whole colony appearance (circular, rhizoid, irregular, filamentous, spindle); (ii) margin (entire, undulate, lobate, filiform, rhizoid, curled); (iii) surface (smooth, glistening, rough, wrinkled, dry/powdery); (iv) elevation (flat, raised, convex, pulvinate, umbonate, crateriform); (v) pigmentation (pigmented – purple, red, yellow; or nonpigmented – white, cream) and (vi) optical property (opaque, translucent, transparent). In YPD and MEA solid media, growth of isolates was observed after one to twelve days, whereas in SM from two to ten days of growth, all at 30°C (isolate-colony pictures were taken upon 3 days of growth). For the micromorphology analysis, morphological traits of asexual cells in distinct growth media were evaluated by microscopic observation using the epifluorescence microscope (Olympus BX51) coupled with a DP71 digital camera (Olympus Portugal SA, Porto, Portugal). All images were acquired using the Olympus cellSens software. Spore formation was evaluated using *Schaeffer-Fulton* staining method for coloration of spores with malachite green and safranin from the isolates grown in SM for four days, at 30°C [41].

4.2.5 WHOLE GENOME RE-SEQUENCING AND *DE NOVO* ASSEMBLY OF *C. JADINII* STRAINS

Yeast DNA was extracted from two *C. jadinii* isolates (Isolate 7 and TB105 - internal codes for isolates) and two collection strains (PYCC2578, PYCC3092) using the DNeasy PowerMax® soil kit (Qiagen), following the manufacturer's instructions. DNA concentration was determined by Nano drop ND-1000 spectrometer. The whole genome sequence was obtained by Illumina next-generation sequencing, according to the manufacture's protocols (Illumina 2009), in paired end 104 bp mode, using an Illumina HiSeq2000 analyser (NOVOGENE Company, Beijing, China). Genome sequencing yielded a total of 3.14 Gb raw data and produced between 3.9 and 6.4 million sequence reads per strain. Sequence reads were quality controlled with FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc). Reads were firstly aligned to the publicly available genomes of *C. jadinii* (CBS1600, NBRC 0988 and NRRL Y-1542), using bowtie2 with default parameters [42]. After results validation, the genome of the strain NBRC 0988 was used as a reference for further analysis, due to higher homology detected (data available in the Supplementary data section). Sequences were aligned with SAMtools v.1.11 [43] using the commands view, sort, index and mpileup. All possible variants including frameshift insertions/deletions (Indels) and single nucleotide polymorphisms (SNPs) were then called from the aligned sequences, using Annovar

[44], with the following filters: QUAL > 40 and DP > 10. In a second phase, the genomes of the four sequenced *C. jadinii* strains were de-novo assembled using Spades v. 3.14.1 [45], testing k-mer lengths 21, 33, 55, 77, 99 and 121.

4.2.6 PCA ANALYSIS

After alignment of *C. jadinii* genomes to the reference using miniclip2 [46], principal component analysis was used to assess differences between isolates. The function "Full genome/assembly alignment" (-ax asm10) was used with the divergence level at 1% for better alignments. Having all samples aligned to the reference, duplicate reads were excluded using SAMtools [43, 47] *markdup* function and later samples had their unmapped reads removed also using SAMtools. With the samples filtered, these were converted into genotype array using pileupCaller (SequenceTools – pileupCaller assessed in https://github.com/stschiff/sequenceTools), with the final format being EIGENSTRAT. After this, the array was converted to PLINK format where samples were further trimed to remove unnecessary information. With the files set up, the final PCA visualization was obtained [48].

4.2.7 GENOME AND FUNCTIONAL ANNOTATION

The seven *C. jadinii* genomes were annotated with the AUGUSTUS [49] software, in order to establish potential coding regions in each of the *C. jadinii*' s chromosomes, using S. cerevisiae S288c as training set. Relevant information was extracted, in particular start and end positions of coding regions, strain orientation and known homologs in the reference genome of *S. cerevisiae* (strain S288c). The potential coding regions reported by AUGUSTUS were extracted from the complete *C. jadinii* genome into a FASTA file. The functional genomic annotation was done by EggNOG-mapper v5.0 [50], in view of the predicted proteins using AUGUSTUS, and the results were outlined in the light of the Ontology (GO) terms, KEGG pathways [51] and clusters of orthologous groups – COG with their associated functional categories [52].

4.3 RESULTS

4.3.1 PHYSIOLOGICAL CHARACTERIZATION OF ISOLATES

In this study, ten isolates from vegetable-biowastes acidic niches were grown on solid media containing monocarboxylic (acetic and lactic acids), dicarboxylic (malic, succinic and fumaric acids) and tricarboxylic acids (citric acid), as sole carbon and energy sources at 30°C (Figure 1).

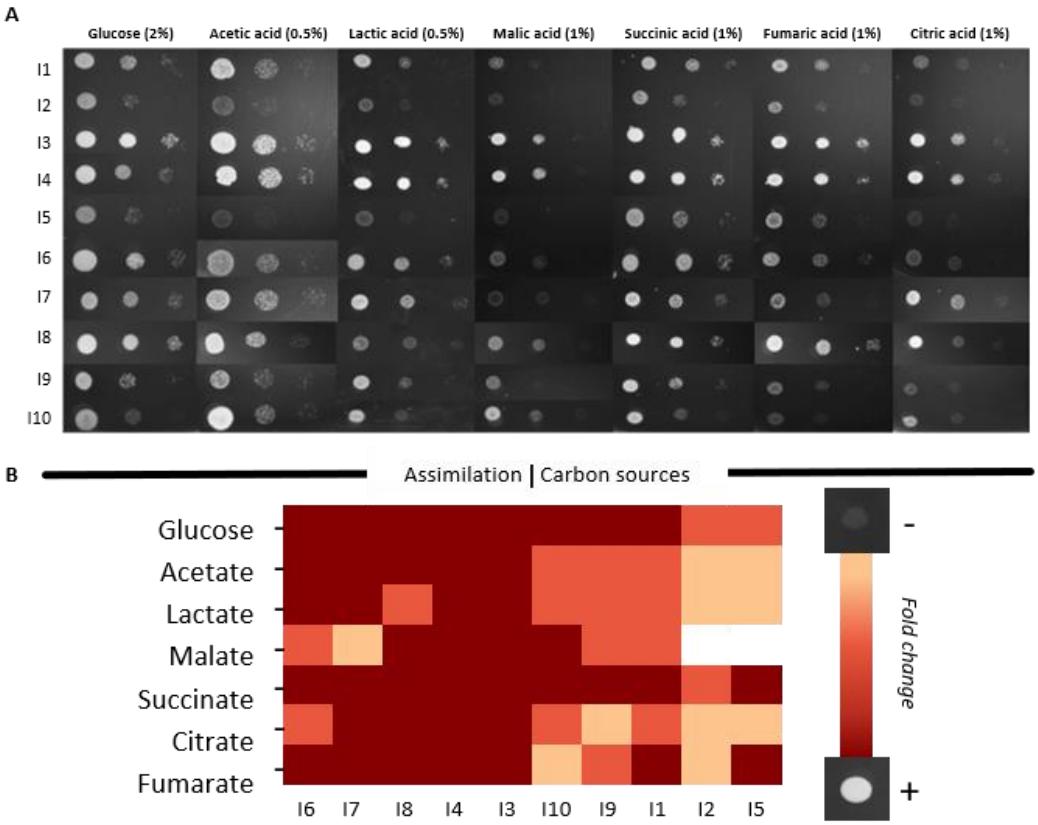


Figure 1. Growth pattern of ten wild yeast isolates. A) Assimilation profiles on YNB containing glucose (2% w/v), acetic acid (0.5% v/v; pH 6.0), lactic acid (0.5% v/v; pH 5.0), malic acid (1% w/v; pH 5.0), succinic acid (1% w/v; pH 5.0), fumaric acid (1% w/v; pH 5.0) and citric acid (1% w/v; pH 5.5) as sole carbon and energy source. B) Heatmap representing the capacity of each isolate (I1 to I10) to assimilate the carbon sources shown in A. The color reflects the normalized fitness component measure: white=no growth; salmon=weak, light rose=moderate and red=strong growth using a particular carbon source.

All isolates assimilated effectively glucose as a sole carbon source at 30°C. Isolates 3 and 4 presented a strong growth in all tested carbon sources. Isolates 6, 7, and 8, also showed an improved growth on acetate, succinate and fumarate media. Succinate was the carbon source most widely utilized. The isolates I2 and I5 were unable to grow in malate, and in particular the I2, presented poor growth in the all the carbon sources (Figure 1A). The isolates were grouped according to their phenotypic profile demonstrating the great biodiversity found among these ten isolates (Figure 1B).

4.3.2 Species identification

The sequence of amplified fragments from the ITS1-5.8S-ITS2 region of the ten isolates was compared with sequences available at NCBI and MYCOBank public databases. The PCR-ITS analysis revealed the

identification of 8 species (Table 1) belonging to the genera: *Candida* (three isolates), *Pichia* (five isolates), *Cyberlindnera* (one isolate), and *Apiotrichum* (one isolate). Nine of the identified species belongs to the phylum Ascomycota and one, *Apiotrichum brassicae*, to the phylum Basidiomycota.

Table 1. Identification of the wild yeast isolates by molecular typing using internal transcribed sequencing (ITS)

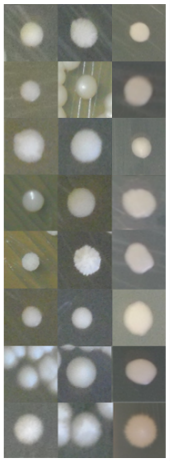
Isolates	ITS match	Sequence length (bp)	Similarity
I1	*Pichia fermentans*	588	99.7
I2	*Candida tetrigidarum*	478	99.2
I3	*Pichia kudriavzevii*	406	100
I4	*Pichia kudriavzevii*	434	100
I5	*Candida inconspicua*	387	97.5
I6	*Pichia barkeri*	368	98.9
I7	*Cyberlindnera jadinii*	478	100
I8	*Candida tropicalis*	494	99.8
I9	*Pichia fermentans*	355	100
I10	*Apiotrichum brassicae*	442	99.8

PCR products of ITS polymerized region revealed different band sizes for different species varying between 355 and 588 nucleotides. The cross-results from the two databases showed that isolates at least 97.5% of identity with known species, with a total 99% coverage (Table 1). If distinct taxa were assigned by the two databases, we considered the designation attributed by MYCOBank with a more updated nomenclature for yeasts.

4.3.3 MORPHOLOGICAL CHARACTERIZATION

Eight isolates, one per yeast species, were grown on three culture media for 3-days at 30°C (Figure 2) aiming at analyzing colony and cellular morphology. The guidelines provided by Kurtzman, *et al.* (2011) [39] were used for colony analysis. The results observed for each species are summarized in Figure 2.

A

Yeast isolates		Macromorphology parameters												Colony appearance		
		Whole colony			*Margin*			*Surface*			*Elevation*					
		MEA	YPD	SM	MEA	YPD	SM	MEA	YPD	SM	MEA	YPD	SM	MEA	YPD	SM
P. fermentans	I1	circular	irregular	circular	entire	undulate	entire	smooth	rough	smooth	convex	raised	raised			
C. tetrigidarum	I2	circular	circular	circular	entire	entire	entire	smooth	smooth	smooth	convex	pulvinate	flat			
P. kudriavzevii	I3	circular	circular	circular	entire	entire	entire	smooth	smooth	smooth	umbonate	pulvinate	pulvinate			
C. inconspicua	I5	circular	circular	circular	entire	entire	entire	smooth	smooth	smooth	convex	pulvinate	flat			
P. barkeri	I6	circular	rhizoid	irregular	undulate	undulate	entire	rough	rough	smooth	umbonate	umbonate	flat			
C. jadinii	I7	circular	circular	circular	entire	entire	undulate	smooth	smooth	smooth	raised	convex	raised			
C. tropicalis	I8	circular	circular	circular	entire	entire	entire	smooth	smooth	smooth	growth into medium	pulvinate	raised			
A. brassicae	I10	circular	rhizoid	irregular	rhizoid	undulate	filiform	rough	wrinkled	rough	umbonate	umbonate	growth into medium			

B

Yeast isolates		*Pigmentation*	*Optical property*
P. fermentans	I1	All colonies were non-pigmented in tested media	All colonies had opaque appearance over tested media
C. tetrigidarum	I2		
P. kudriavzevii	I3		
C. inconspicua	I5		
P. barkeri	I6		
C. jadinii	I7		
C. tropicalis	I8		
A. brassicae	I10		

Figure 2. Colony morphology of the eight yeast species grown on solid media after 3 days of growth at 30°C. A) Macromorphology of colonies considering whole colony, margin, surface, elevation and colony appearance onto distinct growth media. B) Analysis of pigmentation and optical properties of the colonies. Initials stand for MEA – malt extract media; YPD – yeast extract-peptone-dextrose and SM – sporulation media.

Whole colony, margin and surface were more evident among *P. fermentans*, *P. barkeri* and *A. brassicae* isolates. The colony elevation was dependent on the culture media. Considering the pigmentation and optical properties, all species were nonpigmented and had opaque appearance. The micromorphological characterization was conducted through a microscopic analysis of the isolates grown in the different culture media (Figure 3).

Pichia fermentans cells showed distinct growth profiles on YPD and MEA media (12 days of growth at 30°C). Cell size varied between 2.0-4.5 x 3.0-6.0 µm, and cells occur isolated and in pairs with ovoid form but also apparently are with ability to pseudohyphae. The cell morphology of *Candida tetrigidarum* is ellipsoidal, 2.5-3.5 x 3.7-7.5 µm, and may occur individually and in pairs, with holoblastic bipolar budding. *Pichia kudriavzevii* cells present ellipsoidal but also elongate shape occurring in single or in pairs with dimensions 2.3-3.0 by 4.5-8.0 µm, with bipolar budding on a wide base. *C. inconspicua* cells present a cylindrical and ovoid form, a cell size of 2.0-3.0 x 3.0-4.5 µm, with presence of asexual cells that reproduce via holoblastic polar budding on a narrow base. *Pichia barkeri* cells have elongate form (2.3-4.0 x 6.0-8.3 µm) with ability to produce pseudohyphae on MEA and YPD media. *C. jadinii* cells display an ellipsoidal and elongate shape, a cell size of 3.0-4.5 x 4.5-7.5 µm, occurring in individually and in pairs with associated polar budding on a narrow base. *Candida tropicalis* cells are subglobose, round to ovoid forms, with a cell size of 4.5-6.8 x 4.0-6.8 µm, occurring singly and with associated polar budding. The basidiomycetous yeast *A. brassicae* exhibited cells with branched shape and also septate hyphae.

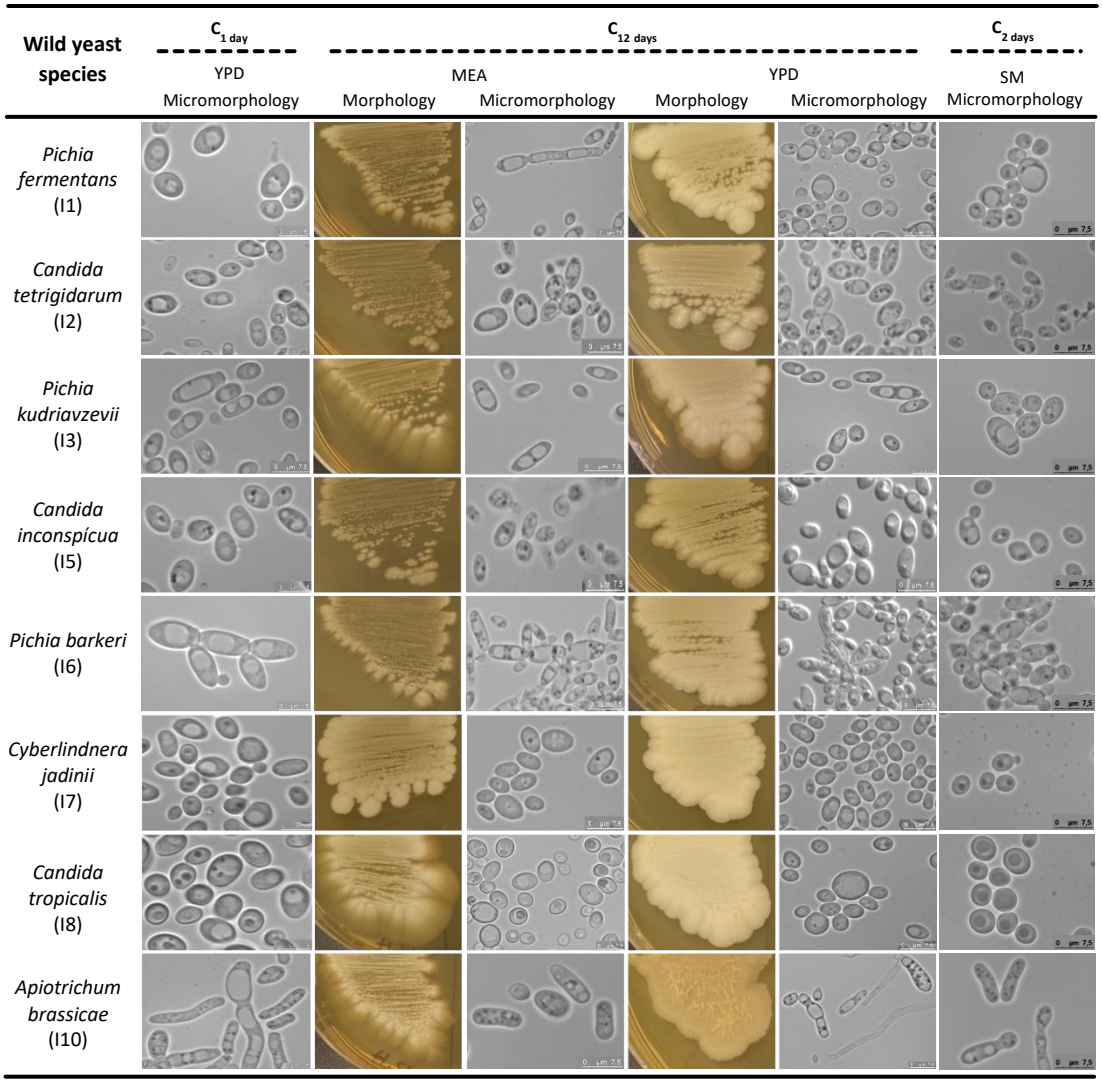


Figure 3. Morphology and micromorphology of the wild yeast species. Microscopic images were obtained by differential interference contrast (DIC) after one, two, four and twelve days of growth on MEA, YPD and SM media. Scale bars = 5-7.5 µm. MEA – malt extract media; YPD – yeast extract-peptone-dextrose; SM – sporulation media.

4.3.4 Morphological and Physiological characterization of *Cyberlindnera jadinii* strains

The *C. jadinii* laboratory strains (DSM 2361; PYCC 2578; PYCC 3092) were characterized considering the macro- and micromorphological traits (Figure 4).

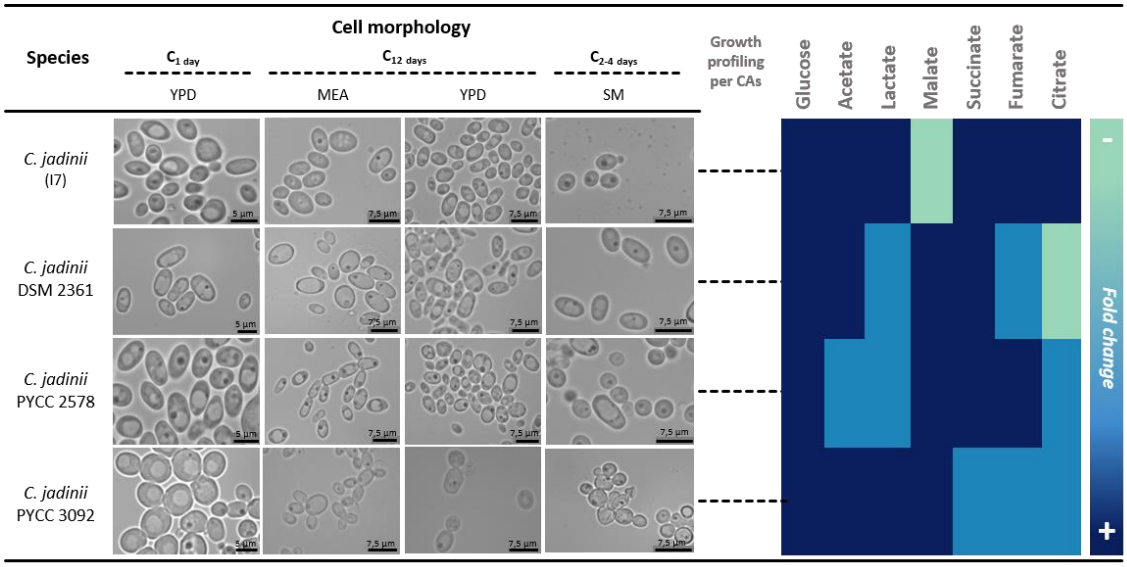


Figure 4. Characterization of *Cyberlindnera jadinii* strains DSM 2361, PYCC 3092, PYCC 2578 and isolate I7. Right panel - Cells were grown at 30°C on MEA, YPD and SM. Microscopic images were obtained by differential interference contrast (DIC) with scale bars of 5 µm for 1 day-YPD and 7.5 µm for the other samples. MEA – malt extract media; YPD – yeast extract-peptone-dextrose; SM – sporulation media. Left panel - Heatmap of the capacity of each strain to grown on YMB supplemented with glucose (2% w/v), acetic acid (0.5% v/v; pH 6.0), lactic acid (0.5% v/v, pH 5.0), malic acid (1% w/v, pH 5.0), succinic acid (1% w/v, pH 5.0), fumaric acid (1% w/v, pH 5.0) and citric acid (1% w/v, pH 5.5) at 30°C. Color scheme reflects the normalized fitness component measure: white=weak, light blue=moderate and dark blue=strong growth using a particular carbon source.

Cells of *C. jadinii* DSM 2361 with ovoid and ellipsoidal shape, after twelve days of growth in rich media with sizes between 3.0-6.0 x 5.0-8.0 µm, reproduced by polar budding. *C. jadinii* PYCC 3092 cells revealed a rounded to ellipsoidal shape, cell size of 3.0-4.5 x 4.5-6.0 µm, and reproduce by multilateral budding. Helmet-shaped ascospores released from deliquescent asci were also detected in sporulation media for this strain and for *C. jadinii* I7 (Figure S1). The cells of PYCC 2578 strain exhibited rounded and ovoid forms, with size of 3.0-4.5 x 6.0-7.5 µm, with polar budding. Compared to the laboratory strains, the wild *C. jadinii* cells displayed a pronounced ellipsoid shape. All laboratory strains assimilated glucose as well as carboxylic acids as sole carbon and energy sources (Figure 4, heatmap). The isolate I7 presented strong growth in all carboxylates tested, with exception for malate. Acetate (monocarboxylate) and succinate (dicarboxylate) were the two carboxylates better assimilated by all species.

4.3.5 GENOME ANALYSIS OF *C. JADINII*

To understand the genomic features that can be related to the phenotypic variations observed in *C. jadinii* strains, we sequenced the genomes of four *C. jadinii* strains, the I7 wild isolate, two lab-collection strains (PYCC 2578 and PYCC 3092) and another isolate (identified as TB105) belonging to the Department of Biology collection, also derived from the TransBio project. The genomes were sequenced with a high (73-132x) coverage using Illumina *de novo* assembly technology (Table 2).

Table 2. Data characteristics and sequencing statistics from the *de-novo* assembly of *C. jadinii* wild and collection genomes

Components	***Cyberlindnera jadinii* strains**			
	I7	**TB105**	**PYCC 2578**	**PYCC 3092**
Number of reads	6,406,326	5,081,695	3,981,001	6,243,058
Coverage	131.4x	73.4x	81.57x	127x
GC-content (%)	43.0	43.0	43.0	43.0
	***De novo* Assembly with Spades**			
Assembly length (bp)	21,305,676	20,802,315	18,116,029	18,381,892
Number of scaffolds	12,198	11,873	8,510	9,776
N50 (bp)	2863	2903	4165	3765
Number of Ns	36,420	40,900	36,740	38,310
Number of scaffolds > 1000pb	5,824	5,532	3,975	3,884
Assembly length > 1000pb	17,339,108	16,929,544	15,252,290	14,887,584
Proteins predicted	5579	5365	4687	4617

All the genomes presented a predicted GC-content of 43%. *De novo* assembly of the filtered high-quality sequences from all genomes resulted in a variable number of scaffolds being higher in wild-isolates than in PYCC strains. The N50 was smaller for the wild isolates (2900 bp) than for and PYCC 2578 and PYCC 3092 (~4000bp). The total genome length size was higher in wild isolates (~ 20Mbp) than in PYCC strains (~18Mbp). The four strains presented a highly variable number of predicted proteins, around 4600 in the PYCC strains, and up to 5580 in the I7. The N50 value and the high number of scaffolds obtained above 1000 bp suggests a high quality of both the sequencing process and the *de-novo* assembly. For a more comprehensive study of the *C. jadinii* genome, data from genomes already deposited in NCBI belonging to strains NBRC0988, CBS 1600 and NRRL Y-1542 [53] were also analyzed. These two last *C. jadinii* strains are described as being alias, however we decided to keep the results from both genome sequencing projects in our analysis as the data reported for each sequencing project is different. The results

from the overall alignment rate, using strain NBRC0988 as reference, showed the best alignment rate, with values above 96% (table S1). Thus, the *C. jadinii* NBRC0988 genome was used to make a comparative genomic study between strains and to generate the principal component analysis (PCA) (Figure 5).

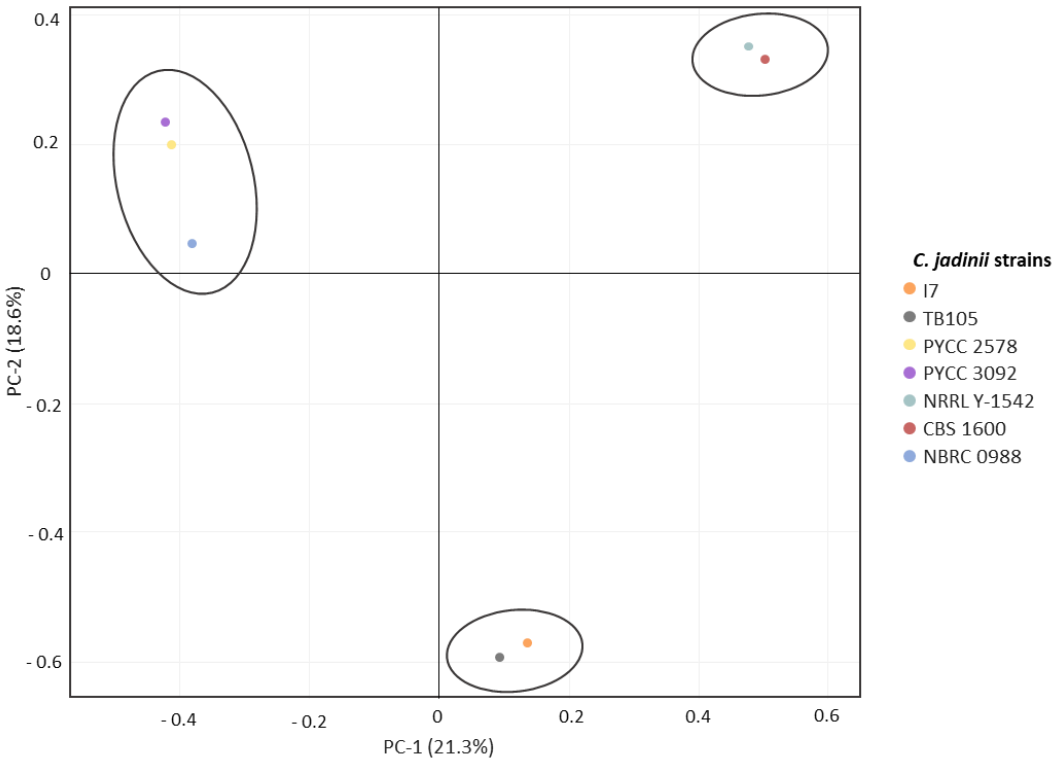


Figure 5. Genetic diversity obtained for seven *Cyberlindnera jadinii* strains as disclosed by the principal component analysis (PCA) plot. Scores attributed are based on the *C. jadinii* strain distribution.

The PCA diagram depicts the variability of *C. jadinii* strains, where 40% of the genetic diversity is explained by the two first components (PC-1 – 21.3%, PC-2 – 18.6%). The strains were distributed into three quadrants/directions in the PCA visualization and three major groups are present; i) NRRL Y-1542 and CBS 1600 strains; ii) PYCC 2578, PYCC 3092 and NBRC 0988 strains; iii) wild isolates I7 and TB105. In this analysis, the wild isolates are the most distant group, which is in accordance to their higher genetic variability, e.g. higher number of ORFs and assembly length. In addition, further genetic relatedness was found among strains considering their origin as the closest strains share the same habitat (isolation sources), such is the case of the wild isolates originated from biowastes (mix of salad). The PYCC 3092 and NBRC 0988 strains were isolated from fodder yeast, although in this group is also included the PYCC 2578 strain, isolated from the digestive tract of a cow. The genomes from the NRRL Y-1542 and CBS 1600 strain are also clustered in this

analysis, as expected since these resulted from sequencing projects of the same strain. However, these two genomes showed slight variance, as detected by the first and second principal components, probably due to differences in the sequencing projects and/or artifacts in the genome assembly and annotation pipeline used since it was done by different groups. PCA visualization using other components was also obtained (figure S2 in supplementary data). However, even though the inclusion of the third component (PC-3) still accounts for a relevant data variance, we could not find any biological relevance in the analysis.

In this work, we also analyzed the functional clustering of *C. jadinii* predicted proteins (Figure 6). The basal analysis focused on the functional annotation of the *C. jadinii* strain NBRC0988 (Figure 6A), used as reference. From a total of 6435 proteins predicted by AUGUSTUS, egg-NOG software attributed 24 COG categories to 5370 proteins (83.4%). The same prediction was performed for the other *C. jadinii* strains (see table S2 in supplementary data). The candidate categories include: 1) RNA processing/modification, including translation, transcription, replication, recombination and repair; 2) proteins associated to energy production and conversion, amino acid/carbohydrate/coenzyme/lipid transport, inorganic ion transport and metabolism, and secondary metabolites biosynthesis, transport and catabolism; 3) cell processes and signaling, such as cell division, cell wall/membrane/envelope biogenesis to intracellular trafficking, secretion, and vesicular transport, protein turnover and chaperones functions; 4) poorly characterized or without an associated function. Among the 5370 predicted genes, the majority of functional genes are classified in the following categories: metabolism (29%; 1552), cellular processes and signaling (25%; 1365), information storage and processing (24%; 1291) and unknown function (22%; 1162). A comparative analysis between the *S. cerevisiae* S288c genome and the seven genomes from *C. jadinii* strains, revealed slight differences in the genes clustered in each COG category (fig.6-B). The *S. cerevisiae* strain was used herein as the reference due to the extensive gene functional analysis, and its biotechnological relevance. The PYCC 2578 strain presents a considerable smaller number of genes involved in the cell energy production and conversion, compared to the other *C. jadinii* and *S. cerevisiae* strains. As for the clustered genes involved in amino acid transport/metabolism (COG-E), all *C. jadinii* strains present a higher number of genes associated to this category, when compared to *S. cerevisiae*. The same happened for the catabolism of sugars (COG-G category: carbohydrate transport and metabolism), secondary metabolites biosynthesis, transport, and catabolism category (COG-Q).

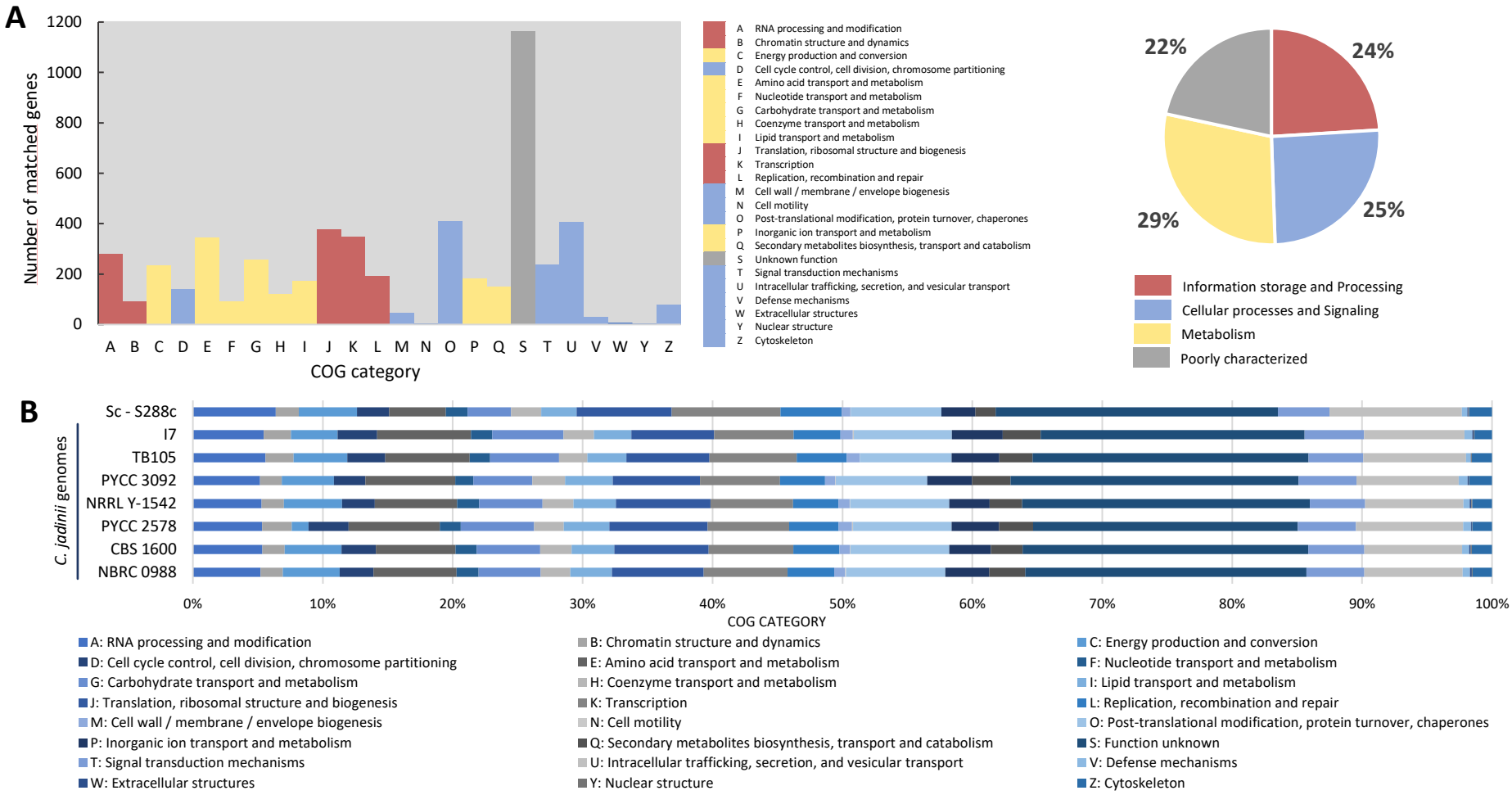


Figure 6. Functional annotation via EggNOG software of *Cyberlindnera jadinii* and *Saccharomyces cerevisiae* genes. Annotations were organized in 24 main categories belonging to clusters of orthologous groups (COG). A) Number of genes clustered in each of the 24 COG categories for the *C. jadinii* NBRC0988. Colors stand for the functional categories analyzed. B) Comparison of the clustered genes (%) for each COG category between the *S. cerevisiae* S288c and the seven *C. jadinii* genomes.

4.4 Discussion

Bioprocesses display a promising and environmentally friendly option to take over conventional processes currently used to produce added-value products such as fine chemicals, fuels and other useful end-products for commercial purposes. The improvement of bioconversion processes can be achieved through the development of robust stress-resistant strains able to convert efficiently sugars into valuable compounds, e.g. organic acids [17]. In the present study, we identified and characterized nine yeast species isolated from acidic environments. This characterization included a morphological analysis (colony and cell morphology), as well as an evaluation of the capacity to use organic acids as sole carbon and energy sources.

For *C. inconspicua, C. tropicalis* and the basidiomycetous *A. brassicae,* all the morpho- and physiological evaluation here provided contribute to enrich the characterization of these species as the information available in the literature is very scarce [54-57].

Pichia fermentans and *Pichia barkeri*, whose morphology was described by Kurtzman *et al.* (2011) [55] and reported to utilize glucose, D-xylose, DL-lactate, succinate, citrate (weak growth) is also able to utilize acetate, fumarate, and malate.

The morphological data here reported for *C. tetrigidarum* is in accordance with literature [58]. Regarding the capacity to use carboxylic acids as carbon sources, this was extended to dicarboxylates, fumarate and malate [59].

Physiological screening revealed that *P. kudriavzevii*, *C. tropicalis* and *C. jadinii* wild isolates showed a strong growth in all tested organic acids as sole carbon and energy sources. Taking into account the data from the drop-test, among the isolates, *P. kudriavzevii* presented the highest ability to growth on carboxylic acids. These observation points to the existence of specific transport systems for carboxylates in the plasma membrane, since at the pH values tested, the acids are mainly in the dissociated form and cannot diffuse through the plasma membrane, requiring a transport system [25]. Previous studies already reported the use of *P. kudriavzevii* as a host for lactate and succinate production, with titers of 135 g/L and 11.63 g/L respectively, due to its tolerance to low pH values [60-63]. From these data it is possible to speculate that this species may present interesting traits for its exploitation for the industrial bioproduction of carboxylic acids. Considering the ability to assimilate carboxylates, *P. kudriavzevii*, *C. tropicalis* and *C. jadinii* might also become promising cell factories for the utilization of organic acids present in biowastes.

The yeast *C. utilis,* the anamorph of *C. jadinii* [53], display activity for carboxylic acid transport in the plasma membrane [64, 65]. Other yeast species within the *Cyberlindnera* clade evidenced

relevant features for biotechnological applications [66-68]. *Cyberlindnera* species were reported to easily assimilate sugars, organic acids and nitrogen sources [66, 67, 69]. In particular *Cyberlindnera jadinii* is able to grow in in substrates like biomass-derived wastes (as hardwood hydrolysates from the pulp industry), has a high tolerance to the temperature range between 19 to 37 °C and to long-term mild acid pH [53]. The wild isolates of *C. jadinii* presented an improved growth on carboxylic acids, when compared to the other strains, which suggests that these substrates might be present in mixed salad vegetables. Concerning these physiological results, the wild I7 isolate was the most promising for the utilization of carboxylic acids.

At the cell morphology level, similar cell shapes were found in all strains, and DSM 2361 colonies were more similar to I7. Reproduction occurred by polar to multilateral budding and, hat- and helmet-shaped ascospores were detected in the I7 and PYCC 3092 strains, which is accordance with type of spores reported for this yeast [39, 66]. The ascospore morphology is often used as tool for genus delimitation, as recognized by Boekhout and Kurtzman [70]. Several ploidy levels were reported in *C. jadinii* strains. A diploid state was detected in the *C. jadinii* NRRL Y-1542 strain [71, 72], the ploidy of the *C. jadinii* ATCC 9950/DSM 2361/NBRC 0988 varied from 2n, 3n to 4n according to the different methodologies used [73-76], and the *C. jadinii* CCY 39-38-18 strain a tetraploid state was reported [77].

In this work a comparative analysis of four strains of *C. jadinii* was performed namely: the DSM 2361 (NBRC 0988), which genome is already sequenced and available at NCBI; the PYCC 3092 and the PYCC 2578 strains characterized in previous studies for the transport of carboxylic acids [32, 64, 65] and the wild isolate I7. Overall, the *C. jadinii* I7 and PYCC strains were more similar between them. The genomic features, e.g. GC-content and number of proteins, of the previously sequenced *C. jadinii* strains reanalyzed in this study are in agreement with the published data [53]. The PCA visualization suggests the existence of three-clusters. The present landscape of high intra-strain diversity might be explained by genetic rearrangements due to selective pressure in specific ecological niches [78]. Several studies reported the phenotypic evolution driven by environmental adaptation [78, 79]. Moreover, this variation may be associated with different ploidy levels, a phenomena already reported in *C. jadinii* strains that can range from 2n to 5n [53], although we were unable to explore this aspect in this study.

The functional annotation of the seven *C. jadinii* genomes revealed the presence of a high number of genes assigned to categories such as (i) amino acid transport/metabolism (COG-E), (ii) carbohydrate transport and metabolism (COG-G) and (iii) secondary metabolites biosynthesis,

transport, and catabolism category (COG-Q). This reinforces in part the data that reports that this yeast is able to assimilate a considerable amount of compounds when compared to *Saccharomyces* species [53]. In addition, the presence of a high number of genes assigned to the COG-G category reflected the ability of *C. jadinii* to assimilate several carbohydrates such as sucrose, glucose, raffinose, glycerol, xylose or maltose [80-82]. The same happens for COG-Q category, correspondent to genes involved in secondary metabolites biosynthesis, transport and catabolism, in which all seven *C. jadinii* strains had a higher number of genes assigned than *S. cerevisiae.* The predominance of genes required for the transport of metabolites, such as organic acids, can justify the improved growth of *C. jadinii* strains using these compounds as the sole carbon and energy source [32, 64, 65]. As for the annotated proteins with unknown function, it is necessary to perform a more detailed study based on the search for homologs and conserved domains that can help to predict the function of these proteins.

4.5 Conclusions

Currently, the sustainable production of organic acids is a major concern of the industrial biotechnology. Product toxicity remains one of the major bottlenecks to achieve a cost-effective fermentative process for production of carboxylic acids. The biorefinery industry requires hosts that exhibit an excellent tolerance to acidic pH values and high product titer conditions, being able to grow and produce the target organic acid at high rates. This study explores the morphological and physiological characterization of wild yeast, isolated from biowastes, providing novel insights on their potential to utilize or even produce carboxylic acids, due to their tolerance to acidic environments, with emphasis for *Pichia kudriavzevii*, *Candida tropicalis* and *Cyberlindnera jadinii*. In particular, *C. jadinii* with its GRAS status and its known biotechnological applications, is a promising microbial cell factory. Nevertheless, further studies on these wild isolates are required to understand its metabolic routes, transport systems and regulatory networks. Concerning the species *Pichia kudriavzevii* and *Candida tropicalis,* a more in-depth study, including whole genome sequencing, may reveal interesting phenotypic and genomic variations occurred due to the selection on these specific niches.

This study, correlating physiological features and the genomic information on *C. jadinii* strains, provide the basis for the development of genetic engineering strategies for the improvement of this biotechnologically relevant yeast.

SUPPLEMENTARY DATA

Table S1. Comparison for quality validation over the alignment of sequenced strains deposited in NCBI against the Cj-strains in analysis by different parameter setting in variant calling.

C. jadinii strains	Alignment with NRRL Y-1542			Alignment with CBS1600			Alignment with NBRC 0988		
	Overall alignment rate	Variant calling	Filtering (QUAL< 40; DP <10)	Overall alignment rate	Variant calling	Filtering (QUAL< 40; DP <10)	Overall alignment rate	Variant calling	Filtering (QUAL< 40; DP <10)
I7	89.08 %	268,142	256,339	88.93%	263,822	250,728	96.61 %	250,728	267,649
TB105	87.28 %	268,802	254,511	87.28 %	261,932	247,785	96.35 %	247,785	262,875
PYCC2578	88.99 %	272,363	257,357	88.62 %	271,663	255,744	96.77 %	231,821	217,024
PYCC3092	89.08 %	257,734	244,220	88.47%	262,145	247,548	96.98 %	214,917	202,277

Table S2. Predicted proteins by AUGUSTUS versus functionally annotated proteins by Cluster of Orthologous Groups (COG)

C. jadinii strains	Proteins predictions	
	Augustus	COG
I7	5579	4421
TB105	5365	4198
PYCC 2578	4687	3637
PYCC 3092	4617	3658
NRRL Y-1542	5919	5028
CBS 1600	5952	5045
NBRC0988	6435	5370

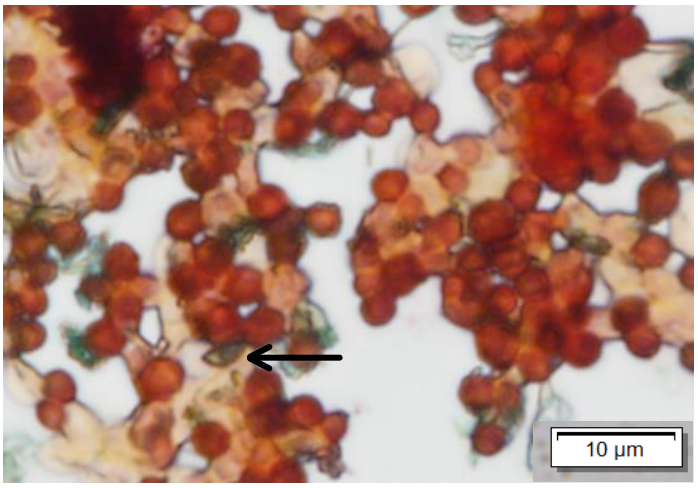
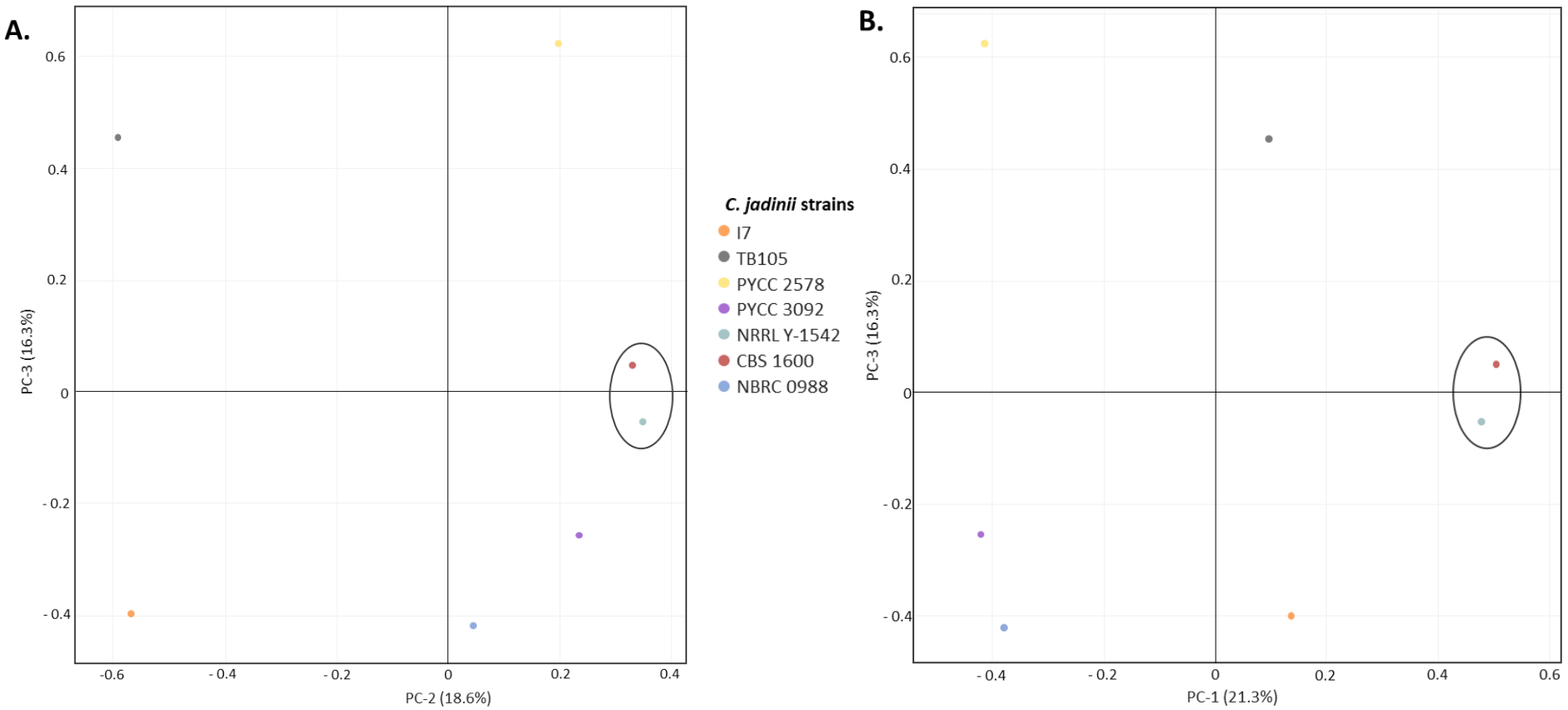


Figure S1. Helmet-shaped ascospores released from deliquescent asci observed for *C. jadinii* wild isolate (I7). Microscopic image obtained after four days of growth in sporulation media using bright field with scale bars = 10 µm, black arrow indicates the spore identified. In the BF images: green color – spore and red – vegetative cells.



Figures S2. Genetic diversity obtained for the seven *Cyberlindnera jadinii* strains as disclosed by the PCA plot of A) PC-2 vs PC-3 and B) PC-1 vs PC-3 components. Scores attributed are based on the Cj-strains distribution.

4.6 REFERENCES

1. Gustavsson, J., Cederberg, C., Sonesson, U., Van Otterdijk, R.,Meybeck, A., *Global food losses and food waste.* **2011**
2. Mavropoulos, A., *Waste management 2030: Waste Management World Magazine.* PenWell Corporation, Oklahoma, US, **2013**
3. Ren, Y., Yu, M., Wu, C., Wang, Q., Gao, M., Huang, Q.,Liu, Y., *A comprehensive review on food waste anaerobic digestion: Research updates and tendencies.* Bioresource technology, **2018**. 247: p. 1069-1076
4. FAO, *Food wastage footprint: Impacts on natural resources - Summary Report*, **2013**, Food and Agriculture Organization of the United Nations
5. Six, L., Velghe, F., Verstichel, S.,De Meester, S., *Sustainability considerations on the valorization of organic waste.* Biotransformation of agricultural waste and by-products, **2016**: p. 287-307
6. Trivedi, J., Bhonsle, A.,Atray, N., *Processing food waste for the production of platform chemicals.* Refining Biomass Residues for Sustainable Energy and Bioproducts, **2020**: p. 427-448
7. Arancon, R., Lin, C., Chan, K., Kwan, T.,Luque, R., *Advances on waste valorization: new horizons for a more sustainable society.* Energy Science & Engineering, **2013**. 1(2): p. 53-71
8. Mohan, S., Dahiya, S., Amulya, K., Katakojwala, R.,Vanitha, T., *Can circular bioeconomy be fueled by waste biorefineries–A closer look.* Bioresource Technology Reports, **2019**. 7: p. 100277
9. Fisgativa, H., Tremier, A., Le Roux, S., Bureau, C.,Dabert, P., *Understanding the anaerobic biodegradability of food waste: relationship between the typological, biochemical and microbial characteristics.* Journal of environmental management, **2017**. 188: p. 95-107
10. Dugmore, T., Clark, J., Bustamante, J., Houghton, J.,Matharu, A., *Valorisation of biowastes for the production of green materials using chemical methods.* Chemistry and Chemical Technologies in Waste Valorization, **2017**: p. 73-121
11. Biotechnology, E.F.o., *Biowaste biorefiney in Europe: opportunities and R&D needs - Report*, **2012**. p. 1-23,
12. Gavrilescu, M., *Biorefinery systems: an overview.* Bioenergy research: advances and applications, **2014**: p. 219-241
13. Maina, S., Kachrimanidou, V.,Koutinas, A., *A roadmap towards a circular and sustainable bioeconomy through waste valorization.* Current Opinion in Green and Sustainable Chemistry, **2017**. 8: p. 18-23
14. Querol, A., Belloch, C., Fernández-Espinar, M.,Barrio, E., *Molecular evolution in yeast of biotechnological interest.* International Microbiology, **2003**. 6(3): p. 201-205
15. Erkmen, O.,Bozoglu, T., *Factors Affecting Microbial Growth in Foods.* Food Microbiology: Principles into Practice, **2016**: p. 91-106, 10.1002/9781119237860.ch5
16. Attfield, P., *Stress tolerance: the key to effective strains of industrial baker's yeast.* Nature biotechnology, **1997**. 15(13): p. 1351-1357

17. Guan, N.,Liu, L., *Microbial response to acid stress: mechanisms and applications.* Applied microbiology and biotechnology, **2020**. 104(1): p. 51-65
18. Sun, W., Vila-Santa, A., Liu, N., Prozorov, T., Xie, D., Faria, N., Ferreira, F., Mira, N.,Shao, Z., *Metabolic engineering of an acid-tolerant yeast strain Pichia kudriavzevii for itaconic acid production.* Metabolic engineering communications, **2020**: p. e00124
19. Lodi, T., Diffels, J., Goffeau, A.,Baret, P.V., *Evolution of the carboxylate Jen transporters in fungi.* FEMS Yeast Res, **2007**. 7(5): p. 646-656
20. Abbott, D., Zelle, R., Pronk, J.,Van Maris, A., *Metabolic engineering of Saccharomyces cerevisiae for production of carboxylic acids: current status and challenges.* FEMS Yeast Res, **2009**. 9(8): p. 1123-1136
21. Kim, H., Yoo, S.,Kang, H., *Yeast synthetic biology for the production of recombinant therapeutic proteins.* FEMS Yeast Res, **2015**. 15(1): p. 1-16
22. Walker, R.,Pretorius, I., *Applications of yeast synthetic biology geared towards the production of biopharmaceuticals.* Genes, **2018**. 9(7): p. 340
23. Liu, Q., Liu, Y., Chen, Y.,Nielsen, J., *Current state of aromatics production using yeast: achievements and challenges.* Current opinion in biotechnology, **2020**. 65: p. 65-74
24. Kell, D., Swainston, N., Pir, P.,Oliver, S., *Membrane transporter engineering in industrial biotechnology and whole cell biocatalysis.* Trends in biotechnology, **2015**. 33(4): p. 237-246
25. Casal, M., Paiva, S., Queirós, O.,Soares-Silva, I., *Transport of carboxylic acids in yeasts.* FEMS microbiology reviews, **2008**. 32(6): p. 974-994
26. Show, P., Oladele, K., Siew, Q., Aziz Zakry, F., Lan, J.,Ling, T., *Overview of citric acid production from Aspergillus niger.* Frontiers in Life Science, **2015**. 8(3): p. 271-283
27. Soares-Silva, I., Ribas, D., Foskolou, I., Barata, B., Bessa, D., Paiva, S., Queirós, O.,Casal, M., *The Debaryomyces hansenii carboxylate transporters Jen1 homologues are functional in Saccharomyces cerevisiae.* FEMS Yeast Res, **2015**. 15(8), 10.1093/femsyr/fov094
28. Goldberg, I., Rokem, J.,Pines, O., *Organic acids: old metabolites, new themes.* Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology, **2006**. 81(10): p. 1601-1611
29. Sauer, M., Porro, D., Mattanovich, D.,Branduardi, P., *Microbial production of organic acids: expanding the markets.* Trends in biotechnology, **2008**. 26(2): p. 100-108
30. Agler, M., Wrenn, B., Zinder, S.,Angenent, L., *Waste to bioproduct conversion with undefined mixed cultures: the carboxylate platform.* Trends in biotechnology, **2011**. 29(2): p. 70-78
31. Straathof, A., *Transformation of biomass into commodity chemicals using enzymes or cells.* Chemical reviews, **2014**. 114(3): p. 1871-1908
32. Cássio, F.,Leão, C., *A comparative study on the transport of L (-) malic acid and other short-chain carboxylic acids in the yeast Candida utilis: Evidence for a general organic acid permease.* Yeast, **1993**. 9(7): p. 743-752
33. Queirós, O., Pereira, L., Paiva, S., Moradas-Ferreira, P.,Casal, M., *Functional analysis of Kluyveromyces lactis carboxylic acids permeases: heterologous expression of KlJEN1 and*

KIJEN2 genes. Current Genetics, **2007**. 51(3): p. 161-169, 10.1007/s00294-006-0107-9

34. Mendes, I., Franco-Duarte, R., Umek, L., Fonseca, E., Drumonde-Neves, J., Dequin, S., Zupan, B.,Schuller, D., *Computational models for prediction of yeast strain potential for winemaking from phenotypic profiles.* PLoS One, **2013**. 8(7): p. e66523
35. Demšar, J., Zupan, B., Leban, G.,Curk, T., *Orange: From experimental machine learning to interactive data mining.* European conference on principles of data mining and knowledge discovery, **2004**: p. 537-539
36. Lõoke, M., Kristjuhan, K.,Kristjuhan, A., *Extraction of genomic DNA from yeasts for PCR-based applications.* Biotechniques, **2011**. 50(5): p. 325-328, 10.2144/000113672
37. White, T., Bruns, T., Lee, S.,Taylor, J., *Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics.* PCR protocols: a guide to methods and applications, **1990**. 18(1): p. 315-322
38. Pincus, D., Orenga, S.,Chatellier, S., *Yeast identification – past, present, and future methods.* Medical Mycology, **2007**. 45(2): p. 97-121, 10.1080/13693780601059936
39. Kurtzman, C., Fell, J., Boekhout, T.,Robert, V., *Methods for isolation, phenotypic characterization and maintenance of yeasts*, in *The yeasts***2011**, Elsevier. p. 87-110
40. Wickerham, L.J., *Taxonomy of yeasts.* **1951**(1029)
41. Schaeffer, A.,Fulton, M., *A simplified method of staining endospores.* Science, **1933**. 77(1990): p. 194-194
42. Langmead, B.,Salzberg, S., *Fast gapped-read alignment with Bowtie 2.* Nature methods, **2012**. 9(4): p. 357
43. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G.,Durbin, R., *The sequence alignment/map format and SAMtools.* Bioinformatics, **2009**. 25(16): p. 2078-2079
44. Wang, K., Li, M.,Hakonarson, H., *ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.* Nucleic acids research, **2010**. 38(16): p. e164-e164
45. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A., Dvorkin, M., Kulikov, A., Lesin, V., Nikolenko, S., Pham, S.,Prjibelski, A., *SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing.* Journal of computational biology, **2012**. 19(5): p. 455-477
46. Li, H., *Minimap2: pairwise alignment for nucleotide sequences.* Bioinformatics, **2018**. 34(18): p. 3094-3100
47. Li, H., *A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data.* Bioinformatics, **2011**. 27(21): p. 2987-2993
48. Price, A., Patterson, N., Plenge, R., Weinblatt, M., Shadick, N.,Reich, D., *Principal components analysis corrects for stratification in genome-wide association studies.* Nature genetics, **2006**. 38(8): p. 904-909
49. Stanke, M.,Morgenstern, B., *AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints.* Nucleic acids research, **2005**. 33(Web Server issue): p. W465-W467, 10.1093/nar/gki458

50. Huerta-Cepas, J., Forslund, K., Coelho, L., Szklarczyk, D., Jensen, L., Von Mering, C.,Bork, P., *Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper.* Molecular biology and evolution, **2017**. 34(8): p. 2115-2122
51. Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M.,Tanabe, M., *Data, information, knowledge and principle: back to metabolism in KEGG.* Nucleic acids research, **2014**. 42(D1): p. D199-D205
52. Galperin, M., Makarova, K., Wolf, Y.,Koonin, E., *Expanded microbial genome coverage and improved protein family annotation in the COG database.* Nucleic acids research, **2015**. 43(D1): p. D261-D269
53. Sousa-Silva, M., Vieira, D., Soares, P., Casal, M.,Soares-Silva, I., *Expanding the Knowledge on the Skillful Yeast Cyberlindnera jadinii.* Journal of Fungi, **2021**. 7(1): p. 36
54. Kurtzman, C., Fell, J.,Boekhout, T., *The yeasts: a taxonomic study.***2011**: Elsevier
55. Kurtzman, C., *Chapter 57 - Pichia E.C. Hansen (1904)*, in *The Yeasts (Fifth Edition)*, C.P. Kurtzman, J.W. Fell, and T. Boekhout, Editors. **2011**, Elsevier: London. p. 685-707, https://doi.org/10.1016/B978-0-444-52149-1.00057-4
56. Yarrow, D.,Meyer, S., *Proposal for Amendment of the Diagnosis of the Genus Candida Berkhout nom. cons. .* International Journal of Systematic Bacteriology, **1978**. 28(4): p. 611-615
57. Liu, X., Wang, Q., Göker, M., Groenewald, M., Kachalkin, A., Lumbsch, H., Millanes, A., Wedin, M., Yurkov, A., Boekhout, T.,Bai, F., *Towards an integrated phylogenetic classification of the Tremellomycetes.* Studies in mycology, **2015**. 81: p. 85-147, 10.1016/j.simyco.2015.12.001
58. Suh, S., Nguyen, N.,Blackwell, M., *Yeasts isolated from plant-associated beetles and other insects: seven novel Candida species near Candida albicans.* FEMS Yeast Res, **2008**. 8(1): p. 88-102, 10.1111/j.1567-1364.2007.00320.x
59. Lachance, M., Boekhout, T., Scorzetti, G., Fell, J.,Kurtzman, C., *Candida Berkhout (1923).* The yeasts, **2011**: p. 987-1278
60. Xiao, H., Shao, Z., Jiang, Y., Dole, S.,Zhao, H., *Exploiting Issatchenkia orientalis SD108 for succinic acid production.* Microbial cell factories, **2014**. 13(1): p. 121, 10.1186/s12934-014-0121-4
61. Rajgarhia, V., Koivuranta, K., Penttila, M., Ilmen, M., Suominen, P., Aristidou, A., Miller, C., Olson, S.,Ruohonen, L., *Genetically modified yeast species, and fermentation processes using genetically modified yeast.* **2017**
62. Suominen, P., Aristidou, A., Pentilla, M., Ilmen, M., Ruohonen, L., Koivuranta, K.,Roberg-Perez, K., *Genetically modified yeast of the species Issatchenkia orientalis and closely relates species, and fermentation processes using same*, **2012**, Google Patents,
63. Park, H., Bae, J., Ko, H., Lee, S., Sung, B., Han, J.,Sohn, J., *Low-pH production of d-lactic acid using newly isolated acid tolerant yeast Pichia kudriavzevii NG7.* Biotechnology and bioengineering, **2018**. 115(9): p. 2232-2242
64. Leão, C.,Van Uden, N., *Transport of lactate and other short-chain monocarboxylates in the yeast Candida utilis.* Applied microbiology and biotechnology, **1986**. 23(5): p. 389-393

65. Cássio, F.,Leão, C., *Low-and high-affinity transport systems for citric acid in the yeast Candida utilis.* Appl. Environ. Microbiol., **1991**. 57(12): p. 3623-3628
66. Cadete, R.M., Cheab, M.A., Santos, R.O., Safar, S.V., Zilli, J.E., Vital, M.J., Basso, L.C., Lee, C.-F., Kurtzman, C.P.,Lachance, M.-A., *Cyberlindnera xylosilytica sp. nov., a xylitol-producing yeast species isolated from lignocellulosic materials.* International journal of systematic and evolutionary microbiology, **2015**. 65(9): p. 2968-2974
67. Kamat, S., Gaikwad, S., Kumar, A.,Gade, W., *Xylitol production by Cyberlindnera (Williopsis) saturnus, a tropical mangrove yeast from xylose and corn cob hydrolysate.* Journal of applied microbiology, **2013**. 115(6): p. 1357-1367
68. Junyapate, K., Jindamorakot, S.,Limtong, S., *Yamadazyma ubonensis fa, sp. nov., a novel xylitol-producing yeast species isolated in Thailand.* Antonie van Leeuwenhoek, **2014**. 105(3): p. 471-480
69. Buerth, C., Tielker, D.,Ernst, J., *Candida utilis and Cyberlindnera (Pichia) jadinii: yeast relatives with expanding applications.* Applied microbiology and biotechnology, **2016**. 100(16): p. 6981-6990, https://doi.org/10.1007/s00253-016-7700-8
70. Boekhout, T.,Kurtzman, C., *Principles and methods used in yeast classification, and an overview of currently accepted yeast genera.* Non-conventional yeasts in biotechnology, **1996**: p. 1-81
71. Kurtzman, C., Johnson, C.,Smiley, M., *Determination of conspecificity of Candida utilis and Hansenula jadinii through DNA reassociation.* Mycologia, **1979**. 71(4): p. 844-847
72. Riley, R., Haridas, S., Wolfe, K., Lopes, M., Hittinger, C., Göker, M., Salamov, A., Wisecaver, J., Long, T.,Calvey, C., *Comparative genomics of biotechnologically important yeasts.* Proceedings of the National Academy of Sciences, **2016**. 113(35): p. 9882-9887
73. Kondo, K., Saito, T., Kajiwara, S., Takagi, M.,Misawa, N., *A transformation system for the yeast Candida utilis: use of a modified endogenous ribosomal protein gene as a drug-resistant marker and ribosomal DNA as an integration target for vector DNA.* Journal of bacteriology, **1995**. 177(24): p. 7171-7177
74. Ikushima, S., Fujii, T.,Kobayashi, O., *Efficient gene disruption in the high-ploidy yeast Candida utilis using the Cre-loxP system.* Bioscience, biotechnology, and biochemistry, **2009**. 73(4): p. 879-884
75. Ikushima, S., Fujii, T., Kobayashi, O., Yoshida, S.,Yoshida, A., *Genetic engineering of Candida utilis yeast for efficient production of L-lactic acid.* Bioscience, biotechnology, and biochemistry, **2009**. 73(8): p. 1818-1824
76. Rupp, O., Brinkrolf, K., Buerth, C., Kunigo, M., Schneider, J., Jaenicke, S., Goesmann, A., Pühler, A., Jaeger, K.,Ernst, J., *The structure of the Cyberlindnera jadinii genome and its relation to Candida utilis analyzed by the occurrence of single nucleotide polymorphisms.* Journal of biotechnology, **2015**. 211: p. 20-30, https://doi.org/10.1016/j.jbiotec.2015.06.423
77. Krahulec, J., Lišková, V., Boňková, H., Lichvariková, A., Šafranek, M.,Turňa, J., *The ploidy determination of the biotechnologically important yeast Candida utilis.* Journal of Applied Genetics, **2020**: p. 1-12

78. Yue, J., Li, J., Aigrain, L., Hallin, J., Persson, K., Oliver, K., Bergström, A., Coupland, P., Warringer, J.,Lagomarsino, M., *Contrasting evolutionary genome dynamics between domesticated and wild yeasts.* Nature genetics, **2017**. 49(6): p. 913-924
79. Eldarov, M., Beletsky, A., Tanashchuk, T., Kishkovskaya, S., Ravin, N.,Mardanov, A., *Whole-genome analysis of three yeast strains used for production of sherry-like wines revealed genetic traits specific to flor yeasts.* Frontiers in microbiology, **2018**. 9: p. 965
80. Peinado, J., Cameira-Dos-Santos, P.,Loureiro-Dias, M., *Regulation of glucose transport in Candida utilis.* Microbiology, **1989**. 135(1): p. 195-201
81. Boze, H., Moulin, G.,Galzy, P., *Production of food and fodder yeasts.* Critical reviews in biotechnology, **1992**. 12(1-2): p. 65-86, https://doi.org/10.3109/07388559209069188
82. Flores, C., Rodríguez, C., Petit, T.,Gancedo, C., *Carbohydrate and energy-yielding metabolism in non-conventional yeasts.* FEMS microbiology reviews, **2000**. 24(4): p. 507-529

CHAPTER V

Combined *in silico* approaches to uncover plasma membrane carboxylate transporters in *Cyberlindnera jadinii*

Adapted from:

Sousa-Silva, M.[1,2], Vieira, D.[1,2], Alves, J.[1,2], Soares, P.[1,2], Casal, M.[1,2], Soares-Silva, I.[1,2] (2021). Combined in silico approaches to uncover plasma membrane carboxylate transporters in *Cyberlindnera jadinii.*

[1] Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, Portugal
[2] Centre of Molecular and Environmental Biology (CBMA), Department of Biology, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal

The work presented in this chapter is in preparation for publication.

CHAPTER V

Combined *in silico* approaches to uncover plasma membrane carboxylate transporters in *Cyberlindnera jadinii*

ABSTRACT

The yeast *Cyberlindnera jadinii* has great potential in the biotechnology industry due to its ability to produce a variety of compounds of interest, such as carboxylic acids. In this work, we aimed at identifying the genes encoding carboxylate transporters of this yeast species to explore their efficient use to export organic acids by microbial cell factories. Putative plasma membrane carboxylate transporters were identified with a bioinformatics pipeline analyzing the inferred proteome of the *C. jadinii* NRRL Y-1542 strain. The functional characterization of the relevant hits by heterologous expression in *Saccharomyces cerevisiae* unveiled the existence of sixteen plasma membrane carboxylate transporters belonging to the AceTr, SHS, TDT, MCT, SSS, and DASS transporter families. The newly identified *C. jadinii* transporters presented a diverse range of specificity, being able to transport mono-, di- and tricarboxylates alone or in combination: sixteen transport monocarboxylates, five transport dicarboxylates (CjAto2p, CjAto5p, CjJen6p, CjSlc5p, and CjSlc13-1p) and four transport citrate (CjAto5p, CjJen6p, CjSlc5p, and CjSlc13-1p). In this study, a detailed characterization of these transporters was performed which also included their phylogenetic reconstruction, 3D-structure prediction, and molecular docking studies.

KEYWORDS: *Cyberlindnera jadinii*; Carboxylic acids; Plasma membrane transporters; Phylogeny; 3D structure model; Molecular docking

5.1. INTRODUCTION

The yeast *Cyberlindnera jadinii* is an attractive platform for industrial utilization due to its intrinsic robust fermentation characteristics and its ability to utilize a wide range of substrates, including hexoses, pentoses, organic acids (including mono-, di- and tricarboxylic acids) and amino acids [1-4]. Membrane transporters are powerful tools to improve yeast cell factories aimed at the production of valuable chemical compounds [5]. With few exceptions, neither substrates nor products of cell metabolism can cross the plasma membrane freely. Instead, these molecules require the activity of membrane transporter proteins to cross the cell membrane outwards or inwards.

The productivity of carboxylic acids by microbial cell factories is significantly affected by the influx of substrates and efflux of products and by-products. The expression of suitable exporters avoids the internal cell´s toxicity of carboxylic acids and facilitates their purification from the culture broth [6-10]. A unique transporter can recognize distinct substrates, and multiple transporters may be associated with the transport of a particular substrate across cell membranes [11, 12]. The recent publication of Soares-Silva *et al.* (2020) reviews the most relevant cell membrane transporters (importers and exporters) used for the improvement of microbial cell factories' production of carboxylic acids [5]. The Sialate:H^+ symporter (SHS) family (TC 2.A.1.12) is a well-characterized family able to transport carboxylates and sugar acids [5] to which belongs the Jen1p of *S. cerevisiae*, the first the monocarboxylate:H^+ symporter transporter found in fungi [13]. Later, the same group of researchers described the first member of the acetate uptake transporter (AceTr) family (TC 2.A.96) in *S. cerevisiae*, the Ady2/Ato1 [14], which includes both importers and exporters of carboxylic acids [15]. Furthermore, the solute carrier (SLC) superfamily, distributed throughout all domains of Life, includes well-known transporter families, like the Monocarboxylate Transporter (MCT)/SLC16 Family (TC 2.A.1.13) found in *Homo sapiens* or the Divalent Anion:Na^+ Symporter (DASS)/SLC13 Family (TC 2.A.47) which can either import and export mono-, di- and tricarboxylates [16-18].

In the 80s, distinct plasma membrane transporter systems for organic acids were physiologically characterized in *C. jadinii* strains [19-21]: (i) a proton-symporter accepting the monocarboxylates L-lactate and D-lactate, pyruvate, propionate, and acetate [20]; (ii) a dicarboxylate proton-symporter for L-malate, succinate, fumarate, oxaloacetate, and α-ketoglutarate [21, 22]; (iii) a tricarboxylate proton-symporter for citrate and isocitrate [19] and (iv) a general organic permease for the facilitated diffusion of amino acids, and mono-, di- and tricarboxylic acids with low affinity for the substrate [19, 23]. The genes encoding these transporters remain until now unidentified. Hence, in this work, we searched for *S. cerevisiae* Ato1

and Jen1 homologs present in the *C. jadinii* genome by sequence homology. In parallel, we have developed a bioinformatic pipeline to identify and annotate genes coding for carboxylate transporters. The most promising hits were expressed in *S. cerevisiae* to assess their physiological function. For those revealing carboxylic acid transport activity, a phylogenetic and an *in silico* structural 3D analysis was performed. Together these data bring novel insights on the *C. jadinii* inferred proteome, shedding light on the processes of solute plasma membrane transport in this yeast species.

5.2 Materials and Methods

5.2.1 Yeast cultivations and plasmid constructs

Yeast strains and plasmids used in this study are listed in tables 1 and 2. The *S. cerevisiae* W303-1A *jen1*Δ *ady2*Δ (Soares-Silva et al., 2007) and *S. cerevisiae* IMX1000 strains (Mans et al., 2017), lacking carboxylate uptake under the conditions tested, were used to express putative carboxylate transporters [24, 25]. Cultures were maintained on YPD medium, yeast extract (1%, w/v), peptone (1%, w/v), glucose (2%, w/v) and agar (2%, w/v), or minimal media with the required supplements for growth of strains with auxotrophies. Yeast cells were grown in yeast nitrogen base (YNB, Difco), 0.67%, w/v (YNB medium), enriched with adequate requirements for prototrophic growth. Carbon sources used were glucose (2%, w/v), acetic acid (0.5%, v/v, pH 6.0), lactic acid (0.5%, v/v, pH 5.0), pyruvic acid (0.5% v/v, pH 5.0), fumaric acid (1%, w/v, pH 5.0), succinic acid (1%, w/v, pH 5.0), malic acid (1%, w/v, pH 5.0) and citric acid (1%, w/v, pH 5.5). Cell growth was carried out at 30°C, both in liquid and solid media. For growth phenotypes, cells were grown on YNB Glu–Ura media until mid-exponential phase and adjusted to an OD_{640nm} of 0.1. A set of three 1:10 serial dilutions were performed and 3 μL of each suspension was spotted in the desired media, using YNB Glu–Ura as a control. Cells were incubated at 18°C for 22 days. At 18°C, carboxylic acid uptake by diffusion is drastically reduced so that growth on CA as sole carbon and energy source is directly dependent on a functional transporter [24]. For the transport assays, YNB containing (YNB Glu-Ura) media was used for growth of yeast cells under repression conditions. Cultures were harvested during the exponential growth phase. For derepression conditions glucose-grown cells were centrifuged, washed twice in ice-cold deionized water and cultivated into fresh YNB media supplemented with the suitable carbon source.

5.2.2 TRANSPORT ASSAYS

Transport assays were performed as previously described by Ribas *et al.* (2017) [26]. Cells were harvested by centrifugation (5000 rpm, 2 minutes), washed twice in ice-cold deionized water and resuspended in ice-cold deionized water to a final concentration of about 25–35 mg cell dry weight/mL. The reaction mixtures were prepared in 1.5 mL microtubes tubes containing 60 μL of KH_2PO_4 (0.1 M, pH 5.0), and 30 μL of the yeast cell suspension. After 2 minutes of incubation at 30°C, the reaction was started by the addition of 10 μL of a solution of radiolabelled substrate, at the desired pH and concentration, rapidly mixed by vortex and incubated at 30°C. After 15-30s the reaction was stopped by adding 100 μL of non-labelled substrate tested (in a 100-fold the substrate concentration in the reaction mixture), quickly mixed by vortex and chilled on ice. The suspension was centrifuged for 7 minutes at 13200 rpm. The supernatant was carefully rejected, the pellet was resuspended in 1 mL of desionized cold water and centrifuged for 10 minutes at 13200 rpm. The resulting pellet was resuspended in 1 mL of scintillation liquid (Opti-Phase HiSafe II; LKB FSA Laboratory Supplies). The labelled carboxylates used were [1-^{14}C] acetic acid (Perkin Elmer, Massachusetts, USA); L-[U-^{14}C] lactic acid (Perkin Helmer, Massachusetts, USA); [2,3-^{14}C] succinic acid (Moravek Biochemicals, California, USA) and [1,5-^{14}C] citric acid (Perkin Elmer, Massachusetts, USA). Working solutions presented a specific activity ranging from 300 to 4000 dpm, depending on the final carboxylate concentrations. The best-fitting for initial uptake rates was determined through computer-assisted non-linear regression analysis, performed by GraphPad Prism (California, USA) version 4.0 for Windows. The values for kinetic parameters were obtained with a significance level $p<0.05$. The data shown are mean values of at least three independent experiments, with three replicas each.

Table 1. Yeast strains used in this study.

Strain	Genotype	Source
***Cyberlindnera jadinii* DSM 2361**	Type strain DSM 2361	DSM collection
***Saccharomyces cerevisiae* W303-1A**	MATα *ade2 leu2 his3 trp1 ura3*	Thomas and Rothstein (1989)
S. cerevisiae jen1Δ ato1Δ	W303-1A; *JEN1::KanMX4 ATO1::HphMX4*	Soares-Silva *et al.* (2007)
***S. cerevisiae* CEN.PK113-7D**	MATα *URA3 TRP1 LEU2 HIS3*	Entian and Kötter (2007)
IMX1000 (parental strain CEN.PK113-7D)	*MAT*a *ura3-52 trp1-289 leu2-3112 his3Δ can1Δ::cas9-natNT2 mch1Δ mch2Δ mch5Δ aqy1Δ itr1Δ pdr12Δ mch3Δ mch4Δ yil166cΔ hxt1Δ jen1Δ ato1Δ aqr1Δ thi73Δ fps1Δ aqy2Δ yll053cΔ ato2Δ ato3Δ aqy3Δ tpo2Δ yro2Δ azr1Δ yhl008cΔ tpo3Δ*	Mans *et al.* (2017)
***S. cerevisiae jen1Δ ato1Δ* p416GPD**	*jen1*Δ *ato1*Δ transformed with p416GPD	Soares-Silva *et al.* (2007)
***S. cerevisiae jen1Δ ato1Δ* pCaJen2**	*jen1*Δ *ato1*Δ transformed with pCaJen2	Vieira *et al.* (2010)
***S. cerevisiae jen1Δ ato1Δ* pScAto1**	*jen1*Δ *ato1*Δ transformed with pScAto1	This work
***S. cerevisiae jen1Δ ato1Δ* pCjAto1**	*jen1*Δ *ato1*Δ transformed with pCjAto1	This work
***S. cerevisiae jen1Δ ato1Δ* pCjAto2**	*jen1*Δ *ato1*Δ transformed with pCjAto2	This work
***S. cerevisiae jen1Δ ato1Δ* pCjAto3**	*jen1*Δ *ato1*Δ transformed with pCjAto3	This work
***S. cerevisiae jen1Δ ato1Δ* pCjAto4**	*jen1*Δ *ato1*Δ transformed with pCjAto4	This work
***S. cerevisiae jen1Δ ato1Δ* pCjJen1**	*jen1*Δ *ato1*Δ transformed with pCjJen1	This work
***S. cerevisiae jen1Δ ato1Δ* pCjJen2**	*jen1*Δ *ato1*Δ transformed with pCjJen2	This work
***S. cerevisiae jen1Δ ato1Δ* pCjJen3**	*jen1*Δ *ato1*Δ transformed with pCjJen3	This work
***S. cerevisiae jen1Δ ato1Δ* pCjJen4**	*jen1*Δ *ato1*Δ transformed with pCjJen4	This work
***S. cerevisiae jen1Δ ato1Δ* pCjJen5**	*jen1*Δ *ato1*Δ transformed with pCjJen5	This work
***S. cerevisiae jen1Δ ato1Δ* pCjJen6**	*jen1*Δ *ato1*Δ transformed with pCjJen6	This work
IMX1000 pCjAto1	IMX1000 transformed with pCjAto1	This work
IMX1000 pCjAto2	IMX1000 transformed with pCjAto2	This work
IMX1000 pCjAto3	IMX1000 transformed with pCjAto3	This work
IMX1000 pCjAto4	IMX1000 transformed with pCjAto4	This work
IMX1000 pCjJen1	IMX1000 transformed with pCjJen1	This work
IMX1000 pCjJen2	IMX1000 transformed with pCjJen2	This work
IMX1000 pCjJen3	IMX1000 transformed with pCjJen3	This work
IMX1000 pCjJen4	IMX1000 transformed with pCjJen4	This work
IMX1000 pCjJen5	IMX1000 transformed with pCjJen5	This work
IMX1000 pCjJen6	IMX1000 transformed with pCjJen6	This work
IMX1000 pCjMch4	IMX1000 transformed with pCjMch4	This work
IMX1000 pCjSlc5	IMX1000 transformed with pCjSlc5	This work
IMX1000 pCjTDT	IMX1000 transformed with pCjTDT	This work
IMX1000 pCjSlc13-1	IMX1000 transformed with pCjSlc13-1	This work
IMX1000 pCjSlc13-2	IMX1000 transformed with pCjSlc13-2	This work
IMX1000 pCjAto5	IMX1000 transformed with pCjAto5	This work
***S. cerevisiae jen1Δ ato1Δ* pScJen1-GFP**	*jen1*Δ *ato1*Δ transformed with pScJen1-GFP	Soares-Silva *et al.* (2003)

Table 2. Plasmids used in this study.

Plasmid	Description	Reference
p416GPD	Glyceraldehyde-3-phosphate dehydrogenase (GPD) promoter	Mumberg, Muller and Funk (1995)
pJen1-GFP	*ScJen1* cloned in p416 under the control of GPD promoter with the GFP gene	Soares-Silva *et al.* (2003)
pCaJen2	*CaJen2* cloned in p416 under the control of GPD promoter	Vieira *et al.* (2010)
pScAto1	*ScAto1* cloned in p416 under the control of GPD promoter	This work
pCjAto1	*CEP24587* cloned in p416 under the control of GPD promoter	This work
pCjAto2	*CEP20823* cloned in p416 under the control of GPD promoter	This work
pCjAto3	*CEP20822* cloned in p416 under the control of GPD promoter	This work
pCjAto4	*CEP20690* cloned in p416 under the control of GPD promoter	This work
pCjJen1	*CEP23088.1* cloned in p416 under the control of GPD promoter	This work
pCjJen2	*CEP21966.1* cloned in p416 under the control of GPD promoter	This work
pCjJen3	*CEP22358.1* cloned in p416 under the control of GPD promoter	This work
pCjJen4	*CEP21989.1* cloned in p416 under the control of GPD promoter	This work
pCjJen5	*CEP21602.1* cloned in p416 under the control of GPD promoter	This work
pCjJen6	*CEP25129.1* cloned in p416 under the control of GPD promoter	This work
pCjSlc16	XP_020067635.1 cloned in p416 under the control of GPD promoter	This work
pCjSlc5	XP_020068154.1 cloned in p416 under the control of GPD promoter	This work
pCjTDT	XP_020068891.1 cloned in p416 under the control of GPD promoter	This work
pCjSlc13-1	XP_020069270.1 cloned in p416 under the control of GPD promoter	This work
pCjSlc13-2	XP_020073044.1 cloned in p416 under the control of GPD promoter	This work
pCjAto5	XP_020067765.1 cloned in p416 under the control of GPD promoter	This work

5.2.3 Pipeline for the identification of carboxylate transporter genes

A computational analysis of the *Cyberlindnera jadinii* transportome was carried out to identify novel carboxylate transporters. The proteome from *C. jadinii* NRRL Y-1542 was downloaded from the NCBI database and analyzed using a bioinformatic pipeline (figure 2). This tool was designed to retrieve data from a specific database that contains: a) a single representative genome/proteome on the species level; b) where multiple matches within a species directly reflects the presence of orthologues, and c) the e-values from BLAST searches are statistically more reliable. To avoid redundancy of results we build our in-house genetic information database by using the inferred proteome of 13099 species with

full-annotated genomes in the NCBI assembly database (RefSeq). In situations where multiple reference strains were present for the same species, only the one presenting the higher number of proteins uncovered was used. The final database contained a total of 2863 genomes with 11969222 protein sequences in which: 30 are from metazoa, 193 from archaea, 2489 from bacteria, 72 from fungi, 34 from plants and 45 from protists. The proteomes were searched for promising putative carboxylate transporters using the following criteria: a) proteins displaying a topology of 4-20 transmembrane segments (TMS) predicted by the TMHMM server v2.0 [78]; b) removal of the partial proteins; c) search for conserved motifs involved in the transport of carboxylate and d) identification of sequences with homology to known di- and tricarboxylate transporters.

5.2.4 CLONING AND EXPRESSION OF HETEROLOGOUS GENES

The target genes selected in 2.3 were amplified by PCR using the genomic DNA from the yeast *Cyberlindnera jadinii* DSM 2361 [27] and cloned in the centromeric plasmid p416GPD [28] under the control of a GPD constitutive promoter (pCj-gene plasmids) [24, 25]. As expression hosts the strains *S. cerevisiae* W303-1A *jen1Δ ato1Δ* and *S. cerevisiae* IMX1000 were used. The primers used for amplification and cloning of pCj-gene are listed in table 3. Sixteen genes were cloned in the centromeric plasmid p416GPD, by PCR amplification with proofreading polymerase ACCUZYME Mix DNA (Bioline, London, UK) [28]. DNA cloning and manipulation were performed according to standard protocols [29]. Primers from *ThermoFisher Scientific* (USA) contain restriction sites for *BamH*I and *EcoR*I (CjAto1-4, CjJen1, CjJen3 and CjJen5); *EcoR*I and *Hind*III (CjJen2 and CjJen4); *Xba*I and *EcoR*I (CjJen6); *Spe*I and *Sal*I (CjSlc16, CjSlc5, CjTDT, CjSlc13-1 genes); *Spe*I and *Xho*I (Slc13-2 gene) *Spe*I and *EcoR*I (CjAto5), in the Forward and Reverse primers respectively. The final products were inserted in the vector p416GPD, previously digested with the same restriction enzymes. The resulting pCjgene plasmids were transformed in *S. cerevisiae* strains.

Table 3. Oligonucleotides used for strain construction, cloning and expression

Name	Sequence
p416-CJAd1_fwd	GCAGGATCCATGTCAGACAAGGAAAGC
p416-CJAd1_rev	GCAGAATTCCTAGGAGTGCACTTGAGC
p416-CJAd2_fwd	GCAGGATCCATGTCTAGTATTAATGAG
p416-CJAd2_rev	GCAGAATTCTCACAAATGGCTAGCACC
p416-CJAd3_fwd	GCAGGATCCATGGCCGCCAACGTTGAC
p416-CJAd3_rev	GCAGAATTCTCAGGCACGCTTTGCACC
p416-CJAd4_fwd	GCAGGATCCATGTCGGACAAGGAAAAC
p416-CJAd4_rev	GCAGAATTCTCAAGAACGCTTTGCACC
Cjad2-I75L_fwd	GCCAATCCAGTCCCTCTGGGTCTCTGTGGGTTC
Cjad2-I75L_rev	GAACCCACAGAGACCCAGAGGGACTGGATTGGC
p416-CJJen1_fwd	GCA GGATCCATGCACAAACTTGAAGAG
p416-CJJen1_rev	GCAGAATTCTCACTTCTTCTCCTGTGG
p416-CJJen2-new_Fwd	GCAGGATCCATGACTTCACCATTGCCT
p416-CJJen2-new_Rev	GCACTCGAGTCACTCACTTGAAGAGCC
p416-CJJen3_fwd	GCAGAATTCATGGCAATGTCTGATGTT
p416-CJJen3_rev	GCAGAATTCTCAGGTTTTTTCATTATG
p416-CJJen4-new_Fwd	GCAGGATCC ATGACTGCTGGGAGATAC
p416-CJJen4-new_Rev	GCACTCGAGCTAGTCTCTAGCAGATTC
p416-CJJen5_fwd	GCAGGATCCATGGACTGGGATGCTTTC
p416-CJJen5_rev	GCAGAATTCTCACTTTGGCTCTATCTT
Cut-Jen6_Fwd	GCCTCTAGAATGGGATTCAAGACGTAC
Cut-Jen6_Rev	GCCGAATTCTCATTTAACCTCAGAAAC
ct1.635_Fwd	GCCACTAGTATGACTGAAATCATCACT
ct1.635_Rev	GCCGTCGACTCAGAATTTACAAATTCT
ct2.154_Fwd	GCCACTAGTATGTTTGCAGAGACCGAG
ct2.154_Rev	GCCGTCGACTCACGAGTTGTCAGCACA
ct3.891_Fwd	GCCACTAGTATGACATCTGATGAGAAT
ct3.891_Rev	CCGTCGACTCAATTCCTTTCACTGTT
ct4.270_Fwd	GCCACTAGTATGAAATTCTCCCTCTCT
ct4.270_Rev	GCCGTCGACTTAACCGTGTAAAGTTGC
ct8.044_fwd	GCCACTAGTATGAAGTTCTCCCATTCG
ct8.044_rev	GCCCTCGAGTCACATCCCTGTTAATCT
ct9.7765_fwd	GCCACTAGTATGTCGACCTCTTCTCTC
ct9.7765_rev	GCCGAATTCCTAAACTCTTGGTGCATG

5.2.5 Phylogenetic reconstructions

A total of over 10000 inferred proteomes from NCBI Assembly platform were downloaded as individual FASTA files, all belonging to the refseq subsection of NCBI and they were converted into a local database. To avoid redundancies, only sequences from a single genome of a given species were considered. In these cases, the specimen with the higher number of proteins described in the database was the selected. A BLAST search, with a cut-off e-value 10^{-10} and an associated query-cover value higher than 65%, was performed on our in-house database using fourteen queries from *C. jadinii.* We analyzed the six Ato1 homologs (CjAto1 XP_020073178.1, CjAto2 XP_020073031.1, CjAto3 XP_020073179.1, CjAto4 XP_020070445.1, CjAto5 XP_020067765.1 and CjAto6 XP_020069005.1) but the results from the six queries were merged into a single file by removing redundant proteins across results. The same approach was taken for the six queries of Jen1 homologs in *C. jadinii* (CjJen1 CEP23088.1, CjJen2 CEP21966.1, CjJen3 CEP22358.1, CjJen4 CEP21989.1, CjJen5 CEP21602.1 and CjJen6 CEP25129.1). The sequences of CjSlc5 (XP_020068154.1) and CjSlc13-1 (XP_020069270.1) were individually used as queries for search of homologs of the respective proteins. Given the fact that the general structure of the phylogeny of ATO1 was already detailed recently [30] in this case the analysis only included Ascomycota. Retrieved protein sequences were aligned using the MAFFT online server [31], that incorporates multiple alignment strategies. Sequences that were not matching extensively across the conserved region of the alignment were further excluded from the phylogenetic analysis. These sequences, in many cases, could represent lower quality of the stretch of the genome where they are located or incomplete annotation of the full gene and they should not necessarily be regarded as non-functional genes.

A phylogenetic reconstruction was performed using Maximum Likelihood, more appropriate for the deeper divergences under analysis here, using MEGA7 [32] and the Jones-Taylor-Thornton (JTT) substitution model. Bootstrap was performed for 1000 repetitions. Obtained phylogenetic trees were displayed and edited in FigTree v.1.4.4. (http://tree.bio.ed.ac.uk/).

5.2.6 Sequence alignment and Percentage of identity shared between the Ato and Jen homologs

For sequence alignment, the sequence data for *C. jadinii* genome was obtained from the NCBI database. Using the BLASTp program, six ORFs were identified revealing homology to the *ATO1* of *S. cerevisiae*, XP_020073178.1 (CjAto1), XP_020073031.1 (CjAto2), XP_020073179.1 (CjAto3), XP_020070445.1 (CjAto4), XP_020067765.1 (CjAto5) and XP_020069005.1 (CjAto6), and six ORFs revealing homology to the *JEN1* of *S. cerevisiae*, *CEP23088p* (CjJen1), *CEP21966p* (CjJen2), *CEP22358p* (CjJen3),

CEP21989p (CjJen4), *CEP21602p* (CjJen5) and *CEP25129p* (CjJen6). Furthermore, the query sequences used for the members of the SLC transporter superfamily over the simple alignments were NP_666018.3 (*Homo sapiens* SLC5 member 8, HsSlc5A8), XP_020068154.1 (*C. jadinii* SLC5 homolog, CjSlc5), NP_001011554.1 (*H. sapiens* SLC13 member 3, HsSlc13A3) and XP_020069270.1 (*C. jadinii* SLC13 homolog, CjSlc13). Retrieved protein sequences, including homologs present in other microorganisms, were aligned with EMBOSS needle and ClustalW through pairwise sequence alignment. The TMHMM server (http://www.cbs.dtu.dk/services/TMHMM-2.0/) was used for topology prediction of TMS. Multiple sequence alignment was also performed with M-Coffee (http://tcoffee.crg.cat/apps/tcoffee/do:mcoffee) for a further validation of identified TMS regions across the protein sequences [33]. The percentage of identity shared by protein sequences was calculated using Matcher (EMBOSS) tool (https://www.ebi.ac.uk/Tools/psa/emboss_matcher/) (Table S1). The Ato homologs selected belong to *C. jadinii, Yarrowia lipolytica, S. cerevisiae, Aspergillus nidulans, Kluyveromyces lactis, Debaryomyces hansenii, Escherichia coli* and *Methanosarcina acetivorans*. The Jen homologs selected belong to *C. jadinii, S. cerevisiae, K. lactis, Candida albicans, D. hansenii* and *Y. lipolytica* species. The reference sequences used were the following: ScAto1 NP_009936.1, ScAto2 NP_014399.3, ScAto3 NP_010672.1, Ca XP_710295.1, CaFrp5p XP_716747.2, CaFrp6p XP_716748.1, CaAto1 XP_710650.1, CaAto2 XP_718515.2, CaAto5 XP_714703.1, CaAto6 XP_714701.1, CaAto7 XP_019330752.1, CaAto9 XP_717951.1, CaAto10 XP_717953.1, YlGpr1 XP_502188.1, Yl1 XP_504461.1, Yl2 XP_505489.1, Yl3 XP_505359.1, Yl4 XP_503877.1, Yl5 XP_502175.1, AnAcpA XP_662830.1, AnAcpB XP_659443.1, AnAcpC XP_680586.1, AnAlcS XP_682250.1, An XP_681659.1, An XP_657970.1, EcSatP XP_000528538.1, Ma4008 WP_011023903.1, Ma0103 AAM03557.1, Ma WP_048066054.1, Ma WP_011024272.1, CjAto1 XP_020073178.1, CjAto2 XP_020073031.1, CjAto3 XP_020073179.1, CjAto4 XP_020070445.1, CjAto5 XP_020067765.1 and CjAto6 XP_020069005.1 (table S1-A). For Jen homologs, ScJen1 NP_012705.1, KlJen1 XP_454682.1, KlJen2 XP_455537.1, CaJen1 XP_716108.1, CaJen2 XP_717031.1, Dh17 XP_461118.2, Dh18 XP_459308.2, Dh24 XP_002770539.1, Dh27 XP_461518.2, YlJen1 XP_503239.1, YlJen2 XP_503058.1, YlJen3 XP_504706.1, YlJen4 XP_501871.1, YlJen5 XP_502090.1, YlJen6 XP_501098.1, CjJen1 CEP23088.1, CjJen2 CEP21966, CjJen3 CEP22358.1, CjJen4 CEP21989.1, CjJen5 CEP21602 and CjJen6 CEP25129.1 (table S1-B).

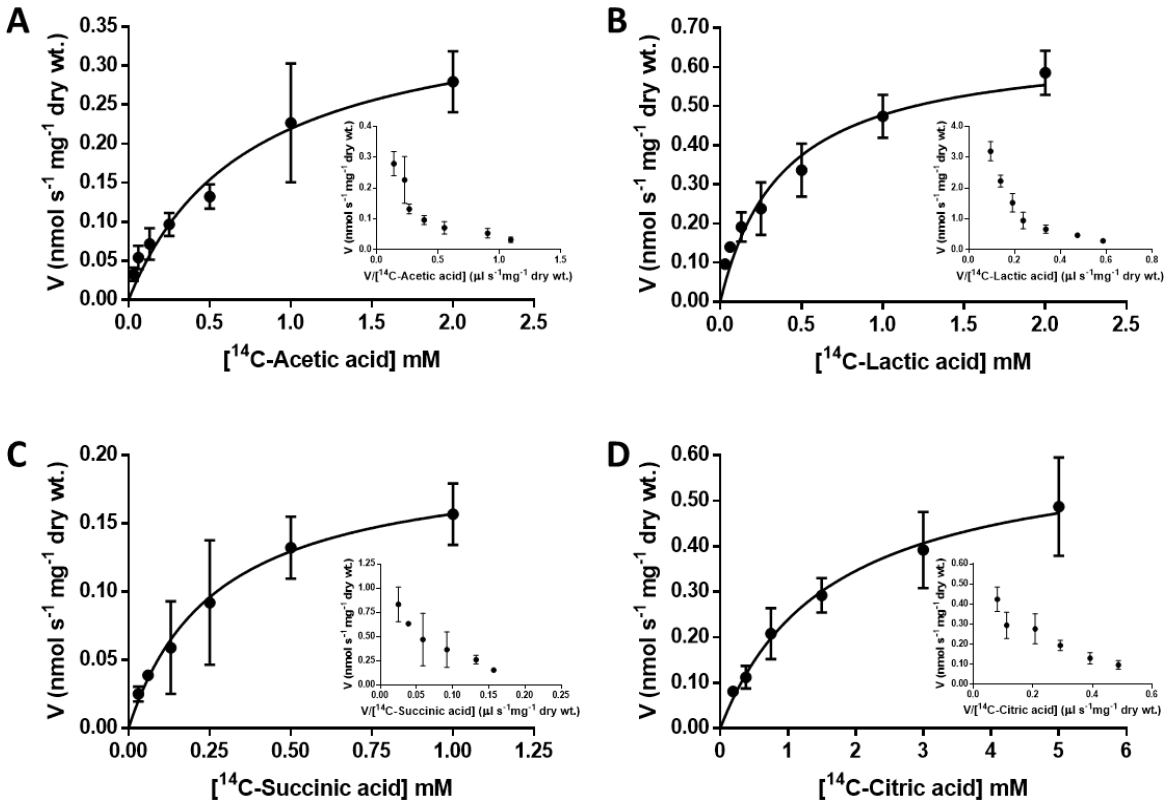
5.2.7 Three-dimensional modelling, Molecular docking studies and pore radius simulations

The predicted transporter 3D structures, were obtained with the HHPred [34] and LOMETS (Local Meta-Threading-Server) [35] softwares. Molecular docking simulations were performed as described by Ribas *et al.* (2017) [26]. Ligand structures of acetic, lactic, succinic and citric acids for all target proteins in the study were downloaded from Zinc database [36]. Only deprotonated forms of each acid were used in the docking prediction, with the protonation states adjusted to match a pH of 5.0-6.0. The substrate 3D structures were built by inputting canonical SMILES (Simplified Molecular Input Line Entry Specification) strings in the UCSF Chimera [37], being minimized before molecular docking in PyRx software [38] using AutoDock Vina. The simulated interactions were analyzed in 2D and 3D pose views using both Chimera and the Maestro v11.2. The HOLE program (2.2.005 Linux) was used to predict the pore radius throughout each transporter studied [39]. The radiuses of the homolog transporters were compared in a graph with the coordinate in the direction of the channel vector serving as the X-axis. The images of the predicted pores were obtained using Visual Molecular Dynamics program (VMD, 1.9.3) [40].

5.3 Results

5.3.1 Characterization of carboxylic acid transport systems in *Cyberlindnera jadinii*

The yeast *Cyberlindenera jadinii* DSM 2361 when grown in acetic (0,5%, v/v; at pH 6.0), lactic (0,5%, v/v; at pH 5.0), succinic (1.0%, w/v; at pH 5.0) and citric (1.0%, w/v; at pH 5.5) acids displays ability to transport these substrates through a Michaelis-Menten kinetics (Figure 1). The Eadie-Hofstee plot of the initial uptake rates of labelled carboxylates revealed complex transporter systems. A non-linear regression analysis of the initial uptake rates allowed to determine the kinetic parameters of these transport systems that are summarized in fig.1.



Acid	K_m (mM)	V_{max} (nmol carboxylate s^{-1} mg^{-1} dry wt.)
Acetate	0.73 ± 0.21	0.38 ± 0.05
Lactate	0.014 ± 0.03	0.12 ± 0.07
	1.16 ± 0.63	0.74 ± 0.10
Succinate	0.27 ± 0.08	0.20 ± 0.02
Citrate	1.62 ± 0.36	0.63 ± 0.06

Figure 1. Initial uptake rates of labelled carboxylic acids as a function of the acid concentration of *C. jadinii* DSM 2361 cells in the presence of (A)^{14}C-acetic acid, pH 6.0; (B)^{14}C-lactic acid, pH 5.0; (C)^{14}C-succinic acid, pH 5.0 and (D)^{14}C-citric acid, pH 5.5, at 30°C.

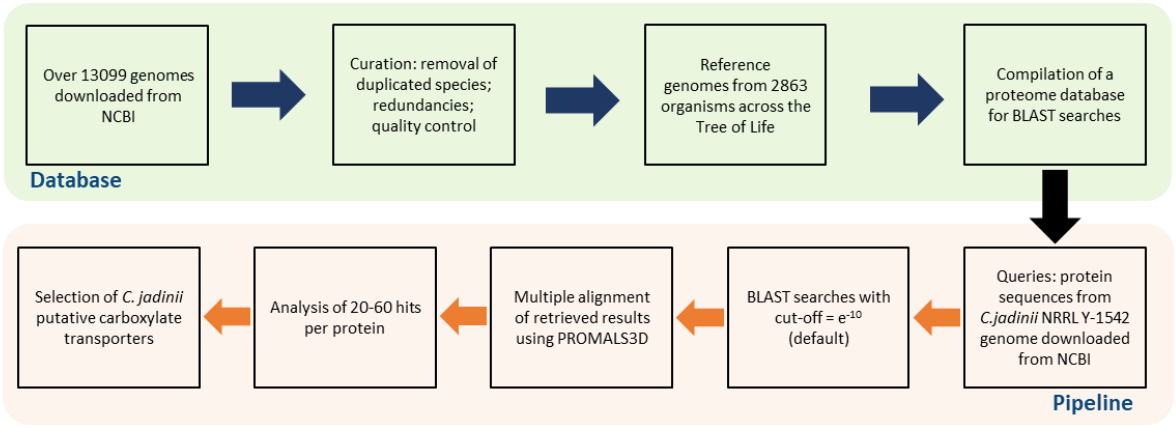
Mid-exponential grown-cells on YNB glucose were collected, washed and cultivated on YNB containing as carbon source acetic acid (A), lactic acid (B), succinic acid (C) or citric acid (D) as described in materials and methods. The inserts represent the Eadie-Hofstee plots of the data presented in the respective main chart. The data shown are mean values of at least three independent experiments and the error bars represent the standard deviation. The bottom-table presents the best fitting transport kinetic values for carboxylate transport systems in *C. jadinii* DSM 2361 cells, obtained using a computer-assisted non-linear regression analysis (GraphPaD Prism software).

5.3.2 IDENTIFICATION OF THE *CYBERLINDNERA JADINII* INFERRED TRANSPORTOME

Currently there are three *C. jadinii* genomes available at the NCBI. Two of them are from the same strain, the CBS 1600 and the NRRL Y-1542 genome, and the other from the NBRC0988 (DSM 2761) strain. The information for this latter is not available for BLASTp searches. In an initial approach the CBS 1600/NRRL Y-1542 genomes were used to search for homologs of Ato1p and Jen1p transporters from *S. cerevisiae*. Four ScAto1p homologs (CjAto1-4) and six ScJen1p homologs (CjJen1-6) were identified having a query cover value higher than 67%, an identity value between 40-58% and a similarity between 58-73%. CjAto3 was the most similar to ScAto1 sharing 61.5% of identity (table S1-A from the supplementary data). The CjAto1 and CjAto3 are more similar (50% identity) with the AcpA acetate transporter from *A. nidulans* while the CjAto2 is more identical to the acetate transporter Gpr1p from *Y. lipolytica* [30]. CjJen1 shares 58.2% identity with the ScJen1p from *S. cerevisiae* (table S1-B). CjJen1 and CjJen2 have a significant 52-57% identity with Jen1 from *K. lactis* (KlJen1) [41]. On the other hand, CjJen5-6 proteins present a higher identity (47-59%) with *C. albicans* Jen2p, KlJen2p, *D. hansenii* DH18p and DH24p, and Jen1-6 from *Y. lipolytica*.

Aiming at identifying other putative plasma membrane transporters encoded in the *C. jadinii* predicted proteome, we designed a new strategy by combining distinct bioinformatic tools (see materials and methods). Figure 2A summarizes the pipeline used to infer the proteome of the *C. jadinii* NRRL Y-1542. The outcome of this search was a list of 6147 protein-coding sequences in which 6032 are system-annotated proteins. From these, 642 display one to three TMSs and, 525 have four to twenty TMSs (Figure 2B). In this last group, several annotated proteins are defined as MFS general substrate transporters, having the great majority system-annotated protein functions. Nonetheless, several are still annotated as hypothetical proteins. Proteins with four to twenty TMSs were screened against the NCBI database in order to identify functionally characterized homologs (Figure 2D; detailed information is presented in table S2). These 525 proteins were distributed into five major categories: (i) transporters, (ii) enzymes, (iii) receptors, (iv) other membrane proteins and (v) unknown proteins. The transporters were annotated as: sugar, amino acid, oligopeptides, ABC-type, ATPases (V-type, A-type and P-type), carboxylic acid transporters, among others. Transporters associated with carboxylic acid (12,3%), polyatomic anions (14,5%), amino acid and oligopeptides (24%), belonging to the Solute carrier (SLC) superfamily (SLC5, SLC13-1, SLC13-2, SLC16), to the Tellurite-resistance/Dicarboxylate transporter (TDT) family (CjTDT) and one member to the Acetate transporter (AceTr) family (Cj-Ato5) were the most promising outcomes of this analysis.

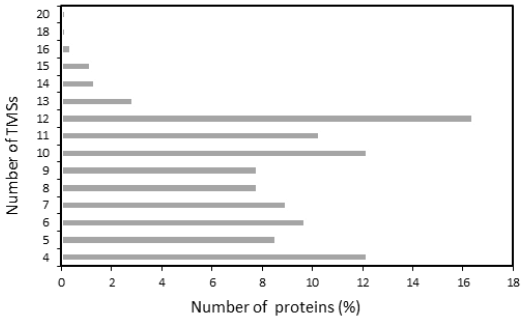
A



B

C. jadinii NRRL Y-1542 genome	Nº hits	Reference
Protein-coding sequences	6147	Rupp *et al.*, 2015
Annotated proteins in genome	6032	NCBI manually curated system
Proteins with 1-3 TMSs	642	This study
Proteins with 4-20 TMSs	562	This study
System-annotated protein functions	257	This study
Hypothetical proteins	203	This study
MFS general substrate transporters	65	This study
Partial proteins	37	This study

C



D

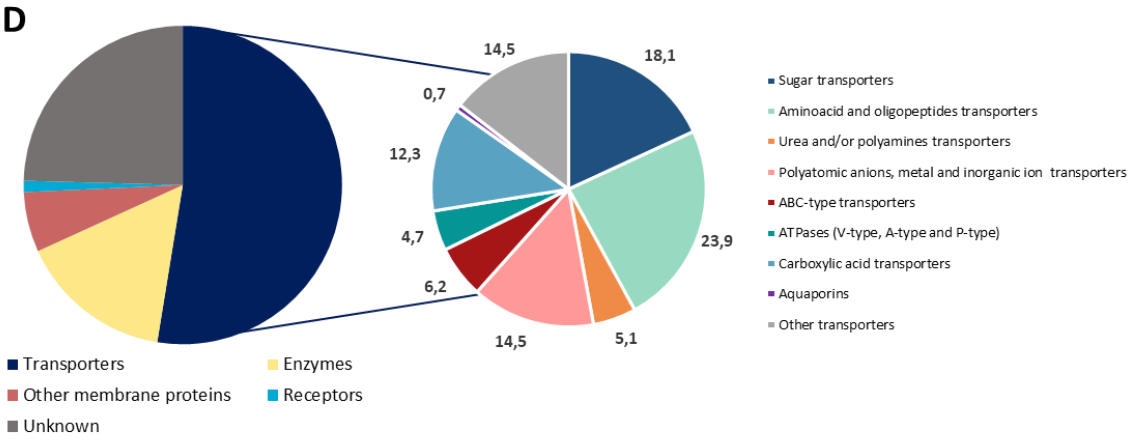


Figure 2. Identification of putative membrane proteins encoded in the *C. jadinii* NRRL Y-1542 genome. A) Establishment of a bioinformatic pipeline to analyze inferred proteomes and identify putative plasma membrane carboxylate transporters. B) Protein categories detected. C) Distribution (%) of the proteins estimated to contain 4 to 20 TMS. D) Estimated function of the 525 proteins (filtered from C) and distribution by categories (%) of those annotated as "transporters".

5.3.3 Heterologous expression of *Cyberlindnera jadinii* putative carboxylate transporters in *S. cerevisiae*

The *S. cerevisiae* IMX1000 strain transformed with the putative *C. jadinii* carboxylate transporters cloned in the plasmid p416GPD [24, 25] was used to evaluate the cells´ ability to grow on carboxylic acids as sole carbon and energy sources (Figure 3). All of the *S. cerevisiae* transformants were able to grow acetic acid, except for cells expressing the CjTDT. The strains transformed with CjAto homologs recovered the ability to grow on lactic acid. Furthermore, cells expressing CjAto2p also have an improved growth on pyruvic, malic, and succinic acids. The expression of CjAto5 allows cells to grow on all the mono-, di- and tricarboxylic acids tested. The *CjJen1* and *CjJen2* transformed cells were also able to grow on pyruvic acid. The expression of *CjJen6* improved cells' growth on dicarboxylic and tricarboxylic acids. The cells expressing SLC homologs CjSlc5 and CjSlc13-1 presented an improved growth, although not very strong, in all carbon sources tested, namely, lactic, pyruvic, fumaric, malic, succinic and citric acids. The cells expressing CjSlc13-2 present slightly decreased growth in lactic and citric acids media when compared to the negative control. The expression of *CjTDT* did not allow an improved growth of the cells in any of the carbon sources tested. Cells of *C. jadinii* DSM 2361 present a strong of growth on all carbon sources tested (Figure 3). Equivalent results were obtained using the *S. cerevisiae jen1*Δ *ato1*Δ strain transformed with the putative *C. jadinii* carboxylate transporters cloned in the plasmid p416GPD (not shown).

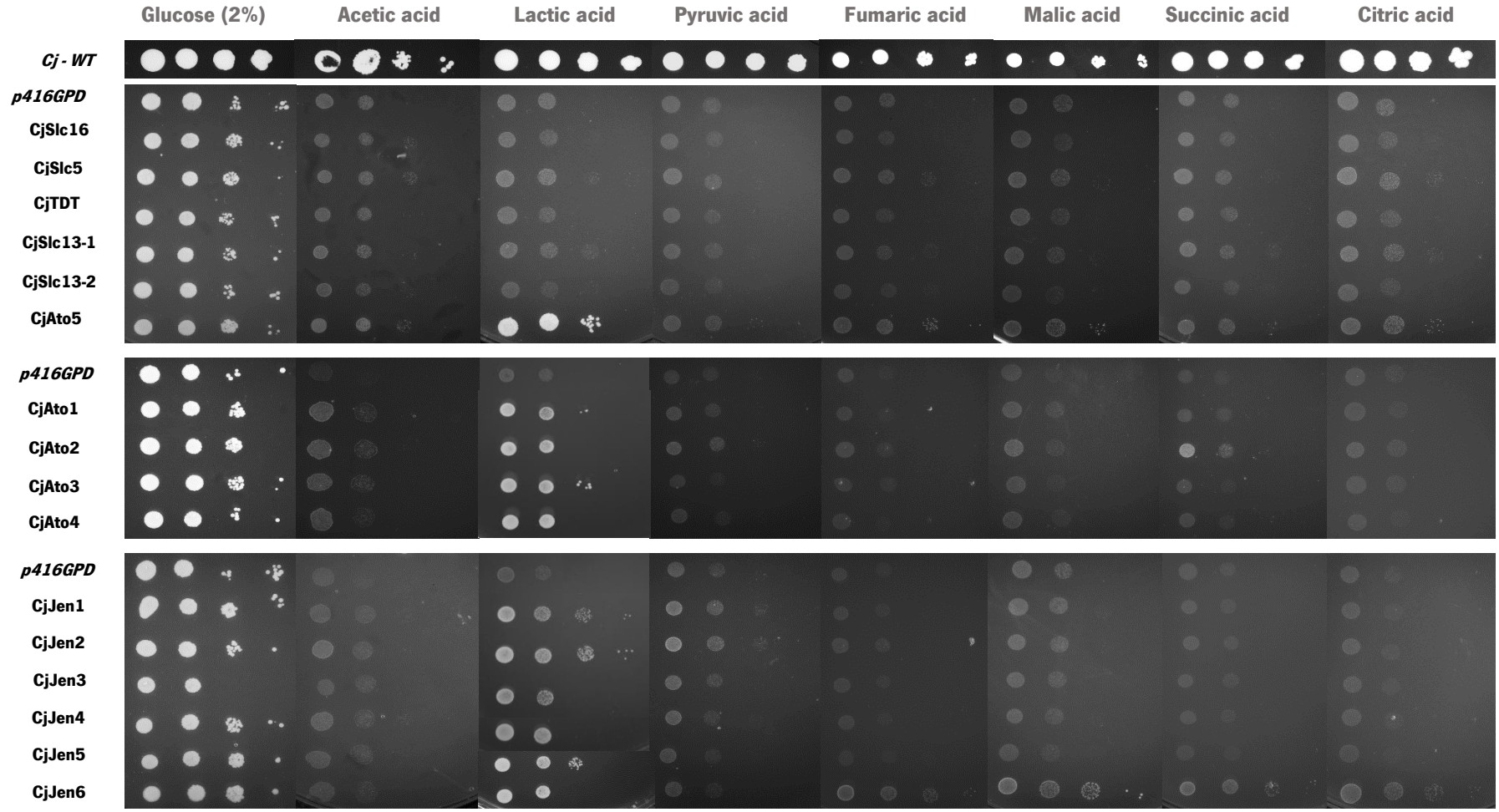


Figure 3. Functional analysis of *C. jadinni* putative carboxylate transporters in *S. cerevisiae* IMX1000. Cells were grown at 18 °C for 22 days on YNB media containing glucose (2% w/v), acetic acid (0.5% v/v; pH 6.0), lactic acid (0.5% v/v; pH 5.0), pyruvic acid (0.5% w/v; pH 5.0), fumaric acid (1% w/v; pH 5.0), malic acid (1% w/v; pH 5.0), succinic acid (1% w/v; pH 5.0) or citric acid (1% w/v; pH 5.5), as sole carbon and energy source. Spot assays were performed as described in materials and methods. The *C. jadinii* DSM 2361 strain was included as a positive control and *S. cerevisiae* IMX1000 cells transformed with p416GPD as a negative control.

Cells of *S. cerevisiae* W303-1A *jen1*Δ *ato1*Δ expressing Cj-transporters were used to evaluate the uptake of [^{14}C]acetic acid (pH 6.0), [^{14}C]lactic acid and [2,3-^{14}C]succinic acid (pH 5.0) (Figures 4 and 5).

Cells expressing the CjAto1-4 homologs presented an increased uptake of acetate and lactate (1mM concentration) when compared to the strain carrying the empty vector, supporting their role as acetate and lactate transporters (Figure 4A). For acetate, the values obtained for CjAto3 and CjAto4, were similar to the native ScAto1, while the lactate uptake was higher than for ScJen1. Cells expressing CjAto2 presented uptake for succinate higher than the uptake for the CaJen2, the dicarboxylate transporter from *Candida albicans*, used as positive control. All the CjJen1-6 homologs presented transport activity for lactate and acetate, confirming their function as acetate and lactate transporters (Figure 4B). Succinate transporter activity was found to be associated with the expression of CjJen6, CjSlc5 and CjAto5 (data not shown) as the estimated velocity (*V*) for the uptake of 1mM succinate (pH 5.0) in cells transformed with the empty vector was 0.039 nmol succinate s^{-1} mg^{-1} dry wt., whereas 0.131, 0.240 and 0.277 nmol succinate s^{-1} mg^{-1} dry wt. was estimated for CjJen6, CjSlc5 and CjAto5, respectively.

The *S. cerevisiae* IMX1000 strain was used for the heterologous expression of CjSLC16, CjSLC5, CjSLC13-1 and CjSLC13-2, the CjTDT and CjAto5p, as this *S. cerevisiae* strain is deleted in all known and putative carboxylate transporters. All the resulting strains revealed activity for an acetate transporter (Figure 4C).

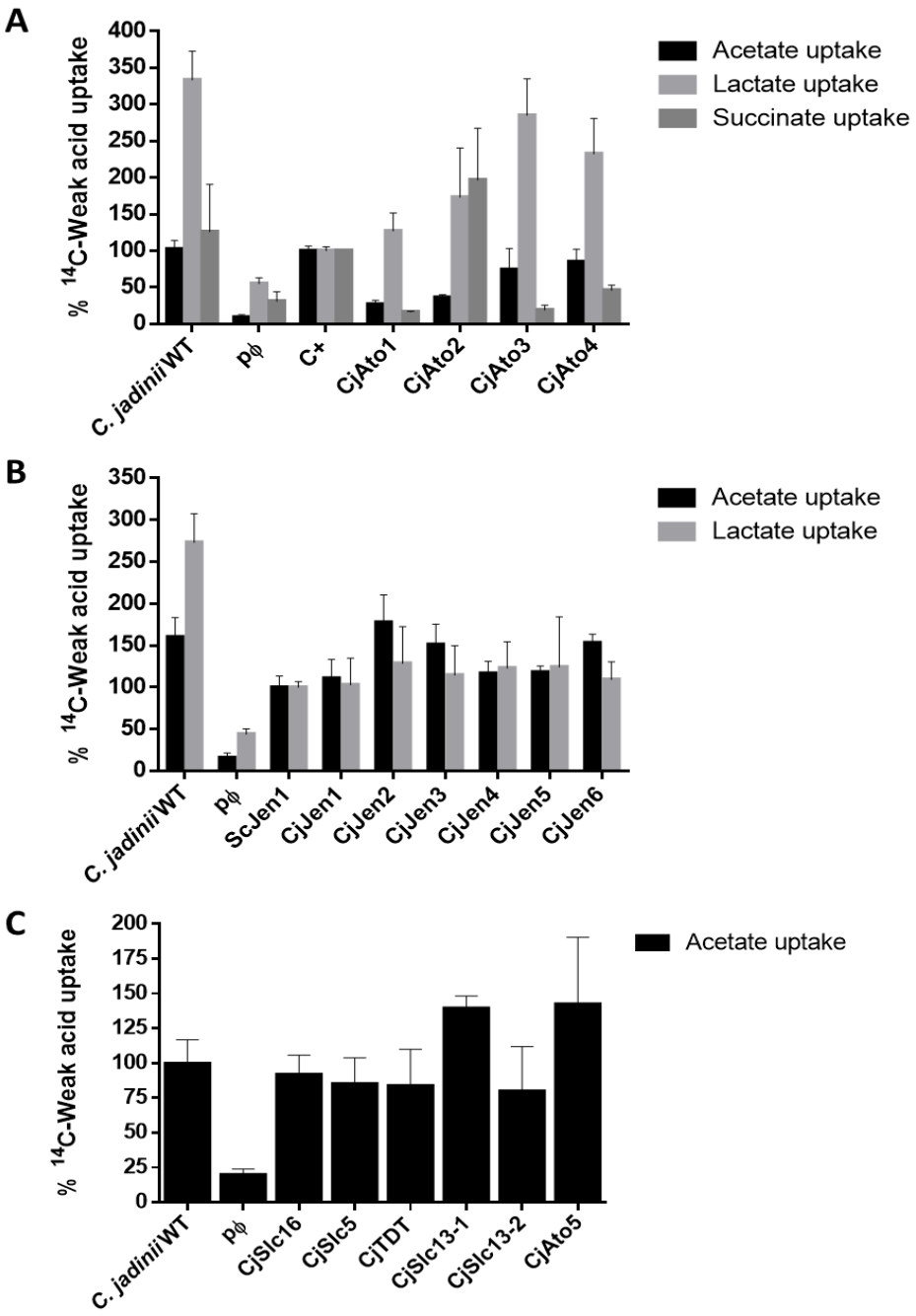


Figure 4. Transport of radiolabelled carboxylates in *S. cerevisiae jen1*Δ *ato1*Δ (A,B) and IMX1000 (C) cells expressing the *C. jadinii* CjAto1-5, CjJen1-6, CjSlc16, CjSlc5, CjTDT, CjSlc13-1, CjSlc13-2 proteins and empty vector (pϕ), as negative control. A) Uptake of 1 mM of ^{14}C-acetic acid (pH 6.0), ^{14}C-lactic acid (pH 5.0) and ^{14}C-succinic (pH 5.0) acid. B) Uptake of 1 mM of ^{14}C-acetic acid (pH 6.0) and 1 mM of ^{14}C- lactic acid (pH 5.0). C) Uptake of 1 mM of ^{14}C-acetic acid (pH 6.0). Cells were grown on YNB-glucose, washed and incubated on YNB containing the respective carbon source (see materials and methods) used in the uptake assay for 6 hours (Acetate 0.5 %, pH 6.0) and 5 hours (Lactate 0.5 %, pH 5.0; Succinate 1 %, pH 5.0). In A) the value of 100% (C+) corresponds to the uptake of acetic acid and lactic acid in *S. cerevisiae* cells transformed with the plasmid p416GPD expressing the native transporters Ato1 and Jen1, respectively. For the uptake of succinic acid, the value 100% (C+) corresponds to isogenic cells expressing the *Candida albicans* Jen2 transporter. In B) the value 100% corresponds to the uptake of acetic and lactic acid displayed by cells expressing the native Jen1. In C) the value 100% corresponds to the uptake of acetic acid by *C. jadinii* DSM 2361 cells.

The kinetic parameters of acetic acid uptake were estimated in the *S. cerevisiae* W303-1A *jen1*Δ *ato1*Δ strain expressing the CjAto and CjJen homologs (table 4).

Table 4. Kinetic parameters for ^{14}C-acetic acid (pH 6.0) transport estimated in cells of *S. cerevisiae* W303-1A *jen1*Δ *ato1*Δ expressing the CjAto1, CjAto3-4 and CjJen1-6 genes. Cells were cultivated in glucose media until mid-exponential growth phase and shifted to YNB acetate 0.5%, pH 6.0 for 6 h.

Plasmid	K_m (mM)	V_{max} (nmol acetate s^{-1} mg^{-1} dry wt.)
pCjAto1	12.18±4.67	4.92±1.32
pCjAto3	9.17±4.70	3.85±1.23
pCjAto4	1.28±0.38	0.69±0.10
pCjJen1	2.15±0.60	0.86±0.12
pCjJen2	1.26±0.35	0.49±0.06
pCjJen3	3.21±1.07	1.34±0.25
pCjJen4	1.62±0.43	0.62±0.08
pCjJen5	5.18±1.51	2.08±0.39
pCjJen6	3.47±1.08	1.52±0.26

The transporters presented a wide range of V_{max}, ranging from 0.49 to 4.92 nmol acetate s^{-1} mg^{-1} dry wt, and K_m, from 1.26 to 12.18 mM acetate.

Cells of *S. cerevisiae* W303-1A *jen1*Δ *ato1*Δ strain expressing the CjAto2 presented a mediated transport system for lactate with the following kinetic parameters for lactate and succinate: K_m 4.33 ± 0.70 mM lactate and V_{max} 1.69 ± 0.17 nmol lactate s^{-1} mg^{-1} dry wt. K_m 3.38 ± 0.82 mM succinate and V_{max} 1.14 ± 0.17 nmol succinate s^{-1} mg^{-1} dry wt.) at pH 5.0 (Figure 5A). The initial uptake rates of ^{14}C-succinic acid (pH 5.0) estimated in *S. cerevisiae* IMX1000 transformed with pCjAto5 revealed a Michaelis-Menten kinetics with the following parameters K_m = 0.45 ± 0.09 mM and V_{max} = 0.33 ± 0.02 nmol s^{-1} mg^{-1} dry wt (Figure 5B). The cells transformed with the empty vector present a first order kinetics associated with the diffusion of the acid (Figure 5B).

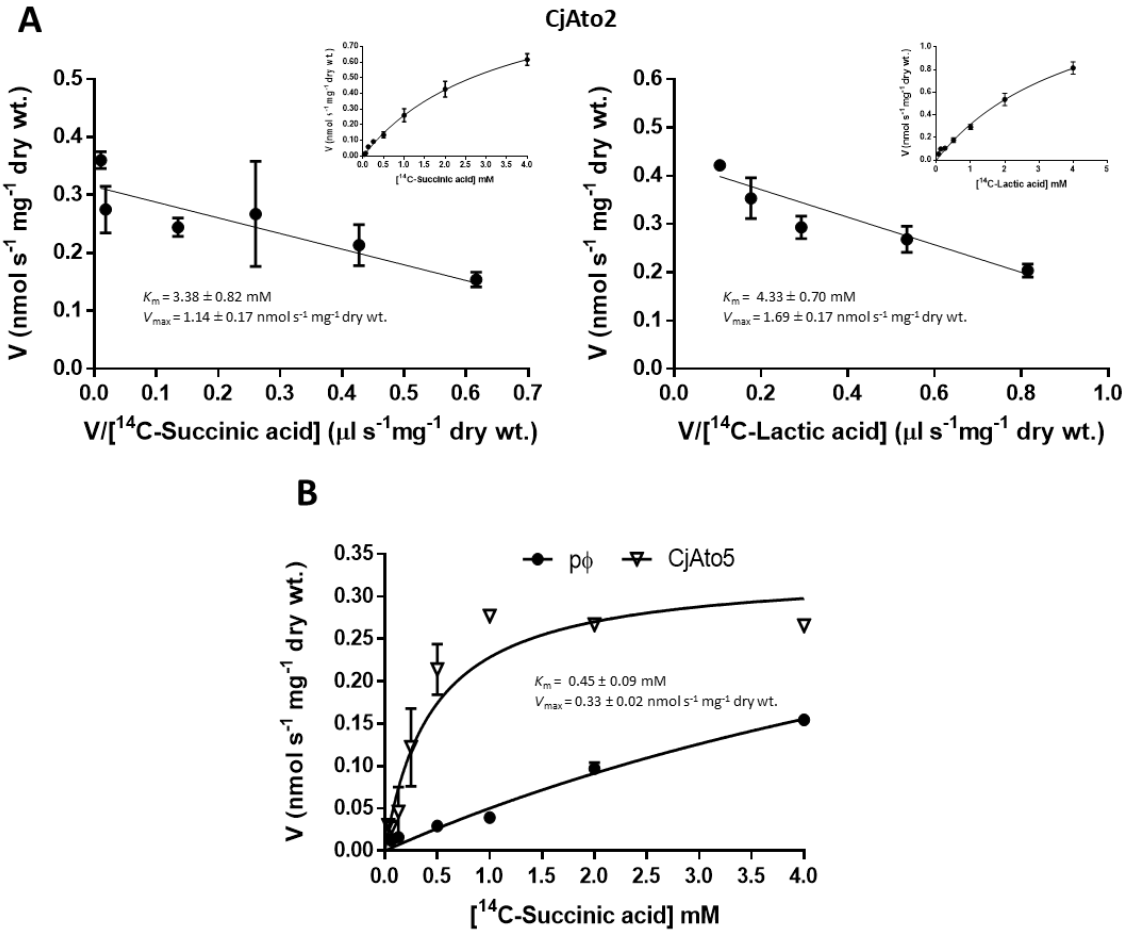
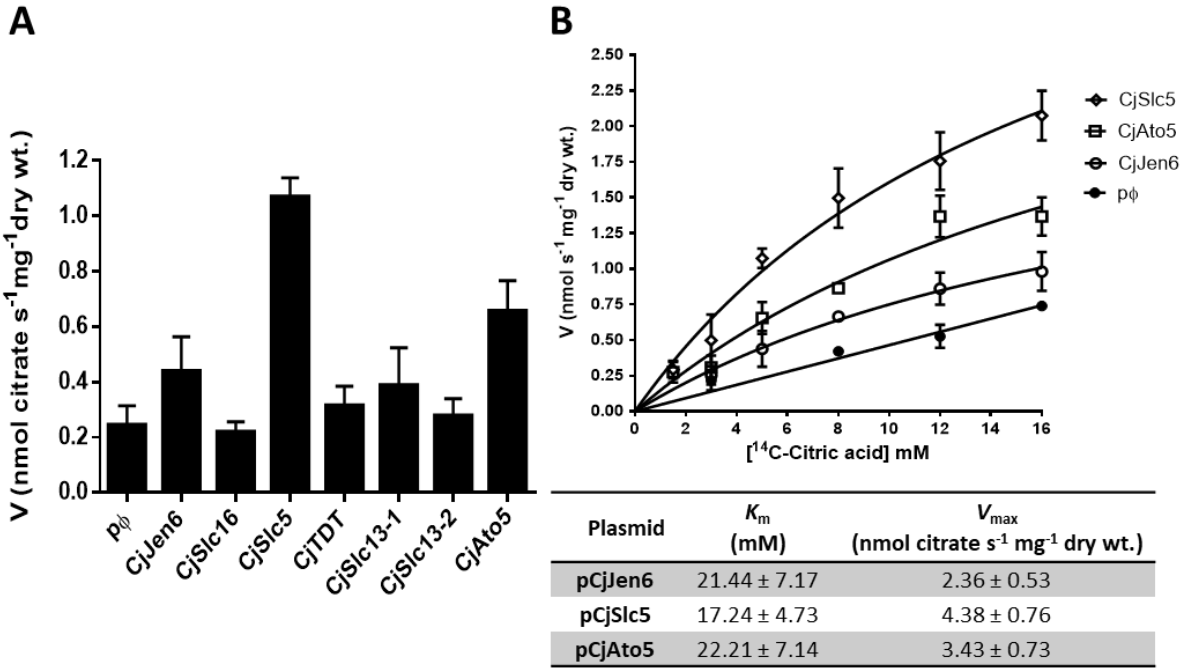


Figure 5. Kinetic parameters for the initial uptake rates of labelled carboxylates of *S. cerevisiae* expressing *CjATO2* and *CjATO5*. A) Eadie-Hofstee plots of labelled carboxylic acids (^{14}C-succinic and ^{14}C-lactic acids) as a function of the acid concentration of the *S. cerevisiae* W303-1A *jen1*Δ *ato1*Δ strain expressing pCjAto2 at pH 5.0, 30°C. B) Initial uptake rates of radiolabelled ^{14}C-succinic acid as a function of the acid concentration of the *S. cerevisiae* IMX1000 strain expressing the pCjAto5 and transformed with the expression vector p416GPD, pH 5.0, 30°C.

5.3.4 CHARACTERIZATION OF *C. JADINII* CITRATE TRANSPORTERS

Citrate uptake was evaluated in *S. cerevisiae* IMX1000 expressing *CjJEN6, CjSLC16, CjSLC5, CjTDT, CjSLC13-1, CjSLC13-2* and *CjATO5* (Figure 6A). A more detailed analysis of citrate transport kinetics was done in selected transporters (Figure 6B). The kinetic parameters of the most promising citrate transporters, i.e. CjAto5p, CjJen6p, CjSlc5p, are displayed in Figure 6B presenting an affinity (K_m) in the same range of values. The transporter with the highest activity is the CjSlc5p followed by CjAto5p. Despite having the lowest transport activity, CjJen6p still allows the growth of *S. cerevisiae* cells on media containing citrate as sole carbon and energy source (see Figure 3).



Plasmid	K_m (mM)	V_{max} (nmol citrate s^{-1} mg^{-1} dry wt.)
pCjJen6	21.44 ± 7.17	2.36 ± 0.53
pCjSlc5	17.24 ± 4.73	4.38 ± 0.76
pCjAto5	22.21 ± 7.14	3.43 ± 0.73

Figure 6. Citrate transport in *S. cerevisiae* IMX1000 cells expressing *C. jadinii* putative transporters. (A) Uptake of ^{14}C-citric acid 5 mM at pH 5.5, 30°C, of cells transformed with CjJen6, CjSLC16, CjSlc5, CjTDT, CjSlc13-1, CjSlc13-2, CjAto5 and empty vector (pϕ). (B) Initial uptake rates of radiolabeled ^{14}C-citric acid as a function of the acid concentration (pH 5.5, 30°C) in cells expressing *C. jadinii* Jen6, Slc5 and Ato5; the table summarizes the estimated kinetic parameters. Cells were grown in YNB-glucose, washed and incubated on YNB-citric acid (1.0 %, pH 5.5) for 5 h. The data shown are mean values of at least three independent experiments and the error bars represent the standard deviation.

5.3.5 Phylogenetic analysis of carboxylate transporter families

A phylogenetic analysis of *C. jadinii* carboxylate transporters belonging to the SHS (TC 2.A.1.12), SSS (TC 2.A.21) and DASS (TC 2.A.47) transporter families was carried out throughout the tree of life. For the AceTr (TC 2.A.96) family the analysis was limited to Ascomycota, since a broad phylogenetic analysis was already presented in a previous study [30]. Homolog nominations and bootstrap values from the phylogenetic reconstructions are all available in figures S3 in the Supplementary material section.

For the AceTr family a total of 319 hits were obtained for the BLAST search in the NCBI's Assembly database, using complete genomes only, using CjAto1-5 homologs as queries. Eight sequences lacking large conserved regions were excluded for a final dataset of 311 sequences. The phylogenetic tree presents five main clades (E1-E5, Figure 7), depicting the functionally characterized Ato transporters

described in literature. Clade E1 includes the Ato1 and Ato2 from *S. cerevisiae* [14], Gpr1 from *Y. lipolytica* [14, 30], and the eight Ato homologs from *C. albicans* [42, 43]; the *S. cerevisiae* Ato3 is in the clade E2 [44]; acetate transporters AcpA and AcpB from *A. nidulans* belongs to E3 clade [45, 46]; the acetate transporter *A. nidulans* AcpC displayed in clade E4 [46] and AlcS from *A. nidulans* in E5 clade [47]. The CjAto homologs were found into two distinct clades, five are located in E1 and CjAto5p appears in E2. Half of the AceTr homologs are located in Ato1/Ato2/Gpr1 clade, the majority belong to the Debaryomycetaceae (*Candida* and *Debaryomyces* species) and Saccharomycetaceae (*Saccharomyces, Kluyveromyces* and *Lachancea* species) families. Some of these species have homologs in the E2 clade too, probably due to the occurrence of a genetic duplication phenomena, as previously reported [30]. Homologs from the genus *Fusarium* spp are dispersed in several clades.

For the phylogenetic analysis the of SHS family a total of 319 hits were obtained from the BLAST search with the six CjJen members. Fourteen sequences that lacked large conserved regions were excluded for a final dataset of 305 sequences. The phylogenetic reconstruction proposes a basal split into prokaryotic and eukaryotic organisms forming two monophyletic clades (Figure 8). The prokaryotic clade gathers two branches, P1 and P2, showing bacterial and archaeal homologs, respectively. P1 clade encloses mostly homologs belonging to Proteobacteria, Acidobacteria and Actinobacteria phylum. The P2 clade only includes two archaeal homologs that belong to the *Thermoplasma* genus. The eukaryotic clade splits into three branches labelled E1, E2 and E3 clades. The E1 clade comprises two subclades: the monocarboxylate transporters Jen1-homologs, namely, ScJen1 from *S. cerevisiae* [13], CaJen1p from *C. albicans* [48], KlJen1p from *K. lactis* [41], and two homologs DH17p and DH27p from *D. hansenii* [49]; the dicarboxylate transporters Jen2-homologs, namely, CaJen2p from *C. albicans* [26, 50], KlJen2p from *K. lactis* [41], DH18p and DH24p from *D. hansenii* [49], the six clustered-homologs characterized from *Y. lipolytica* as functional mono/dicarboxylate transporters [51], and the PkJEN2-1 (a dicarboxylate transporter) and PkJEN2-2 (a di- and tricarboxylate transporter) from *Pichia kudriavzevii* [52]. In E2 clade we find three homologs from Ascomycota phyla *Pochonia chlamydosporia* and two other from the filamentous fungi *Aspergillus fumigatus* and *Zymoseptoria tritici.* The E3 clade contains homologs from Ascomycota and Basidiomycota, namely Jen1 from *Cryptococcus neoformans*, reported as a 3-bromopyruvate transporter [53]. All the CjJen homologs are in the E1 clade, four in the Jen1-cluster (CjJen1-4) and two (CjJen5,6) in the Jen2-cluster.

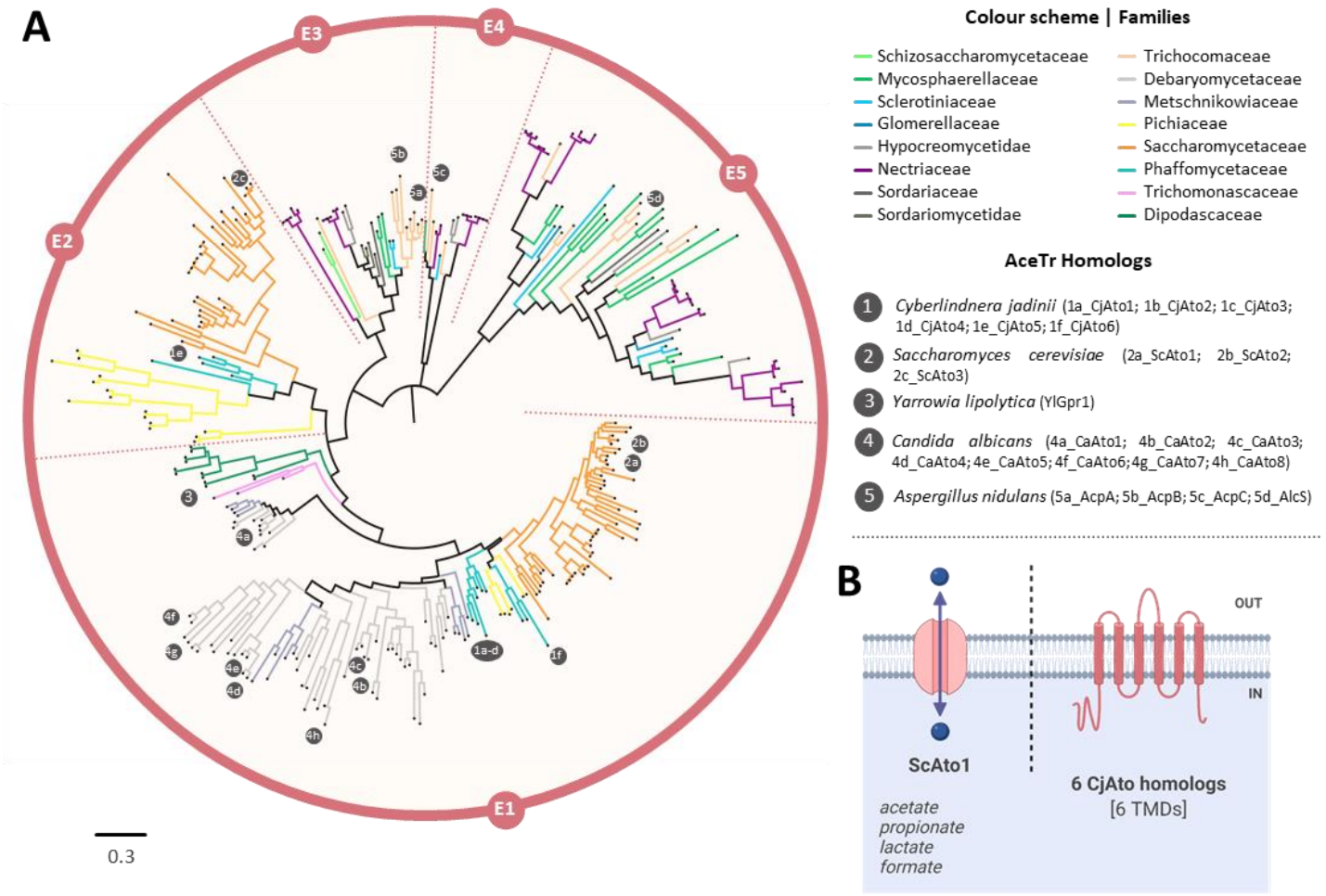


Figure 7. Maximum likelihood phylogenetic tree of AceTr family members (TCDB 2.A.96) present in ascomycetes. A) Branch lengths are proportional to sequence divergence and different colors correspond to the different families. The clades indicated as E1, E2, E3, E4 and E5 were created to facilitate the tree description in the main text and are not meant to provide any type of classification. Homologs relevant for the discussion through the manuscript are highlighted. B) Simplified scheme of ATO1 native and CjAto proteins with inferred substrate specificities is depicted.

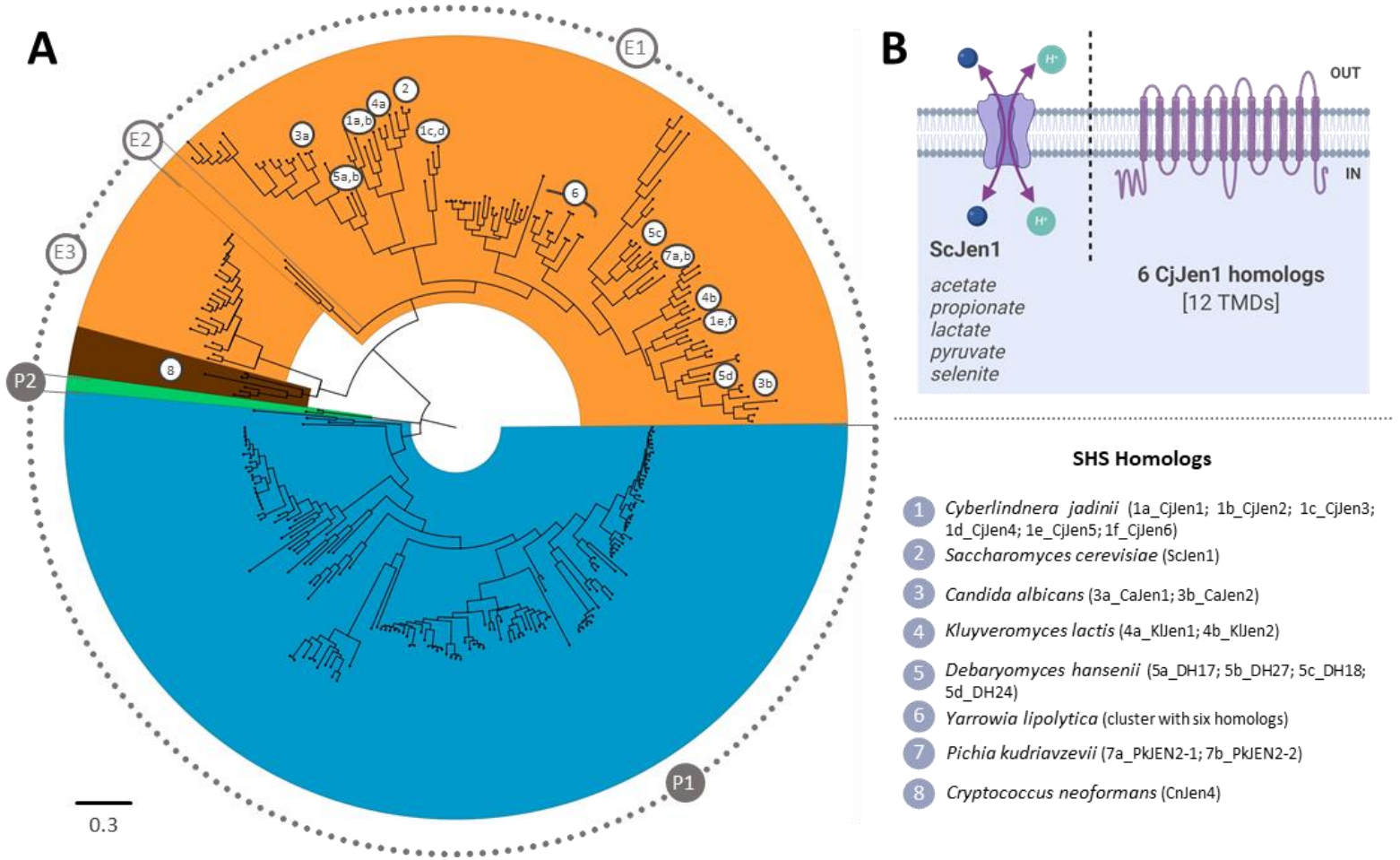


Figure 8. Maximum likelihood phylogenetic tree of SHS family members (TCDB 2.A.1.12). A) Groups indicated as E1, E2 and E3 for eukaryotic clades, P1 and P2 for prokaryotic clades, were established to enable the following of the tree description in the main text and are not meant to provide any type of classification. Major taxonomic groups are indicated in shades of blue – bacteria, orange – ascomycetes, brown – basidiomycetes and green – archaea. Homologs relevant for the discussion through the manuscript are highlighted. B) Simplified scheme of JEN1 native and CjJen proteins with inferred substrate specificities is depicted.

The CjJen homologs were found in groups of two suggesting their origin derived from a genetic duplication phenomenon. Only microbial genomes were found in the SHS-phylogenetic reconstruction, from which around 60% belong to eukaryotic taxa.

Phylogenetic reconstruction of the SSS-tree (SLC5 member) showed a total of 325 hits obtained from the BLAST search. About ten protein sequences that lacked large conserved regions were excluded for a final dataset of 315 sequences (Figure 9). The CjSLC5 is manually annotated as member of solute carrier superfamilies 5 and 6-like. This superfamily includes a specific solute-binding domain of SLC5 proteins, also called the sodium/glucose cotransporters or solute sodium symporters; SLC6 proteins are known as Na^+/Cl-dependent transporters and nucleobase-cation-symport-1 (NCS1) transporters. CjSl5 homologs were only identified in eukaryotic organisms, mostly belonging to Asco- and Basidiomycota phyla. Homologs were distributed across three major clades, E1, E2 and E3. The E1 clade contains the sole *C. jadinii* homolog, CjSlc5, that is phylogenetical closer to the homologs from the Metschnikowiaceae, Debaryomycetaceae and Saccharomycetaceae families. In E1 there is a minor subclade that includes basidiomycetes members from *Cryptococcus*, *Ustilago* and *Sporisorium* genus, as well as the amino acid transporters NcMtr/AAP1 and PcMtr from *Neurospora crassa* [54] and *Penicillium chrysogenum* [55], respectively.

The E2 and E3 contain mainly yeast homologs and few basiodiomycetes. Interestingly, SLC5 homologs are present in several *Fusarium* species, which together with *Aspergillus* and *Penicillium* species, constitute 40% of total members present in the phylogenetic tree, however, most of these remain uncharacterized.

The phylogenetic tree of DASS (SLC13 member) family contains a total of 530 members retrieved from the BLAST search with two the homologs CjSlc13-1 and CjSlc13-2 (Figure 10). A total of eleven sequences that lacked large conserved regions were excluded, resulting in a final dataset of 519 sequences. The DASS family is found in bacteria, archaea, and fungi, in a few single-cell eukaryotes from the *Leishmania* and *Trypanosoma* genus, and in one homolog from the red algae Rhodophyta *Cyanidioschyzon merolae*. Additionally, among the prokaryotic clade, there is an isolated cluster with multicellular eukaryotes. Homologs were found in plants, e.g. *Zea mays* (corn), *Beta vulgaris* (beet) and *Panicum hallii* (punch grass), and in the nematode *Strongyloides ratt.* The sections in the CjSlc13-tree were labelled as E1 and E2 for the eukaryotic clades, and P1 to P3 for the prokaryotic clades. E1 clade contains SLC13 homologs present in ascomycetes and basidiomycetes, namely the *S. cerevisiae* phosphate transporters Pho87p and Pho90p [56, 57], and Pho91p [58], as well as the CjSlc13 member. The E2 clade is defined by the microbial eukaryote species and red algae species. P1 clade splits into two branches.

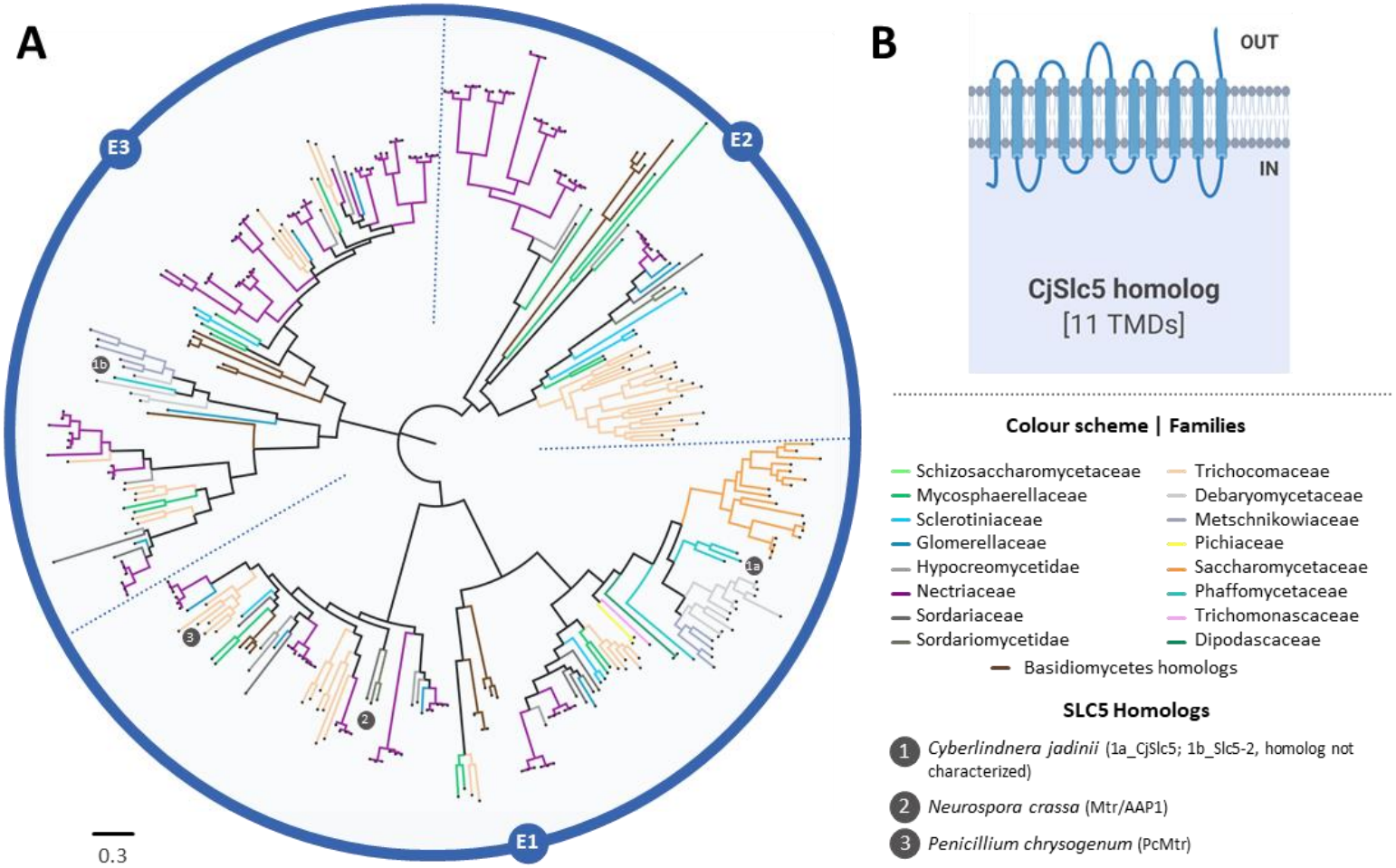


Figure 9. Maximum likelihood phylogenetic tree of SSS (SLC5 homolog) family members (TCDB 2.A.21). A) Branch lengths are proportional to sequence divergence. Groups indicated as E1, E2 and E3 were defined to enable the following of the tree description in the main text and are not meant to provide any type of classification. Homologs relevant for the discussion through the manuscript are highlighted. B) Simplified scheme with predicted topology for CjSlc5 protein.

P1 major branch contains two other sub-branches: one extensive containing mostly homologs present in the bacterial phylum from Actinobacteria, Firmicutes, Proteobacteria and others, some archaeal homologs and others from plant and animal species; the other small sub-branch holds homologs belonging to Firmicutes phyla. In the larger P1 sub-branch, the SdcL, a sodium-dependent succinate and malate transporter from *Bacillus licheniformis*, is found [59]. The P1 minor branch holds a high percentage of bacterial homologs from Proteobacteria, Spirochaetes phyla and from other bacterial categories. Homologs from Proteobacteria species, one from *Oleiphilus messinensis* (marine bacteria) and another from Spirochaetes phyla constitute the P2 clade. Finally, the P3 clade includes homologs from Proteobacteria and Firmicutes phyla. Slc13 homologs display a ubiquitous distribution, being mostly found in prokaryotic genomes but this could also reflect the higher number of prokaryotic genomes present in the database, when compared to eukaryotic and archaeal.

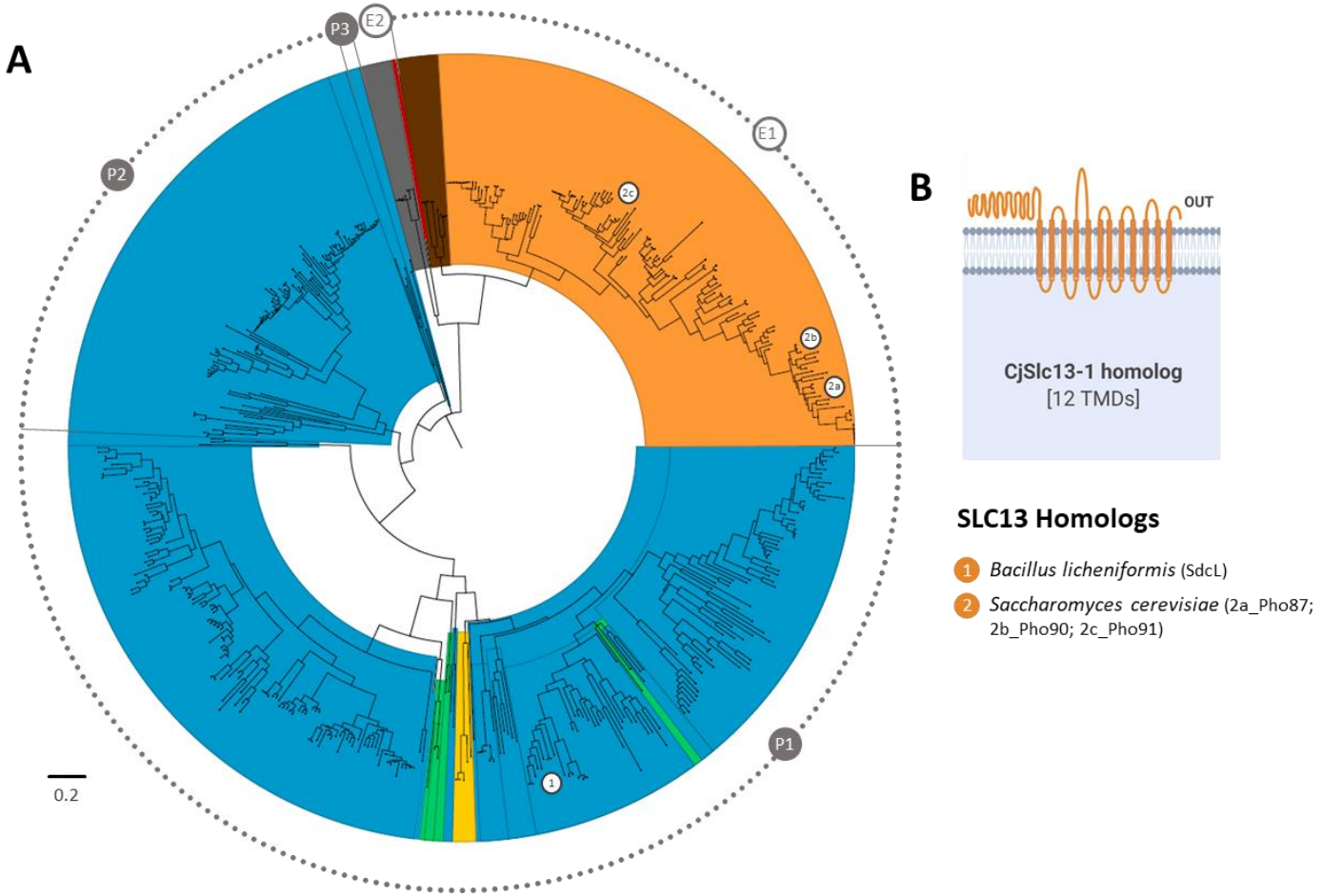


Figure 10. Maximum likelihood phylogenetic tree of DASS (SLC13 homolog) family members (TC 2.A.47). A) Branch lengths are proportional to sequence divergence. Groups indicated as E1 and E2 for eukaryotic clades, P1, P2 and P3 for prokaryotic clades, were defined to facilitate the following of the tree description in the main text and are not meant to provide any type of classification. Major taxonomic groups are indicated in shades of blue – bacteria, orange – ascomycetes, green – archaea, yellow – higher eukaryotes, brown – basidiomycetes, grey – Leishmania/Trypanosoma species, red – algae. Homologs relevant for the discussion through the manuscript are highlighted. B) Simplified scheme with predicted topology for CjSlc13 protein.

5.3.6 *IN SILICO* STRUCTURAL ANALYSIS

The structure of *C. jadinii* carboxylate transporters was analyzed using (i) conserved residues found in the multiple alignment of functionally characterized homologs (Figure S1), (ii) HHPred for 3D structure prediction (Table S3.1) and (iii) molecular docking studies with substrates (Figures 11-14).

5.3.6.1 *IN SILICO* STRUCTURAL ANALYSIS OF THE CJATO2, CJATO5 AND CJATO6 TRANSPORTERS

Multiple-sequence alignments revealed that CjAto1-5, ScAto1-3 and EcSatp share several conserved regions (Figure S1). The signature NPAPLGL(M/S) motif of the AceTr family [30], previously reported as crucial for substrate uptake, has substitutions in six positions, in which four of them are verified in CjAto6. The N89-ScAto1 residue is replaced by D83 in CjAto6p, A91-ScAto1 is substituted by G85 in CjAto6p, and the highly conserved L93-ScAto1 is substituted by the hydrophobic residues I75 in CjAto4, I87 in CjAto6 and by M85 in CjAto5p. As for L95-ScAto1, a single substitution was found in CjAto6p, to M89. Other residues substitutions, such as M107-CjAto5 (L99-ScAto1), both in CjAto6 and ScAto3 were found. In ScAto3 and CjAto5, the S89 residue located close to the pore constriction site was also found not to be conserved considering G/A97-ScAto1. All these substitutions are located in the first TMS and are likely to alter the affinity for the substrates, and as a consequence, can affect transporter specific and activity. In the *Citrobacter koseri* SatP crystal structure the central and narrowest hydrophobic constriction of the anion pathway is formed by F17, Y72 and L131 residues [60]. This narrowest constriction in ScAto1 corresponds to F98-Y155-L219. The F98 residue is highly conserved among the homologs analyzed here, but the Y155 is replaced by F147 in CjAto5p, ScAto2 and ScAto3, whose alterations can play a vital role in substrate passage through the pore. In the position L219 a neutral substitution occurs in CjAto5 (V210) [61]. In the third TMS, there is another signature motif SYG(X)FW that includes the Y155 of the FYL constriction site. Here F158 was substituted by Y152 in CjAto6 [62]. This substitution by tyrosine that contains a reactive hydroxyl group turns it much more likely to be involved in interactions with the substrates.

The predicted 3D structure was obtained by homology threading using the crystal structure of a bacterial homolog, the succinate-acetate permease Ck_SatP (PDB 5YS3) [60], as a template. The identity presented with the different homologs was 35% for ScAto1, 32% for CjAto2, 26% for CjAto5 and 30% for CjAto6 (Figure 11). Substrate docking studies were carried out for lactate, succinate and citrate. Several residues seem to be involved in carboxylate binding in the CjAto protein templates (see table S3.2). In the predicted 3-D model, succinate deprotonated forms (mono- and dianionic) present strong interactions

with residues S211 (end of TMS5) and Q219 (beginning of TMS6) from CjAto2, two residues that as aforementioned were not conserved. A higher number of strong interactions were detected for succinate in CjAto2, when compared to ScAto1 that does not transport this acid. Also in CjAto5, succinate is predicted to interact with K221 (end of TMS5) and, with E125 (located in TMS2). This latter residue is also predicted to interact with citrate, in addition to residues Q103 (TMS1), K230 (TMS6) and D253 (TMS6). In fact, both Q103 and D253 are not conserved across all the characterized proteins, having instead an arginine and a tyrosine in these locations, respectively (Figure S1.1).

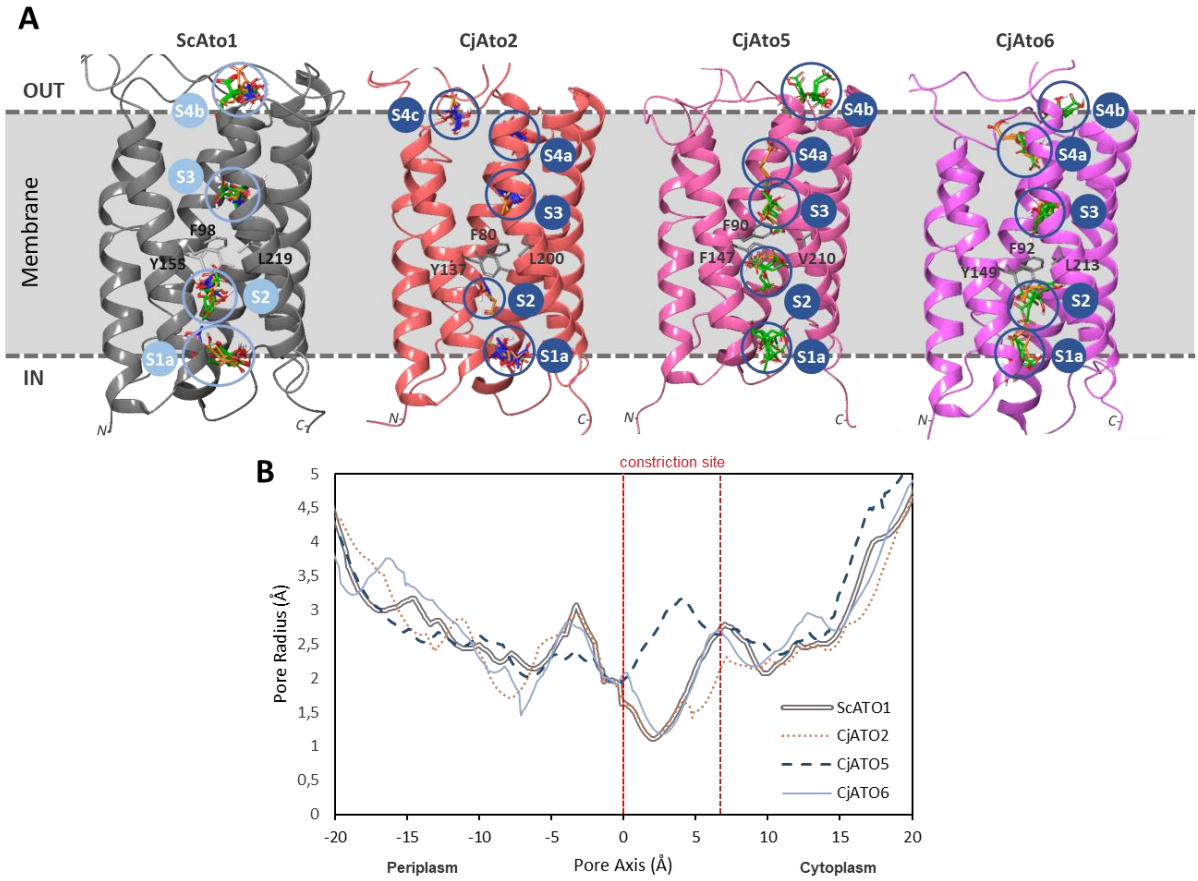


Figure 11. Predicted 3D structure of ScAto1, CjAto2, CjAto5 and CjAto6 transporters. A) Molecular docking of ScAto1, CjAto2, CjAto5 and CjAto6 3D-models, based on SatP_Ck structure (PDB 5YS3), with the substrates lactate (blue ligand), succinate (orange ligand) and citrate (green ligand). The four binding sites (S1-4) described in SatP-Ck were also found in the CjAto2/5/6. Localization of N- and C-terminal of the proteins is depicted. B) Simulations for the pore radius profiles through channel axis in the ScAto1 (grey line), CjAto2 (orange dotted line), CjAto5 (dark blue dashed-line) and CjAto6 (blue line) proteins. The core of the channel where the constriction is found site is labeled by the vertical dashed red lines.

The binding affinities of the residues presenting strong and van-der-walls interactions with the substrates were analyzed (Table S3.3). In accordance to what is reported for acetate in the homolog CkSatP [60], four binding sites were found through the channel axis in all CjAto proteins (Figure 11A). The estimated lactate binding affinity values are similar in ScAto1 and CjAto2. A higher number of interactions for lactate and succinate was detected in CjAto2. An increased citrate binding affinity was found in the S2 binding-site of CjAto5 (residue K221) and CjAto6 (Figure 11A). The pore radius at the FYL-constriction site was measured in the four protein models (Figure 11B). A significant increase of the pore radius was found in CjAto5, compared to the ScAto1 (1.25 Å) that can lead to a structural alteration in the channel pore constriction site (Figures 11B and S4.2) thus allowing the passage of larger molecules. CjAto6 only presents a slight increase of the pore radius. In CjAto2 the constriction pore size is similar to ScAto1, although it is able to transport larger molecules like succinate.

5.3.6.2 *IN SILICO* STRUCTURAL ANALYSIS OF THE CJJEN5 AND CJJEN6 TRANSPORTERS

The multiple-sequence alignment of Jen homologs (using ScJen1 as reference) revealed the conservation of some amino acid residues among functionally characterized homologs. The conserved motif ^{379}NXX[S/T]HX[S/T]QD^{387}, located at TMS7, is conserved in all the CjJen homologs. The L366 residue near to this domain is conserved across the CjJen homologs except for CjJen6p that has been substituted by V260. In TMS5, conserved S266, whose substitution to Leucine abolished CaJen1 activity [48], is a glycine in CjJen3 and CjJen4. On TMS11, Q498 and N501 residues were also found to be conserved in all the Jen homologs. CjJen5 and CjJen6 proteins share 75% identity, however they present a very different substrate specificity. In TMS1, CjJen6p presents a W in the residue 154, that is also present in CaJen2p and KlJen2p which are dicarboxylate transporters [41, 50]. The L-lactate MFS transporter of *Syntrophobacter fumaroxidans* (PDB 6G9X) [63] was used as template for the 3D structure prediction by homology threading of ScJen1p (14% of identity), CaJen2p (13% of identity), and CjJen5p (13% of identity) and CjJen6p (15% of identity) (Figure 12). Docking studies predicted the amino acid residues involved in strong interactions with lactate, succinate and citrate (Table S3.4). Additionally, the results revealed that the highest number of interactions for citrate is found for CjJen6 with an increased binding affinity in three binding-sites for this acid (positions 1, 2 and 4 in figure 12, see Table S3.5).

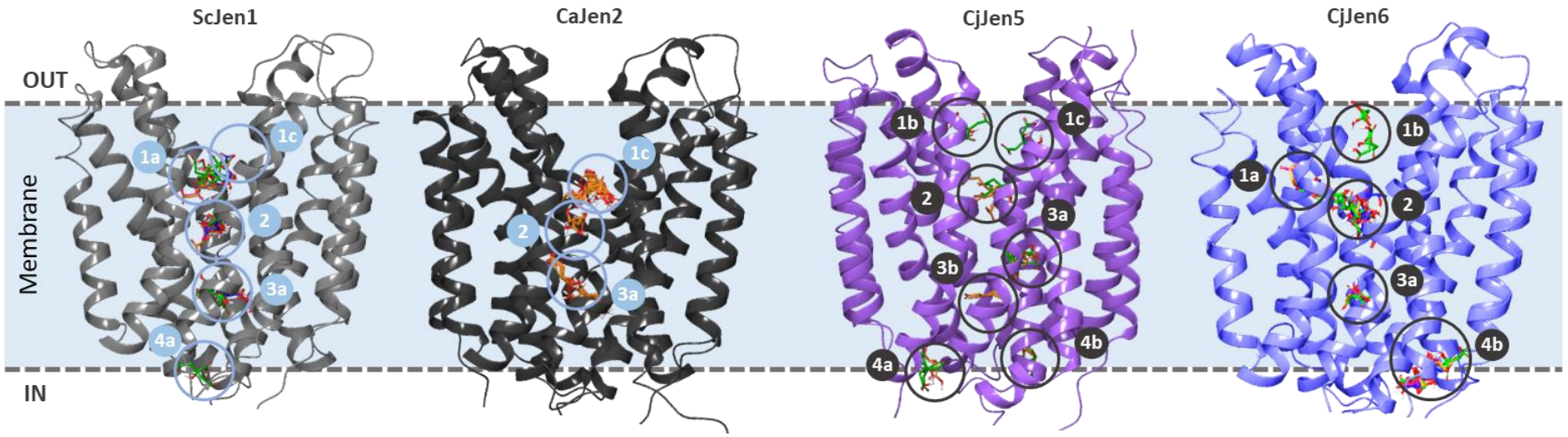


Figure 12. Predicted 3D-structure of ScJen1, CaJen2, CjJen5 and CjJen6. Molecular docking of ScJen1, CaJen2, CjJen5 and CjJen6 3D-models, based on *S. fumaroxidans* crystal structure (PDB 6G9X) of a MFS transporter, with the substrates acetate (pink ligand), lactate (blue ligand), succinate (orange ligand) and citrate (green ligand).

5.3.6.3 *IN SILICO* STRUCTURAL ANALYSIS OF CJSLC5 TRANSPORTER

An alignment between the *Homo sapiens* SLC5 member 8 protein and the *C. jadinii* SLC5 transporter was done revealing the existence of conserved residues ^{32}VLGXPS^{37} on TMS1, ^{116}ALN^{118} in the TMS3 and, in the TMS8 the ^{287}PGLXXXC^{293} (Figure S1.3; topology prediction depicted in Figure S2). The 3D-protein structures of HsSLC5A8 (24% of identity) and CjSLC5 (10%) were predicted based on the crystal structure of a sodium/sugar symporter (PDB 3DH4) of *Vibrio parahaemolyticus* [64]. Docking studies with succinate and citrate identified putative critical residues involved in the establishment of strong interactions with ligands (Figure 13, Tables S3.6 and S3.7). Two conserved residues were identified as interacting with substrates: the F222 in TMS6 (a putative binding-site for succinate and citrate) and the L299, at the end of TMS8 (a putative binding-site for citrate).

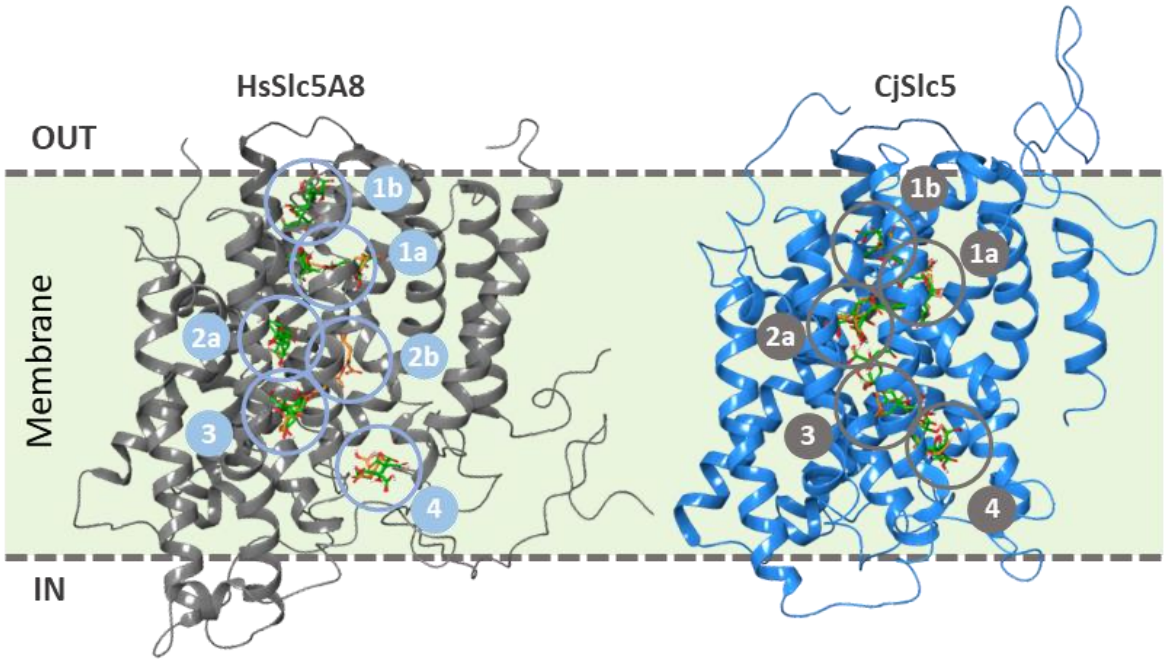


Figure 13. Predicted 3D structure of the *Homo sapiens* SLC5 member 8 and *C. jadinii* SLC5 homolog. Molecular docking of HsSlc5A8 and CjSlc5 3D-models, based on the *Vibrio parahaemolyticus* crystal structure of a sodium/galactose symporter (vSGLT; PDB 3DH4) with succinate (orange ligand) and citrate (green ligand).

5.3.6.4 *IN SILICO* STRUCTURAL ANALYSIS OF SLC13 TRANSPORTER

The alignment of *Homo sapiens* SLC13 member 3 protein with the CjSLC13 transporter (Figure S1.4; topology prediction depicted in Figure S2) revealed the presence of three conserved motifs as follows: in the TMS2, ^{420}TEA[I/L]PLXVTXLL^{431}, in the TMS4, ^{515}FLSMW[I/L]SNXAS^{525} and ^{706}SGL^{708} in the TMS9 [65]. The

crystal structure of the dicarboxylate antiporter *Lactobacillus acidophilus* INDY-transporter (PDB 6WTW) [18] belonging to the DASS family (TC 2.A.47) was used as template to predict the 3D structure of HsSLC13A3 (22% of identity) and CjSLC13 (17% of identity) proteins (Figure 14). Modelling studies uncovered four conserved residues located in beginning of TMS3, Q462 and T467; K701 in TMS9 and T746 in TMS10 involved in citrate binding. None of the aforementioned residues identified by the double-alignment and by the molecular docking analysis, belong to the conserved motifs identified in the TMS (see table S3.8 with residues).

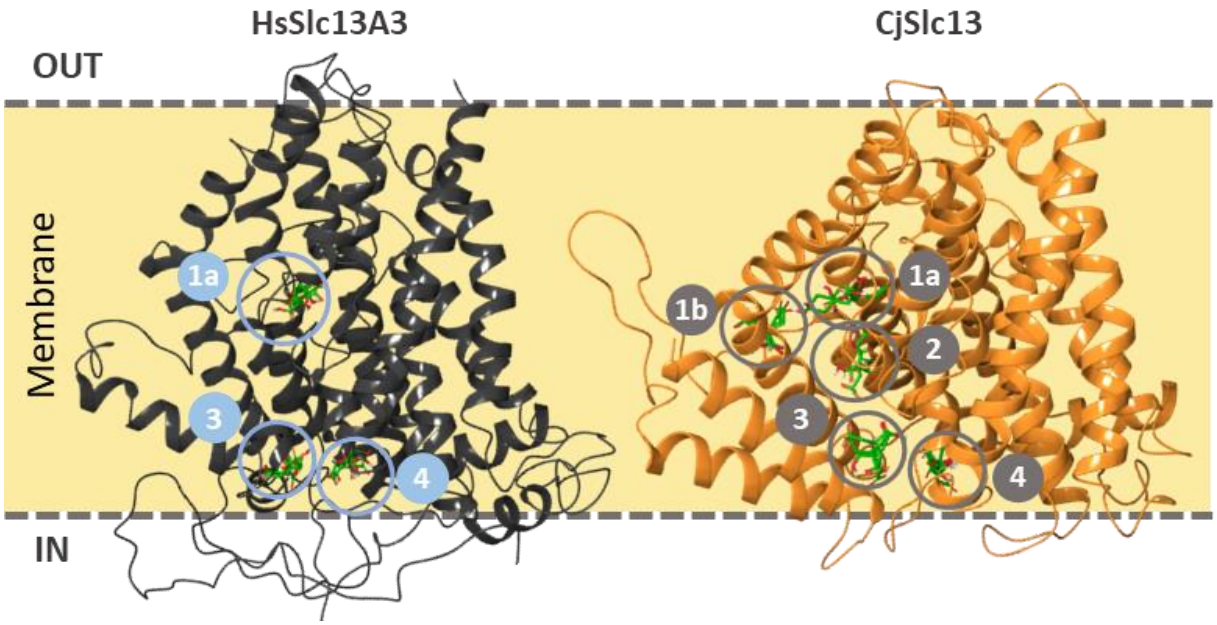


Figure 14. Predicted 3D structure of the *Homo sapiens* SLC13 member 3 and CjSLC13 homolog. Molecular docking of HsSlc13A3 and CjSlc13 3D-models, based on the *Lactobacillus acidophilus* crystal structure of INDY-transporter (dicarboxylate antiporter, belonging to DASS family) (PDB 6WTW) with citrate (green ligand).

The predicted binding energy of citrate in both Slc13 protein templates was similar (table S3.9), supporting the 3-D structural models of these transporters since *C. jadinii* SLC13 and HsSlc13A3 (also known as NaDC3) are citrate transporters [65].

5.4 Discussion

5.4.1 The *C. jadinii* transportome

To uncover a new transporter in a given organism, diverse strategies can be used. One is to search for homologs of already known transporters. Throughout evolution, homologs may retain their activity in the new host or developed new characteristics [41, 49, 66]. This method can rely on bioinformatic approaches for genome-wide scan or in transcriptomic analysis to pinpoint transporters specific for molecules of interest [8].

In this study, we analyzed the transportome of *Cyberlindnera jadinii* and identified several carboxylate transporters belonging to distinct transporter families. The physiological characterization of the *C. jadinii* DSM 2361 strain revealed the existence of mediated transport systems for acetate, lactate, succinate, and citrate. Comparing with the previous studies reported in the C. *jadinii* IGC/PYCC 3092 strain [19-21] differences in the kinetic parameters were found [67], which can be due to the distinct growth conditions used, namely pH, growth temperature, and the substrate concentration in the culture media [68, 69]. The genetic background is also a source of variation since the two *C. jadinii* strains with the full genome sequenced present differences on genome size and in the number of predicted coding genes [67].

The complete *C. jadinii* genomes available in the NCBI database enabled the analysis of the predicted transportome of this species. The inferred *C. jadinii* proteome was determined with BlastP against complete genomes available in the NCBI database and data available at the Saccharomyces Genome Database. Additionally, the analysis was complemented with topology prediction provided by the TMHMM software. This study focused on proteins displaying four to twenty predicted TMSs, as they are more likely to include membrane transporters [5, 70]. These constituted 9% of the total inferred proteome, being classified into major categories such as transporters, enzymes, and receptors. The proteins with unknown function accounted for 2% of the inferred protein coding genes. Our findings are in accordance with previous studies analyzing the repertoire of transporters encoded within the complete genome of a yeast species [71]. A probable function was assigned to ~75% of the putative proteins identified (see description of analyzed genes in the supplementary data, table S2).

To identify the carboxylate transporters previously reported for the C. *jadinii* IGC/PYCC 3092 strain [19-21], and other present in this yeast, we searched for putative carboxylic acid as well as amino acid transporters [72].

Besides the well characterized yeast carboxylate transporters belonging to the AceTr and SHS families, SLC superfamily members, SLC5, SLC16, SLC13 and SLC22, display a broad affinity for carboxylic acids

[73, 74]. For example, SLC5 members possesses a solute-binding domain associated with the transport of small solutes, such as sugars, vitamins, amino acids, monocarboxylates and smaller organic ions [75]. Members of the SLC13 family are reported as sodium/dicarboxylate cotransporters, namely the NaDC1-type present in the renal proximal tubule (apical membrane) and at the small intestine in human cells, that is able to translocate Krebs cycle intermediates as succinate, citrate and α-ketoglutarate across plasma membranes [76].

The heterologous expression of the CjAto, CjJen, CjTDT and CjSlc in *S. cerevisiae* strains improved cell growth on carboxylic acids, suggesting their role as carboxylate transporters (Figure 3). The five CjAto transporters promoted cell growth in acetic and lactic acid as sole carbon and energy source. Surprisingly, cells expressing CjAto2 and CjAto5 also presented improved growth on dicarboxylates (malic and succinic acids) and CjAto5 on tricarboxylates (citrate). CjAto5p presents an affinity for succinate similar to the Jen homologs: Jen2 from *C. albicans* (K_m 0.49 ± 0.27 mM) DH18 (K_m 0.31 mM) and DH24 (K_m 0.16 mM) from *D. hansenii,* and Jen2 (K_m 0.11 mM) from *K. lactis* [5, 50]. Recently Alves *et al.* (2020) reviewed the evolutionary roots of orthologues of ScAto transporters present in twelve *Candida* species, however these were not yet functionally characterized [43]. Besides ScAto1, *Y. lipolytica* Gpr1 is the only AceTr yeast member functionally characterized as an acetate transporter [30]. The bacterial homolog Satp from *Escherichia coli* is able to transport acetate and succinate [77].

S. cerevisiae IMX1000 cells expressing CjJen1-6 homologs promoted cell growth on acetic and lactic acids corroborating the already characterized monocarboxylate specificities for ScJen1 [13, 24]. The CjJen1 and CjJen2 also promoted improved cell growth on pyruvate. Besides improving growth in monocarboxylates, cells expressing CjJen6 grew in all di- and tricarboxylates used as sole carbon and energy sources. A similar growth pattern was uncovered upon CjSlc5 and CjSlc13-1 expression in *S. cerevisiae*, that also presented an improved growth in all carbon sources tested (lactic, pyruvic, fumaric, malic, succinic and citric acid), being the phenotype more evident in CjSlc5 transformant. The evaluation of kinetic parameters for the uptake of mono-, di- and tricarboxylates confirmed the role of CjSlc5 as a carboxylate transporter. The Slc5 members of the SSS family are reported as monocarboxylate transporters [75, 78, 79]. The CjSlc13-1 function is also in accordance with data reported for the DASS family (SLC13) members found in rabbit, rat and humans, described as sodium/dicarboxylate cotransporters with affinity for succinate, citrate and α-ketoglutarate across the cell plasma membranes [76, 80]. With our study, the *C. jadinii* proteins homologs of the SSS and DASS family members revealed to be functional monocarboxylate, dicarboxylate and tricarboxylate transporters in *S. cerevisiae* yeast.

A mild toxic phenotype was observed in *S. cerevisiae* IMX1000 strain expressing the CjSlc13-2 in media containing lactic and citric acids as sole carbon and energy sources. A similar phenotype was found for CjJen5p transformant on succinic acid. These observations can be an indication of its role as carboxylate exporter, however, no experimental assays were performed to confirm such hypothesis.

5.4.2 Unravelling novel citrate transporters in yeasts

The industrial bioproduction of citric acid started in the 19th century using *Aspergillus niger* as workhorse [7, 81]. Since then, the market for this acid had a predicted annual growth of 3.7% untill 2020 due to its broad range of applications [7, 82]. Despite this long-lasting microbial production, until recently it was unclear how citrate could be exported from the cytosol to the extracellular medium. Recent studies uncovered three fungal plasma membrane citrate transporters, the CexA citrate exporter from *A. niger* [83, 84], the Cex1 citrate exporter from *Y. lipolytica* [85] and the PkJEN2-2 importer from *Pichia kudriavzevii* [52], but no data is available for the kinetic parameters of these transporters [84, 85].

In this work, we have found three potential citrate importers in *C. jadinii*, as the expression of CjAto5, CjJen6 and CjSlc5 in *S. cerevisiae* enabled it to metabolize citric acid as sole carbon and energy source. The comparison of the kinetic parameters of these three transformant strains (Figure 6) evidences that CjSlc5 displays the higher capacity and lower affinity for citrate. An intermediate behavior was found for CjAto5 and the CjJen6p encodes the lower capacity transporter for citrate. Considering the kinetic parameters for citrate uptake presented by these transporters, the gene encoding the high-affinity system characterized in *C. jadinii* citrate grown-cells remains unidentified [19] .

Regardless the fact that the present work only measured the import of carboxylates, previous reports demonstrate that carboxylate permeases can work both as importers and exporters of these compounds [10, 86]. In summary, CjSlc5p, CjAto5p and CjJen6 can be regarded as general carboxylate permeases capable of recognizing a wide range of substrates (mono-, di-, tricarboxylates).

5.4.3 Phylogenetic roots of the *C. jadinii* carboxylate transporters

The phylogenetic reconstruction of Ato homologs suggests a that CjAto5p is distant from the other five CjAtos, that are clustered together in a distinct clade (Figure 7-E1, Ato1/Ato2 cluster). CjAto5p shares the same clade with Ato3 from *S. cerevisiae* [44], which function remains unidentified. Nonetheless, the functional analysis of several AceTr family members from bacteria, archaea and fungi evidenced their role as carboxylate transporters, including Ato1 and Gpr1 from *S. cerevisiae* and *Y. lipolytica* yeasts.

The analysis of all the Jen1 homologs so far sequenced, proposes a basal split between fungi, bacteria (high prevalence in Proteobacteria) and archaea, forming three monophyletic clades. *C. jadinii* SHS-homologs were found in the Jen1 and Jen2 clusters, where other functionally characterized homologs from *S. cerevisiae*, *D. hansenii*, *C. albicans* and *Y. lipolytica* are also included (Figure 8-E1). CjJen1-4 homologs are found in Jen1-cluster together with other yeast mono- and dicarboxylate transporters, as the DH17p (malate transporter) and DH27p (acetate transporter) [49], the ScJen1 [13, 24] and the CaJen1 [48]. As for CjJen5 and CjJen6, they present a closer phylogenetic relationship with succinate transporters DH24 from *D. hansenii* [49], KlJen2 from *K. lactis* [26, 41, 87] and the CaJen2, a dicarboxylate and sugar acid permease from *C. albicans* [26, 50] than to the succinate transporter DH18 from *D. hansenii* [49] or the dicarboxylate and di/tricarboxylate permeases PkJen2-1 and PkJen2-2 from *Pichia kudriavzevii* [52]. The divergence of CjJen homologs, particularly for CjJen6, supports the distinct growth phenotype behavior of this transporter that presents a wide substrate range, including citrate. Nevertheless, despite being phylogenetically close, the specificity of CjJen5p and CjJen6p is quite distinct.

The phylogenetic reconstruction of the CjSLC5 homologs only identified members of the eukaryotic domain of life, namely ascomycetes and basidiomycetes. A high prevalence of homologs is found in *Fusarium* species and none was found in *S. cerevisiae*. The majority of SLC5 homologs are annotated as putative neutral amino acid transporters. This fact may be explained by the existence of two homologs functionally characterized as amino acid transporters, the Mtr/AAP1 from *N. crassa* [54] and PcMtr from *P. chrysogenum* [55] (Figure 9, E1 clade). These data suggest that CjSlc5 may also behave as an amino acid transporter, withal no functional studies were made to confirm the hypothesis. Cassio *et al.* (1993) reported the presence of a facilitated diffusion system in *C. jadinii* that is likely to operate as a general organic permease, accepting mono-, di-, and tricarboxylates as well as amino acids, as glycine and glutamic acid [21].

In the CjSLC13-tree, homologs are found in all domains of life, including bacteria, eukarya and archaea, with a strong prevalent of prokaryotic homologs. The two CjSlc13 homologs are distributed in two distinct branches of the eukaryotic clade (Figure 10, E1) and share phylogenetic relationship with Pho87 and Pho90 *S. cerevisiae* transporters [56, 57] suggesting they could also transport phosphate.

5.4.4 Structural features of the *C. jadinii* carboxylate transporters

To obtain insights on the structural-functional properties of the different transporters, we followed several strategies that included multiple-sequence alignments with characterized carboxylate transporters,

molecular docking studies and analysis of protein pore size variations. The alterations in the highly conserved L93 residue of the signature NPAPLGL(M/S) motif of the AceTr family [30] observed in CjAto5 (M85) may play a relevant role on its broader substrates specificity. Also in CjAto5, the substitution of conserved G/A-97 to a Serine located in TMS1 near the narrowest constriction site formed by F17, Y72 and L131 residues [60] might influence transporter function[61]. The docking analysis uncovered a residue located at the end of TMS5 as a binding-site for succinate present in CjAto2 (S211) and CjAto5 (K221). According to the literature, these CjAto homologs are the first described in yeast with affinity for dicarboxylates [30]. In this regard, the latter exposed residues will need to be further explored by site-directed mutagenesis to verify their role on the protein function. In addition, besides dicarboxylates, CjAto5 also accepts tricarboxylic acids. From our analysis the non-conserved residues Q103, K230 and D253, located in TMS1 and TMS6, may act as putative binding-sites for citrate. As for CjAto6p, functional studies are needed to unveil the transporter specificity.

As for Jen protein members, docking results highlighted two residues in CjJen6p that were not conserved across all Jen members. On TMS1, W154 (in ScJen1) was substituted for L65 in CjJen6p and that is also present in the succinate transporters CaJen2p [26, 50] and KlJen2p [26, 41, 87]. In addition, in TMS7 a highly conserved leucine is substituted for valine in residue 260 of CjJen6p.

Considering the *C. jadinii* SLC5 member, the F222 located in the TMS6, is predicted to be involved in succinate and citrate binding, and the L299 present in the end of TMS8 seems to be specific for citrate binding. To infer more accurately the role of these residues in substrate uptake, a site-directed mutagenesis approach will be necessary.

In this study, further insights were also provided for the DASS family members. The conserved motif hereby identified, ^{420}TEA[I/L]PLXVTXLL^{431} located in TMS2, is also conserved in other SLC13 members that included the dicarboxylate transporter from *Vibrio cholerae* (VcINDY) [65]. Also in the highly conserved ^{515}FLSMW[I/L]SNXAS^{525}motif located in TMS4, the S521 and N522 residues were reported as relevant for carboxylate and sodium binding [65]. The *Homo sapiens* SLC13 members, belonging to Na^+-di- and tricarboxylate cotransporters (NaDC), are involved in the transport of succinate, citrate, and α-ketoglutarate across cell membranes in kidney, intestine, placenta, liver, brain and testis [16]. The citrate binding-affinity values obtained in the molecular docking studies were similar for both protein templates which supports the 3D models of these transporters.

5.5 Conclusions

In the present study, novel *C. jadinii* carboxylate transporters belonging to a total of six transporter families with specificity for a set of organic acids with biotechnological relevance were uncovered and analyzed. Two yeast protein members from the AceTr family able to transport succinate, CjAto2, as well as citrate, CjAto5p were identified. One member from SHS transporter family, CjJen6, was the first member described as a broad range carboxylate transporter, e.g. mono-, di- and tricarboxylates. The two homologs from the SSS (SLC member 5) and DASS (SLC member 13) transporter families also showed ability to transport mono-, di- and tricarboxylates. To our knowledge, these are the first members from these families characterized in yeast as carboxylate transporters.

With this study novel directions were provided for a possible application of such membrane transporters as an additional driving force for the bioproduction of carboxylates.

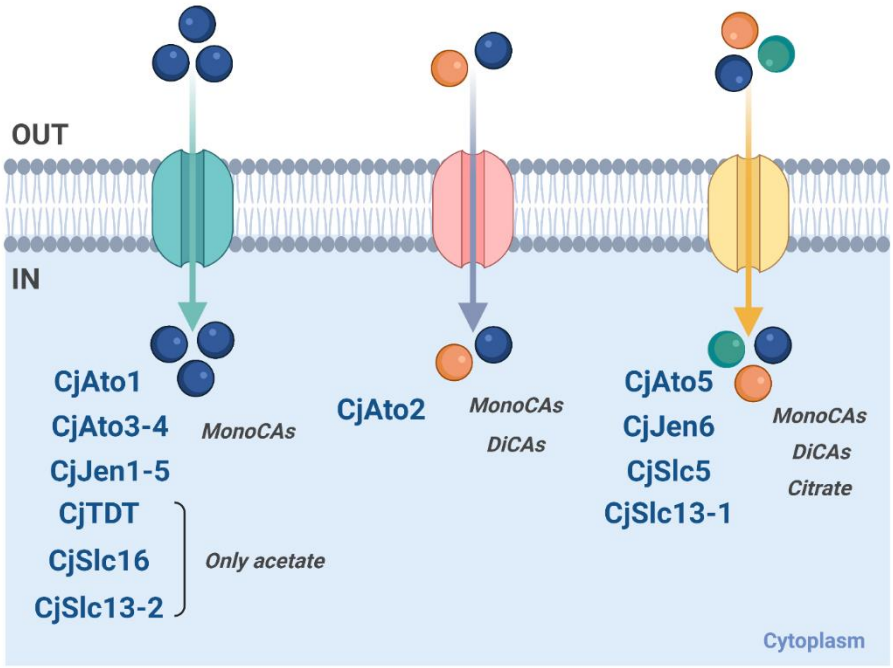


Figure 15. *C. jadinii* plasma membrane transporters functionally characterized within this work by heterologous expression in *S. cerevisiae* strains. Color circles represent the specificities uncovered for each of the protein-system: monocarboxylic acids — blue filled circle, dicarboxylic acids – orange filled circle, and tricarboxylic acids—green filled circle. Initials stand for IN – intracellular space; OUT – extracellular space.

To evaluate the contribution of each of these lactate/succinate/citrate transporters to the overall transport capacity in *C. jadinii*, single and multiple deletions of the respective genes are necessary to verify a possible loss of carboxylate transport activity. Such deletions can be performed using the CRISPR/Cas system that was recently described in a patent application for the knock out and insert of exogenous

genes in the *C. jadinii* ATCC 22023 strain [88]. Additionally, the evaluation of gene expression by RT-PCR will unveil the expression patterns of each transporter.

To assess the role of *C. jadinii* citrate transporters as possible exporters, engineered *S. cerevisiae* citrate-producer strains will be used as expression hosts. Bioreactor cultivations coupled with metabolite quantification in the extracellular media will allow the evaluation of their export capacity. Further studies are also needed to clarify the proton-citrate stoichiometry and determine which of the citrate anion form is being transported.

The expression of these membrane transporters in carboxylate producing cell factories defines one of the crucial steps to develop new and more efficient strains working as biosustainable platforms in the production of biobased compounds. Overall, this work is a contribution to the knowledge on plasma membrane transporter proteins that can allow the development of improved bioprocesses with an impact in industrial biotechnology.

5.6 REFERENCES

1. Boze, H., G. Moulin, and P. Galzy, *Production of food and fodder yeasts.* Critical reviews in biotechnology, 1992. **12**(1-2): p. 65-86.
2. Bekatorou, A., C. Psarianos, and A.A. Koutinas, *Production of food grade yeasts.* Food Technology and Biotechnology, 2006. **44**(3): p. 407-415.
3. Buerth, C., D. Tielker, and J.F. Ernst, *Candida utilis and Cyberlindnera (Pichia) jadinii: yeast relatives with expanding applications.* Applied Microbiology and Biotechnology, 2016. **100**(16): p. 6981-6990.
4. Kurtzman, C.P., C.J. Robnett, and E. Basehoar-Powers, *Phylogenetic relationships among species of Pichia, Issatchenkia and Williopsis determined from multigene sequence analysis, and the proposal of Barnettozyma gen. nov., Lindnera gen. nov. and Wickerhamomyces gen. nov.* FEMS yeast research, 2008. **8**(6): p. 939-954.
5. Soares-Silva, I., et al., *Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications.* FEMS microbiology letters, 2020.
6. Kell, D., et al., *Membrane transporter engineering in industrial biotechnology and whole cell biocatalysis.* Trends in biotechnology, 2015. **33**(4): p. 237-246.
7. Sauer, M., et al., *Microbial production of organic acids: expanding the markets.* Trends in biotechnology, 2008. **26**(2): p. 100-108.
8. Boyarskiy, S. and D. Tullman-Ercek, *Getting pumped: membrane efflux transporters for enhanced biomolecule production.* Current opinion in chemical biology, 2015. **28**: p. 15-19.
9. Zelle, R., et al., *Malic acid production by Saccharomyces cerevisiae: engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export.* Applied and Environmental Microbiology, 2008. **74**(9): p. 2766-2777.
10. Pacheco, A., et al., *Lactic acid production in Saccharomyces cerevisiae is modulated by expression of the monocarboxylate transporters Jen1 and Ady2.* FEMS Yeast Res, 2012. **12**(3): p. 375-381.
11. Lee, T., I. Paulsen, and P. Karp, *Annotation-based inference of transporter function.* Bioinformatics, 2008. **24**(13): p. i259-i267.
12. Mishra, N., J. Chang, and P. Zhao, *Prediction of membrane transport proteins and their substrate specificities using primary sequence information.* PLoS One, 2014. **9**(6): p. e100278.
13. Casal, M., et al., *The lactate-proton symport of Saccharomyces cerevisiae is encoded by JEN1.* Journal of bacteriology, 1999. **181**(8): p. 2620-2623.
14. Paiva, S., et al., *Ady2p is essential for the acetate permease activity in the yeast Saccharomyces cerevisiae.* Yeast, 2004. **21**(3): p. 201-210.
15. Pacheco, A., et al., *Lactic acid production in Saccharomyces cerevisiae is modulated by expression of the monocarboxylate transporters Jen1 and Ady2.* FEMS Yeast Res, 2012. **12**(3): p. 375-81.
16. Bergeron, M., et al., *SLC13 family of Na+-coupled di-and tri-carboxylate/sulfate transporters.* Molecular aspects of medicine, 2013. **34**(2-3): p. 299-312.

17. Bonetti, S., S. Hirschi, and P. Bosshart, *Expression, purification and crystallization of an SLC16 monocarboxylate transporter family homologue specific for l-lactate.* Protein expression and purification, 2020. **165**: p. 105484.
18. Sauer, D., et al., *Structural basis for the reaction cycle of DASS dicarboxylate transporters.* Elife, 2020. **9**: p. e61350.
19. Cássio, F. and C. Leão, *Low-and high-affinity transport systems for citric acid in the yeast Candida utilis.* Appl. Environ. Microbiol., 1991. **57**(12): p. 3623-3628.
20. Leão, C. and N. Van Uden, *Transport of lactate and other short-chain monocarboxylates in the yeast Candida utilis.* Applied microbiology and biotechnology, 1986. **23**(5): p. 389-393.
21. Cássio, F. and C. Leão, *A comparative study on the transport of L (-) malic acid and other short-chain carboxylic acids in the yeast Candida utilis: Evidence for a general organic acid permease.* Yeast, 1993. **9**(7): p. 743-752.
22. Cássio, F., M. Côrte-Real, and C. Leão, *Quantitative analysis of proton movements associated with the uptake of weak carboxylic acids. The yeast Candida utilis as a model.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 1993. **1153**(1): p. 59-66.
23. Roush, A., L. Questiaux, and A. Domnas, *The active transport and metabolism of purines in the yeast, Candida utilis.* Journal of Cellular and Comparative Physiology, 1959. **54**(3): p. 275-286.
24. Soares-Silva, I., et al., *The conserved sequence NXX [S/T] HX [S/T] QDXXXT of the lactate/pyruvate: H+ symporter subfamily defines the function of the substrate translocation pathway.* Molecular membrane biology, 2007. **24**(5-6): p. 464-474.
25. Mans, R., et al., *A CRISPR/Cas9-based exploration into the elusive mechanism for lactate export in Saccharomyces cerevisiae.* FEMS Yeast Res, 2017. **17**(8).
26. Ribas, D., et al., *Yeast as a tool to express sugar acid transporters with biotechnological interest.* FEMS Yeast Res, 2017. **17**(2).
27. Lõoke, M., K. Kristjuhan, and A. Kristjuhan, *Extraction of genomic DNA from yeasts for PCR-based applications.* Biotechniques, 2011. **50**(5): p. 325-328.
28. Mumberg, D., R. Müller, and M. Funk, *Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds.* Gene, 1995. **156**(1): p. 119-122.
29. Sambrook, J., E. Fritsch, and T. Maniatis, *Molecular cloning: a laboratory manual*1989: Cold spring harbor laboratory press.
30. Ribas, D., et al., *The acetate uptake transporter family motif "NPAPLGL (M/S)" is essential for substrate uptake.* Fungal Genetics and Biology, 2019. **122**: p. 1-10.
31. Katoh, K., J. Rozewicki, and K. Yamada, *MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization.* Briefings in Bioinformatics, 2017. **20**(4): p. 1160-1166.
32. Kumar, S., G. Stecher, and K. Tamura, *MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets.* Molecular biology and evolution, 2016. **33**(7): p. 1870-1874.
33. Wallace, I.M., et al., *M-Coffee: combining multiple sequence alignment methods with T-Coffee.* Nucleic acids research, 2006. **34**(6): p. 1692-1699.
34. Zimmermann, L., et al., *A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core.* Journal of molecular biology, 2018. **430**(15): p. 2237-2243.

35. Zheng, W., et al., *LOMETS2: improved meta-threading server for fold-recognition and structure-based function annotation for distant-homology proteins.* Nucleic acids research, 2019. **47**(W1): p. W429-W436.
36. Irwin, J.J. and B.K. Shoichet, *ZINC– A free database of commercially available compounds for virtual screening.* Journal of chemical information and modeling, 2005. **45**(1): p. 177-182.
37. Pettersen, E., et al., *UCSF Chimera—a visualization system for exploratory research and analysis.* Journal of computational chemistry, 2004. **25**(13): p. 1605-1612.
38. Dallakyan, S. and A. Olson, *Small-molecule library screening by docking with PyRx*, in *Chemical biology*2015, Springer. p. 243-250.
39. Smart, O., et al., *HOLE: a program for the analysis of the pore dimensions of ion channel structural models.* Journal of molecular graphics, 1996. **14**(6): p. 354-360.
40. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics.* Journal of molecular graphics, 1996. **14**(1): p. 33-38.
41. Queirós, O., et al., *Functional analysis of Kluyveromyces lactis carboxylic acids permeases: heterologous expression of KlJEN1 and KlJEN2 genes.* Current Genetics, 2007. **51**(3): p. 161-169.
42. Danhof, H. and M. Lorenz, *The Candida albicans ATO gene family promotes neutralization of the macrophage phagolysosome.* Infection and immunity, 2015. **83**(11): p. 4416-4426.
43. Alves, R., et al., *Carboxylic Acid Transporters in Candida Pathogenesis.* MBio, 2020. **11**(3): p. e00156-20.
44. Palková, Z., et al., *Ammonia pulses and metabolic oscillations guide yeast colony development.* Molecular biology of the cell, 2002. **13**(11): p. 3901-3914.
45. Robellet, X., et al., *AcpA, a member of the GPR1/FUN34/YaaH membrane protein family, is essential for acetate permease activity in the hyphal fungus Aspergillus nidulans.* Biochemical journal, 2008. **412**(3): p. 485-493.
46. Sá-Pessoa, J., et al., *Expression and specificity profile of the major acetate transporter AcpA in Aspergillus nidulans.* Fungal Genetics and Biology, 2015. **76**: p. 93-103.
47. Flipphi, M., et al., *Functional analysis of alcS, a gene of the alc cluster in Aspergillus nidulans.* Fungal Genetics and Biology, 2006. **43**(4): p. 247-260.
48. Soares-Silva, I., et al., *The disruption of JEN1 from Candida albicans impairs the transport of lactate.* Molecular membrane biology, 2004. **21**(6): p. 403-411.
49. Soares-Silva, I., et al., *The Debaryomyces hansenii carboxylate transporters Jen1 homologues are functional in Saccharomyces cerevisiae.* FEMS Yeast Res, 2015. **15**(8).
50. Vieira, N., et al., *Functional specialization and differential regulation of short-chain carboxylic acid transporters in the pathogen Candida albicans.* Molecular Microbiology, 2010. **75**(6): p. 1337-1354.
51. Dulermo, R., et al., *The evolution of Jen3 proteins and their role in dicarboxylic acid transport in Yarrowia.* Microbiologyopen, 2015. **4**(1): p. 100-120.
52. Xi, Y., et al., *Characterization of JEN family carboxylate transporters from the acid-tolerant yeast Pichia kudriavzevii and their applications in succinic acid production.* Microbial biotechnology, 2021.

53. Niedźwiecka, K., et al., *The Cryptococcus neoformans monocarboxylate transporter Jen4 is responsible for increased 3-bromopyruvate sensitivity.* FEMS Yeast Res, 2019. **19**(3): p. foz029.
54. Koo, K. and W. Stuart, *Sequence and structure of mtr, an amino acid transport gene of Neurospora crassa.* Genome, 1991. **34**(4): p. 644-651.
55. Trip, H., M. Evers, and A. Driessen, *PcMtr, an aromatic and neutral aliphatic amino acid permease of Penicillium chrysogenum.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 2004. **1667**(2): p. 167-173.
56. Hürlimann, H., et al., *The SPX domain of the yeast low-affinity phosphate transporter Pho90 regulates transport activity.* EMBO reports, 2009. **10**(9): p. 1003-1008.
57. Ghillebert, R., et al., *Differential roles for the low-affinity phosphate transporters Pho87 and Pho90 in Saccharomyces cerevisiae.* Biochemical journal, 2011. **434**(2): p. 243-251.
58. Hürlimann, H., et al., *Pho91 is a vacuolar phosphate transporter that regulates phosphate and polyphosphate metabolism in Saccharomyces cerevisiae.* Molecular biology of the cell, 2007. **18**(11): p. 4438-4445.
59. Strickler, M., et al., *Functional characterization of a Na^+-coupled dicarboxylate transporter from Bacillus licheniformis.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 2009. **1788**(12): p. 2489-2496.
60. Qiu, B., et al., *Succinate-acetate permease from Citrobacter koseri is an anion channel that unidirectionally translocates acetate.* Cell Research, 2018. **28**(6): p. 644-654.
61. Betts, M. and R. Russell, *Amino acid properties and consequences of substitutions.* Bioinformatics for geneticists, 2003. **317**: p. 289.
62. Augstein, A., et al., *Characterization, localization and functional analysis of Gpr1p, a protein affecting sensitivity to acetic acid in the yeast Yarrowia lipolytica.* Microbiology, 2003. **149**(3): p. 589-600.
63. Bosshart, P., et al., *Mechanistic basis of L-lactate transport in the SLC16 solute carrier family.* Nat Commun, 2019. **10**(1): p. 1-11.
64. Faham, S., et al., *The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na+/sugar symport.* Science, 2008. **321**(5890): p. 810-814.
65. Schlessinger, A., et al., *Determinants of substrate and cation transport in the human Na+/dicarboxylate cotransporter NaDC3.* Journal of Biological Chemistry, 2014. **289**(24): p. 16998-17008.
66. Niimi, M., et al., *Functional analysis of fungal drug efflux transporters by heterologous expression in Saccharomyces cerevisiae.* Jpn J Infect Dis, 2005. **58**(1): p. 1-7.
67. Sousa-Silva, M., et al., *Expanding the Knowledge on the Skillful Yeast Cyberlindnera jadinii.* Journal of Fungi, 2021. **7**(1): p. 36.
68. Casal, M., et al., *Transport of carboxylic acids in yeasts.* FEMS microbiology reviews, 2008. **32**(6): p. 974-994.
69. Diallinas, G., *Understanding transporter specificity and the discrete appearance of channel-like gating domains in transporters.* Frontiers in pharmacology, 2014. **5**: p. 207.

70. Paulsen, I., M. Sliwinski, and M. Saier Jr, *Microbial genome analyses: global comparisons of transport capabilities based on phylogenies, bioenergetics and substrate specificities.* Journal of molecular biology, 1998. **277**(3): p. 573-592.
71. Paulsen, I., et al., *Unified inventory of established and putative transporters encoded within the complete genome of Saccharomyces cerevisiae.* FEBS letters, 1998. **430**(1-2): p. 116-125.
72. Letunic, I., et al., *SMART 4.0: towards genomic data integration.* Nucleic acids research, 2004. **32**(suppl_1): p. D142-D144.
73. Colas, C., P. Ung, and A. Schlessinger, *SLC Transporters: Structure, Function, and Drug Discovery.* MedChemComm, 2016. **7**(6): p. 1069-1081.
74. Song, W., et al., *Solute carrier transporters: the metabolic gatekeepers of immune cells.* Acta Pharmaceutica Sinica B, 2020. **10**(1): p. 61-78.
75. Gyimesi, G., et al., *Sodium-coupled glucose transport, the SLC5 family, and therapeutically relevant inhibitors: from molecular discovery to clinical application.* Pflügers Archiv - European Journal of Physiology, 2020. **472**(9): p. 1177-1206.
76. Pajor, A., *Molecular properties of the SLC13 family of dicarboxylate and sulfate transporters.* Pflugers Archiv : European journal of physiology, 2006. **451**(5): p. 597-605.
77. Sá-Pessoa, J., et al., *SATP (YaaH), a succinate–acetate transporter protein in Escherichia coli.* Biochemical journal, 2013. **454**(3): p. 585-595.
78. Srinivas, S., et al., *Cloning and functional identification of slc5a12 as a sodium-coupled low-affinity transporter for monocarboxylates (SMCT2).* Biochemical journal, 2005. **392**(3): p. 655-664.
79. Ganapathy, V., et al., *Sodium-coupled Monocarboxylate Transporters in Normal Tissues and in Cancer.* The AAPS Journal, 2008. **10**(1): p. 193.
80. Pajor, A., *Sodium-coupled dicarboxylate and citrate transporters from the SLC13 family.* Pflügers Archiv-European Journal of Physiology, 2014. **466**(1): p. 119-130.
81. Rehm, H., et al., *Biotechnology. Vol. 9: Enzymes, Biomass, Food and Feed: A Multi-volume Comprehensive Treatise*1995: VCH.
82. Sahu, Y., *Organic acids market by type (Acetic Acid, Citric Acid, Formic Acid, Lactic Acid, Itaconic Acid, Succinic Acid, Gluconic Acid, Ascorbic Acid, Fumaric Acid, and Propionic Acid), Source (Biomass, Molasses, Starch, Chemical Synthesis, Agro-Industrial Residue)*, in *Global Opportunity Analysis and Industry Forecast*, A.F. End-User (Food & Beverage, Chemicals & Industrial, Pharmaceuticals, Personal Care, Agriculture), Editor 2017.
83. Odoni, D., et al., *Aspergillus niger citrate exporter revealed by comparison of two alternative citrate producing conditions.* FEMS microbiology letters, 2019. **366**(7): p. fnz071.
84. Steiger, M.G., et al., *Engineering of the citrate exporter protein enables high citric acid production in Aspergillus niger.* Metabolic engineering, 2019. **52**: p. 224-231.
85. Erian, A.M., et al., *Identification of the citrate exporter Cex1 of Yarrowia lipolytica.* FEMS Yeast Res, 2020.
86. Camarasa, C., et al., *Characterization of Schizosaccharomyces pombe malate permease by expression in Saccharomyces cerevisiae.* Applied and Environmental Microbiology, 2001. **67**(9): p. 4144-4151.

87. Lodi, T., et al., *Carboxylic acids permeases in yeast: two genes in Kluyveromyces lactis.* Gene, 2004. **339**: p. 111-119.
88. Zehuan, L., L. Jianghai, and D. Yu, *Method for verifying feasibility of CRISPR-Cas9 system for knocking out Candida utilis target gene*, 2019: China.

SUPPLEMENTARY MATERIAL

Table S1. Percentage of shared amino acids between A) Ato1 homologs of *C. jadinii, S. cerevisiae, Y. lipolytica, A. nidulans, E. coli* and *M. acetivorans* and B) Jen1 homologs of *C. jadinii, S. cerevisiae, Y. lipolytica, C. albicans, K. lactis* and *D. hansenii.*

(A)

	CjAto1	CjAto2	CjAto3	CjAto4	CjAto5	CjAto6	ScAto1	ScAto2	ScAto3	YlGpr1	AnAcpA	AnAcpB	AnAcpC	AnAlcS	EcSatP	Ma4008	Ma0103
CjAto1	100																
CjAto2	69.7	100															
CjAto3	78.8	72.1	100														
CjAto4	75.3	66.1	81.7	100													
CjAto5	41.1	38.4	43.3	41.0	100												
CjAto6	48.1	49.4	52.1	49.6	48.1	100											
ScAto1	59.0	55.0	61.5	60.0	41.4	50.6	100										
ScAto2	55.6	57.9	65.1	62.5	42.2	50.6	77.7	100									
ScAto3	40.1	33.2	40.7	39.2	40.2	33.3	33.5	34.7	100								
YlGpr1	47.7	52.5	51.7	51.7	39.4	50.4	51.4	52.2	31.5	100							
AnAcpA	49.8	46.5	53.5	51.0	35.8	42.8	50.4	46.9	37.6	46.0	100						
AnAcpB	45.5	44.8	48.3	45.9	33.9	41.8	39.2	41.1	36.4	40.5	55.9	100					
AnAcpC	32.0	34.1	35.2	34.1	29.8	32.7	36.5	35.0	28.4	33.8	36.5	31.6	100				
AnAlcS	31.0	29.7	29.0	28.1	23.6	28.7	33.1	30.6	25.8	28.5	33.2	24.8	29.5	100			
EcSatP	36.3	33.1	33.1	34.3	25.4	33.1	35.6	36.7	33.3	35.8	35.9	30.5	28.0	27.9	100		
Ma4008	34.7	35.5	34.1	35.5	29.9	34.8	37.9	37.9	30.5	34.6	34.4	32.0	36.2	30.4	62.4	100	
Ma0103	30.8	28.4	29.1	30.4	25.1	25.0	31.0	34.1	34.9	28.7	27.8	29.0	30.1	26.4	46.4	45.8	100

Dark blue: identity > 70%; Blue: identity 50-69 %; Dark green: identity 30-49%; Green: identity < 30%.

(B)

	CjJen1	CjJen2	CjJen3	CjJen4	CjJen5	CjJen6	ScJen1	CaJen1	CaJen2	KlJen1	KlJen2	Dh17	Dh18	Dh24	Dh27	YlJen1	YlJen2	YlJen3	YlJen4	YlJen5	YlJen6
CjJen1	100																				
CjJen2	66.4	100																			
CjJen3	42.2	40.5	100																		
CjJen4	44.8	42.4	72.5	100																	
CjJen5	44.1	42.7	41.5	44.0	100																
CjJen6	42.9	44.4	43.8	44.8	74.8	100															
ScJen1	58.2	51.7	40.5	41.9	43.0	40.6	100														
CaJen1	41.8	40.6	42.2	41.9	42.0	44.2	41.2	100													
CaJen2	43.0	40.8	38.2	40.1	51.7	55.7	38.3	43.8	100												
KlJen1	56.8	52.7	38.4	38.9	39.0	40.7	65.4	38.3	40.4	100											
KlJen2	37.9	38.6	41.6	45.6	51.8	55.3	37.9	39.8	48.9	36.9	100										
Dh17	46.4	43.6	44.4	48.3	47.2	49.8	43.9	55.5	45.3	42.4	41.6	100									
Dh18	44.3	39.5	42.7	42.7	58.2	57.1	38.1	41.8	50.0	38.5	51.6	46.4	100								
Dh24	45.0	42.8	42.6	44.0	55.5	57.1	40.7	42.2	57.5	37.0	55.8	48.3	55.4	100							
Dh27	44.1	42.9	44.5	44.9	46.9	45.1	41.0	54.4	43.4	42.3	40.6	64.1	43.3	47.3	100						
YlJen1	43.6	43.8	46.7	48.0	57.4	57.9	43.3	46.0	50.5	37.8	50.2	50.7	58.2	54.2	46.0	100					
YlJen2	45.6	43.7	43.4	46.5	54.1	56.7	42.1	45.9	49.7	38.1	52.4	52.0	58.4	52.8	47.1	73.8	100				
YlJen3	44.0	44.2	41.9	44.2	54.0	56.9	39.2	48.6	47.9	38.4	53.4	50.5	59.8	55.6	47.8	67.3	72.2	100			
YlJen4	44.3	43.1	43.5	46.0	57.5	59.2	40.4	47.2	49.9	39.4	52.7	53.4	62.6	56.0	48.7	74.8	78.9	71.5	100		
YlJen5	40.1	43.4	42.5	43.8	51.7	55.0	41.0	45.7	47.9	37.9	49.0	48.6	53.5	52.2	46.3	63.9	70.8	65.2	76.7	100	
YlJen6	43.3	41.2	42.5	43.4	47.0	50.2	37.4	41.2	44.2	36.4	47.2	45.0	49.1	54.8	42.9	62.8	58.9	59.2	61.7	56.9	100

Dark Red: identity > 90%; Red: identity 70-90%; Orange: identity 50-69%; Yellow: identity 30-49%; White: identity < 30%.

Table S2. Annotation of the inferred transportome of *C. jadinii* NRRL Y-1542 presenting 4 to 20 predicted transmembrane segments (TMSs): (i) annotated function; (ii) function revised after BlastP result with more than 40% identity; (iii) predicted localization, (iv) protein length (bp), (v) protein family and (vi) topology. In this study, partial proteins were removed for a more accurate analysis of *C. jadinii* proteome. Initials stand for: localization, PM – Plasma membrane, VA – Vacuole, ER – Endoplasmatic reticulum, GC – Golgi complex, N – nucleus, M – Mitochondria and for transporter families (listed above, DUF stands for domain unknown function).

Acession number	Annotated Function	Function revised (considered with BlastP result with 40% identity)	Predicted localization	Length (bp)	Family	Topology (TMSs)
XP_020069104.1	MFS general substrate transporter	Sugar transporter	PM	527	SP	10
XP_020073496.1	maltose permease	Maltose permease	PM	542	SP	12
XP_020072909.1	quinate permease-like protein	HXT - Sugar transporter	PM	539	HXT	9
XP_020071255.1	arabinose-proton symporter	arabinose-H$^+$ symporter	PM	598	MFS	8
XP_020068499.1	hypothetical protein CYBJADRAFT_131513	Polyol transporter	PM	567	HXT	10
XP_020067682.1	high-affinity inorganic phosphate transporter and low-affinity manganese transporter	Sugar phosphate permease; putative low-affinity manganese transporter	PM	532	MFS	11
XP_020067723.1	general substrate transporter	HXT - Sugar transporter	PM	535	HXT	12
XP_020073619.1	hypothetical protein CYBJADRAFT_14416	L-rhamnose-proton symporter	PM	352	RhaT	8
XP_020073488.1	general substrate transporter	HXT - Sugar transporter	PM	566	HXT	11
XP_020073390.1	hypothetical protein CYBJADRAFT_165648	High-affinity glucose transporter; glucose sensor	PM	768	SP	12
XP_020072691.1	hypothetical protein CYBJADRAFT_161053	Sugar phosphate permease	PM	474	MFS	12
XP_020072673.1	hypothetical protein CYBJADRAFT_147912	Sugar transporter	PM	616	SP	12
XP_020072427.1	hypothetical protein CYBJADRAFT_46559	nucleotide-sugar transporters	ER	367	SLC	8
XP_020072304.1	general substrate transporter	HXT - Sugar transporter	PM	517	HXT	12
XP_020072062.1	MFS general substrate transporter	glycerophosphoinositol and glycerophosphocholine transporter	PM	485	MFS	9
XP_020072019.1	general substrate transporter	Myo-inositol transporter	PM	548	SP	12
XP_020071539.1	hypothetical protein CYBJADRAFT_138033	Myo-inositol transporter	PM	600	SP	10
XP_020073509.1	hypothetical protein CYBJADRAFT_13180	Sugar phosphate permease; Nicotinic acid permease	PM	518	MFS	12
XP_020071808.1	high affinity nicotinic acid plasma membrane permease	Sugar phosphate permease; Nicotinic acid permease	PM	487	MFS	12
XP_020071592.1	high-affinity glucose transporter of the major facilitator superfamily	galactose transporter	PM	569	MFS	12
XP_020071589.1	high-affinity glucose transporter of the major facilitator superfamily	Glucose transporter	PM	559	MFS	10
XP_020072615.1	high-affinity glucose transporter of the major facilitator superfamily	Glucose transporter	PM	574	MFS	10
XP_020071526.1	hypothetical protein CYBJADRAFT_172465	Myo-inositol transporter	PM	530	SP	4
XP_020071185.1	MFS general substrate transporter	Fucose permease	PM	449	MFS	12
XP_020070843.1	general substrate transporter	HXT4 - Sugar transporter	PM	528	HXT	11
XP_020070629.1	MFS general substrate transporter	Fucose permease	PM	440	MFS	12
XP_020069676.1	General substrate transporter	Sugar transporter	PM	505	SP	12
XP_020068172.1	hypothetical protein CYBJADRAFT_169679	Sucrose/H$^+$ symporter	PM	443	GPH	7
XP_020069855.1	general substrate transporter	Sugar transporter (glucose transporter - HGT)	PM	478	SP	11

XP_020073641.1	general substrate transporter	Maltose permease	PM	537	SP	12
XP_020070205.1	MFS general substrate transporter	Sugar phosphate permease	PM	532	MFS	10
XP_020068136.1	MFS general substrate transporter	Sugar phosphate permease; Nicotinic acid permease	PM	517	MFS	8
XP_020068224.1	MFS general substrate transporter	HXT - Sugar transporter	PM	512	HXT	9
XP_020070790.1	MFS general substrate transporter	Sugar phosphate permease; D-galactonate transporter	PM	520	MFS	8
XP_020071468.1	allantoate permease	Sugar phosphate permease; D-galactonate transporter	PM	517	MFS	10
XP_020068066.1	allantoate permease	Sugar phosphate permease; D-galactonate transporter	PM	505	MFS	11
XP_020069500.1	MFS general substrate transporter	Sugar phosphate permease; D-galactonate transporter	PM	454	MFS	6
XP_020068391.1	MFS general substrate transporter	Sugar phosphate permease	PM	340	MFS	8
XP_020069947.1	MFS general substrate transporter	Sugar phosphate permease; D-galactonate transporter	PM	548	MFS	8
XP_020071662.1	MFS general substrate transporter	Sugar phosphate permease	PM	560	MFS	11
XP_020070288.1	MFS general substrate transporter	Sugar phosphate permease; D-galactonate transporter	PM	519	MFS	12
XP_020068389.1	major facilitator superfamily	Sugar phosphate permease	PM	534	MFS	10
XP_020068494.1	sugar transporter	Maltose permease	PM	537	SP	12
XP_020068676.1	UAA transporter	UDP-galactose transporter; HUT1	-	274	-	6
XP_020070544.1	major facilitator superfamily	Sugar phosphate permease	PM	510	MFS	9
XP_020068608.1	hypothetical protein CYBJADRAFT_164142	Sugar phosphate permease; D-galactonate transporter	PM	510	MFS	10
XP_020069159.1	general substrate transporter	HXT - Sugar transporter	PM	555	HXT	10
XP_020070804.1	UDP-galactose transporter	UDP-galactose transporter	GC/ER	319	-	9
XP_020069967.1	MFS quinate transporter QutD	Quinate transporter	PM	556	SP	12
XP_020068711.1	glycerol proton symporter of the plasma membrane	glycerol proton symporter	PM	623	SP	9
XP_020071189.1	Xanthine/uracil permease	xanthine-uric acid:H^+ symporter (UapA)	PM	507	NAT/NCS2	12
XP_020073285.1	uric acid-xanthine permease	xanthine-uric acid:H^+ symporter (UapA)	PM	544	NAT/NCS2	12
XP_020073583.1	Xanthine/uracil permease	xanthine-uric acid:H^+ symporter (UapA)	PM	599	NAT/NCS2	11
XP_020070234.1	DUF1212-domain-containing protein	Putative threonine/serine exporter	-	648	-	9
XP_020068982.1	hypothetical protein CYBJADRAFT_130504	Amino acid permease	PM	539	SLC 5 and 6-like	12
XP_020073346.1	amino acid permease	arginine permease	PM	563	APC	10
XP_020072765.1	proline specific permease	Proline permease	PM	542	APC	12
XP_020072758.1	amino acid transporter	Methionine permease	PM	577	APC	11
XP_020072698.1	amino acid permease	high-affinity histidine permease	PM	578	APC	12
XP_020072692.1	amino acid permease	high-affinity histidine permease	PM	595	APC	12
XP_020072219.1	PQ-loop-domain-containing protein	Putative vacuolar membrane transporter for cationic amino acids	V	291	APC	6
XP_020072117.1	amino acid transporter	GABA permease	-	549	APC	12
XP_020072096.1	amino acid permease	arginine permease	PM	542	APC	12
XP_020072095.1	amino acid permease	Lysine permease, one of three amino acid permeases (Alp1p/Can1p/Lyp1p)	PM	559	APC	12
XP_020072046.1	mitochondrial carrier	GABA permease	M	865	APC	14
XP_020071993.1	general amino acid permease	Transporter of all L-amino acids, some D-amino acids, related compounds, toxic analogs and polyamines	PM	593	APC	12
XP_020071781.1	very low affinity methionine permease	Methionine permease	PM	541	APC	11

XP_020070906.1	amino acid transporter	Methionine permease	PM	446	APC	10
XP_020070879.1	general amino acid permease	Amino acid permease	PM	499	SLC 5 and 6-like	12
XP_020070372.1	amino acid permease	L-glutamate and L-aspartate transporter	PM	581	SLC	12
XP_020069152.1	high-affinity glutamine permease	high-affinity polyamine permease	PM	557	APC	12
XP_020068708.1	amino acid transporter	High-affinity methionine permease (MUP1)	PM	525	APC	12
XP_020068687.1	gamma-aminobutyric acid transporter	GABA permease	-	554	APC	12
XP_020068589.1	amino acid transporter	GABA permease	-	539	APC	12
XP_020070250.1	gamma-aminobutyric acid transporter	GABA permease	-	576	APC	11
XP_020068588.1	amino acid transporter	GABA permease	-	519	APC	12
XP_020072671.1	amino acid permease	glutamine permease	-	573	-	10
XP_020068591.1	hypothetical protein CYBJADRAFT_191930	Amino acid permease	PM	554	SLC 5 and 6-like	13
XP_020068336.1	high affinity methionine permease	Methionine permease	PM	562	APC	12
XP_020073538.1	hypothetical protein CYBJADRAFT_165774	uptake of cationic amino acids	-	550	SLC 5 and 6-like	11
XP_020072859.1	hypothetical protein CYBJADRAFT_122835	Amino acid permease	V	406	SLC 5 and 6-like	10
XP_020072635.1	hypothetical protein CYBJADRAFT_124199	neutral amino acid permease	-	495	SLC 5 and 6-like	11
XP_020072543.1	hypothetical protein CYBJADRAFT_166235	Proline permease	-	557	APC	11
XP_020072526.1	hypothetical protein CYBJADRAFT_123422	Amino acid permease	-	569	-	11
XP_020072469.1	hypothetical protein CYBJADRAFT_166163	Aspartate and glutamate exporter	PM	332	SLC	7
XP_020072189.1	hypothetical protein CYBJADRAFT_165903	Amino acid transporter	PM	1127	SLC12	11
XP_020072047.1	hypothetical protein CYBJADRAFT_166772	purine-cytosine permease	PM	353	NCS1	9
XP_020071104.1	hypothetical protein CYBJADRAFT_167449	large neutral amino acids transporter	V	605	AAAP	10
XP_020071886.1	hypothetical protein CYBJADRAFT_193543	Asparagine/glutamine permease	-	555	-	12
XP_020072054.1	hypothetical protein CYBJADRAFT_166778	purine-cytosine permease	PM	520	NCS1	12
XP_020071645.1	hypothetical protein CYBJADRAFT_124945	large neutral amino acids transporter	V	432	AAAP	11
XP_020071283.1	hypothetical protein CYBJADRAFT_193726	Amino acid permease	-	584	APC	12
XP_020070552.1	hypothetical protein CYBJADRAFT_167553	GABA-glycine transporter	V	468	APC	10
XP_020070247.1	hypothetical protein CYBJADRAFT_177594	Proline permease	-	567	APC	11
XP_020067707.1	hypothetical protein CYBJADRAFT_187372	Uridine permease	-	554	SLC 5 and 6-like	12
XP_020067734.1	hypothetical protein CYBJADRAFT_133015	Amino acid permease	-	524	APC	12
XP_020068118.1	hypothetical protein CYBJADRAFT_169726	GABA permease	-	523	APC	11
XP_020067965.1	amino acid transporter	GABA permease	-	538	APC	12
XP_020070370.1	hypothetical protein CYBJADRAFT_184971	Amino acid permease	-	350	APC	8
XP_020070752.1	hypothetical protein CYBJADRAFT_88763	Amino acid permease	V	597	-	13
XP_020068154.1	hypothetical protein CYBJADRAFT_178724	Amino acid permease	-	451	SLC 5 and 6-like	11
XP_020068313.1	MFS general substrate transporter	putative aminoacid permease	V	558	-	13
XP_020068388.1	MFS general substrate transporter	putative aminoacid trasnporter	V	547	-	12
XP_020069720.1	hypothetical protein CYBJADRAFT_168226	nucleobase-cation-symporter-1	-	595	NCS1	12

XP_020069721.1	hypothetical protein CYBJADRAFT_168227	nucleobase-cation-symporter-1	-	577	NCS1	11
XP_020070877.1	hypothetical protein CYBJADRAFT_150020	carnitine transporter	PM	555	OCTN	12
XP_020070203.1	hypothetical protein CYBJADRAFT_168214	putative nucleobase-cation-symporter-1	-	545	NCS1	12
XP_020073450.1	hypothetical protein CYBJADRAFT_165694	large neutral amino acids transporter	V	702		10
XP_020071097.1	hypothetical protein CYBJADRAFT_126342	nucleobase-cation-symporter-1	-	570	NCS1	12
XP_020071352.1	hypothetical protein CYBJADRAFT_167001	Thiamine transporter	PM	574	NCS1	12
XP_020068915.1	small oligopeptide transporter	Oligopeptide transporter	PM	809	OPT	13
XP_020069810.1	OPT-domain-containing protein	Oligopeptide transporter	PM	879	OPT	14
XP_020069892.1	OPT oligopeptide transporter	Oligopeptide transporter	PM	748	OPT	14
XP_020070895.1	PTR2-domain-containing protein	Oligopeptide transporter	PM	520	OPT	10
XP_020071584.1	oligopeptide transporter	Oligopeptide transporter	PM	671	OPT	13
XP_020072614.1	small oligopeptide transporter	Oligopeptide transporter	PM	884	OPT	13
XP_020068934.1	synaptic vesicle transporter	Polyamine transporter	PM	550	DHA1	10
XP_020068307.1	MFS general substrate transporter	Polyamine transporter	PM	590	DHA1	12
XP_020070925.1	MFS general substrate transporter	Polyamine transporter	PM	637	DHA1	13
XP_020073548.1	MFS general substrate transporter	Polyamine transporter	PM	584	DHA1	12
XP_020068504.1	MFS general substrate transporter	Polyamine transporter	PM	608	DHA1	12
XP_020071598.1	hypothetical protein CYBJADRAFT_193886	Na:urea-polyamine cotransporter	PM	684	SSS	15
XP_020071597.1	hypothetical protein CYBJADRAFT_177232	Na:urea-polyamine cotransporter	PM	682	SSS	15
XP_020071599.1	hypothetical protein CYBJADRAFT_161964	Na:urea-polyamine cotransporter	PM	687	SSS	15
XP_020071587.1	hypothetical protein CYBJADRAFT_167195	Na:urea-polyamine cotransporter	PM	662	SSS	15
XP_020071600.1	urea transport protein	Na:urea-polyamine cotransporter	PM	605	SSS	12
XP_020067703.1	hypothetical protein CYBJADRAFT_156520	Urea transporter	PM	667	SSS	14
XP_020068110.1	hypothetical protein CYBJADRAFT_195367	Urea transporter	PM	672	SSS	16
XP_020068713.1	Na+/solute symporter	Na:urea cotransporter	PM	719	Solute carrier families 5 and 6-like	15
XP_020067650.1	hypothetical protein CYBJADRAFT_180897	Urea transporter	PM	389	Solute carrier families 5 and 6-like	9
XP_020071975.1	ZIP zinc/iron transport family	Zinc/iron transporter	PM	385	ZIP	8
XP_020069858.1	Zip-domain-containing protein	Zinc transporter	PM	401	ZIP	7
XP_020073040.1	ZIP zinc/iron transport family	Zinc/iron transporter	PM	373	ZIP	7
XP_020068969.1	hypothetical protein CYBJADRAFT_130476	putative Zinc transporter	PM	443	ZIP	7
XP_020069481.1	DUF803-domain-containing protein	putative magnesium transporter	PM	361	-	8
XP_020073120.1	hypothetical protein CYBJADRAFT_145818	Manganese transporter	PM	324	OST3/OST6	4
XP_020068314.1	natural resistance-associated macrophage protein	Manganese transporter	PM	585	OST3/OST6	11
XP_020071115.1	natural resistance-associated macrophage protein	Manganese transporter	PM	519	OST3/OST6	11
XP_020068637.1	hypothetical protein CYBJADRAFT_164167	Sulphate permease	-	520	SulP	6
XP_020068935.1	hypothetical protein CYBJADRAFT_130450	Sulphate permease	V	805	SulP	10
XP_020069987.1	high affinity sulfate transporter 2	Sulphate permease	-	703	SulP	9
XP_020069499.1	hypothetical protein CYBJADRAFT_168745	Sulphur transporter	?	327	-	5

XP_020067917.1	nitrate transporter	high-affinity nitrate transporter	PM	541	NRT/MFS	10
XP_020068695.1	Rh-like protein/ammonium transporter	ammonium transporter	PM	378	AMT	8
XP_020071845.1	ammonium transporter	ammonium transporter	PM	482	AMT	9
XP_020071846.1	ammonium transporter	ammonium transporter	PM	472	AMT	9
XP_020071849.1	ammonium transporter	ammonium transporter	PM	494	AMT	9
XP_020070915.1	hypothetical protein CYBJADRAFT_150094	Siderophore-iron transporter	PM	596	ARN	13
XP_020068393.1	MFS general substrate transporter	Siderophore-iron transporter	PM	621	ARN	12
XP_020069643.1	hypothetical protein CYBJADRAFT_168527	Siderophore-iron transporter	V	308	ARN	5
XP_020068874.1	MFS general substrate transporter	siderophore-iron chelates transporter	PM	622	ARN	12
XP_020067684.1	MFS general substrate transporter	Siderophore-iron transporter	PM	634	MFS	14
XP_020067952.1	natural resistance-associated macrophage protein	Putative divalent metal ion transporter involved in iron homeostasis	-	486	-	10
XP_020068163.1	iron permease FTR1	Putative high-affinity iron transporter	V?	400	ILT	7
XP_020068199.1	iron permease FTR1	Iron transporter	PM?	424	ILT	7
XP_020069246.1	FTR1-domain-containing protein	High-affinity Fe^{2+}/Pb^{2+} permease	PM?	394	ILT	7
XP_020070898.1	FTR1-domain-containing protein	High-affinity Fe^{2+}/Pb^{2+} permease	PM?	454	ILT	7
XP_020071914.1	MFS general substrate transporter	Boron transporter	PM	426	DAG	11
XP_020068717.1	hypothetical protein CYBJADRAFT_131112	Boron transporter	PM	522	DAG	11
XP_020072519.1	hypothetical protein CYBJADRAFT_123387	putative calcium channel	V	657	-	7
XP_020073394.1	hypothetical protein CYBJADRAFT_146349	high-affinity calcium channel	-	1606	-	20
XP_020070691.1	TRP-domain-containing protein	Putative calcium channel	V	735	-	9
XP_020070750.1	hypothetical protein CYBJADRAFT_167722	Calcium-activated chloride channel	?	1009	-	5
XP_020070809.1	hypothetical protein CYBJADRAFT_167207	High-affinity copper transporter	PM	255	CTR	4
XP_020071056.1	Cu-transporting P-type ATPase	Copper transporte		1155	CTR	8
XP_020069151.1	phosphate permease	Phosphate permease	PM	573	PIT	9
XP_020069206.1	phosphate transporter	Na^+-phosphate transport protein	PM	583	PIT	9
XP_020069270.1	SPX-domain-containing protein	Phosphate permease	PM	843	PIT	12
XP_020069943.1	high-affinity inorganic phosphate transporter and low-affinity manganese transporter	phosphate:H+ symporter	PM	524	MFS	11
XP_020071038.1	MFS general substrate transporter	phosphate:H+ symporter	PM	431	MFS	10
XP_020069515.1	hypothetical protein CYBJADRAFT_168415	Macrolide transporter/ATP-binding cassette permease	ER?	999	AAA	8
XP_020069233.1	hypothetical protein CYBJADRAFT_168857	ABC-type bacteriocin/lantibiotic exporter	V	1634	PDR	13
XP_020071812.1	P-loop containing nucleoside triphosphate hydrolase protein	ABC-type bacteriocin/lantibiotic exporter	-	1427	PDR	8
XP_020068423.1	multidrug resistance protein	ATP-binding cassette permease	-	1510	PDR	12
XP_020068721.1	hypothetical protein CYBJADRAFT_178513	ABC-type bacteriocin/lantibiotic exporter	M	655	PDR	6
XP_020068180.1	pleiotropic drug resistance protein PDR	ABC transporter actively exports various drugs expression regulated by Pdr1p	PM	1495	PDR	12
XP_020072697.1	hypothetical protein CYBJADRAFT_161059	ABC-type bacteriocin/lantibiotic exporter	V	1128	MRP/CFTR family	12
XP_020068178.1	hypothetical protein CYBJADRAFT_192230	Macrolide transporter/ATP-binding cassette permease	PM	416	PDR	6
XP_020068427.1	hypothetical protein CYBJADRAFT_169435	Macrolide transporter/ATP-binding cassette permease	PM	1114	PDR	11
XP_020072078.1	hypothetical protein CYBJADRAFT_161514	ABC-type bacteriocin/lantibiotic exporter	-	895	ABC	4

XP_020068426.1	Orfs10c	ATP-binding cassette permease	-	463	PDR	6
XP_020068428.1	pleiotropic drug resistance protein PDR	ATP-binding cassette permease	-	1522	PDR	12
XP_020068503.1	plasma membrane ATP-binding cassette transporter	ATP-binding cassette permease	PM	1370	PDR	10
XP_020069109.1	ABC transporter	Macrolide transporter/ATP-binding cassette permease	PM	1262	PDR	12
XP_020070621.1	mitochondrial ABC transporter	ATP-binding cassette permease responsible to synthesize the precursors of iron-sulfur (Fe/S) clusters to the cytosol	M	665	PDR	5
XP_020071933.1	P-loop containing nucleoside triphosphate hydrolase protein	ABC transporter actively exports various drugs	PM	1381	PDR	10
XP_020072139.1	P-loop containing nucleoside triphosphate hydrolase protein	ABC transporter actively exports various drugs	PM	1352	PDR	9
XP_020072401.1	V0 domain of vacuolar H+ATPase	Protein of unknown function	V	793	-	7
XP_020070368.1	hypothetical protein CYBJADRAFT_127994	Alpha-factor-transporting ATPase	-	1034	ABC	9
XP_020070779.1	V-type ATPase	ATPase V-type proteolipid subunit	V	160	-	4
XP_020069220.1	potassium/sodium eff	P-type ATPase sodium pump	-	1068	-	10
XP_020069697.1	Cu-transporting P-type ATPase	Cation transport ATPase	-	916	-	8
XP_020070039.1	plasma membrane H+-ATPase, pumps protons out of the cell	H-transporting ATPase	PM	897	-	8
XP_020070620.1	potassium/sodium eff	ATPase sodium pump	-	1062	-	10
XP_020071470.1	K, P-type ATPase	Magnesium-transporting ATPase (P-type)	GC?	1069	-	10
XP_020071820.1	membrane protein	Ca^{2+}-transporting ATPase	V	1199	-	10
XP_020071856.1	cation-transporting ATPase	Ca^{2+}-transporting ATPase	V	1326	-	11
XP_020072990.1	calcium-transporting ATPase	Ca^{2+}-transporting ATPase	ER	994	-	8
XP_020069052.1	calcium-transporting P	High-affinity Ca^{2+}/Mn^{2+} P-type ATPase	GC	915	-	8
XP_020070551.1	V0/A0 complex, 116-kDa subunit of ATPase	Protein of unknown function	-	829	-	7
XP_020070887.1	MFS general substrate transporter	mch1 transporter		348	MCT	9
XP_020070160.1	MFS general substrate transporter	mch2 transporter	PM	468	MCT	12
XP_020070165.1	MFS general substrate transporter	mch2 transporter	PM	780	MCT	12
XP_020072088.1	hypothetical protein CYBJADRAFT_172054	mch4 transporter	PM	612	MCT	11
XP_020067635.1	MFS general substrate transporter	mch4 transporter	PM	470	MCT	12
XP_020067736.1	MFS general substrate transporter	monocarboxylate transporter; mch	PM	487	MCT	11
XP_020067765.1	hypothetical protein CYBJADRAFT_178915	putative Ato3	PM or VA	272	AceTr	6
XP_020069696.1	MFS general substrate transporter	putative mch3	PM	505	MCT	10
XP_020068891.1	hypothetical protein CYBJADRAFT_130575	Dicarboxylate Transporter	PM	384	TDT	8
XP_020068647.1	MFS general substrate transporter	oxalate/formate antiporter	PM	470	OFA	12
XP_020072451.1	Formate/nitrite transporter	Formate/nitrite transporter	V	597	OFA	6
XP_020069263.1	MFS general substrate transporter	Putative riboflavin transporter; mch5 transporter	PM	491	MCT	12
XP_020069005.1	transmembrane protein	ATO2	PM	280	AceTr	6
XP_020070869.1	MFS general substrate transporter	monocarboxylate transporter	PM	504	SHS	10
XP_020069382.1	MFS general substrate transporter	monocarboxylate transporter	PM	511	SHS	12
XP_020071769.1	hypothetical protein CYBJADRAFT_124710	monocarboxylate transporter	PM	504	SHS	9
XP_020070445.1	hypothetical protein CYBJADRAFT_128088	monocarboxylate transporter	PM	279	AceTr	4
XP_020073031.1	hypothetical protein CYBJADRAFT_121868	monocarboxylate transporter	PM	269	AceTr	5

XP_020073178.1	hypothetical protein CYBJADRAFT_123057	monocarboxylate transporter	PM	271	AceTr	4
XP_020073179.1	transmembrane protein	ATO2 Ammonia transporter	PM	262	AceTr	6
XP_020069826.1	MFS general substrate transporter	Allantoate transporter	PM	491	MFS	10
XP_020070263.1	MFS general substrate transporter	Allantoate transporter	PM	485	MFS	10
XP_020071152.1	MFS general substrate transporter	Allantoate transporter	PM	524	MFS	10
XP_020071813.1	MFS general substrate transporter	Allantoate transporter	PM	528	MFS	10
XP_020069269.1	allantoin permease	Allantoate transporter	PM	506	MFS	10
XP_020071707.1	plasma membrane permease proposed to be	Allantoate transporter	PM	486	MFS	10
XP_020072609.1	dicarboxylic amino acid permease	Dicarboxylic amino acid permease	PM	564	-	12
XP_020073044.1	hypothetical protein CYBJADRAFT_165372	sodium/dicarboxylate cotransporter	PM	821	SLC13	12
XP_020068615.1	hypothetical protein CYBJADRAFT_131197	Dicarboxylic amino acid (L-glutamate and L-aspartate) permease	PM	524	YAT	11
XP_020069468.1	hypothetical protein CYBJADRAFT_168724	Dicarboxylic amino acid permease	PM	567	-	12
XP_020069774.1	MFS general substrate transporter	putative oxalate/formate antiporter	PM	476	MFS	12
XP_020071456.1	MFS general substrate transporter	putative sialic acid transporter; Jen1 homolog	PM	548	SHS	11
XP_020069652.1	Tricarboxylate/iron carrier	Tricarboxylate carrier	M?	327		4
XP_020072234.1	MFS general substrate transporter	putative sialic acid transporter; Jen1 homolog	PM	506	SHS	12
XP_020069659.1	aquaporin	aquaporin	PM	344	MIP	5
XP_020073474.1	aquaporin-like protein	aquaporin	PM	223	MIP	5
XP_020068161.1	dityrosine transporter A Q resistance	dityrosine transporter	?	488	MFS	12
XP_020067705.1	MFS general substrate transporter	dityrosine transporter	?	436	MFS	12
XP_020068827.1	MFS general substrate transporter	Quinidine resistance protein	PM	624	DHA1	12
XP_020070272.1	MFS general substrate transporter	Aminotriazole resistance protein	-	534	MFS	13
XP_020071110.1	MATE efflux family protein	Oligosaccharidyl-lipid/polysaccharide (MOP) exporter	PM	545	MATE	10
XP_020073439.1	MATE efflux family protein	Na+-driven efflux pump	PM	543	MATE	10
XP_020073380.1	MATE efflux family protein	Na+-driven efflux pump	PM	642	MATE	11
XP_020071098.1	hypothetical protein CYBJADRAFT_189880	Na+-driven efflux pump	PM	486	MATE	11
XP_020069868.1	MATE efflux family protein	Na+-driven efflux pump	PM	480	MATE	11
XP_020067763.1	MATE efflux family protein	Na+-driven efflux pump	PM	490	MATE	11
XP_020070598.1	hypothetical protein CYBJADRAFT_167592	Cation efflux transporter	PM	525	?	5
XP_020069514.1	hypothetical protein CYBJADRAFT_173923	Cation efflux transporter	PM	527	?	6
XP_020069686.1	hypothetical protein CYBJADRAFT_168553	Na:Ca exchanger	PM	710	CaCA	11
XP_020070439.1	hypothetical protein CYBJADRAFT_167962	Na:Ca exchanger	PM	538	CaCA	9
XP_020072276.1	endosomal Na+/H+ exchanger	Na+/H+ exchanger	-	596	CaCA	10
XP_020069295.1	hypothetical protein CYBJADRAFT_163451	Na+/H+ exchanger	-	887	CaCA	9
XP_020071440.1	hypothetical protein CYBJADRAFT_167074	Sodium-solute symporter	PM	517	SSS	13
XP_020071640.1	hypothetical protein CYBJADRAFT_124719	Ca^{2+}:H^{+} antiporter	V	924	CaCA	13
XP_020069658.1	calcium/proton exchanger	H+/Ca2+ exchanger	V	398	CaCA	11
XP_020071450.1	UPF0016-domain-containing protein	Putative Ca2+/H+ antiporter	GC	284	CaCA	5

XP_020070929.1	hypothetical protein CYBJADRAFT_126498	Pyridoxine (vitamin B6) transporter	PM	536	NCS1	12
XP_020071249.1	hypothetical protein CYBJADRAFT_177061	Pyridoxine transporter	PM	494	NCS1	12
XP_020070771.1	hypothetical protein CYBJADRAFT_167740	Voltage-gated chloride channel	GC	765	-	9
XP_020069557.1	MFS general substrate transporter	H-pantothenate symporter	PM	514	ACS	11
XP_020067902.1	MFS general substrate transporter	Pantothenate transporter	PM	524	P-RFT	9
XP_020067992.1	MFS general substrate transporter	Pantothenate transporter	PM	518	P-RFT	8
XP_020070852.1	MFS general substrate transporter	Pantothenate transporter	PM	531	P-RFT	10
XP_020071176.1	MFS general substrate transporter	Pantothenate transporter	PM	525	P-RFT	10
XP_020070206.1	pantothenate transporter	Pantothenate transporter	PM	509	P-RFT	10
XP_020070764.1	pantothenate transporter	Pantothenate transporter	PM	471	P-RFT	10
XP_020073174.1	DUF6-domain-containing protein	Permease of the drug/metabolite transporter (DMT) superfamily	-	371	DMT	10
XP_020068269.1	hypothetical protein CYBJADRAFT_175365	Acetyl-coenzyme A transporter	?	527	SLC	7
XP_020069754.1	voltage-gated potassium channel	voltage-gated potassium channel	-	638	-	11
XP_020069511.1	PQ-loop-domain-containing protein	Lysosomal Cystine Transporter	V	266	-	6
XP_020068308.1	multidrug resistance protein 4	Putative spermidine export pump	-	616	MFS	12
XP_020071094.1	MFS general substrate transporter	putative transporter	PM	498	DHA1	10
XP_020072755.1	hypothetical protein CYBJADRAFT_165142	Nucleoside transporter	V	414	ENT	10
XP_020068972.1	MFS general substrate transporter	Protein of unknown function	-	575	MFS	12
XP_020069015.1	UAA transporter	UDP-N-acetylglucosamine transporter	GC	374	-	10
XP_020073060.1	hypothetical protein CYBJADRAFT_160020	putative Zinc-regulated transporter 3	V	519	ZIP	7
XP_020073523.1	hypothetical protein CYBJADRAFT_165762	putative sugar transferase	-	353	-	6
XP_020073237.1	dolichyl pyrophosphate Glc1Man9GlcNAc2 alpha-1,3-glucosyltransferase	Glycosyltransferase	-	577	-	12
XP_020072571.1	MBOAT-domain-containing protein	lysophospholipid acyltransferase	-	559	MBOAT	4
XP_020070926.1	alkaline ceramidase	alkaline ceramidase	-	283	-	7
XP_020073043.1	ALG6, ALG8 glycosyltransferase	glycosyltransferase	-	543	-	10
XP_020069072.1	Choline/ethanolaminephosphotransferase	Choline/ethanolaminephosphotransferase	-	387	-	7
XP_020073198.1	hypothetical protein CYBJADRAFT_165490	Glucosylceramide synthase	-	515	GT-A	4
XP_020069966.1	hypothetical protein CYBJADRAFT_128263	Glycerophosphocholine acyltransferase	-	377	-	8
XP_020067654.1	hypothetical protein CYBJADRAFT_195583	Acyltransferase	-	532	MBOAT	12
XP_020067666.1	hypothetical protein CYBJADRAFT_170103	Acyltransferase	-	395	membrane-bound O-acyltransferase family	6
XP_020067667.1	acyl-CoA:sterol acyltransferase, isozyme of Are1p	Acyltransferase	-	530	membrane-bound O-acyltransferase family	8
XP_020067672.1	hypothetical protein CYBJADRAFT_170090	Ferric reductase, NADH/NADPH oxidase	-	493	Ferric reductase, NADH/NADPH oxidase	4
XP_020068011.1	hypothetical protein CYBJADRAFT_169818	Dolichol kinase	ER	541	-	10
XP_020067844.1	hypothetical protein CYBJADRAFT_169968	Transferase	-	541	Cytidylyltransferase family	10

XP_020070174.1	hypothetical protein CYBJADRAFT_163012	Dolichyl-phosphate-mannose-protein mannosyltransferase	ER	453	GT39	6
XP_020070754.1	hypothetical protein CYBJADRAFT_167725	Dolichyl-phosphate-mannose-protein mannosyltransferase	ER	599	GT39	8
XP_020068361.1	hypothetical protein CYBJADRAFT_164511	Dolichyl-phosphate-mannose-protein mannosyltransferase	ER	695	GT39	9
XP_020069681.1	hypothetical protein CYBJADRAFT_129185	Ferric reductase	-	704	Ferric reductase, NADH/NADPH oxidase	5
XP_020069876.1	hypothetical protein CYBJADRAFT_128930	UDP-N-acetyl-glucosamine-1-P-transferase	ER	448	-	10
XP_020069995.1	hypothetical protein CYBJADRAFT_173475	putative diacylglycerol pyrophosphate (DGPP) phosphatase	V	336	PBRBC	6
XP_020070244.1	hypothetical protein CYBJADRAFT_127575	Ferric reductase, NADH/NADPH oxidase	-	581	Ferric reductase, NADH/NADPH oxidase	6
XP_020070438.1	hypothetical protein CYBJADRAFT_167960	putative oxidoreductase/ferric-chelate reductase	-	548	Ferric reductase, NADH/NADPH oxidase	7
XP_020071131.1	hypothetical protein CYBJADRAFT_167474	NADPH oxidase	ER	493	-	4
XP_020070569.1	hypothetical protein CYBJADRAFT_172940	galactosyltransferase	GC	648	GT7	5
XP_020071669.1	hypothetical protein CYBJADRAFT_166429	Phospholipid methyltransferase	-	297	-	7
XP_020071832.1	hypothetical protein CYBJADRAFT_124665	Phospholipid methyltransferase	-	803	-	8
XP_020071840.1	hypothetical protein CYBJADRAFT_166581	Chitin synthase	-	1055	-	7
XP_020072038.1	hypothetical protein CYBJADRAFT_166767	phosphoryltransferase	ER	978	-	15
XP_020073397.1	hypothetical protein CYBJADRAFT_165654	iron/copper reductase	-	687	-	8
XP_020072598.1	hypothetical protein CYBJADRAFT_147754	putative serine protease	GC	274	-	5
XP_020073266.1	hypothetical protein CYBJADRAFT_165542	Chitin synthase	-	1006	GT-A	7
XP_020073402.1	hypothetical protein CYBJADRAFT_146367	alpha-1,2 glucosyltransferase	ER	510	-	11
XP_020067856.1	Zn-dependent exopeptidase	metalloprotease	ER	979	-	9
XP_020068928.1	hypothetical protein CYBJADRAFT_174544	glycerol-3-phosphate O-acyltransferase	-	668	-	4
XP_020068041.1	phospholipid-translocating P-type ATPase	Aminophospholipid translocases	-	1555	-	7
XP_020068436.1	phosphatidate cytidylyltransferase (CDP-diglyceride synthetase)	phosphatidate cytidylyltransferase	-	425	-	6
XP_020068085.1	PMT-domain-containing protein	Dolichyl-phosphate-mannose-protein mannosyltransferase	ER	714	GT39	9
XP_020068550.1	dolichyl-phosphate-mannose-protein mannosyltransferase	Dolichyl-phosphate-mannose-protein mannosyltransferase	ER	742	GT39	7
XP_020068856.1	Alg9-like mannosyltransferase	Dolichyl-phosphate-mannose-protein mannosyltransferase	ER	465	GT39	6
XP_020068875.1	ICMT-domain-containing protein	protein-S-isoprenylcysteine O-methyltransferase	-	323	-	4
XP_020069045.1	COX15-CtaA-domain-containing protein	Cytochrome c OXidase	-	585	-	8
XP_020069148.1	PMT-domain-containing protein	Dolichyl-phosphate-mannose-protein mannosyltransferase	ER	765	GT39	9
XP_020069357.1	sphingolipid delta8-desaturase	Fatty acid desaturase	-	559	-	5
XP_020069415.1	phospholipid-translocating P-type ATPase	Phospholipid-translocating P-type ATPase, Flippase	-	1489	-	10
XP_020071674.1	phospholipid-translocating P-type ATPase	Phospholipid-translocating P-type ATPase, Flippase	-	1105	-	8
XP_020069444.1	delta 12-fatty acid desaturase	Fatty acid desaturase	-	419	-	5
XP_020069456.1	fatty acid elongase 3	Elongase involved in fatty acid biosynthesis	-	335	-	7
XP_020069522.1	rhomboid-domain-containing protein	rhomboid protease	-	336	-	7

XP_020069555.1	MBOAT-domain-containing protein	D-alanyl-lipoteichoic acid acyltransferase	-	537	MBOAT	11
XP_020069757.1	3-hydroxyacyl-CoA dehydratase	3-hydroxyacyl-CoA dehydratase	-	225	-	5
XP_020069939.1	dolichyl-phosphate-mannose-protein mannosyltransferase	Dolichyl-phosphate-mannose-protein mannosyltransferase	ER	739	GT39	11
XP_020070320.1	chitin synthase 3	Chitin synthase III	-	1179	-	6
XP_020070582.1	Dol-P-Man:Man(5)GlcNAc(2)-PP-Dol alpha-1,3-mannosyltransferase	Dolichol-P-Man dependent alpha(1-3) mannosyltransferase	-	473	GT39	7
XP_020070354.1	mannosyltransferase	Dolichyl-phosphate-mannose-protein mannosyltransferase	ER	368	GT39	8
XP_020070367.1	phospholipid methyltransferase	phospholipid methyltransferase	-	211	-	4
XP_020070814.1	CDP-diacylglycerol–glycerol-3-phosphate 3-phosphatidyltransferase	CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase	-	210	-	4
XP_020072610.1	CDP-diacylglycerol–serine O-phosphatidyltransferase	phosphatidylserine synthase	-	349	-	5
XP_020071022.1	membrane-spanning Ca-ATPase	phospholipid-translocating P-type ATPase, Flippase	-	1150	-	8
XP_020071074.1	rhomboid-domain-containing protein	Serine protease	M	326	-	6
XP_020071075.1	ERG4/ERG24 ergosterol biosynthesis protein	Phospholipid methyltransferase; also similar to C-14 sterol reductase	-	435	-	8
XP_020071233.1	ERG4/ERG24 ergosterol biosynthesis protein	Phospholipid methyltransferase	-	478	-	7
XP_020071538.1	dolichyl-P-mannose-protein mannosyltransferase	Dolichyl-phosphate-mannose-protein mannosyltransferase	-	761	GT39	9
XP_020071563.1	alkaline phosphatase-like protein	alkaline phosphatase-like protein	-	799	-	12
XP_020071693.1	fatty acid elongase	Elongase involved in fatty acid biosynthesis	-	332	-	5
XP_020071907.1	3-hydroxy-3-methylglutaryl-coenzyme a reductase	3-hydroxy-3-methylglutaryl-coenzyme A reductase	-	934	-	6
XP_020071872.1	phosphatidylinositol N-acetylglucosaminyltransferase	phosphatidylinositol N-acetylglucosaminyltransferase	-	285	-	5
XP_020072131.1	PAP2-domain-containing protein	diacylglycerol pyrophosphate phosphatase	V	256	-	4
XP_020072734.1	para-hydroxybenzoate–polyprenyltransferase	4-hydroxybenzoate polyprenyl transferase	-	352	-	4
XP_020072858.1	longevity assurance proteins LAG1/LAC1	Ceramide synthase	-	318	-	6
XP_020073019.1	omega3 fatty acid desaturase	Fatty acid desaturase	-	423	-	5
XP_020073121.1	phospholipid-translocating P-type ATPase	Phospholipid-translocating P-type ATPase, Flippase	-	1313	-	10
XP_020073131.1	protoheme IX farnesyltransferase	putative Cytochrome c Oxidase	-	425	-	6
XP_020073177.1	phosphatidylinositol synthase	phosphatidylinositol synthase	-	224	-	4
XP_020069085.1	hypothetical protein CYBJADRAFT_179724	C-5 sterol desaturase	-	359	-	4
XP_020069333.1	hypothetical protein CYBJADRAFT_168610	Dolichyl-phosphate-mannose-protein mannosyltransferase	-	400	-	9
XP_020069427.1	hypothetical protein CYBJADRAFT_168692	Dolichol kinase	-	319	-	6
XP_020070789.1	sphingolipid long-chain base-1-phosphate phosphatase	Membrane-associated phospholipid phosphatase	ER	482	-	7
XP_020072026.1	Glycosylphosphatidylinositol:protein transamidase, GAA1 component	Glycosylphosphatidylinositol:protein transamidase	-	571	-	6
XP_020071270.1	hypothetical protein CYBJADRAFT_183891	Membrane-embedded ubiquitin-protein ligase	ER	570	-	9
XP_020068970.1	zf-DHHC-domain-containing protein	Subunit of a palmitoyltransferase	-	334	-	4
XP_020068922.1	STT3 subunit of Oligosaccharyl transferase	STT3 subunit of Oligosaccharyl transferase		728	-	11
XP_020070011.1	1,3-beta glucan synthase	Catalytic subunit of 1,3-beta-glucan synthase	-	1899	-	13
XP_020069944.1	MFS general substrate transporter	Protein of unknown function	-	425	MFS	12

XP_020069446.1	hypothetical protein CYBJADRAFT_129925	Protein of unknown function	-	610	-	8
XP_020069441.1	MFS general substrate transporter	Protein of unknown function	-	1056	-	10
XP_020069424.1	hypothetical protein CYBJADRAFT_174246	Protein of unknown function	-	457	-	4
XP_020069423.1	hypothetical protein CYBJADRAFT_168690	Protein of unknown function	-	453	-	5
XP_020069375.1	hypothetical protein CYBJADRAFT_185812	Protein of unknown function	-	255	-	5
XP_020069339.1	hypothetical protein CYBJADRAFT_19366	Protein of unknown function	-	232	-	5
XP_020069318.1	hypothetical protein CYBJADRAFT_168600	Protein of unknown function	-	750	-	4
XP_020069304.1	hypothetical protein CYBJADRAFT_153097	Protein of unknown function	-	495	SLC 5 and 6-like	10
XP_020069268.1	MFS general substrate transporter	Protein of unknown function	-	499	MFS	11
XP_020069131.1	hypothetical protein CYBJADRAFT_130128	Protein of unknown function	-	579	-	11
XP_020068998.1	hypothetical protein CYBJADRAFT_130426	Protein of unknown function	-	1513	-	14
XP_020073529.1	Yip1-domain-containing protein	Protein of unknown function	GC	217	-	5
XP_020073472.1	DUF396-domain-containing protein	Protein of unknown function	-	202	-	4
XP_020073418.1	Got1-domain-containing protein	Protein of unknown function	GC	139	-	4
XP_020073395.1	Gpi1-domain-containing protein	Protein of unknown function	-	618	-	5
XP_020073281.1	DUF1774-domain-containing protein	Protein of unknown function	-	288	-	8
XP_020073132.1	DUF221-domain-containing protein	Protein of unknown function	-	891	-	10
XP_020073056.1	mannose-P-dolichol utilization defect 1 protein	Protein of unknown function	-	282	-	5
XP_020072913.1	Rft-1-domain-containing protein	unknown	-	541	-	11
XP_020072672.1	Yip1-domain-containing protein	unknown	GC	294	-	5
XP_020072644.1	DUF221-domain-containing protein	unknown	-	606	-	7
XP_020071932.1	SURF4-domain-containing protein	unknown	ER	302	-	6
XP_020071745.1	pali-domain-containing protein	unknown	-	654	-	4
XP_020071732.1	Yip1-domain-containing protein	unknown	ER	310	-	5
XP_020069657.1	hypothetical protein CYBJADRAFT_163392	unknown	-	356	-	7
XP_020067730.1	hypothetical protein CYBJADRAFT_170007	unknown	-	261	-	4
XP_020067771.1	hypothetical protein CYBJADRAFT_156392	vacuolar membrane protein	VA	314	?	5
XP_020068055.1	hypothetical protein CYBJADRAFT_169764	EamA-like transporter family	PM	444	DUF6	10
XP_020068109.1	MFS general substrate transporter	unknown	PM	445	MFS	12
XP_020068372.1	hypothetical protein CYBJADRAFT_169548	Membrane-domain of unknow function	-	168	-	4
XP_020068228.1	hypothetical protein CYBJADRAFT_175325	CrcB-like protein, Camphor Resistance (CrcB)	?	332	FLUC	8
XP_020070087.1	hypothetical protein CYBJADRAFT_168114	Putative protein of unknown function	?	274	-	6
XP_020070587.1	hypothetical protein CYBJADRAFT_167581	Putative protein of unknown function	?	719	-	9
XP_020069469.1	MFS general substrate transporter	Putative protein of unknown function	?	562	-	13
XP_020071321.1	hypothetical protein CYBJADRAFT_161724	Protein of unknown function	-	494	-	5
XP_020071960.1	hypothetical protein CYBJADRAFT_166697	Protein of unknown function	-	278	-	7
XP_020069510.1	hypothetical protein CYBJADRAFT_15204	LMBR1-like membrane protein	-	616	-	9
XP_020069726.1	hypothetical protein CYBJADRAFT_168232	Transmembrane 9 superfamily member 1	-	627	-	9

XP_020070096.1	hypothetical protein CYBJADRAFT_168118	Protein of unknown function	-	265	-	4
XP_020070098.1	hypothetical protein CYBJADRAFT_173572	Protein of unknown function	-	394	-	4
XP_020070246.1	hypothetical protein CYBJADRAFT_90748	Protein of unknown function	-	278	-	4
XP_020070123.1	hypothetical protein CYBJADRAFT_173594	Domain of unknown function	-	157	-	4
XP_020070398.1	hypothetical protein CYBJADRAFT_167928	Domain of unknown function (DUF4149)	-	233	-	5
XP_020069044.1	DUF747-domain-containing protein	Domain of unknown function (DUF4149)	-	514	-	4
XP_020068741.1	MFS general substrate transporter	Protein of unknown function	-	587	-	11
XP_020073415.1	hypothetical protein CYBJADRAFT_170824	Putative protein of unknown function	-	339	-	5
XP_020073336.1	hypothetical protein CYBJADRAFT_170748	Putative protein of unknown function	-	333	-	4
XP_020071689.1	hypothetical protein CYBJADRAFT_193449	Putative protein of unknown function	-	516	-	7
XP_020070440.1	hypothetical protein CYBJADRAFT_167964	Putative protein of unknown function	-	291	-	6
XP_020072875.1	MFS general substrate transporter	Putative protein of unknown function	-	511	DHA1	12
XP_020070450.1	hypothetical protein CYBJADRAFT_181657	Putative protein of unknown function	-	316	-	4
XP_020070451.1	hypothetical protein CYBJADRAFT_181659	Putative protein of unknown function	-	366	-	4
XP_020070452.1	hypothetical protein CYBJADRAFT_190419	Putative protein of unknown function	-	405	-	4
XP_020070758.1	hypothetical protein CYBJADRAFT_173117	Domain of unknown function (DUF92)	-	293	-	6
XP_020068905.1	Zip-domain-containing protein	Domain of unknown function (DUF92)	-	986	-	6
XP_020068853.1	hypothetical protein CYBJADRAFT_169171	Putative protein of unknown function	-	659	-	7
XP_020068699.1	hypothetical protein CYBJADRAFT_30155	Putative protein of unknown function	-	352	-	5
XP_020070856.1	hypothetical protein CYBJADRAFT_126658	Putative protein of unknown function	-	406	-	9
XP_020070989.1	hypothetical protein CYBJADRAFT_184365	Protein of unknown function	-	192	-	6
XP_020071006.1	hypothetical protein CYBJADRAFT_80850	Protein of unknown function	-	159	-	4
XP_020073637.1	hypothetical protein CYBJADRAFT_182754	Protein of unknown function	-	410	-	6
XP_020071084.1	hypothetical protein CYBJADRAFT_162218	Protein of unknown function	-	416	-	4
XP_020068836.1	DUF221-domain-containing protein	Protein of unknown function	-	832	-	11
XP_020068772.1	TPT-domain-containing protein	Protein of unknown function	-	415	TPT	8
XP_020072309.1	TPT-domain-containing protein	Protein of unknown function	-	377	TPT	9
XP_020068462.1	PQ-loop-domain-containing protein	Protein of unknown function	-	286	-	6
XP_020073493.1	hypothetical protein CYBJADRAFT_170893	Protein of unknown function	-	239	-	4
XP_020071371.1	hypothetical protein CYBJADRAFT_172314	Protein of unknown function	-	621	-	6
XP_020071372.1	hypothetical protein CYBJADRAFT_167026	Protein of unknown function	-	549	-	6
XP_020068508.1	MFS general substrate transporter	Protein of unknown function	-	511	-	12
XP_020071659.1	hypothetical protein CYBJADRAFT_166419	Protein of unknown function	-	229	-	6
XP_020071414.1	hypothetical protein CYBJADRAFT_167052	Protein of unknown function	-	528	-	6
XP_020068808.1	hypothetical protein CYBJADRAFT_169127	Protein of unknown function	-	260	-	7
XP_020071446.1	MFS general substrate transporter	Protein of unknown function	-	446	-	11

XP_020068823.1	hypothetical protein CYBJADRAFT_27884	Protein of unknown function	-	568	-	7
XP_020071502.1	hypothetical protein CYBJADRAFT_73751	Protein of unknown function	-	407	-	8
XP_020068909.1	MFS general substrate transporter	Protein of unknown function	M	509	-	10
XP_020071660.1	hypothetical protein CYBJADRAFT_166420	NSG1, Phosphoprotein	ER	264	-	4
XP_020073458.1	hypothetical protein CYBJADRAFT_165701	Protein of unknown function	-	231	-	4
XP_020068743.1	hypothetical protein CYBJADRAFT_169070	Protein of unknown function	-	268	-	5
XP_020071984.1	hypothetical protein CYBJADRAFT_183638	Protein of unknown function	-	549	-	7
XP_020071798.1	hypothetical protein CYBJADRAFT_171773	Protein of unknown function	-	422	-	8
XP_020071987.1	MFS general substrate transporter	Protein of unknown function	V	548	-	10
XP_020068553.1	hypothetical protein CYBJADRAFT_169362	Protein of unknown function	-	181	-	4
XP_020072127.1	hypothetical protein CYBJADRAFT_149001	Protein of unknown function	-	682	-	5
XP_020072848.1	hypothetical protein CYBJADRAFT_122780	Protein of unknown function	-	178	-	4
XP_020073059.1	MFS general substrate transporter	Protein of unknown function	-	607	-	11
XP_020070923.1	auxin efflux carrier	Protein of unknown function	-	493	-	9
XP_020073046.1	hypothetical protein CYBJADRAFT_182207	Protein of unknown function	M	723	-	12
XP_020072505.1	hypothetical protein CYBJADRAFT_160886	Protein of unknown function	-	230	-	6
XP_020073165.1	hypothetical protein CYBJADRAFT_7427	Protein of unknown function	-	443	-	10
XP_020068636.1	hypothetical protein CYBJADRAFT_191752	Protein of unknown function	-	254	-	4
XP_020069259.1	RTA1-domain-containing protein	Protein of unknown function	-	413	-	7
XP_020069261.1	RTA1-domain-containing protein	Protein of unknown function	-	312	-	7
XP_020069299.1	zf-DHHC-domain-containing protein	Protein of unknown function	-	334	-	5
XP_020072441.1	zf-DHHC-domain-containing protein	Protein of unknown function	-	376	-	4
XP_020069445.1	TPT-domain-containing protein	Protein of unknown function	-	275	TPT	5
XP_020069518.1	pali-domain-containing protein	Protein of unknown function	-	623	-	5
XP_020069683.1	DUF887-domain-containing protein	Protein of unknown function	-	269	-	7
XP_020070089.1	UPF0005-domain-containing protein	Protein of unknown function	-	252	-	6
XP_020070094.1	auxin efflux carrier	Protein of unknown function	ER?	460	-	8
XP_020070109.1	membrane fusion mating protein FIG1	Protein of unknown function	-	267	-	4
XP_020070151.1	DUF580-domain-containing protein	Protein of unknown function	PM	505	-	11
XP_020070237.1	UNC-50-domain-containing protein	Protein of unknown function	GC	295	-	5
XP_020070279.1	TRP-domain-containing protein	Protein of unknown function	-	774	-	9
XP_020070336.1	transport protein SFT2	Protein of unknown function	GC	128	-	4
XP_020070426.1	auxin efflux carrier	Protein of unknown function	-	549	-	8
XP_020070454.1	PQ-loop-domain-containing protein	Protein of unknown function	-	240	-	6
XP_020070521.1	EXS-domain-containing protein	Protein of unknown function	-	334	-	5
XP_020071379.1	TMS membrane protein/tumor differentially expressed protein	Protein of unknown function	V	467	-	11
XP_020071381.1	EMP70 Emp70p	Protein of unknown function	-	617	-	9
XP_020071403.1	DUF846-domain-containing protein	Protein of unknown function	GC	187	-	4
XP_020071511.1	DUF221-domain-containing protein	Protein of unknown function	-	865	-	11

XP_020071823.1	mitochondrial carrier	Protein of unknown function	M	397	-	4
XP_020072346.1	SFT2-domain-containing protein	Protein of unknown function	-	213	-	4
XP_020072412.1	UPF0220-domain-containing protein	Protein of unknown function	V	174	-	4
XP_020072476.1	membrane protein	Protein of unknown function	-	296	-	4
XP_020072542.1	membrane protein	Protein of unknown function	-	390	-	7
XP_020072762.1	auxin efflux carrier	Protein of unknown function	-	421	-	10
XP_020073062.1	COPI associated	Protein of unknown function	GC	135	-	4
XP_020073192.1	mitochondrial carrier	Protein of unknown function	M	320	-	5
XP_020073287.1	HPP-domain-containing protein	Protein of unknown function	-	267	-	5
XP_020073303.1	DUF300-domain-containing protein	Protein of unknown function	-	421	-	7
XP_020073610.1	UPF0016-domain-containing protein	Protein of unknown function	V	358	-	5
XP_020068938.1	hypothetical protein CYBJADRAFT_174555	Protein of unknown function	-	178	-	4
XP_020069016.1	hypothetical protein CYBJADRAFT_168998	Protein of unknown function	-	304	-	8
XP_020069096.1	hypothetical protein CYBJADRAFT_21477	Protein of unknown function	-	223	-	4
XP_020069103.1	hypothetical protein CYBJADRAFT_168762	Protein of unknown function	-	497	-	4
XP_020070874.1	general substrate transporter	Low glucose sensor; SNF3	PM	538	HXT	12
XP_020069134.1	hypothetical protein CYBJADRAFT_130034	Chitin synthase III catalytic subunit	-	303	-	7
XP_020073183.1	transmembrane 9 superfamily member 4	Protein with a role in cellular adhesion and filamentous growth	-	631	-	9
XP_020069864.1	hypothetical protein CYBJADRAFT_190796	putative synovial apoptosis inhibitor	-	529	-	6
XP_020071045.1	longevity assurance proteins LAG1/LAC1	Ceramide synthase component	-	426	-	7
XP_020073057.1	PigN-domain-containing protein	Protein involved in GPI anchor synthesis	ER	930	-	16
XP_020071703.1	hypothetical protein CYBJADRAFT_166452	Regulator of phospholipase D	-	368	-	4
XP_020072134.1	hypothetical protein CYBJADRAFT_123391	Protein involved in spore wall assembly	-	1806	-	18
XP_020072277.1	hypothetical protein CYBJADRAFT_165992	Nucleoporin protein	N	578	-	7
XP_020072544.1	hypothetical protein CYBJADRAFT_123371	Protein required for normal cell wall, plasma membrane, cytoskeletal organization, endocytosis	PM	274	-	4
XP_020072719.1	MFS general substrate transporter	protein required for efflux of amino acids	V	524	MFS	9
XP_020073485.1	hypothetical protein CYBJADRAFT_165727	Inositol phospholipid synthesis	ER	280	-	6
XP_020073486.1	hypothetical protein CYBJADRAFT_165728	Inositol phospholipid synthesis	ER	254	-	6
XP_020071116.1	HlyIII-domain-containing protein	Membrane protein involved in zinc metabolism	-	513	ZIP	6
XP_020067853.1	protein CWH43	putative sensor/transporter protein involved in cell wall biogenesis		930	-	14
XP_020069976.1	hypothetical protein CYBJADRAFT_128528	putative zinc finger protein	ER	1023	-	10
XP_020068846.1	Sec61 protein	Essential subunit of Sec61 complex	ER	479		5
XP_020067910.1	Cu-transporting P-type ATPase	Element required for export of copper from the cytosol into an extracytosolic compartment	-	1169	-	8
XP_020068112.1	SecY protein	Subunit of the Ssh1 translocon complex	-	482	-	10
XP_020067922.1	batten's disease protein Cln3	Protein required for the ATP-dependent transport of arginine; cell homeostasis	V	377	-	7

XP_020068412.1	HSP70-domain-containing protein	Heat shock protein that is highly induced upon stress	V?	920	-	4
XP_020068718.1	Per1-like protein	Protein required for GPI-phospholipase A2 activity	ER	328	-	5
XP_020069175.1	Shr3 amino acid permease chaperone	Packaging chaperone required for incorporation of amino acid permeases into COPII coated vesicles for transport to the cell surface	ER	211	-	4
XP_020069758.1	GPI-anchored wall transfer protein 1	GPI-anchored wall transfer protein 1	-	500	-	11
XP_020069221.1	PIG-U-domain-containing protein	GPI transamidase subunit PIG-U	-	388	-	7
XP_020069597.1	DUF1753-domain-containing protein	Inositolphosphorylceramide synthase subunit	-	207	-	4
XP_020070589.1	vacuolar membrane protein that transits through the biosynthetic vacuolar protein sorting pathway	the biosynthetic vacuolar protein sorting pathway	V	1219	-	13
XP_020068140.1	hypothetical protein CYBJADRAFT_169650	Proteolipid membrane potential modulator	PM	150	-	4
XP_020068268.1	hypothetical protein CYBJADRAFT_169627	Protein for Inositol phospholipid synthesis	ER	276	-	6
XP_020070322.1	hypothetical protein CYBJADRAFT_91416	pH sensor component of the RIM101 pathway	PM	503	-	6
XP_020068298.1	hypothetical protein CYBJADRAFT_175196	ASI1 - anti-silencing protein (predicted)	N	597	-	6
XP_020071012.1	vacuolar protein sorting 55	unknown	V	129	-	4
XP_020071166.1	ER lumen protein retaining receptor	integral protein, a HDEL receptor	ER	212	-	5
XP_020067890.1	family A G protein-coupled receptor-like protein	Receptor-like protein	PM	312	-	7
XP_020067891.1	family A G protein-coupled receptor-like protein	Receptor-like protein	PM	271	-	7
XP_020070346.1	hypothetical protein CYBJADRAFT_173305	Putative G-protein coupled receptor	GC	514	-	8
XP_020072496.1	hypothetical protein CYBJADRAFT_188575	putative G protein-coupled glucose receptor	-	447	-	6
XP_020073219.1	fungal pheromone mating factor STE2G-protein-coupled receptor	Putative receptor	-	386	-	6

Note, initials stand for:

MBOAT	Membrane-bound O-acyltransferase family
APC	Amino acid-polyamine-organocation family
AceTr	Acetate Transporter family
P-RFT	The prokaryotic riboflavin transporter (p-rft) family
GPH	Glycoside-Pentoside-Hexuronide (GPH):Cation Symporter Family
PDR	The Pleiotropic Drug Resistance (PDR) Family
SP	Sugar porter family
FLUC	The camphor resistance or fluoride exporter (fluc) family
GT39	The integral membrane glycosyltransferase family 39 (gt39) family
ARN	Siderophore family
AAA	The ATP:ADP Antiporter (AAA) Family
ACS	The Anion:Cation Symporter (ACS) Family
CaCA	The Ca2+:Cation Antiporter (CaCA) Family
NCS1	Nucleobase:Cation Symporter-1 (NCS1) Family
PBRBC	The lead resistance fusion protein (pbrbc) family
CTR	Copper transporter family
OCTN	Organic cation transporter (OCTN) family
DHA1	Drug:H(+) antiporter DHA1 family
MATE	Multi antimicrobial extrusion (MATE) family
AAAP	The amino acid/auxin permease (aaap) family
SSS	Sodium:solute symporter (SSS) family
DAG	DHA2/ARN/GEX, DHA2 drug:H+ antiporters of family 2 , ARN siderophore transporters, GEX - glutathione exchangers family
ABC	ATP-binding cassette transporters
SLC	Solute carrier family
ENT	Equilibrative nucleoside transporter (ENT) family member
ZIP	The zinc (zn(2+))-iron (fe(2+)) permease (zip) family

OST	Oligosaccharyl transferase complex transporter family
GT-A	Glycosyltransferase family A (GT-A
RhaT	L-rhamnose transporter family
YAT	The Yeast Amino Acid Transporter (YAT) Family
OFA	Oxalate/formate antiporter family
TDT	Tellurite-resistance/Dicarboxylate Transporter Family
NRT	Nitrate transporter family
ILT	The Iron/Lead Transporter (ILT) Family
AMT	The ammonium transporter channel (amt) family
TPT	Triose-phosphate Transporter family
OPT	Oligopeptide transporter family
PIT	The inorganic phosphate transporter (pit) family
NAT/NCS2	Nucleobase-Ascorbate Transporter/Nucleobase-Cation Symporter family

Table S3. Docking parameters and data from the molecular docking studies

Table S3.1 Parameters obtained with HHPred for 3D-model construction.

Protein templates	PDB Hit	Organism	E-value	Similarity	Identity	Score	Probability (%)
ScJen1	6G9X	*Syntrophobacter fumaroxidans*	$2.4e^{-33}$	0.136	14%	281.6	100.0
CaJen2			$9.5e^{-37}$	0.116	13%	286.6	100.0
CjJen5			$1e^{-35}$	0.112	13%	270.2	100.0
CjJen6			$2e^{-36}$	0.171	15%	277.3	100.0
ScAto1	5YS3	*Citrobacter koseri*	$3.7e^{-32}$	0.566	35%	236.5	100.0
EcSatP			$1.5e^{-36}$	1.498	92%	235.4	100.0
CjAto2			$6.7e^{-33}$	0.538	32%	237.2	100.0
CjAto5			$7e^{-35}$	0.525	26%	252.0	100.0
CjAto6			$6e^{-35}$	0.568	30%	254.3	100.0
HsSLC5A8	3DH4	*Vibrio parahaemolyticus*	$4.6e^{-48}$	0.337	24%	425.4	100.0
CjSlc5			$2.2e^{-10}$	0.063	10%	109.2	99.6
HsSlc13A3	6WTW	*Lactobacillus acidophilus*	$7.5e^{-26}$	0.306	22%	245.0	99.9
CjSlc13			$3.9e^{-30}$	0.211	17%	298.5	100.0

Table S3.2 Residues of Ato homologs presenting strong intramolecular interactions with lactate, succinate and citrate identified by molecular docking studies.

3D-Protein templates	Lactate	Succinate		Citrate
	(-1)	*(-1)*	*(-2)*	*(-3)*
ScATO1	R111; T238; N179; T222; T102; S106; E140; T209; N255	T238; N179; R111; T222; T102; Q133; S208; N145; N255; T209; E140	R111; N179; T238; T222; T209; S208; N145	N255; T209; S208; N89; N145; Q133; T238; H230; R111; N179
CjATO2	K244; S189; N127; N91; T203; I97; T99; Y158	N236; K244; S189; C185; T203; N91; T84; T99; D150; R214; Q219; S211; W144; Y158	S189; K244; N127; T203; W144; T99; Q219; R214; Y158	*ND*
CjATO5	*ND*	E125; K221	K221	N137; S199; K198; N81; K221; Q103; T229; K230; G220
CjATO6	*ND*	N139; E134; T209; Q127; N103; Y170; K224; T111	N139; N103; K224; Y170; T111	N139; C202; N103; K224; T111

Note: *ND* - not determined.

Table S3.3 Average of the binding affinity values [kcal/mol] calculated with PyRx software for the docking of Ato proteins with the distinct charged substrates tested.

3D-Protein templates	Average of binding affinities (kcal/mol) at different binding sites						
	Lactate (-1)						
	S4			**S3**	**S2**	**S1**	
	a	**b**	**c**			**a**	**b**
ScAto1	-	-3.8	-	-4.1	-3.9	-3.3	-3.1
EcSatP	-3	-	-	-3.6	-3.8	-3	-
CjAto2	-3.4	-	-3.4	-4.2	-2.6	-3.4	-3.1
	Succinate (-1)						
	S4			**S3**	**S2**	**S1**	
	a	**b**	**c**			**a**	**b**
ScAto1	-	-4.0	-	-5	-4.5	-3.9	-3.8
EcSatP	-3.6	-	-	-4.5	-4.6	-4.2	-
CjAto2	-4.2	-	-4.2	-4.7	-3.2	-3.9	-4
CjAto5	-3.8	-	-	-3.9	**-4.7**	-	-
CjAto6	-3.7	-	-	-4.3	**-5.3**	-3.9	-
	Succinate (-2)						
	S4			**S3**	**S2**	**S1**	
	a	**b**	**c**			**a**	**b**
ScAto1	-	-4.0	-	-5	-4.4	-3.7	-
EcSatP	-3.4	-	-	-4.4	**-4.7**	-4.3	-
CjAto2	-4	-	-4.1	-4.8	-3.1	**-4.1**	-4

CjAto5	-3.8	-	-	-3.9	**-4.6**	-	-
CjAto6	-3.7	-	-	-4.4	**-5.3**	**-3.9**	-

	Citrate (-3)						
	S4			**S3**	**S2**	**S1**	
	a	**b**	**c**			**a**	**b**
ScAto1	-	-5.0	-	-6.1	-3.7	-4.6	-
EcSatP	-4.2	-	-	-5	**-4.3**	-4.2	-
CjAto5	-4	-4	-	-3.5	**-5.5**	-4.5	-4.2
CjAto6	-4.5	-	-	-4.9	**-4.5**	**-4.7**	-

Table S3.4 Residues of Jen homologs presenting strong intramolecular interactions with lactate, succinate and citrate identified by molecular docking studies.

3D-Protein templates	**Lactate**	**Succinate**		**Citrate**
	(-1)	*(-1)*	*(-2)*	*(-3)*
ScJEN1	N379; W473; Y497; S189; S504; G500; A274; Y250; Y273; Y246; F277; R188; A493	N379; W473; R188; H383; S508; G500; S189; F277; Y273; Y246; A493; S250; A490; T178	N379; S189; W473; R188; Q386; S508; Y466; S250; Y246	S508; H383; T178; Y273; N412; Q386; F277; D387; S504; W473; Y497; N379; R188; S189; S250; R188; S189; S250; R486; Y296; Y246; R262
CjJEN5	*ND*	R56; E310; P309; A313; E107; R114; S115; R312; K171; T111; P110; S319; S102; W196; A200; Q293; Y129; A126; Y125; V285; Y129	R312; K171; R56; R114; S115; S102; S319; Y98; G323; Q293; K27; S335; S331	R114; A113; S115; P110; S307; R383; R255; G323; A200; N203; Q293; S331; Y129; T215; N236
CjJEN6	N273; W366; N394; S397; R99; Y184; Y188; Y390; P379; R115; E166; A383; R382	W366; N394; N273; S397; R99; Y157; W60; R115; R382; A383; S161; K116	N273; N394; W366; R99; W60; Y157; R115; A383; S161; T163	N273; W366; R99; S397; Y184; S89; T405; N394; A383; R115; T163; L159; S161

Note: *ND* - not determined.

Table S3.5 Average of the binding affinity values [kcal/mol] calculated with PyRx software for the docking of Jen proteins with the distinct charged substrates tested.

3D-Protein templates	Average of binding affinities (kcal/mol) at different binding sites							
	Acetate (-1)							
	1			**2**	**3**		**4**	
	a	**b**	**c**		**a**	**b**	**a**	**b**
ScJen1	-2.9	-	-	-3.6	-3.3	-	-	-
CjJen6	-3.1	-	-	-3.4	-2.7	-	-	-2.8
	Lactate (-1)							
	1			**2**	**3**		**4**	
	a	**b**	**c**		**a**	**b**	**a**	**b**
ScJen1	-4.0	-	-3.8	-4.4	-3.6	-	-	-
CjJen6	-	-	-	-4.4	-3.2	-	-	-3.8
	Succinate (-1)							
	1			**2**	**3**		**4**	
	a	**b**	**c**		**a**	**b**	**a**	**b**
ScJen1	-4.5	-	-4.2	-5.5	-4.9	-	-	-
CaJen2	-4.3	-	-	-3.4	-5.1	-	-	-
CjJen5	-	-4.0	-3.9	-4.2	-4.6	-4.5	-3.4	-3.7
CjJen6	**-4.6**	-	-	-5.1	-4.1	-	-	**-4.7**
	Succinate (-2)							
	1			**2**	**3**		**4**	
	a	**b**	**c**		**a**	**b**	**a**	**b**
ScJen1	-	-	-4.4	-5.7	-4.8	-	-	-
CaJen2	-4.0	-	-	-3.5	**-5.1**	-	-	-
CjJen5	-	-4.1	-3.8	-4.0	-4.5	-4.7	-3.1	-3.3
CjJen6	**-4.1**	-	-	-5.0	-4.1	-	-	**-4.2**
	Citrate (-3)							
	1			**2**	**3**		**4**	
	a	**b**	**c**		**a**	**b**	**a**	**b**
ScJen1	-4.8	-	-	-4.2	-3.0	-	-2.0	-
CjJen5	-	-4.4	-4.4	-4.8	-5.7	-	-3.7	-
CjJen6	**-5.7**	**-5.1**	-	**-5.9**	-2.3	-	-	**-4.4**

Table S3.6 Residues of CjSlc5 protein presenting strong intramolecular interactions with succinate and citrate identified by molecular docking studies.

3D-Protein template	Succinate		Citrate		
	(-1)	*(-2)*	*(-1)*	*(-2)*	*(-3)*
CjSLC5	G26; S221; F222; T38; Q212; S50; G31; Q27; Y102; Y303; L351; D79; L295	S221; F222; Q212; T38; S50; Q27; Y102	S221; F222; G219; G26; T38; D349; L351; S50; S56; Q212; Y303; Y102; D79; G219; G26; Q27; G31; K302; L299	F222; S221; G219; Q27; S50; T38; Q212; L351; G31; Y102; D79	F222; S221; T38; S50; L351; Q212; Q27; Y102; G219; K302

Table S3.7 Average of the binding affinity values [kcal/mol] calculated with PyRx software for the docking of Slc5 member proteins with the distinct charged substrates tested.

3D-Protein templates	Average of binding affinities (kcal/mol) at different binding sites					
	Succinate (-1)					
	1		**2**		**3**	**4**
	a	**b**	**a**	**b**		
HsSlc5A8	-4.1	-	-	-4.4	-4.5	-4.3
CjSlc5	-4.3	-4.0	-3.7	-	-3.4	-3.8
	Succinate (-2)					
	1		**2**		**3**	**4**
	a	**b**	**a**	**b**		
HsSlc5A8	-4.2	-	-	-4.3	-4.5	-4.1
CjSlc5	-4.3	-	-3.8	-	-3.4	-3.7
	Citrate (-1)					
	1		**2**		**3**	**4**
	a	**b**	**a**	**b**		
HsSlc5A8	-4.7	-4.7	-4.8	-	-5.8	-5.0
CjSlc5	-5.0	-4.3	-4.9	-	-4.1	-5.0
	Citrate (-2)					
	1		**2**		**3**	**4**
	a	**b**	**a**	**b**		
HsSlc5A8	-4.3	-4.4	-4.8	-	-5.8	-5.0
CjSlc5	-5.0	-	-4.7	-	-4.1	-4.7
	Citrate (-3)					
	1		**2**		**3**	**4**
	a	**b**	**a**	**b**		
HsSlc5A8	-4.2	-4.1	-4.7	-	-5.8	-4.9
CjSlc5	-5.2	-4.4	-4.4	-	-4.3	-4.7

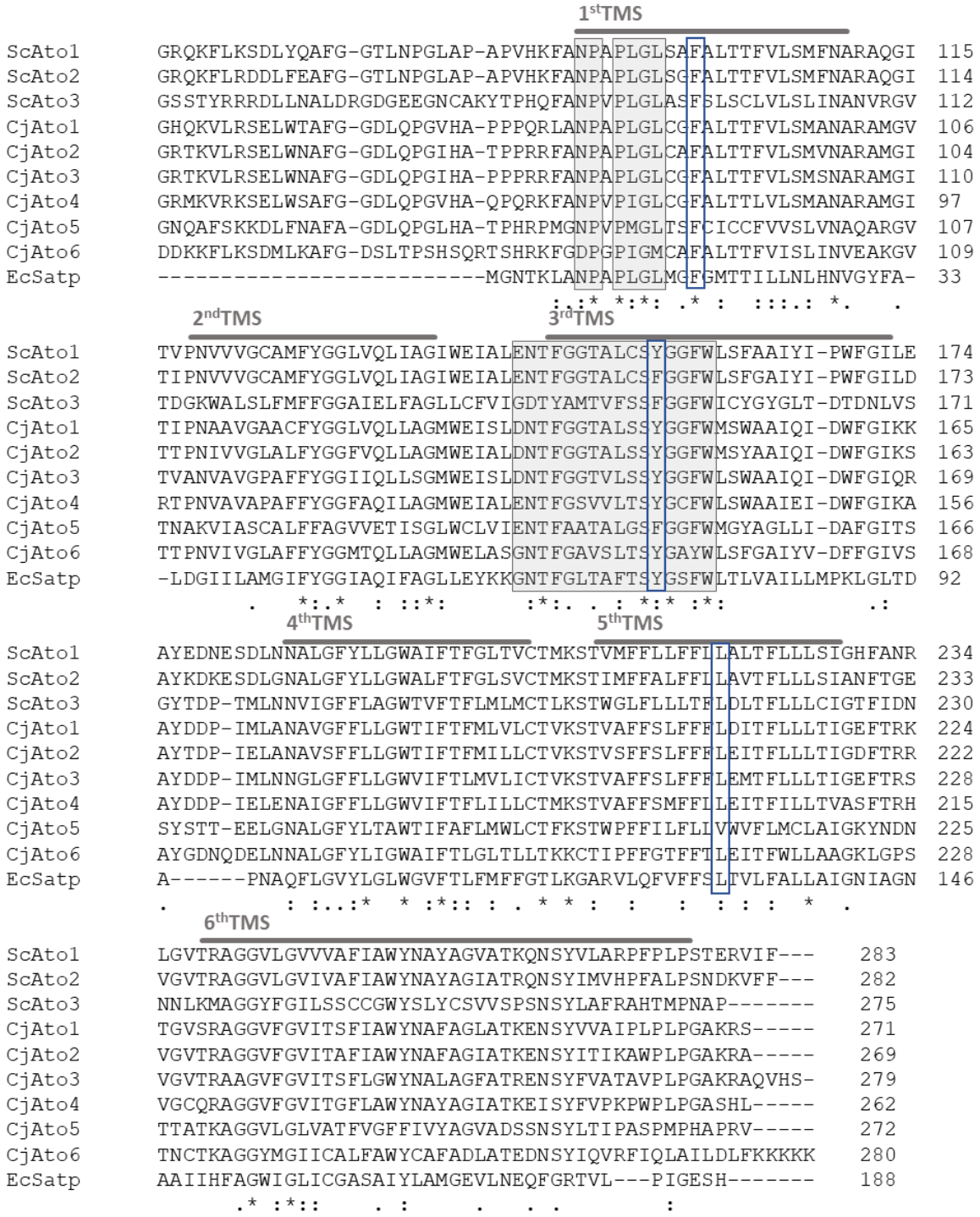
Table S3.8 Residues of CjSlc13 protein presenting strong intramolecular interactions with citrate identified by molecular docking studies.

3D-Protein templates	Citrate		
	(-1)	*(-2)*	*(-3)*
CjSLC13	K701; S461; S457; Q462; T467; S466; T746; V743; A693; S696; S799; H745; V795; N684; Y530; Q534; F685; R538; S465; S482	T746; Q462; S461; K701; S457; T467; S466; H745; V743; S799; S696; Y530; N684; F475; Q534; R538; F685	S466; K701; S457; Q462; T467; T746; H745; S696; M800; Y530; N684; R538; Q534; F685

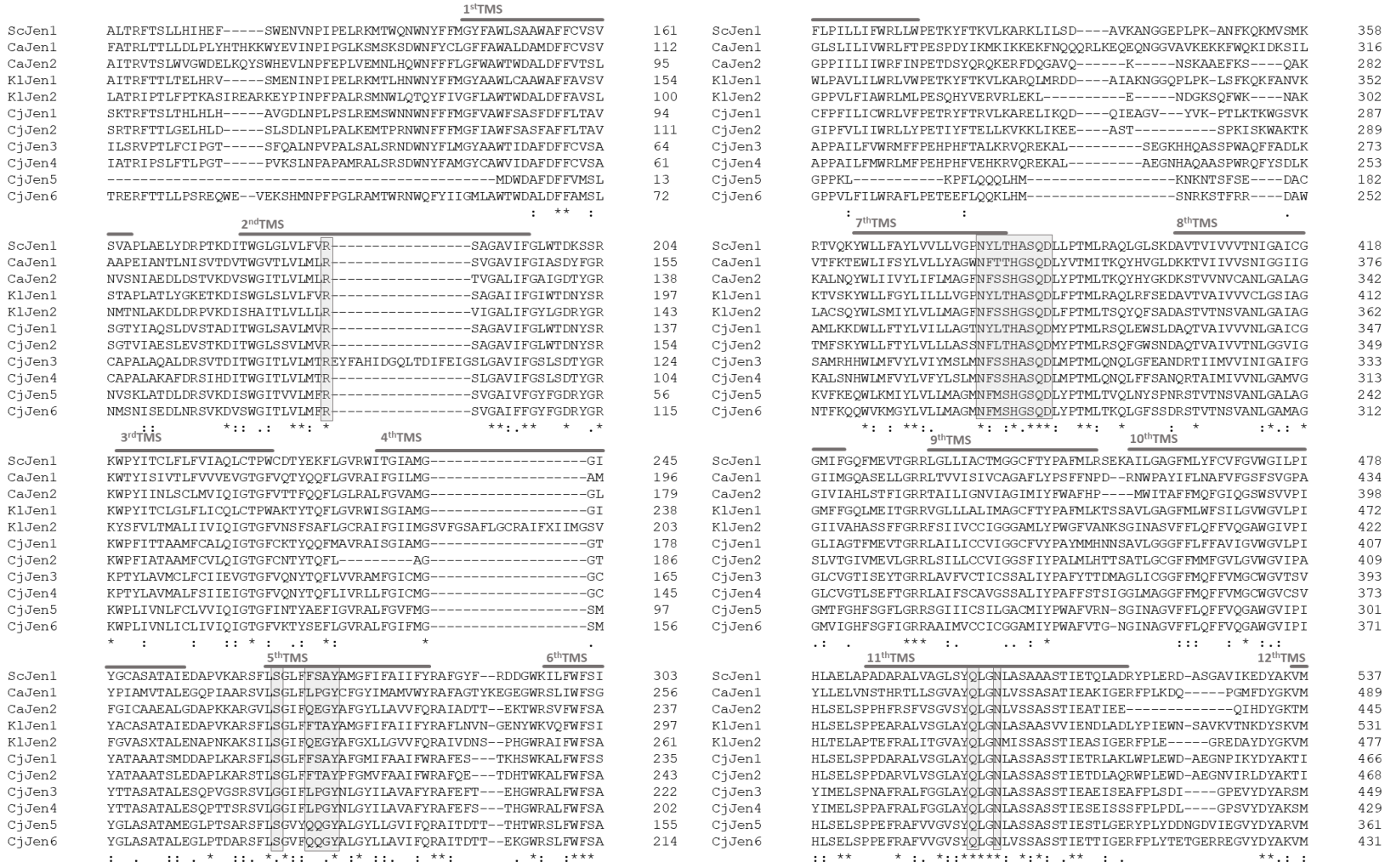
Table S3.9 Average of the binding affinity values [kcal/mol] calculated with PyRx software for the docking of Slc13 proteins with citrate.

3D-Protein templates	Average of binding affinities (kcal/mol) at different binding sites				
	Citrate (-1)				
	1		2	3	4
	a	b			
HsSlc13A3	-5.8	-	-	-4.6	-4.8
CjSlc13	-4.6	-4.7	-5.4	-4.5	-5.3
	Citrate (-2)				
	1		2	3	4
	a	b			
HsSlc13A3	-5.6	-	-	-4.9	-4.8
CjSlc13	-4.4	-4.4	-4.6	-4.1	-5.0
	Citrate (-3)				
	1		2	3	4
	a	b			
HsSlc13A3	-5.7	-	-	-4.9	-4.8
CjSlc13	-4.3	-4.3	-	-4.0	-4.8

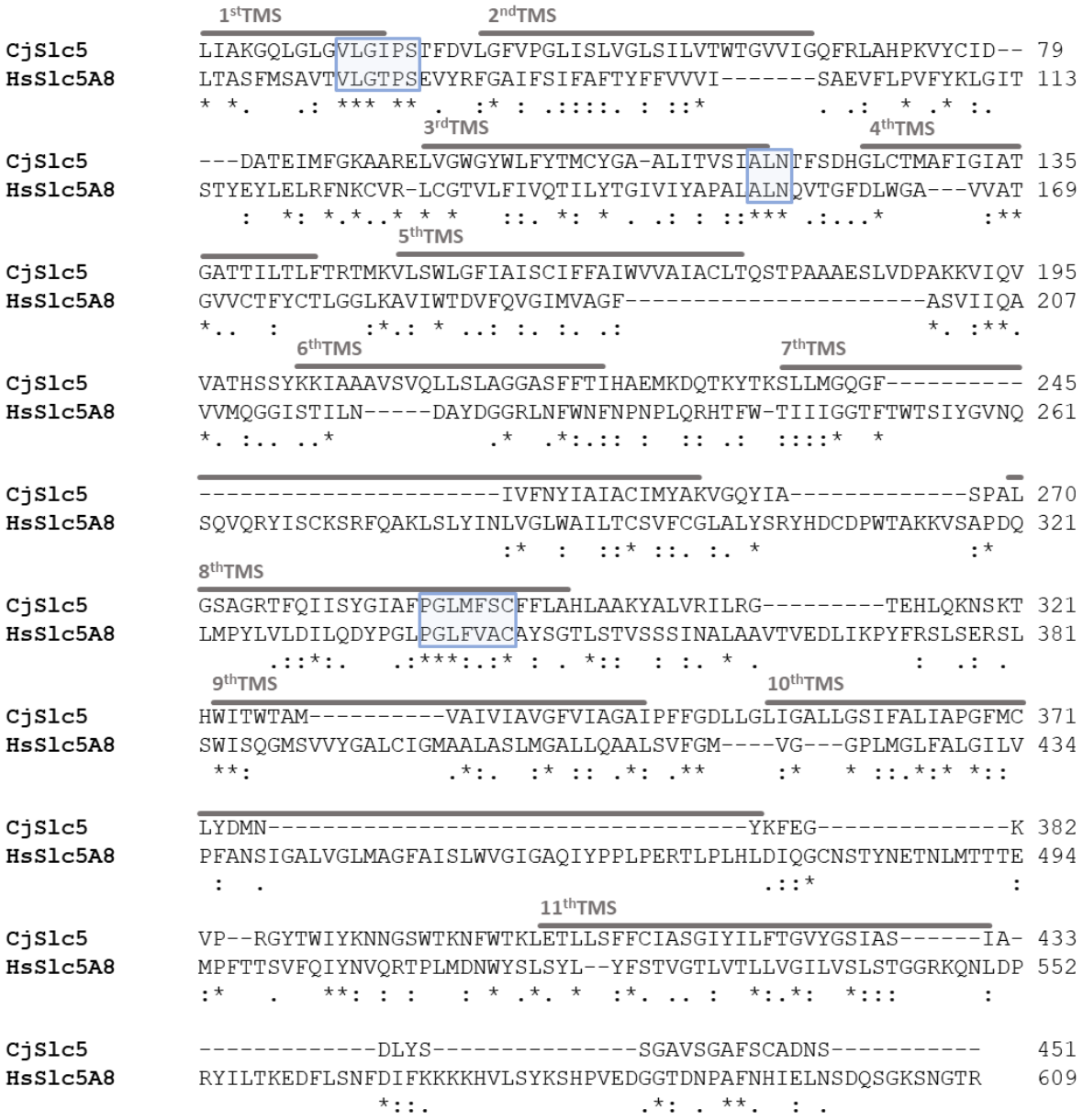
Figure S1. Multiple-sequence alignments



Supplementary figure S1.1 Multiple sequence alignment of Ato homologs in *Saccharomyces cerevisiae* (Ato1, Ato2, Ato3), *C. jadinii* (six homologs) and the *Escherichia coli* SatP. The sequence alignment was built with ClustalOmega (https://www.ebi.ac.uk/Tools/msa/clustalo/). Localization of transmembrane segments (TMSs) was predicted by the PSI/TM-Coffee (http://tcoffee.crg.cat/apps/tcoffee/do:tmcoffee). Grey background highlight the previously identified and characterized signature motifs of the AceTr family. Blue rectangles indicate residues from the narrowest hydrophobic constriction site F98-Y155-L219 (refer to ScAto1p) (Qiu et al., 2018).



Supplementary figure S1.2 Multiple sequence alignment of Jen homologs in *Saccharomyces cerevisiae*, *Candida albicans* (CaJen1 and CaJen2), *Kluyveromyces lactis* (KlJen1 and KlJen2) and C. jadinii (six homologs). The sequence alignment was built with ClustalOmega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and the localization of transmembrane segments (TMSs) predicted by the PSI/TM-Coffee (http://tcoffee.crg.cat/apps/tcoffee/do:tmcoffee). Grey background highlights the previously identified and characterized motifs of the SHS family for carboxylate transporter affinity.



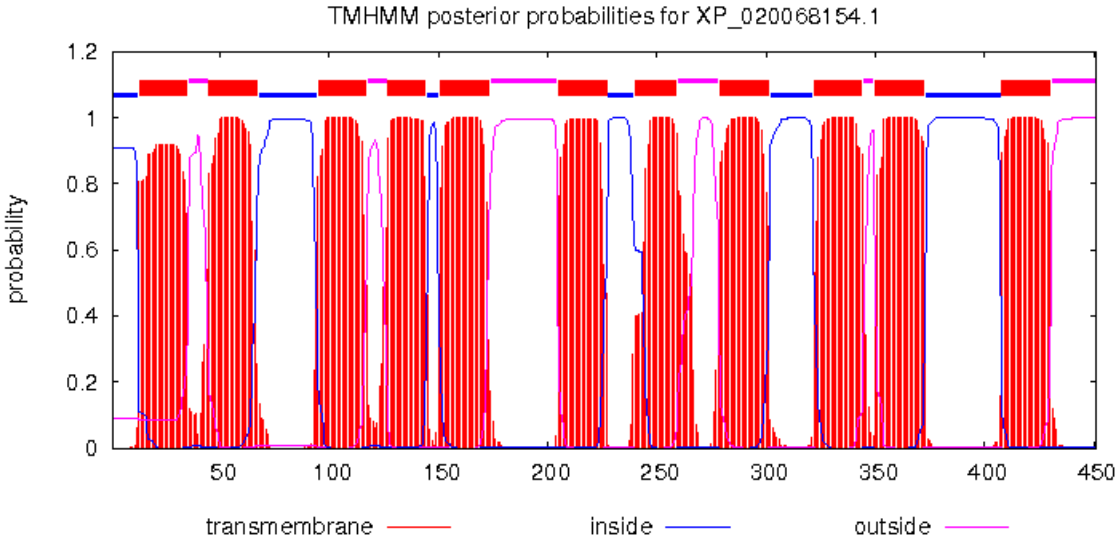
Supplementary figure S1.3 Double alignment of the protein sequences of the *Homo sapiens* SLC5 member 8 (NP_666018.3) and *Cyberlindnera jadinii* SLC5 homolog (XP_020068154.1) was built with ClustalOmega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and the localization of transmembrane segments (TMSs) predicted by the PSI/TM-Coffee (http://tcoffee.crg.cat/apps/tcoffee/do:tmcoffee).

Supplementary figure S1.4 Double alignment of the protein sequences of the *Homo sapiens* SLC13 member 3 (NP_001011554.1) and *Cyberlindnera jadinii* SLC13 homolog (XP_020069270.1) was built with ClustalOmega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and the localization of transmembrane segments (TMSs) predicted by the PSI/TM-Coffee (http://tcoffee.crg.cat/apps/tcoffee/do:tmcoffee). Blue background highlight the identified and/or characterized motifs of DASS transporter family with relevance for ligand and sodium binding.

Figure S2. Topology prediction

- Membrane topology predicted by TMHMM for SLC members

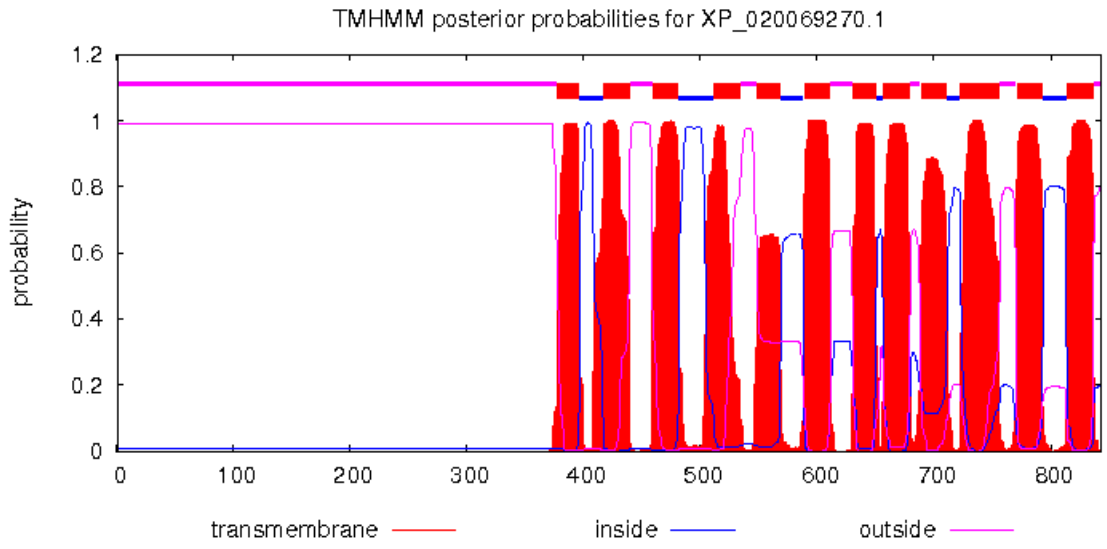
```
# XP_020068154.1 Length: 451
# XP_020068154.1 Number of predicted TMHs:  11
# XP_020068154.1 Exp number of AAs in TMHs: 242.31863
# XP_020068154.1 Exp number, first 60 AAs:  37.33048
# XP_020068154.1 Total prob of N-in:        0.90998
# XP_020068154.1 POSSIBLE N-term signal sequence
XP_020068154.1  TMHMM2.0       inside       1    12
XP_020068154.1  TMHMM2.0       TMhelix     13    35
XP_020068154.1  TMHMM2.0       outside     36    44
XP_020068154.1  TMHMM2.0       TMhelix     45    67
XP_020068154.1  TMHMM2.0       inside      68    94
XP_020068154.1  TMHMM2.0       TMhelix     95   117
XP_020068154.1  TMHMM2.0       outside    118   126
XP_020068154.1  TMHMM2.0       TMhelix    127   144
XP_020068154.1  TMHMM2.0       inside     145   150
XP_020068154.1  TMHMM2.0       TMhelix    151   173
XP_020068154.1  TMHMM2.0       outside    174   204
XP_020068154.1  TMHMM2.0       TMhelix    205   227
XP_020068154.1  TMHMM2.0       inside     228   239
XP_020068154.1  TMHMM2.0       TMhelix    240   259
XP_020068154.1  TMHMM2.0       outside    260   278
XP_020068154.1  TMHMM2.0       TMhelix    279   301
XP_020068154.1  TMHMM2.0       inside     302   321
XP_020068154.1  TMHMM2.0       TMhelix    322   344
XP_020068154.1  TMHMM2.0       outside    345   349
XP_020068154.1  TMHMM2.0       TMhelix    350   372
XP_020068154.1  TMHMM2.0       inside     373   407
XP_020068154.1  TMHMM2.0       TMhelix    408   430
XP_020068154.1  TMHMM2.0       outside    431   451
```



Supplementary figure S2.1 CjSLC5 protein topology predicted.

```
# XP_020069270.1 Length: 843
# XP_020069270.1 Number of predicted TMHs:  12
# XP_020069270.1 Exp number of AAs in TMHs: 261.70215
# XP_020069270.1 Exp number, first 60 AAs:  0
# XP_020069270.1 Total prob of N-in:        0.00723
XP_020069270.1  TMHMM2.0       outside      1   376
```

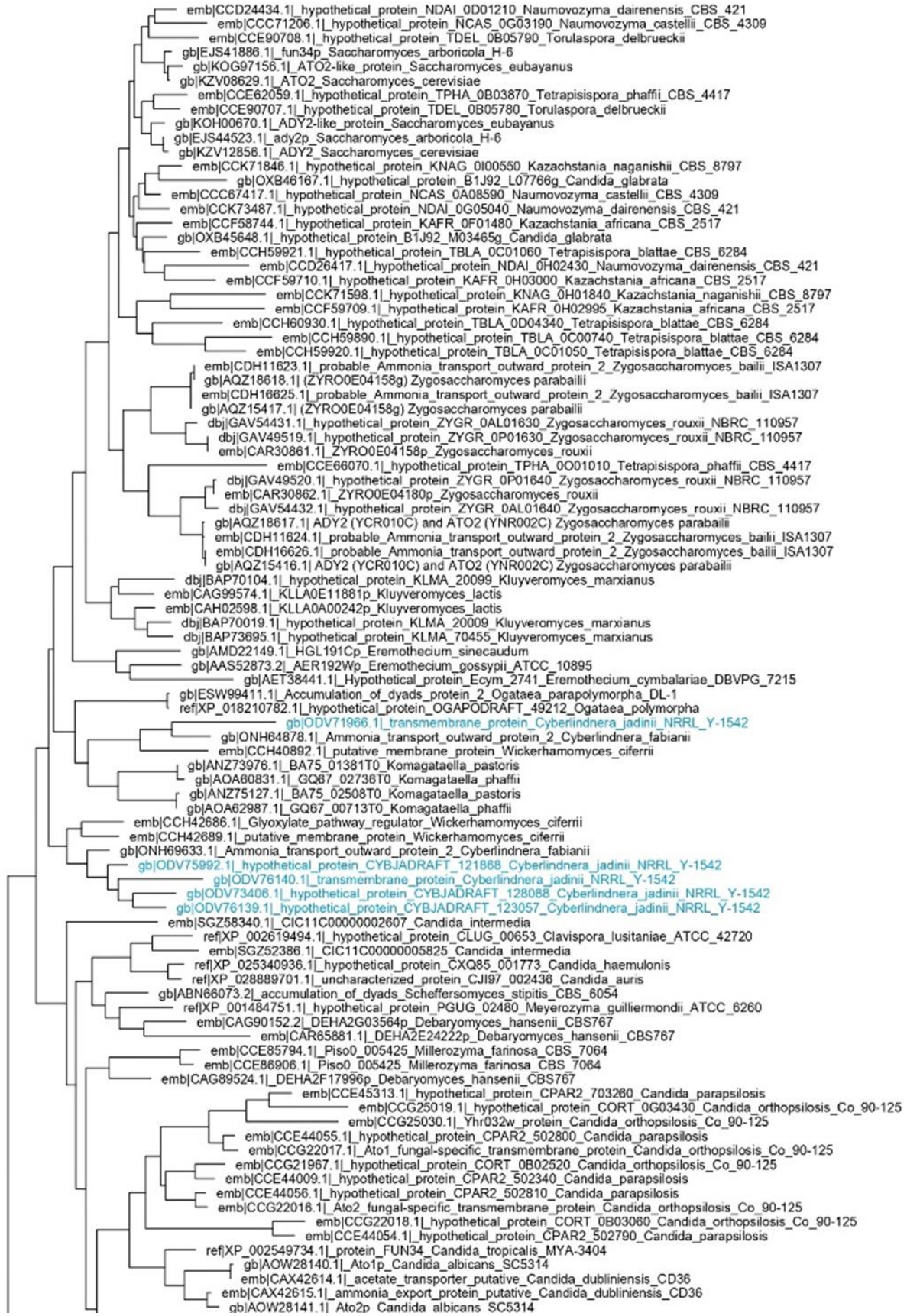
XP_020069270.1	TMHMM2.0	TMhelix	377	396
XP_020069270.1	TMHMM2.0	inside	397	416
XP_020069270.1	TMHMM2.0	TMhelix	417	439
XP_020069270.1	TMHMM2.0	outside	440	458
XP_020069270.1	TMHMM2.0	TMhelix	459	481
XP_020069270.1	TMHMM2.0	inside	482	511
XP_020069270.1	TMHMM2.0	TMhelix	512	534
XP_020069270.1	TMHMM2.0	outside	535	548
XP_020069270.1	TMHMM2.0	TMhelix	549	568
XP_020069270.1	TMHMM2.0	inside	569	588
XP_020069270.1	TMHMM2.0	TMhelix	589	611
XP_020069270.1	TMHMM2.0	outside	612	630
XP_020069270.1	TMHMM2.0	TMhelix	631	650
XP_020069270.1	TMHMM2.0	inside	651	656
XP_020069270.1	TMHMM2.0	TMhelix	657	679
XP_020069270.1	TMHMM2.0	outside	680	688
XP_020069270.1	TMHMM2.0	TMhelix	689	711
XP_020069270.1	TMHMM2.0	inside	712	722
XP_020069270.1	TMHMM2.0	TMhelix	723	756
XP_020069270.1	TMHMM2.0	outside	757	770
XP_020069270.1	TMHMM2.0	TMhelix	771	793
XP_020069270.1	TMHMM2.0	inside	794	813
XP_020069270.1	TMHMM2.0	TMhelix	814	836
XP_020069270.1	TMHMM2.0	outside	837	843

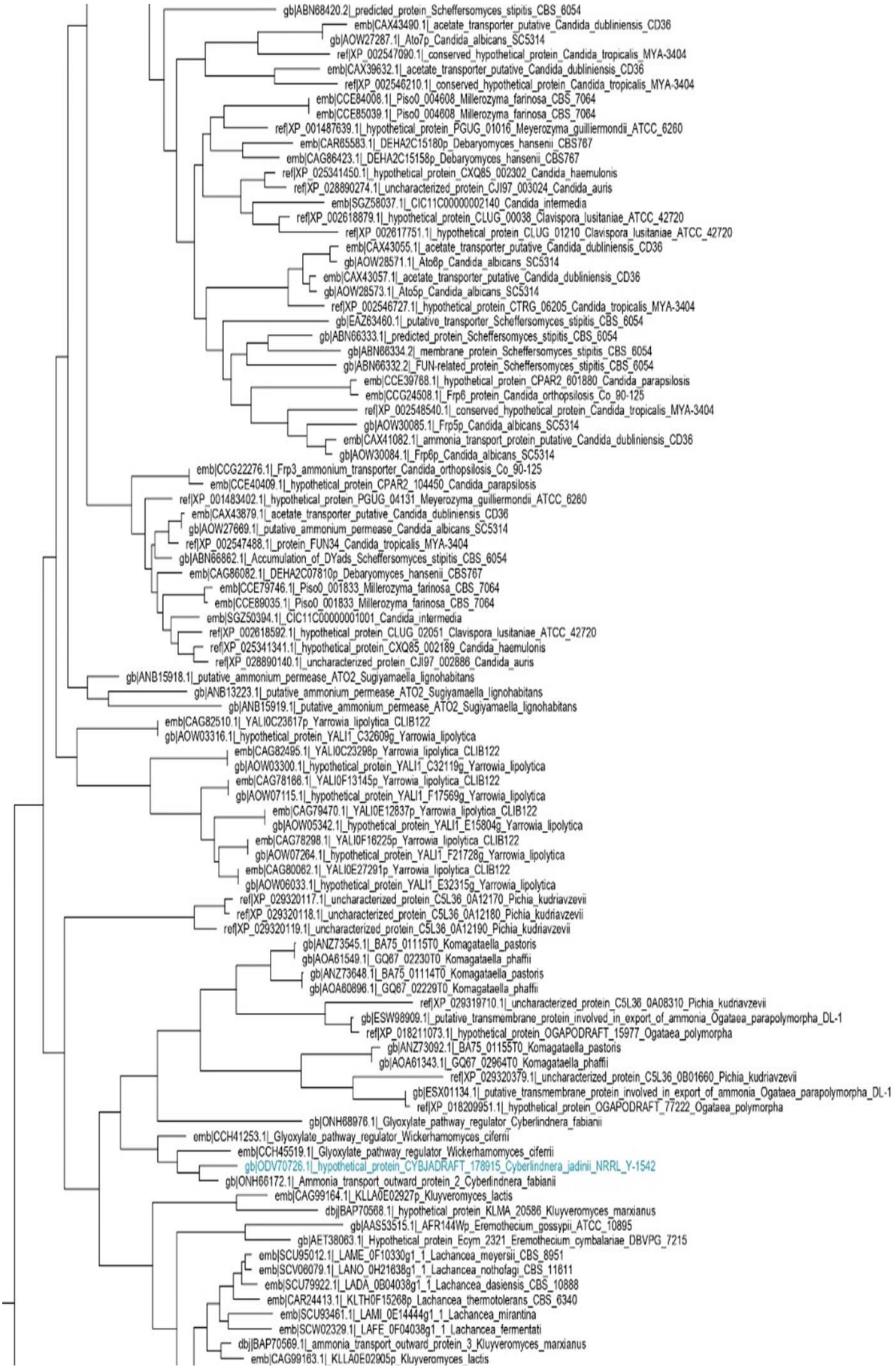


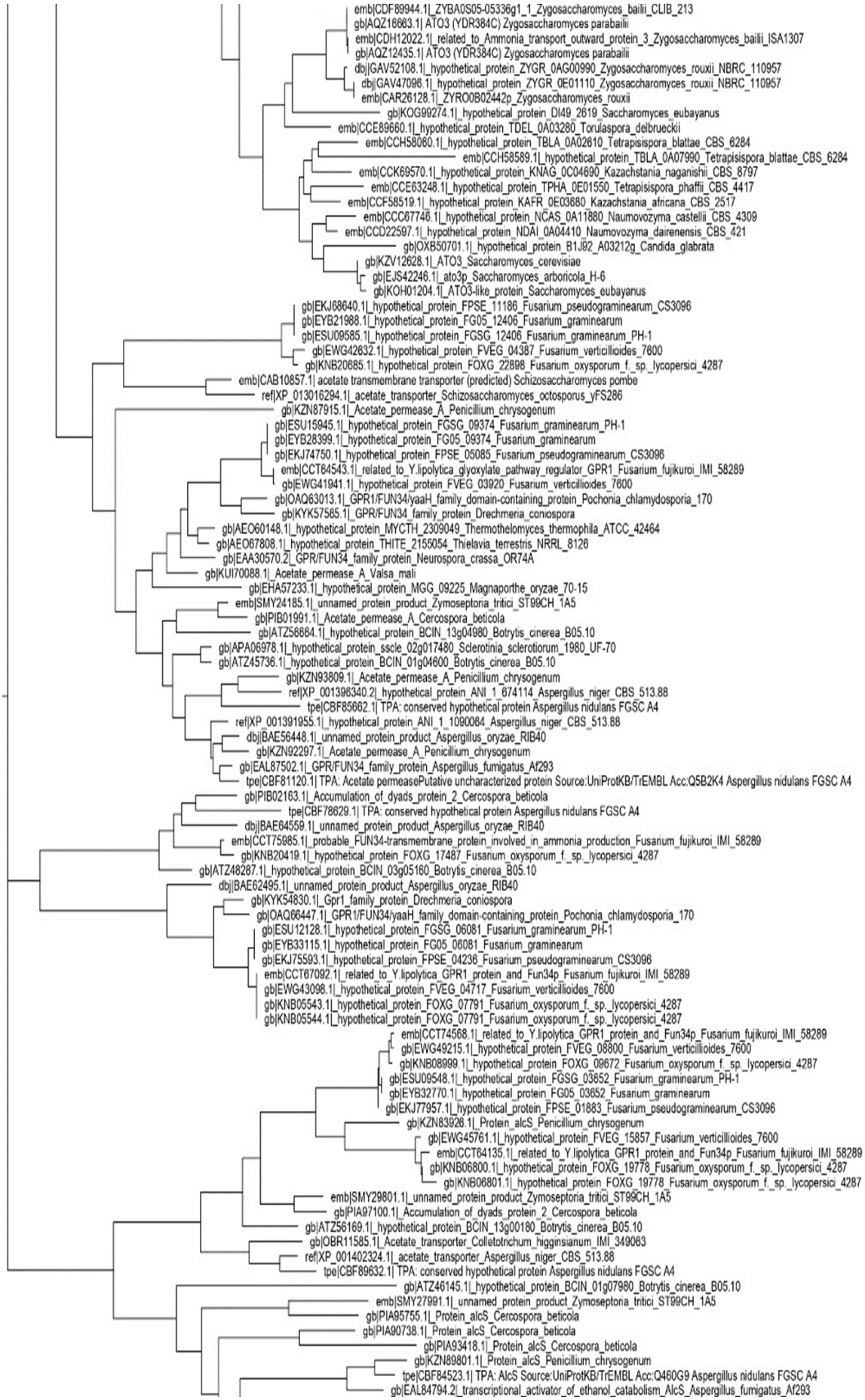
Supplementary figure S2.2 CjSLC13 protein topology predicted.

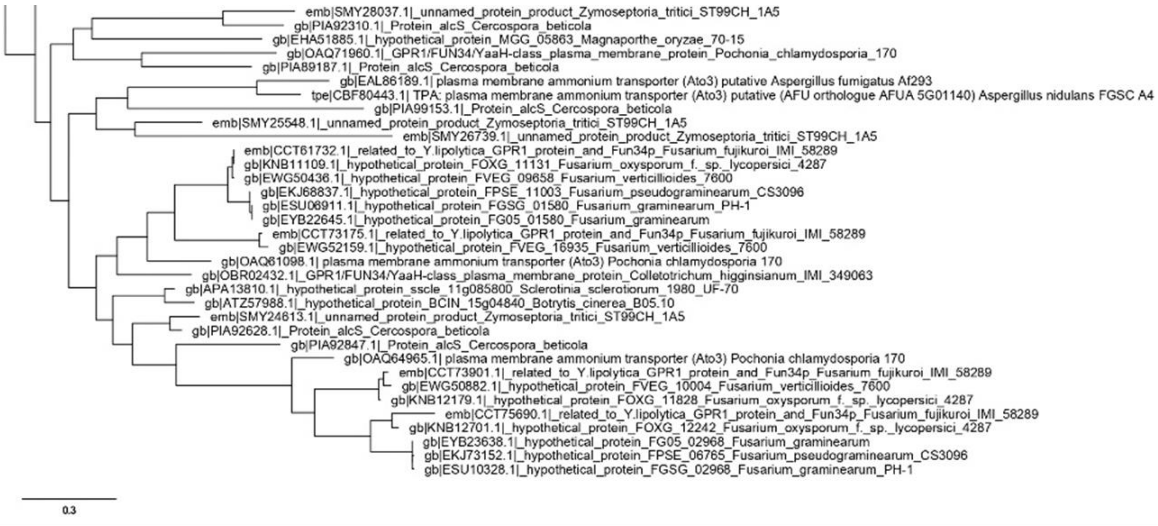
Figure S3. Phylogeny data belonging to AceTr, SHS, SSS and DASS transporter families

S3.1 Detailed CjATO-phylogenetic tree with all homolog nominations and bootstrap values (Ascomycota)

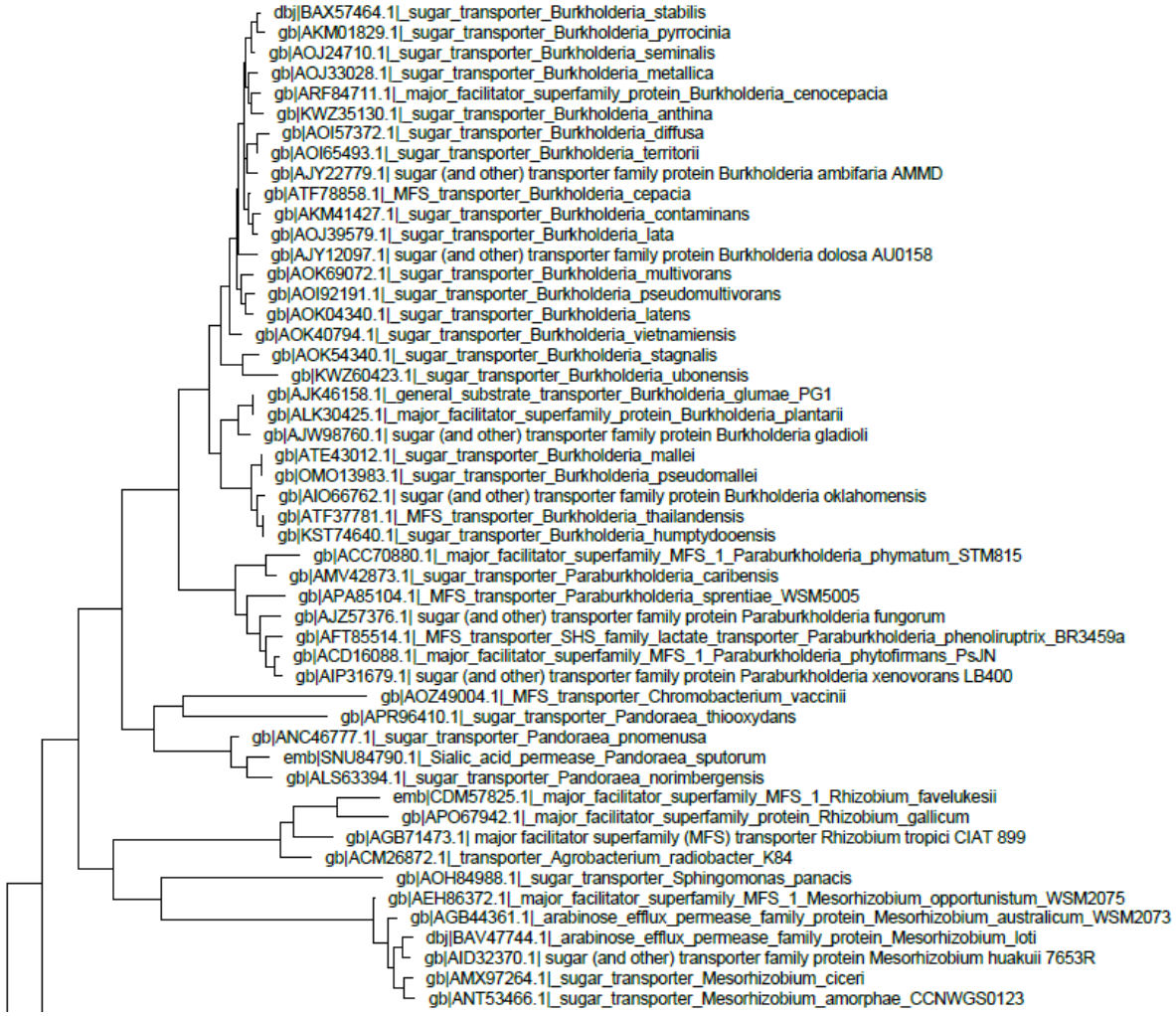



gb|ABN68420.2|_predicted_protein_Scheffersomyces_stipitis_CBS_6054
emb|CAX43490.1|_acetate_transporter_putative_Candida_dubliniensis_CD36
gb|AOW27287.1|_Ato7p_Candida_albicans_SC5314
ref|XP_002547090.1|_conserved_hypothetical_protein_Candida_tropicalis_MYA-3404
emb|CAX39632.1|_acetate_transporter_putative_Candida_dubliniensis_CD36
ref|XP_002546210.1|_conserved_hypothetical_protein_Candida_tropicalis_MYA-3404
emb|CCE84008.1|_Piso0_004608_Millerozyma_farinosa_CBS_7064
emb|CCE85039.1|_Piso0_004608_Millerozyma_farinosa_CBS_7064
ref|XP_001487639.1|_hypothetical_protein_PGUG_01016_Meyerozyma_guilliermondii_ATCC_6260
emb|CAR65583.1|_DEHA2C15180p_Debaryomyces_hansenii_CBS767
emb|CAG86423.1|_DEHA2C15158p_Debaryomyces_hansenii_CBS767
ref|XP_025341450.1|_hypothetical_protein_CXQ85_002302_Candida_haemulonis
ref|XP_028890274.1|_uncharacterized_protein_CJI97_003024_Candida_auris
emb|SGZ58037.1|_CIC11C00000002140_Candida_intermedia
ref|XP_002618879.1|_hypothetical_protein_CLUG_00038_Clavispora_lusitaniae_ATCC_42720
ref|XP_002617751.1|_hypothetical_protein_CLUG_01210_Clavispora_lusitaniae_ATCC_42720
emb|CAX43055.1|_acetate_transporter_putative_Candida_dubliniensis_CD36
gb|AOW28571.1|_Ato8p_Candida_albicans_SC5314
emb|CAX43057.1|_acetate_transporter_putative_Candida_dubliniensis_CD36
gb|AOW28573.1|_Ato5p_Candida_albicans_SC5314
ref|XP_002546727.1|_hypothetical_protein_CTRG_06205_Candida_tropicalis_MYA-3404
gb|EAZ63460.1|_putative_transporter_Scheffersomyces_stipitis_CBS_6054
gb|ABN66333.1|_predicted_protein_Scheffersomyces_stipitis_CBS_6054
gb|ABN66334.2|_membrane_protein_Scheffersomyces_stipitis_CBS_6054
gb|ABN66332.2|_FUN-related_protein_Scheffersomyces_stipitis_CBS_6054
emb|CCE39788.1|_hypothetical_protein_CPAR2_601880_Candida_parapsilosis
emb|CCG24508.1|_Frp6_protein_Candida_orthopsilosis_Co_90-125
ref|XP_002548540.1|_conserved_hypothetical_protein_Candida_tropicalis_MYA-3404
gb|AOW30085.1|_Frp5p_Candida_albicans_SC5314
emb|CAX41082.1|_ammonia_transport_protein_putative_Candida_dubliniensis_CD36
gb|AOW30084.1|_Frp6p_Candida_albicans_SC5314
emb|CCG22276.1|_Frp3_ammonium_transporter_Candida_orthopsilosis_Co_90-125
emb|CCE40409.1|_hypothetical_protein_CPAR2_104450_Candida_parapsilosis
ref|XP_001483402.1|_hypothetical_protein_PGUG_04131_Meyerozyma_guilliermondii_ATCC_6260
emb|CAX43879.1|_acetate_transporter_putative_Candida_dubliniensis_CD36
gb|AOW27669.1|_putative_ammonium_permease_Candida_albicans_SC5314
ref|XP_002547488.1|_protein_FUN34_Candida_tropicalis_MYA-3404
gb|ABN66862.1|_Accumulation_of_DYads_Scheffersomyces_stipitis_CBS_6054
emb|CAG86082.1|_DEHA2C07810p_Debaryomyces_hansenii_CBS767
emb|CCE79746.1|_Piso0_001833_Millerozyma_farinosa_CBS_7064
emb|CCE89035.1|_Piso0_001833_Millerozyma_farinosa_CBS_7064
emb|SGZ50394.1|_CIC11C00000001001_Candida_intermedia
ref|XP_002618592.1|_hypothetical_protein_CLUG_02051_Clavispora_lusitaniae_ATCC_42720
ref|XP_025341341.1|_hypothetical_protein_CXQ85_002189_Candida_haemulonis
ref|XP_028890140.1|_uncharacterized_protein_CJI97_002886_Candida_auris
gb|ANB15918.1|_putative_ammonium_permease_ATO2_Sugiyamaella_lignohabitans
gb|ANB13223.1|_putative_ammonium_permease_ATO2_Sugiyamaella_lignohabitans
gb|ANB15919.1|_putative_ammonium_permease_ATO2_Sugiyamaella_lignohabitans
emb|CAG82510.1|_YALI0C23617p_Yarrowia_lipolytica_CLIB122
gb|AOW03316.1|_hypothetical_protein_YALI1_C32609g_Yarrowia_lipolytica
emb|CAG82495.1|_YALI0C23298p_Yarrowia_lipolytica_CLIB122
gb|AOW03300.1|_hypothetical_protein_YALI1_C32119g_Yarrowia_lipolytica
emb|CAG78186.1|_YALI0F13145p_Yarrowia_lipolytica_CLIB122
gb|AOW07115.1|_hypothetical_protein_YALI1_F17569g_Yarrowia_lipolytica
emb|CAG79470.1|_YALI0E12837p_Yarrowia_lipolytica_CLIB122
gb|AOW05342.1|_hypothetical_protein_YALI1_E15804g_Yarrowia_lipolytica
emb|CAG78298.1|_YALI0F16225p_Yarrowia_lipolytica_CLIB122
gb|AOW07264.1|_hypothetical_protein_YALI1_F21728g_Yarrowia_lipolytica
emb|CAG80082.1|_YALI0E27291p_Yarrowia_lipolytica_CLIB122
gb|AOW06033.1|_hypothetical_protein_YALI1_E32315g_Yarrowia_lipolytica
ref|XP_029320117.1|_uncharacterized_protein_C5L36_0A12170_Pichia_kudriavzevii
ref|XP_029320118.1|_uncharacterized_protein_C5L36_0A12180_Pichia_kudriavzevii
ref|XP_029320119.1|_uncharacterized_protein_C5L36_0A12190_Pichia_kudriavzevii
gb|ANZ73545.1|_BA75_01115T0_Komagataella_pastoris
gb|AOA61549.1|_GQ67_02230T0_Komagataella_phaffii
gb|ANZ73648.1|_BA75_01114T0_Komagataella_pastoris
gb|AOA60896.1|_GQ67_02229T0_Komagataella_phaffii
ref|XP_029319710.1|_uncharacterized_protein_C5L36_0A08310_Pichia_kudriavzevii
gb|ESW98909.1|_putative_transmembrane_protein_involved_in_export_of_ammonia_Ogataea_parapolymorpha_DL-1
ref|XP_018211073.1|_hypothetical_protein_OGAPODRAFT_15977_Ogataea_polymorpha
gb|ANZ73092.1|_BA75_01155T0_Komagataella_pastoris
gb|AOA61343.1|_GQ67_02964T0_Komagataella_phaffii
ref|XP_029320379.1|_uncharacterized_protein_C5L36_0B01660_Pichia_kudriavzevii
gb|ESX01134.1|_putative_transmembrane_protein_involved_in_export_of_ammonia_Ogataea_parapolymorpha_DL-1
ref|XP_018209951.1|_hypothetical_protein_OGAPODRAFT_77222_Ogataea_polymorpha
gb|ONH68976.1|_Glyoxylate_pathway_regulator_Cyberlindnera_fabianii
emb|CCH41253.1|_Glyoxylate_pathway_regulator_Wickerhamomyces_ciferrii
emb|CCH45519.1|_Glyoxylate_pathway_regulator_Wickerhamomyces_ciferrii
gb|ODV70726.1|_hypothetical_protein_CYBJADRAFT_178915_Cyberlindnera_jadinii_NRRL_Y-1542
gb|ONH66172.1|_Ammonia_transport_outward_protein_2_Cyberlindnera_fabianii
emb|CAG99164.1|_KLLA0E02927p_Kluyveromyces_lactis
dbj|BAP70568.1|_hypothetical_protein_KLMA_20586_Kluyveromyces_marxianus
gb|AAS53515.1|_AFR144Wp_Eremothecium_gossypii_ATCC_10895
gb|AET38063.1|_Hypothetical_protein_Ecym_2321_Eremothecium_cymbalariae_DBVPG_7215
emb|SCU95012.1|_LAME_0F10330g1_1_Lachancea_meyersii_CBS_8951
emb|SCV06079.1|_LANO_0H21638g1_1_Lachancea_nothofagi_CBS_11611
emb|SCU79922.1|_LADA_0B04038g1_1_Lachancea_dasiensis_CBS_10888
emb|CAR24413.1|_KLTH0F15268p_Lachancea_thermotolerans_CBS_6340
emb|SCU93461.1|_LAMI_0E14444g1_1_Lachancea_mirantina
emb|SCW02329.1|_LAFE_0F04038g1_1_Lachancea_fermentati
dbj|BAP70569.1|_ammonia_transport_outward_protein_3_Kluyveromyces_marxianus
emb|CAG99163.1|_KLLA0E02905p_Kluyveromyces_lactis


emb|CDF89944.1|_ZYBA0S05-05336g1_1_Zygosaccharomyces_bailii_CLIB_213
gb|AQZ18663.1|_ATO3 (YDR384C) Zygosaccharomyces parabailii
emb|CDH12022.1|_related_to_Ammonia_transport_outward_protein_3_Zygosaccharomyces_bailii_ISA1307
gb|AQZ12435.1|_ATO3 (YDR384C) Zygosaccharomyces parabailii
dbj|GAV52108.1|_hypothetical_protein_ZYGR_0AG00990_Zygosaccharomyces_rouxii_NBRC_110957
dbj|GAV47096.1|_hypothetical_protein_ZYGR_0E01110_Zygosaccharomyces_rouxii_NBRC_110957
emb|CAR26128.1|_ZYRO0B02442p_Zygosaccharomyces_rouxii
gb|KOG99274.1|_hypothetical_protein_DI49_2619_Saccharomyces_eubayanus
emb|CCE89660.1|_hypothetical_protein_TDEL_0A03280_Torulaspora_delbrueckii
emb|CCH58060.1|_hypothetical_protein_TBLA_0A02610_Tetrapisispora_blattae_CBS_6284
emb|CCH58589.1|_hypothetical_protein_TBLA_0A07990_Tetrapisispora_blattae_CBS_6284
emb|CCK69570.1|_hypothetical_protein_KNAG_0C04690_Kazachstania_naganishii_CBS_8797
emb|CCE63248.1|_hypothetical_protein_TPHA_0E01550_Tetrapisispora_phaffii_CBS_4417
emb|CCF58519.1|_hypothetical_protein_KAFR_0E03880_Kazachstania_africana_CBS_2517
emb|CCC67746.1|_hypothetical_protein_NCAS_0A11880_Naumovozyma_castellii_CBS_4309
emb|CCD22597.1|_hypothetical_protein_NDAI_0A04410_Naumovozyma_dairenensis_CBS_421
gb|OXB50701.1|_hypothetical_protein_B1J92_A03212g_Candida_glabrata
gb|KZV12628.1|_ATO3_Saccharomyces_cerevisiae
gb|EJS42246.1|_ato3p_Saccharomyces_arboricola_H-6
gb|KOH01204.1|_ATO3-like_protein_Saccharomyces_eubayanus
gb|EKJ68640.1|_hypothetical_protein_FPSE_11186_Fusarium_pseudograminearum_CS3096
gb|EYB21988.1|_hypothetical_protein_FG05_12406_Fusarium_graminearum
gb|ESU09585.1|_hypothetical_protein_FGSG_12406_Fusarium_graminearum_PH-1
gb|EWG42832.1|_hypothetical_protein_FVEG_04387_Fusarium_verticillioides_7600
gb|KNB20685.1|_hypothetical_protein_FOXG_22898_Fusarium_oxysporum_f._sp._lycopersici_4287
emb|CAB10857.1|_acetate transmembrane transporter (predicted) Schizosaccharomyces pombe
ref|XP_013016294.1|_acetate_transporter_Schizosaccharomyces_octosporus_yFS286
gb|KZN87915.1|_Acetate_permease_A_Penicillium_chrysogenum
gb|ESU15945.1|_hypothetical_protein_FGSG_09374_Fusarium_graminearum_PH-1
gb|EYB28399.1|_hypothetical_protein_FG05_09374_Fusarium_graminearum
gb|EKJ74750.1|_hypothetical_protein_FPSE_05085_Fusarium_pseudograminearum_CS3096
emb|CCT64543.1|_related_to_Y.lipolytica_glyoxylate_pathway_regulator_GPR1_Fusarium_fujikuroi_IMI_58289
gb|EWG41941.1|_hypothetical_protein_FVEG_03920_Fusarium_verticillioides_7600
gb|OAQ63013.1|_GPR1/FUN34/yaaH_family_domain-containing_protein_Pochonia_chlamydosporia_170
gb|KYK57565.1|_GPR/FUN34_family_protein_Drechmeria_coniospora
gb|AEO60148.1|_hypothetical_protein_MYCTH_2309049_Thermothelomyces_thermophila_ATCC_42464
gb|AEO67808.1|_hypothetical_protein_THITE_2155054_Thielavia_terrestris_NRRL_8126
gb|EAA30570.2|_GPR/FUN34_family_protein_Neurospora_crassa_OR74A
gb|KUI70088.1|_Acetate_permease_A_Valsa_mali
gb|EHA57233.1|_hypothetical_protein_MGG_09225_Magnaporthe_oryzae_70-15
emb|SMY24185.1|_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
gb|PIB01991.1|_Acetate_permease_A_Cercospora_beticola
gb|ATZ56664.1|_hypothetical_protein_BCIN_13g04980_Botrytis_cinerea_B05.10
gb|APA06978.1|_hypothetical_protein_sscle_02g017480_Sclerotinia_sclerotiorum_1980_UF-70
gb|ATZ45736.1|_hypothetical_protein_BCIN_01g04600_Botrytis_cinerea_B05.10
gb|KZN93809.1|_Acetate_permease_A_Penicillium_chrysogenum
ref|XP_001396340.2|_hypothetical_protein_ANI_1_674114_Aspergillus_niger_CBS_513.88
tpe|CBF85662.1|_TPA:_conserved_hypothetical_protein_Aspergillus_nidulans_FGSC_A4
ref|XP_001391955.1|_hypothetical_protein_ANI_1_1090064_Aspergillus_niger_CBS_513.88
dbj|BAE56448.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|KZN92297.1|_Acetate_permease_A_Penicillium_chrysogenum
gb|EAL87502.1|_GPR/FUN34_family_protein_Aspergillus_fumigatus_Af293
tpe|CBF81120.1|_TPA:_Acetate_permeasePutative_uncharacterized_protein_Source:UniProtKB/TrEMBL_Acc:Q5B2K4_Aspergillus_nidulans_FGSC_A4
gb|PIB02163.1|_Accumulation_of_dyads_protein_2_Cercospora_beticola
tpe|CBF78629.1|_TPA:_conserved_hypothetical_protein_Aspergillus_nidulans_FGSC_A4
dbj|BAE64559.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
emb|CCT75985.1|_probable_FUN34-transmembrane_protein_involved_in_ammonia_production_Fusarium_fujikuroi_IMI_58289
gb|KNB20419.1|_hypothetical_protein_FOXG_17487_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|ATZ48287.1|_hypothetical_protein_BCIN_03g05160_Botrytis_cinerea_B05.10
dbj|BAE62495.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|KYK54830.1|_Gpr1_family_protein_Drechmeria_coniospora
gb|OAQ66447.1|_GPR1/FUN34/yaaH_family_domain-containing_protein_Pochonia_chlamydosporia_170
gb|ESU12128.1|_hypothetical_protein_FGSG_06081_Fusarium_graminearum_PH-1
gb|EYB33115.1|_hypothetical_protein_FG05_06081_Fusarium_graminearum
gb|EKJ75593.1|_hypothetical_protein_FPSE_04236_Fusarium_pseudograminearum_CS3096
emb|CCT67092.1|_related_to_Y.lipolytica_GPR1_protein_and_Fun34p_Fusarium_fujikuroi_IMI_58289
gb|EWG43098.1|_hypothetical_protein_FVEG_04717_Fusarium_verticillioides_7600
gb|KNB05543.1|_hypothetical_protein_FOXG_07791_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|KNB05544.1|_hypothetical_protein_FOXG_07791_Fusarium_oxysporum_f._sp._lycopersici_4287
emb|CCT74568.1|_related_to_Y.lipolytica_GPR1_protein_and_Fun34p_Fusarium_fujikuroi_IMI_58289
gb|EWG49215.1|_hypothetical_protein_FVEG_08800_Fusarium_verticillioides_7600
gb|KNB08999.1|_hypothetical_protein_FOXG_09672_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|ESU09548.1|_hypothetical_protein_FGSG_03652_Fusarium_graminearum_PH-1
gb|EYB32770.1|_hypothetical_protein_FG05_03652_Fusarium_graminearum
gb|EKJ77957.1|_hypothetical_protein_FPSE_01883_Fusarium_pseudograminearum_CS3096
gb|KZN83926.1|_Protein_alcS_Penicillium_chrysogenum
gb|EWG45761.1|_hypothetical_protein_FVEG_15857_Fusarium_verticillioides_7600
emb|CCT64135.1|_related_to_Y.lipolytica_GPR1_protein_and_Fun34p_Fusarium_fujikuroi_IMI_58289
gb|KNB06800.1|_hypothetical_protein_FOXG_19778_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|KNB06801.1|_hypothetical_protein_FOXG_19778_Fusarium_oxysporum_f._sp._lycopersici_4287
emb|SMY29801.1|_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
gb|PIA97100.1|_Accumulation_of_dyads_protein_2_Cercospora_beticola
gb|ATZ56169.1|_hypothetical_protein_BCIN_13g00180_Botrytis_cinerea_B05.10
gb|OBR11585.1|_Acetate_transporter_Colletotrichum_higginsianum_IMI_349063
ref|XP_001402324.1|_acetate_transporter_Aspergillus_niger_CBS_513.88
tpe|CBF89632.1|_TPA:_conserved_hypothetical_protein_Aspergillus_nidulans_FGSC_A4
gb|ATZ46145.1|_hypothetical_protein_BCIN_01g07980_Botrytis_cinerea_B05.10
emb|SMY27991.1|_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
gb|PIA95755.1|_Protein_alcS_Cercospora_beticola
gb|PIA90738.1|_Protein_alcS_Cercospora_beticola
gb|PIA93418.1|_Protein_alcS_Cercospora_beticola
gb|KZN89801.1|_Protein_alcS_Penicillium_chrysogenum
tpe|CBF84523.1|_TPA:_AlcS Source:UniProtKB/TrEMBL Acc:Q460G9 Aspergillus nidulans FGSC A4
gb|EAL84794.2|_transcriptional_activator_of_ethanol_catabolism_AlcS_Aspergillus_fumigatus_Af293



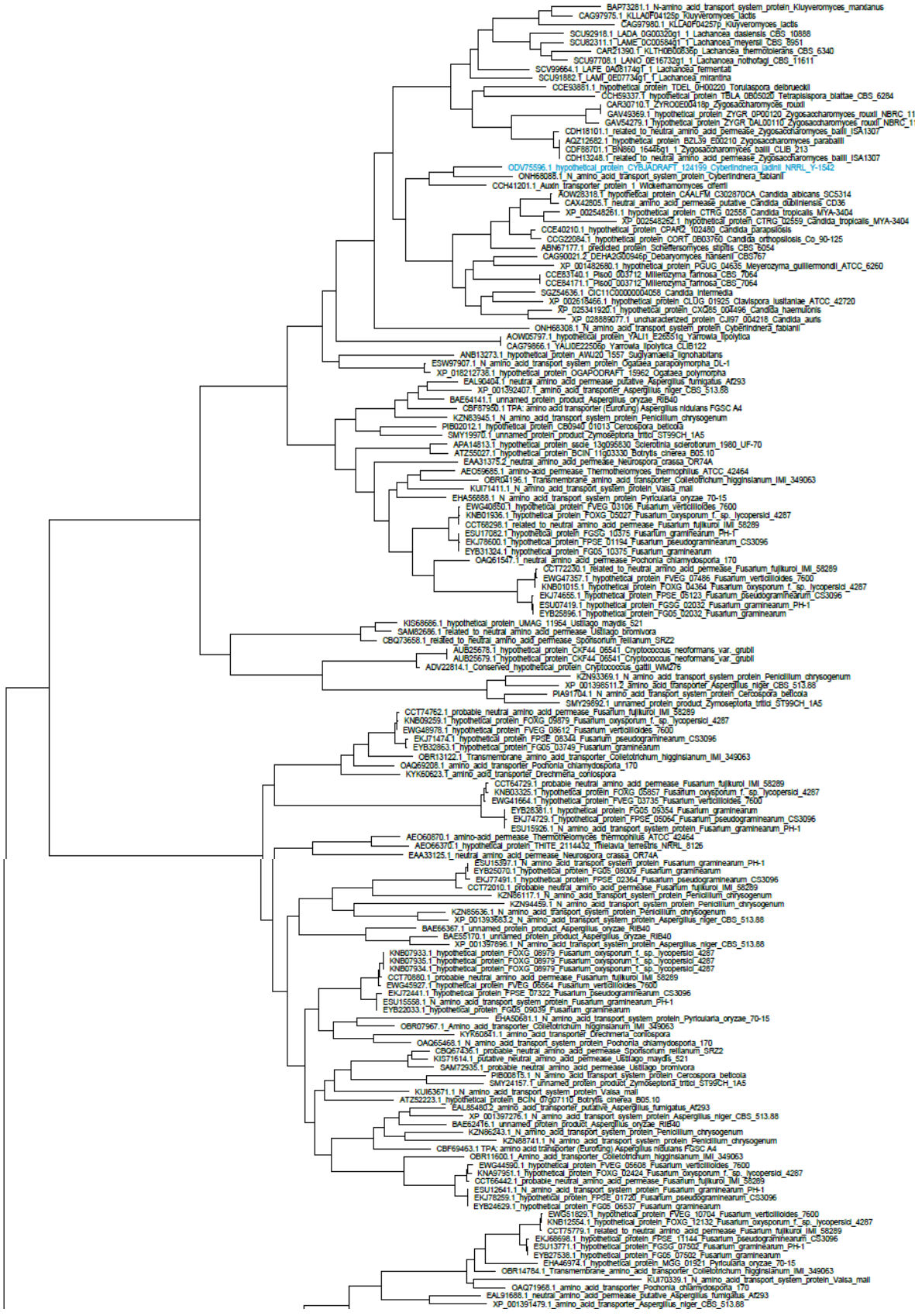
S3.2 Detailed CjJEN-phylogenetic tree with all homolog nominations and bootstrap values

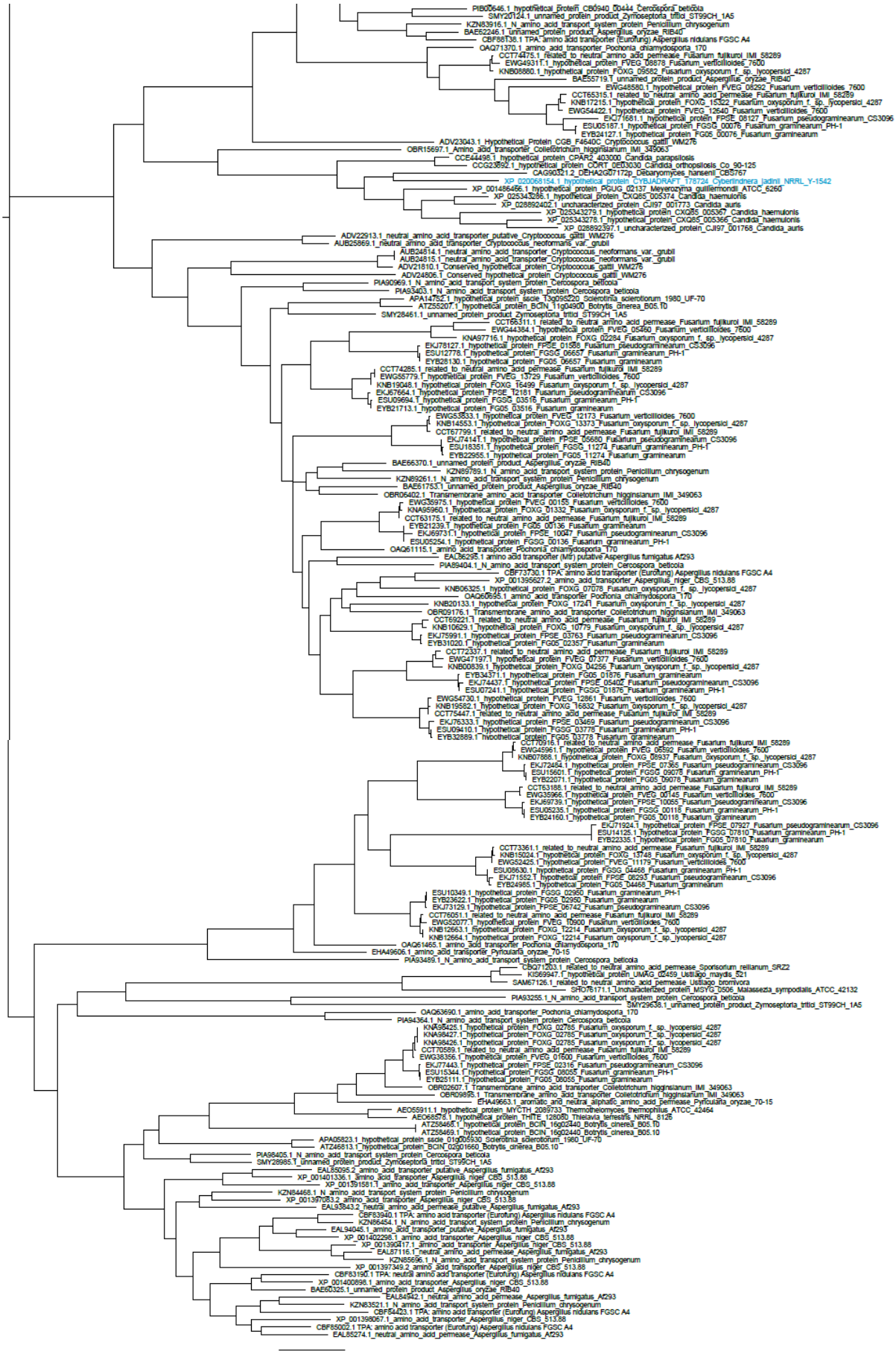



emb|SDT31978.1|_MFS_transporter_SHS_family_lactate_transporter_Bradyrhizobium_canariense
emb|SHN85244.1|_MFS_transporter_SHS_family_lactate_transporter_Bradyrhizobium_erythrophlei
gb|ASG67911.1|_MFS_transporter_Francisella_halioticida
gb|AIR87513.1|_sugar_transporter_Pantoea_rwandensis
gb|ALV93557.1|_sugar_transporter_Pantoea_vagans
gb|ARF48665.1|_MFS_transporter_Pantoea_stewartii_subsp._stewartii_DC283
gb|ASN16569.1|_MFS_transporter_Pantoea_ananatis
gb|APQ13130.1|_MFS_transporter_Pseudomonas_psychrotolerans
gb|ARU99275.1|_sugar_transporter_Tatumella_citrea
gb|AFJ48446.1|_putative_transporter_Shimwellia_blattae_DSM_4481_NBRC_105725
gb|AGR57494.1|_Putative_metabolite_transport_protein_Salmonella_bongori_N268-08
gb|ATQ13171.1|_MFS_transporter_Salmonella_enterica
gb|AIR00748.1|_sugar_transporter_Pluralibacter_gergoviae
gb|ALR77335.1|_sugar_transporter_Enterobacter_lignolyticus
emb|CAX58098.1|_metabolite_transport_protein_Erwinia_billingiae_Eb661
gb|AHF76084.1|_Metabolite_transport_protein_Sodalis_praecaptivus
emb|CAO95649.1|_Putative_metabolite_transport_protein_Erwinia_tasmaniensis_Et1/99
emb|CAY73001.1|_putative_metabolite_transport_protein_Erwinia_pyrifoliae_DSM_12163
gb|AEX54136.1|_arabinose_efflux_permease_family_protein_Rahnella_aquatilis_CIP_78.65_ATCC_33071
gb|ADN00481.1|_Sialic_acid_transporter_Dickeya_dadantii_3937
gb|ATZ96257.1|_MFS_transporter_Dickeya_fangzhongdai
gb|AUH13321.1|_sugar_transporter_Dickeya_solani
gb|ATO35197.1|_Putative_metabolite_transport_protein_Dickeya_dianthicola_RNS04.9
gb|ACZ78749.1|_major_facilitator_superfamily_MFS_1_Dickeya_zeae_Ech586
gb|AIJ08294.1|_Putative_metabolite_transport_protein_Edwardsiella_anguillarum_ET080813
gb|ARD18219.1|_MFS_transporter_Edwardsiella_piscicida
gb|ARD39923.1|_MFS_transporter_Edwardsiella_ictaluri
gb|ATI64127.1|_MFS_transporter_Edwardsiella_tarda
emb|CZQ47060.1|_metabolite_transport_protein_Klebsiella_pneumoniae
gb|ASR29163.1|_sugar_transporter_Klebsiella_quasipneumoniae
gb|ASV20279.1|_MFS_transporter_Klebsiella_quasivaricola
gb|AQL27032.1|_sugar_transporter_Klebsiella_varicola
gb|ATY07766.1|_MFS_transporter_Klebsiella_aerogenes
gb|AML56437.1|_Putative_metabolite_transport_protein_Serratia_rubidaea
gb|AMO79998.1|_sugar_transporter_Obesumbacterium_proteus
gb|ANC39972.1|_sugar_transporter_Hafnia_alvei
gb|ATA20051.1|_sugar_transporter_Gibbsiella_quercinecans
gb|ATM76349.1|_MFS_transporter_Serratia_fonticola
gb|ABV43854.1|_major_facilitator_superfamily_MFS_1_Serratia_proteamaculans_568
gb|AMH00034.1|_sugar_transporter_Serratia_liquefaciens
gb|ANK00953.1|_sugar_transporter_Serratia_plymuthica
emb|SNW06167.1|_Sialic_acid_permease_Serratia_ficaria
gb|PIC09289.1|_MFS_transporter_Serratia_marcescens
dbj|BAM07968.1|_putative_major_facilitator_superfamily_MFS_1_Leptospirillum_ferrooxidans_C2-3
gb|APT56701.1|_sugar_transporter_Roseomonas_gilardii
gb|AQS88572.1|_sugar_transporter_Neoasaia_chiangmaiensis
dbj|BAK82911.1|_major_facilitator_superfamily_sugar_transporter_Komagataeibacter_medellinensis_NBRC_3288
gb|ATU72886.1|_MFS_transporter_Komagataeibacter_xylinus
gb|AOX21240.1|_sugar_transporter_Kozakia_baliensis
gb|EFG83337.1|_major_facilitator_superfamily_MFS_1_Komagataeibacter_hansenii_ATCC_23769
emb|CAP53980.1|_putative_sugar_transporter_protein_Gluconacetobacter_diazotrophicus_PA1_5
emb|CEF42035.1|_putative_metabolite_transport_protein_Acetobacter_senegalensis
gb|ATJ92769.1|_MFS_transporter_Acetobacter_tropicalis
emb|CEF57258.1|_major_facilitator_superfamily_MFS_1_Acetobacter_ghanensis
gb|ASL39367.1|_MFS_transporter_Acetobacter_pasteurianus
gb|ATI11955.1|_MFS_transporter_Acetobacter_pomorum
gb|AND69138.1|_hypothetical_protein_ATSB10_16840_Dyella_thiooxydans
gb|APG02609.1|_MFS_transporter_Luteibacter_rhizovicinus_DSM_16549
gb|AGG88315.1|_arabinose_efflux_permease_family_protein_Rhodanobacter_denitrificans
gb|AFC85638.1|_arabinose_efflux_permease_family_protein_Frateuria_aurantia_DSM_6220
gb|AHX12439.1|_sugar_transporter_Dyella_jiangningensis
gb|AIF47448.1|_sugar_transporter_Dyella_japonica_A8
gb|ACO32581.1|_transporter_major_facilitator_family_Acidobacterium_capsulatum_ATCC_51196
gb|AEU37075.1|_major_facilitator_superfamily_MFS_1_Granulicella_mallensis_MP5ACTX8
emb|SDF26040.1|_MFS_transporter_SHS_family_lactate_transporter_Terriglobus_roseus
gb|ADV82482.1|_major_facilitator_superfamily_MFS_1_Terriglobus_saanensis_SP1PR4
gb|ADW69145.1|_major_facilitator_superfamily_MFS_1_Granulicella_tundricola_MP5ACTX9
gb|AGT84772.1|_MFS_transporter_lactate_transporter_Amycolatopsis_mediterranei_RB
emb|CCB76045.1|_Sialic_acid-transport_integral_membrane_protein_nanT_Streptomyces_cattleya_NRRL_8057_DSM_46488
gb|ANS70284.1|_sialic_acid-transport_integral_membrane_protein_NanT_Streptomyces_lincolnensis
emb|SDT44803.1|_MFS_transporter_SHS_family_lactate_transporter_Microlunatus_soli
gb|ADG97034.1|_major_facilitator_superfamily_MFS_1_Segniliparus_rotundus_DSM_44985
emb|SDU87187.1|_MFS_transporter_SHS_family_lactate_transporter_Friedmanniella_sagamiharensis
gb|AFA74056.1|_sialic_acid_transporter_NanT_Gordonia_polyisoprenivorans_VH2
gb|AHH94452.1|_major_facilitator_superfamily_protein_Kutzneria_albida_DSM_43870
gb|AEA28298.1|_major_facilitator_superfamily_MFS_1_Pseudonocardia_dioxanivorans_CB1190
gb|AQT83166.1|_MFS_transporter_Mycobacterium_litorale
gb|AHH15773.1|_putative_sialic_acid_transporter_Nocardia_nova_SH22a
gb|ATL65781.1|_MFS_transporter_Nocardia_terpenica
emb|SNV65696.1|_sialic_acid-transport_integral_membrane_protein_NanT_Mycobacterium_terrae
gb|AEF35473.1|_sialic_acid-transport_integral_membrane_protein_NanT_Mycobacterium_sinense
gb|ANG90471.1|_sugar_transporter_Mycobacterium_bovis
gb|APU27885.1|_sugar_transporter_Mycobacterium_caprae
gb|AMQ40657.1|_sugar_transporter_Mycobacterium_africanum
emb|CCC44250.1|_putative_sialic_acid-transport_integral_membrane_protein_NANT_Mycobacterium_canettii_CIPT_140010059
gb|AMC59540.1|_sialic_acid-transport_integral_membrane_protein_NanT_Mycobacterium_microti
gb|OMH59831.1|_putative_sialic_acid_transporter_Mycobacterium_tuberculosis
gb|ARG95382.1|_MFS_transporter_Mycobacterium_kansasii
dbj|BAV41584.1|_sialic_acid-transport_integral_membrane_protein_Mycobacterium_ulcerans_subsp._shinshuense
gb|ACC41239.1|_sialic_acid-transport_integral_membrane_protein_NanT_Mycobacterium_marinum_M
dbj|BAX92835.1|_MFS_transporter_Mycobacterium_shigaense
gb|AMP25793.1|_sugar_transporter_Mycobacterium_simiae
gb|AKN18680.1|_sugar_transporter_Mycobacterium_haemophilum_DSM_44634
gb|ANR94190.1|_sugar_transporter_Mycobacterium_avium
gb|ATA26526.1|_sialic_acid_transporter_NanT_Mycobacterium_lepraemurium
gb|ASW90495.1|_MFS_transporter_Mycobacterium_marseillense
gb|ARR83234.1|_sugar_transporter_family_protein_Mycobacterium_yongonense
gb|AFC54001.1|_MFS_transporter_sialate:H+_symporter_(SHS)_family_protein_Mycobacterium_parairtacelluiare
gb|AFS14776.1|_Carboxylic_acid_transporter-like_protein_Mycobacterium_indicus_pranii_MTCC_9506
gb|ASW95559.1|_MFS_transporter_Mycobacterium_intracellulare
gb|ASX00696.1|_MFS_transporter_Mycobacterium_chimaera
gb|ABF41560.1|_major_facilitator_superfamily_(MFS)_transporter_Candidatus_Koribacter_versatilis_Ellin345
gb|AHG88245.1|_major_facilitator_superfamily_MFS_1_Gemmatirosa_kalamazoonesis
gb|ABJ83693.1|_major_facilitator_superfamily_MFS_1_Candidatus_Solibacter_usitatus_Ellin6076
gb|ADW71043.1|_major_facilitator_superfamily_MFS_1_(plasmid)_Granulicella_tundricola_MP5ACTX9
dbj|BAB59736.1|_sugar_transporter_Thermoplasma_volcanium_GSS1
emb|CAC11721.1|_sugar_transport_protein_related_protein_Thermoplasma_acidophilum

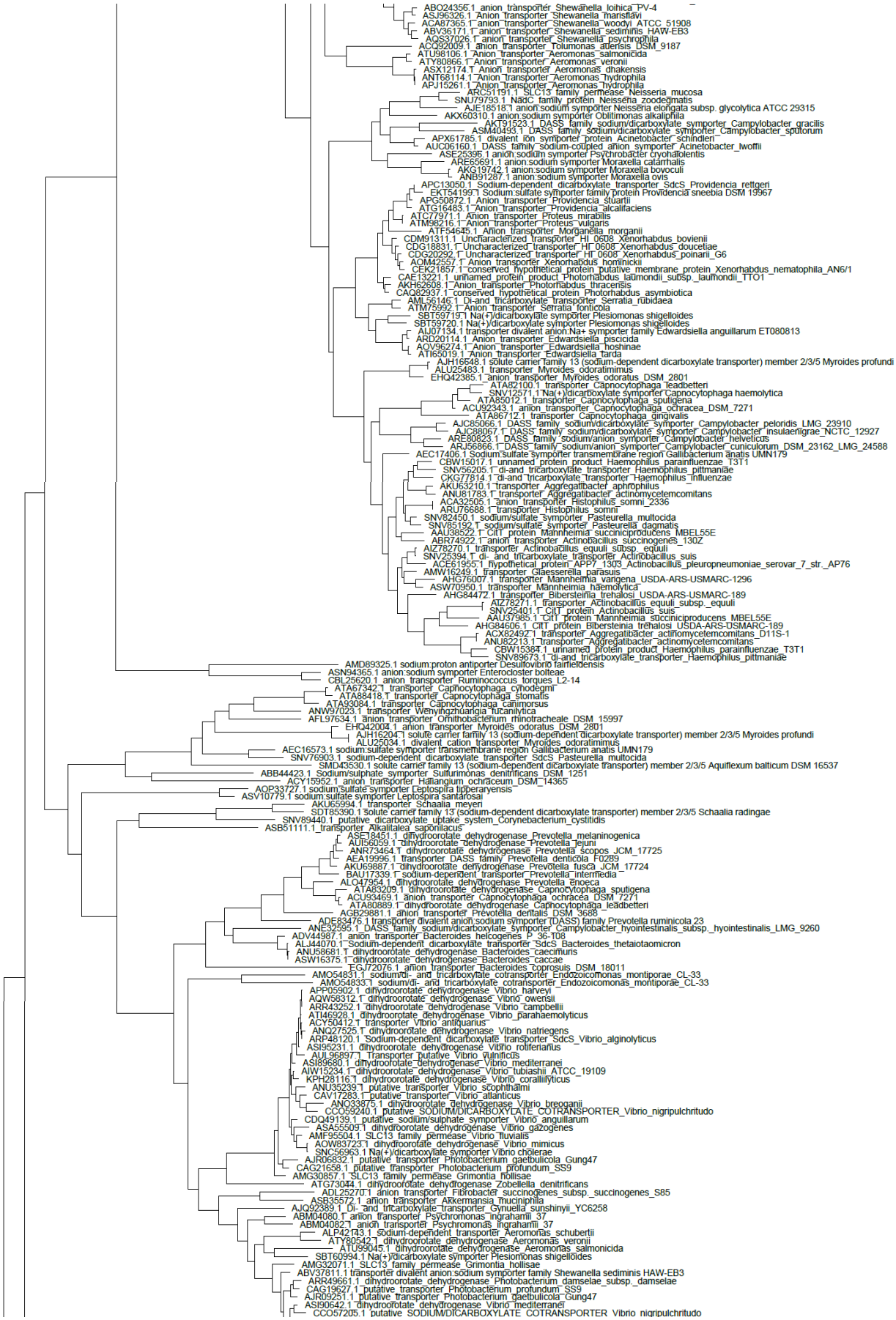

gb|ESU11911.1|_hypothetical_protein_FGSG_05882_Fusarium_graminearum_PH-1
gb|EYB33959.1|_hypothetical_protein_FG05_05882_Fusarium_graminearum
gb|EKJ67583.1|_hypothetical_protein_FPSE_12257_Fusarium_pseudograminearum_CS3096
gb|KNB05311.1|_hypothetical_protein_FOXG_07631_Fusarium_oxysporum_f._sp._lycopersici_4287
emb|CCT66944.1|_related_to_carboxylic_acid_transport_protein_JEN1_Fusarium_fujikuroi_IMI_58289
gb|EWG42839.1|_hypothetical_protein_FVEG_04557_Fusarium_verticillioides_7600
gb|KYK61705.1|_carboxylic_acid_transport_protein_Drechmeria_coniospora
gb|OAG65171.1|_sialate:H+ symporter (SHS) family MFS transporter Pochonia chlamydosporia 170
gb|OBR15164.1|_Major_facilitator_superfamily_transporter_Colletotrichum_higginsianum_IMI_349063
gb|KZN92938.1|_Carboxylic_acid_transporter_protein-like_protein_Penicillium_chrysogenum
tpe|CBF70194.1|_TPA:_carboxylic_acid_transport_protein_(AFU_orthologue_AFUA_2G09450)_Aspergillus_nidulans_FGSC_A4
gb|EAL93241.1|_carboxylic_acid_transport_protein_Aspergillus_fumigatus_Af293
dbj|BAE65195.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
ref|XP_001397997.1|_carboxylic_acid_transport_protein_Aspergillus_niger_CBS_513.88
gb|EHA50270.1|_carboxylic_acid_transporter_Magnaporthe_oryzae_70-15
gb|EAA32361.2|_carboxylic_acid_transporter_Neurospora_crassa_OR74A
gb|AEO56506.1|_general_substrate_transporter_Thermothelomyces_thermophila_ATCC_42464
gb|AEO62951.1|_hypothetical_protein_THITE_2042844_Thielavia_terrestris_NRRL_8126
emb|SMY24056.1|_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
gb|PIA97966.1|_Carboxylic_acid_transporter_protein_Cercospora_beticola
emb|SAM68120.1|_probable_carboxylic_acid_transport_protein_JEN1_Ustilago_bromivora
gb|KIS72107.1|_putative_carboxylic_acid_transport_protein_JEN1_Ustilago_maydis_521
gb|ADV22083.1|_carboxylic_acid_transport_protein_putative_Cryptococcus_gattii_WM276
gb|AUB27544.1|_hypothetical_protein_CKF44_004214_Cryptococcus_neoformans_var._grubii
emb|SHO76643.1|_Similar_to_S.cerevisiae_protein_JEN1_(Monocarboxylate/proton_symporter_of_the_plasma_membrane)_Malassezia_sympodialis_ATCC_42132
emb|SAM81792.1|_related_to_carboxylic_acid_transport_protein_JEN1_Ustilago_bromivora
emb|CBQ72319.1|_related_to_carboxylic_acid_transport_protein_JEN1_Sporisorium_reilianum_SRZ2
gb|KIS70048.1|_hypothetical_protein_UMAG_05122_Ustilago_maydis_521
gb|EAL85090.1|_sugar_transporter_family_protein_Aspergillus_fumigatus_Af293
gb|OAQ60656.1|_carboxylic_acid_transporter_Pochonia_chlamydosporia_170
emb|SMY22408.1|_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
emb|CCE44927.1|_hypothetical_protein_CPAR2_407290_Candida_parapsilosis
emb|CCE44928.1|_hypothetical_protein_CPAR2_407300_Candida_parapsilosis
emb|CCE44923.1|_hypothetical_protein_CPAR2_407250_Candida_parapsilosis
emb|CCE44586.1|_hypothetical_protein_CPAR2_403890_Candida_parapsilosis
emb|CCE44209.1|_hypothetical_protein_CPAR2_400100_Candida_parapsilosis
emb|CCE44929.1|_hypothetical_protein_CPAR2_407310_Candida_parapsilosis
emb|CCE42284.1|_hypothetical_protein_CPAR2_808330_Candida_parapsilosis
emb|CCG25461.1|_tRNA-Arg_Candida_orthopsilosis_Co_90-125
emb|CCE42290.1|_hypothetical_protein_CPAR2_808390_Candida_parapsilosis
emb|CCG25455.1|_hypothetical_protein_CORT_0C00780_Candida_orthopsilosis_Co_90-125
emb|CCE42292.1|_hypothetical_protein_CPAR2_808410_Candida_parapsilosis
emb|CCG25453.1|_Jen1_protein_Candida_orthopsilosis_Co_90-125
ref|XP_002551411.1|_hypothetical_protein_CTRG_05709_Candida_tropicalis_MYA-3404
emb|CAX43151.1|_carboxylic_acid_transporter_putative_Candida_dubliniensis_CD36
gb|AOW28664.1|_Jen1p_Candida_albicans_SC5314
gb|ABN65148.2|_carboxylic_acid_transporter_protein_Scheffersomyces_stipitis_CBS_6054
ref|XP_001484059.1|_hypothetical_protein_PGUG_03440_Meyerozyma_guilliermondii_ATCC_6260
emb|CAG89500.2|_DEHA2F17402p_Debaryomyces_hansenii_CBS767
emb|CAG89946.2|_DEHA2F27126p_Debaryomyces_hansenii_CBS767
emb|CCH43335.1|_Carboxylic_acid_transporter_protein_Wickerhamomyces_ciferrii
emb|CCH43336.1|_Carboxylic_acid_transporter_protein_Wickerhamomyces_ciferrii
emb|CCH43337.1|_putative_membrane_protein_Wickerhamomyces_ciferrii
gb|ONH70146.1|_Carboxylic_acid_transporter_protein_Cyberlindnera_fabianii
gb|ODV74417.1|_MFS_general_substrate_transporter_Cyberlindnera_jadinii_NRRL_Y-1542
CEP21966.1_Carboxylic_acid_transporter_protein_homolog_Cyberlindnera_jadinii
dbj|BAP72814.1|_carboxylic_acid_transporter_protein_homolog_Kluyveromyces_marxianus
emb|CAG99769.1|_KLLA0E16259p_Kluyveromyces_lactis
emb|SCV99651.1|_LAFE_0A07888g1_1_Lachancea_fermentati
emb|CCE93840.1|_hypothetical_protein_TDEL_0G04730_Torulaspora_delbrueckii
gb|KOG98125.1|_hypothetical_protein_DI49_3168_Saccharomyces_eubayanus
gb|KZV09719.1|_JEN1_Saccharomyces_cerevisiae
emb|CCH46564.1|_Carboxylic_acid_transporter_protein_Wickerhamomyces_ciferrii
gb|ONH67566.1|_Carboxylic_acid_transporter_protein_Cyberlindnera_fabianii
gb|ODV72343.1|_MFS_general_substrate_transporter_Cyberlindnera_jadinii_NRRL_Y-1542
gb|ODV75195.1|_MFS_general_substrate_transporter_Cyberlindnera_jadinii_NRRL_Y-1542
emb|CCT73545.1|_related_to_carboxylic_acid_transporter_protein_Fusarium_fujikuroi_IMI_58289
gb|KNB11574.1|_MFS_transporter_SHS_family_lactate_transporter_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|EWG51422.1|_MFS_transporter_SHS_family_lactate_transporter_Fusarium_verticillioides_7600
gb|EKJ69418.1|_hypothetical_protein_FPSE_10408_Fusarium_pseudograminearum_CS3096
gb|ESU08925.1|_hypothetical_protein_FGSG_04204_Fusarium_graminearum_PH-1
gb|EYB21469.1|_hypothetical_protein_FG05_04204_Fusarium_graminearum
gb|OAQ62176.1|_sugar_transporter_family_protein_Pochonia_chlamydosporia_170
gb|OBR12832.1|_H+_symporter_family_protein_Colletotrichum_higginsianum_IMI_349063
gb|EAA33605.3|_sugar_transporter_Neurospora_crassa_OR74A
gb|EHA48046.1|_sugar_transporter_Magnaporthe_oryzae_70-15
gb|APA12974.1|_hypothetical_protein_sscle_10g077440_Sclerotinia_sclerotiorum_1980_UF-70
gb|ATZ55700.1|_Bcin1_Botrytis_cinerea_B05.10
gb|PIA99438.1|_Carboxylic_acid_transporter_protein_Cercospora_beticola
gb|EAL86916.1|_sugar_transporter_family_protein_Aspergillus_fumigatus_Af293
tpe|CBF71283.1|_TPA:_sugar_transporter_family_protein_(AFU_orthologue_AFUA_7G05550)_Aspergillus_nidulans_FGSC_A4
gb|KZN84915.1|_Carboxylic_acid_transporter_protein-like_protein_Penicillium_chrysogenum
dbj|BAE55486.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
ref|XP_001391338.1|_carboxylic_acid_transporter_Aspergillus_niger_CBS_513.88
gb|OAQ63702.1|_MFS_transporter_Pochonia_chlamydosporia_170
emb|CAG81250.1|_YALI0D20108p_Yarrowia_lipolytica_CLIB122
gb|AOW04349.1|_hypothetical_protein_YALI1_D25400g_Yarrowia_lipolytica
emb|CAG81440.1|_YALI0D24607p_Yarrowia_lipolytica_CLIB122
gb|AOW04593.1|_hypothetical_protein_YALI1_D32568g_Yarrowia_lipolytica
emb|CAG82184.1|_YALI0C15486p_Yarrowia_lipolytica_CLIB122
gb|AOW02928.1|_hypothetical_protein_YALI1_C21946g_Yarrowia_lipolytica
emb|CAG82410.1|_YALI0C21406p_Yarrowia_lipolytica_CLIB122
gb|AOW03198.1|_hypothetical_protein_YALI1_C29536g_Yarrowia_lipolytica
emb|CAG80310.1|_YALI0E32901p_Yarrowia_lipolytica_CLIB122
gb|AOW06316.1|_hypothetical_protein_YALI1_E39128g_Yarrowia_lipolytica
emb|CAG83351.1|_YALI0B19470p_Yarrowia_lipolytica_CLIB122
gb|AOW01941.1|_hypothetical_protein_YALI1_B25506g_Yarrowia_lipolytica
gb|AAS50559.1|_AAR192Cp_Eremothecium_gossypii_ATCC_10895
gb|AAS53704.1|_AFR335Wp_Eremothecium_gossypii_ATCC_10895
gb|AAS50561.1|_ABL210Cp_Eremothecium_gossypii_ATCC_10895
gb|AMD22015.1|_HGL325Wp_Eremothecium_sinecaudum
gb|AET40429.1|_hypothetical_protein_Eqym_6024_Eremothecium_cymbalariae_DBVPG_7215
emb|SCU92745.1|_LAMI_0E11848g1_1_Lachancea_mirantina
emb|SCW00203.1|_LAFE_0B11716g1_1_Lachancea_fermentati
emb|CCE85945.1|_Piso0_005587_Millerozyma_farinosa_CBS_7064
emb|CCE87050.1|_Piso0_005587_Millerozyma_farinosa_CBS_7064
emb|CAG87482.2|_DEHA2D18920p_Debaryomyces_hansenii_CBS767
ref|XP_025341422.1|_hypothetical_protein_CXQ85_002273_Candida_haemulonis
emb|SGZ58571.1|_CIC11C00000005764_Candida_intermedia
ref|XP_002617818.1|_hypothetical_protein_CLUG_01274_Clavispora_lusitaniae_ATCC_42720
gb|EWM98136.1|_putative_sialic_acid_transporter_Ogataea_parapolymorpha_DL-1
ref|XP_018211824.1|_hypothetical_protein_OGAPODRAFT_16373_Ogataea_polymorpha
ref|XP_025319245.1|_uncharacterized_protein_CSL36_0A03680_Pichia_kudriavzevii
ref|XP_029319440.1|_uncharacterized_protein_CSL36_0A05580_Pichia_kudriavzevii
emb|SCU88406.1|_LANO_0D02014g1_1_Lachancea_nothofagi_CBS_11611
emb|SCU95322.1|_LADA_0G14950g1_1_Lachancea_dasiensis_CBS_10888
emb|SCV04563.1|_LAME_0H19416g1_1_Lachancea_meyersii_CBS_8951
emb|CAR24692.1|_KLTH0G02024p_Lachancea_thermotolerans_CBS_6340
emb|SCW02888.1|_LAFE_0F16490g1_1_Lachancea_fermentati
emb|SCU97836.1|_LAMI_0F11650g1_1_Lachancea_mirantina
emb|CAG98245.1|_KLLA0F10043p_Kluyveromyces_lactis
dbj|BAP70260.1|_putative_sialic_acid_transporter_Kluyveromyces_marxianus
gb|ODV73830.1|_MFS_general_substrate_transporter_Cyberlindnera_jadinii_NRRL_Y-1542
gb|ODV74730.1|_hypothetical_protein_CYBJADRAFT_124710_Cyberlindnera_jadinii_NRRL_Y-1542
gb|ONH69072.1|_Carboxylic_acid_transporter_protein_Cyberlindnera_fabianii
emb|CCH43458.1|_Carboxylic_acid_transporter_protein_Wickerhamomyces_ciferrii
emb|CCH43459.1|_Carboxylic_acid_transporter_protein_Wickerhamomyces_ciferrii
emb|SGZ53326.1|_CIC11C00000002797_Candida_intermedia
ref|XP_002614689.1|_hypothetical_protein_CLUG_05468_Clavispora_lusitaniae_ATCC_42720
ref|XP_025341865.1|_hypothetical_protein_CXQ85_004438_Candida_haemulonis
ref|XP_028889124.1|_uncharacterized_protein_CJI97_004268_Candida_auris
gb|ANZ73229.1|_BA75_00456T0_Komagataella_pastoris
gb|AOA63681.1|_CQ6T_03508T0_Komagataella_phaffii
ref|XP_001484060.1|_hypothetical_protein_PGUG_03441_Meyerozyma_guilliermondii_ATCC_6260
emb|CAR65879.1|_DEHA2E24024p_Debaryomyces_hansenii_CBS767
gb|ABN65003.2|_Carboxylic_acid_transporter_protein_homolog_partial_Scheffersomyces_stipitis_CBS_6054
emb|CAX42258.1|_carboxylic_acid_transporter_protein_putative_Candida_dubliniensis_CD36
gb|AOW29153.1|_Jen2p_Candida_albicans_SC5314
ref|XP_002545520.1|_hypothetical_protein_CTRG_00301_Candida_tropicalis_MYA-3404
emb|CCE40688.1|_hypothetical_protein_CPAR2_107230_Candida_parapsilosis
emb|CCE44404.1|_hypothetical_protein_CPAR2_402040_Candida_parapsilosis
emb|CCG23796.1|_Jen2_protein_Candida_orthopsilosis_Co_90-125
0.3

S3.3 Detailed CjSLC5-phylogenetic tree with all homolog nominations and bootstrap values

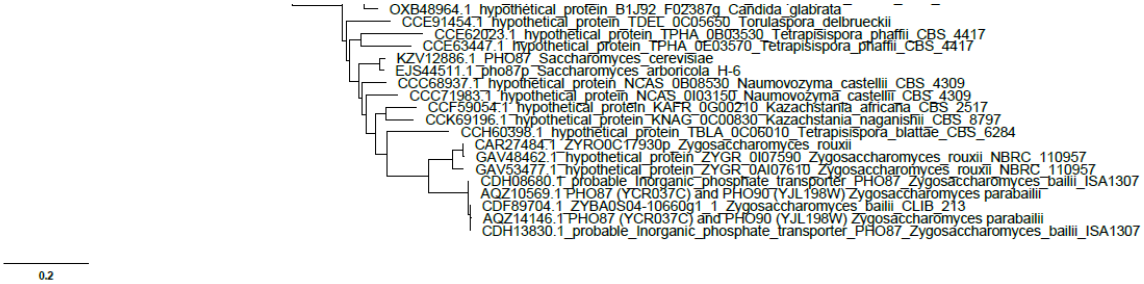



PIB06646.1 hypothetical_protein_CB0940_09444_Cercospora_beticola
SMY20124.1_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
KZN83916.1_N_amino_acid_transport_system_protein_Penicillium_chrysogenum
BAE62246.1_unnamed_protein_product_Aspergillus_oryzae_RIB40
CBF88736.1_TPA:_amino_acid_transporter_(Eurofung)_Aspergillus_nidulans_FGSC_A4
OAQ71370.1_amino_acid_transporter_Pochonia_chlamydosporia_170
CCT74475.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
EWG49311.1_hypothetical_protein_FVEG_08878_Fusarium_verticillioides_7600
KNB08860.1_hypothetical_protein_FOXG_05652_Fusarium_oxysporum_f._sp._lycopersici_4287
BAE55719.1_unnamed_protein_product_Aspergillus_oryzae_RIB40
EWG48580.1_hypothetical_protein_FVEG_08292_Fusarium_verticillioides_7600
CCT66315.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
KNB17215.1_hypothetical_protein_FOXG_13522_Fusarium_oxysporum_f._sp._lycopersici_4287
EWG54422.1_hypothetical_protein_FVEG_12640_Fusarium_verticillioides_7600
EKJ71681.1_hypothetical_protein_FPSE_08127_Fusarium_pseudograminearum_CS3096
ESU05187.1_hypothetical_protein_FGSG_00076_Fusarium_graminearum_PH-1
EYB24127.1_hypothetical_protein_FG05_00076_Fusarium_graminearum
ADV23043.1_Hypothetical_Protein_CGB_F4640C_Cryptococcus_gattii_WM276
OBR15697.1_Amino_acid_transporter_Colletotrichum_higginsianum_IMI_349063
CCE44496.1_hypothetical_protein_CPAR2_403000_Candida_parapsilosis
CCG23892.1_hypothetical_protein_CORT_0B03030_Candida_orthopsilosis_Co_90-125
CAG99321.2_DEHA2G07172p_Debaryomyces_hansenii_CBS767
XP_020068154.1_hypothetical_protein_CYBJADRAFT_178724_Cyberlindnera_jadinii_NRRL_Y-1542
XP_001486456.1_hypothetical_protein_PGUG_02137_Meyerozyma_guilliermondii_ATCC_6260
XP_025343286.1_hypothetical_protein_CXQ85_005374_Candida_haemulonis
XP_028892402.1_uncharacterized_protein_CJI97_001773_Candida_auris
XP_025343279.1_hypothetical_protein_CXQ85_005367_Candida_haemulonis
XP_025343278.1_hypothetical_protein_CXQ85_005366_Candida_haemulonis
XP_028892397.1_uncharacterized_protein_CJI97_001768_Candida_auris
ADV22913.1_neutral_amino_acid_transporter_putative_Cryptococcus_gattii_WM276
AUB25869.1_neutral_amino_acid_transporter_Cryptococcus_neoformans_var._grubii
AUB24814.1_neutral_amino_acid_transporter_Cryptococcus_neoformans_var._grubii
AUB24815.1_neutral_amino_acid_transporter_Cryptococcus_neoformans_var._grubii
ADV21810.1_Conserved_hypothetical_protein_Cryptococcus_gattii_WM276
ADV24806.1_Conserved_hypothetical_protein_Cryptococcus_gattii_WM276
PIA90969.1_N_amino_acid_transport_system_protein_Cercospora_beticola
PIA93403.1_N_amino_acid_transport_system_protein_Cercospora_beticola
APA14752.1_hypothetical_protein_sscle_13g095220_Sclerotinia_sclerotiorum_1980_UF-70
ATZ55207.1_hypothetical_protein_BCIN_11g04900_Botrytis_cinerea_B05.10
SMY28461.1_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
CCT65311.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
EWG44384.1_hypothetical_protein_FVEG_05460_Fusarium_verticillioides_7600
KNA97716.1_hypothetical_protein_FOXG_02284_Fusarium_oxysporum_f._sp._lycopersici_4287
EKJ78127.1_hypothetical_protein_FPSE_01588_Fusarium_pseudograminearum_CS3096
ESU12778.1_hypothetical_protein_FGSG_06657_Fusarium_graminearum_PH-1
EYB28130.1_hypothetical_protein_FG05_06657_Fusarium_graminearum
CCT74285.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
EWG55779.1_hypothetical_protein_FVEG_13729_Fusarium_verticillioides_7600
KNB19048.1_hypothetical_protein_FOXG_16499_Fusarium_oxysporum_f._sp._lycopersici_4287
EKJ67664.1_hypothetical_protein_FPSE_12181_Fusarium_pseudograminearum_CS3096
ESU09694.1_hypothetical_protein_FGSG_03518_Fusarium_graminearum_PH-1
EYB21713.1_hypothetical_protein_FG05_03516_Fusarium_graminearum
EWG53533.1_hypothetical_protein_FVEG_12173_Fusarium_verticillioides_7600
KNB14553.1_hypothetical_protein_FOXG_13373_Fusarium_oxysporum_f._sp._lycopersici_4287
CCT67799.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
EKJ74141.1_hypothetical_protein_FPSE_05680_Fusarium_pseudograminearum_CS3096
ESU18351.1_hypothetical_protein_FGSG_11274_Fusarium_graminearum_PH-1
EYB22955.1_hypothetical_protein_FG05_11274_Fusarium_graminearum
BAE66370.1_unnamed_protein_product_Aspergillus_oryzae_RIB40
KZN89789.1_N_amino_acid_transport_system_protein_Penicillium_chrysogenum
KZN89261.1_N_amino_acid_transport_system_protein_Penicillium_chrysogenum
BAE61763.1_unnamed_protein_product_Aspergillus_oryzae_RIB40
OBR06402.1_Transmembrane_amino_acid_transporter_Colletotrichum_higginsianum_IMI_349063
EWG35975.1_hypothetical_protein_FVEG_00155_Fusarium_verticillioides_7600
KNA95960.1_hypothetical_protein_FOXG_01332_Fusarium_oxysporum_f._sp._lycopersici_4287
CCT63175.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
EYB21239.1_hypothetical_protein_FG05_00136_Fusarium_graminearum
EKJ69731.1_hypothetical_protein_FPSE_10047_Fusarium_pseudograminearum_CS3096
ESU05254.1_hypothetical_protein_FGSG_00136_Fusarium_graminearum_PH-1
OAQ61115.1_amino_acid_transporter_Pochonia_chlamydosporia_170
EAL86295.1_amino_acid_transporter_(Mtr)_putative_Aspergillus_fumigatus_Af293
PIA89404.1_N_amino_acid_transport_system_protein_Cercospora_beticola
CBF73730.1_TPA:_amino_acid_transporter_(Eurofung)_Aspergillus_nidulans_FGSC_A4
XP_001396627.2_amino_acid_transporter_Aspergillus_niger_CBS_513.88
KNB06325.1_hypothetical_protein_FOXG_07078_Fusarium_oxysporum_f._sp._lycopersici_4287
OAQ60695.1_amino_acid_transporter_Pochonia_chlamydosporia_170
KNB20133.1_hypothetical_protein_FOXG_17241_Fusarium_oxysporum_f._sp._lycopersici_4287
OBR09176.1_Transmembrane_amino_acid_transporter_Colletotrichum_higginsianum_IMI_349063
CCT69221.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
KNB10629.1_hypothetical_protein_FOXG_10779_Fusarium_oxysporum_f._sp._lycopersici_4287
EKJ75991.1_hypothetical_protein_FPSE_03763_Fusarium_pseudograminearum_CS3096
EYB31020.1_hypothetical_protein_FG05_02357_Fusarium_graminearum
CCT72337.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
EWG47197.1_hypothetical_protein_FVEG_07377_Fusarium_verticillioides_7600
KNB00839.1_hypothetical_protein_FOXG_04256_Fusarium_oxysporum_f._sp._lycopersici_4287
EYB34371.1_hypothetical_protein_FG05_01876_Fusarium_graminearum
EKJ74437.1_hypothetical_protein_FPSE_05402_Fusarium_pseudograminearum_CS3096
ESU07241.1_hypothetical_protein_FGSG_01876_Fusarium_graminearum_PH-1
EWG54730.1_hypothetical_protein_FVEG_12861_Fusarium_verticillioides_7600
KNB19562.1_hypothetical_protein_FOXG_16632_Fusarium_oxysporum_f._sp._lycopersici_4287
CCT75447.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
EKJ76333.1_hypothetical_protein_FPSE_03469_Fusarium_pseudograminearum_CS3096
ESU09410.1_hypothetical_protein_FGSG_03778_Fusarium_graminearum_PH-1
EYB32889.1_hypothetical_protein_FG05_03778_Fusarium_graminearum
CCT70916.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
EWG45961.1_hypothetical_protein_FVEG_06592_Fusarium_verticillioides_7600
KNB07868.1_hypothetical_protein_FOXG_09937_Fusarium_oxysporum_f._sp._lycopersici_4287
EKJ72454.1_hypothetical_protein_FPSE_07365_Fusarium_pseudograminearum_CS3096
ESU15601.1_hypothetical_protein_FGSG_09078_Fusarium_graminearum_PH-1
EYB22071.1_hypothetical_protein_FG05_09078_Fusarium_graminearum
CCT63188.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
EWG35966.1_hypothetical_protein_FVEG_00145_Fusarium_verticillioides_7600
EKJ69739.1_hypothetical_protein_FPSE_10065_Fusarium_pseudograminearum_CS3096
ESU05235.1_hypothetical_protein_FGSG_00118_Fusarium_graminearum_PH-1
EYB24160.1_hypothetical_protein_FG05_00118_Fusarium_graminearum
EKJ71924.1_hypothetical_protein_FPSE_07927_Fusarium_pseudograminearum_CS3096
ESU14125.1_hypothetical_protein_FGSG_07810_Fusarium_graminearum_PH-1
EYB22335.1_hypothetical_protein_FG05_07810_Fusarium_graminearum
CCT73361.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
KNB15024.1_hypothetical_protein_FOXG_13748_Fusarium_oxysporum_f._sp._lycopersici_4287
EWG52425.1_hypothetical_protein_FVEG_11179_Fusarium_verticillioides_7600
ESU08630.1_hypothetical_protein_FGSG_04468_Fusarium_graminearum_PH-1
EKJ71552.1_hypothetical_protein_FPSE_08293_Fusarium_pseudograminearum_CS3096
EYB24985.1_hypothetical_protein_FG05_04468_Fusarium_graminearum
ESU10349.1_hypothetical_protein_FGSG_02950_Fusarium_graminearum_PH-1
EYB23622.1_hypothetical_protein_FG05_02950_Fusarium_graminearum
EKJ73129.1_hypothetical_protein_FPSE_06742_Fusarium_pseudograminearum_CS3096
CCT76051.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
EWG52077.1_hypothetical_protein_FVEG_10900_Fusarium_verticillioides_7600
KNB12663.1_hypothetical_protein_FOXG_12214_Fusarium_oxysporum_f._sp._lycopersici_4287
KNB12664.1_hypothetical_protein_FOXG_12214_Fusarium_oxysporum_f._sp._lycopersici_4287
OAQ61465.1_amino_acid_transporter_Pochonia_chlamydosporia_170
EHA49606.1_amino_acid_transporter_Pyricularia_oryzae_70-15
PIA93489.1_N_amino_acid_transport_system_protein_Cercospora_beticola
CBX71203.1_related_to_neutral_amino_acid_permease_Sporisorium_reilianum_SRZ2
KIS69947.1_hypothetical_protein_UMAG_02459_Ustilago_maydis_521
SAM67126.1_related_to_neutral_amino_acid_permease_Ustilago_bromivora
SHO76171.1_Uncharacterised_protein_MSYG_0506_Malassezia_sympodialis_ATCC_42132
PIA93255.1_N_amino_acid_transport_system_protein_Cercospora_beticola
SMY29638.1_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
OAQ63690.1_amino_acid_transporter_Pochonia_chlamydosporia_170
PIA94364.1_N_amino_acid_transport_system_protein_Cercospora_beticola
KNA98425.1_hypothetical_protein_FOXG_02785_Fusarium_oxysporum_f._sp._lycopersici_4287
KNA98427.1_hypothetical_protein_FOXG_02785_Fusarium_oxysporum_f._sp._lycopersici_4287
KNA98426.1_hypothetical_protein_FOXG_02785_Fusarium_oxysporum_f._sp._lycopersici_4287
CCT70589.1_related_to_neutral_amino_acid_permease_Fusarium_fujikuroi_IMI_58289
EWG38356.1_hypothetical_protein_FVEG_01600_Fusarium_verticillioides_7600
EKJ77443.1_hypothetical_protein_FPSE_02316_Fusarium_pseudograminearum_CS3096
ESU15344.1_hypothetical_protein_FGSG_08055_Fusarium_graminearum_PH-1
EYB25111.1_hypothetical_protein_FG05_08055_Fusarium_graminearum
OBR02607.1_Transmembrane_amino_acid_transporter_Colletotrichum_higginsianum_IMI_349063
OBR09895.1_Transmembrane_amino_acid_transporter_Colletotrichum_higginsianum_IMI_349063
EHA49663.1_aromatic_and_neutral_alpha66_amino_acid_permease_Pyricularia_oryzae_70-15
AEO55911.1_hypothetical_protein_MYCTH_2089733_Thermothelomyces_thermophilus_ATCC_42464
AEO66578.1_hypothetical_protein_THITE_126080_Thielavia_terrestris_NRRL_8126
ATZ59465.1_hypothetical_protein_BCIN_16g02440_Botrytis_cinerea_B05.10
ATZ58469.1_hypothetical_protein_BCIN_16g02440_Botrytis_cinerea_B05.10
APA05823.1_hypothetical_protein_sscle_01g005930_Sclerotinia_sclerotiorum_1980_UF-70
ATZ46813.1_hypothetical_protein_BCIN_02g01660_Botrytis_cinerea_B05.10
PIA98405.1_N_amino_acid_transport_system_protein_Cercospora_beticola
SMY28985.1_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
EAL85095.2_amino_acid_transporter_putative_Aspergillus_fumigatus_Af293
XP_001401336.1_amino_acid_transporter_Aspergillus_niger_CBS_513.88
XP_001391158.1_amino_acid_transporter_Aspergillus_niger_CBS_513.88
KZN84468.1_N_amino_acid_transport_system_protein_Penicillium_chrysogenum
XP_001397083.2_amino_acid_transporter_Aspergillus_niger_CBS_513.88
EAL93843.2_neutral_amino_acid_permease_putative_Aspergillus_fumigatus_Af293
CBF83940.1_TPA:_amino_acid_transporter_(Eurofung)_Aspergillus_nidulans_FGSC_A4
KZN86454.1_N_amino_acid_transport_system_protein_Penicillium_chrysogenum
EAL94045.1_amino_acid_transporter_putative_Aspergillus_fumigatus_Af293
XP_001402296.1_amino_acid_transporter_Aspergillus_niger_CBS_513.88
XP_001390417.1_amino_acid_transporter_Aspergillus_niger_CBS_513.88
EAL87116.1_neutral_amino_acid_permease_Aspergillus_fumigatus_Af293
KZN85696.1_N_amino_acid_transport_system_protein_Penicillium_chrysogenum
XP_001397349.2_amino_acid_transporter_Aspergillus_niger_CBS_513.88
CBF83190.1_TPA:_neutral_amino_acid_transporter_(Eurofung)_Aspergillus_nidulans_FGSC_A4
XP_001400898.1_amino_acid_transporter_Aspergillus_niger_CBS_513.88
BAE60325.1_unnamed_protein_product_Aspergillus_oryzae_RIB40
EAL84542.1_neutral_amino_acid_permease_Aspergillus_fumigatus_Af293
KZN83521.1_N_amino_acid_transport_system_protein_Penicillium_chrysogenum
CBF84423.1_TPA:_amino_acid_transporter_(Eurofung)_Aspergillus_nidulans_FGSC_A4
XP_001398067.1_amino_acid_transporter_Aspergillus_niger_CBS_513.88
CBF85002.1_TPA:_amino_acid_transporter_(Eurofung)_Aspergillus_nidulans_FGSC_A4
EAL85274.1_neutral_amino_acid_permease_Aspergillus_fumigatus_Af293
0.3

S3.4 Detailed CjSLC13-phylogenetic tree with all homolog nominations and bootstrap values


ABO24356.1 anion transporter Shewanella loihica PV-4
ASJ96326.1 Anion transporter Shewanella harrisflavi
ACA87365.1 anion transporter Shewanella woodyi ATCC 51908
ABV36171.1 anion transporter Shewanella sediminis HAW-EB3
AQS37026.1 anion transporter Shewanella psychrophila
ACQ92009.1 anion transporter Tolumonas auensis DSM 9187
ATU98106.1 Anion transporter Aeromonas salmonicida
ATY80866.1 Anion transporter Aeromonas veronii
ASX12174.1 Anion transporter Aeromonas dhakensis
ANT68114.1 Anion transporter Aeromonas hydrophila
APJ15261.1 Anion transporter Aeromonas hydrophila
ATC51191.1 SLC13 family permease Neisseria mucosa
SNU79793.1 NadC family protein Neisseria zoodegmatis
AJE16518.1 anion:sodium symporter Neisseria elongata subsp. glycolytica ATCC 29315
AKX60310.1 anion:sodium symporter Oblitimonas alkaliphila
AKT91523.1 DASS family sodium/dicarboxylate symporter Campylobacter gracilis
ASM40433.1 DASS family sodium/dicarboxylate symporter Campylobacter sputorum
APX61785.1 divalent ion symporter protein Acinetobacter schindleri
AUC06160.1 DASS family sodium-coupled anion symporter Acinetobacter lwoffii
ASE25396.1 anion:sodium symporter Psychrobacter cryohalolentis
ARE65691.1 anion:sodium symporter Moraxella catarrhalis
AKG19742.1 anion:sodium symporter Moraxella bovoculi
ANB91287.1 anion:sodium symporter Moraxella ovis
APC13050.1 Sodium-dependent dicarboxylate transporter SdcS Providencia rettgeri
EKT54199.1 Sodium:sulfate symporter family protein Providencia sneebia DSM 19967
APG50872.1 Anion transporter Providencia stuartii
ATG16483.1 Anion transporter Providencia alcalifaciens
ATG77971.1 Anion transporter Proteus mirabilis
ATM98216.1 Anion transporter Proteus vulgaris
ATF54645.1 Anion transporter Morganella morganii
CDM91311.1 Uncharacterized transporter HI_0608 Xenorhabdus bovienii
CDG18831.1 Uncharacterized transporter HI_0608 Xenorhabdus doucetiae
CDG20292.1 Uncharacterized transporter HI_0608 Xenorhabdus poinarii_G6
AQM42557.1 Anion transporter Xenorhabdus hominickii
CEK21857.1 conserved hypothetical protein putative membrane protein Xenorhabdus nematophila AN6/1
CAE13221.1 unnamed protein product Photorhabdus laumondii subsp. laumondii_TTO1
AKH62608.1 Anion transporter Photorhabdus thracensis
CAQ82937.1 conserved hypothetical protein Photorhabdus asymbiotica
AML56146.1 Di-and tricarboxylate transporter Serratia rubidaea
ATM75992.1 Anion transporter Serratia fonticola
SBT59719.1 Na(+)dicarboxylate symporter Plesiomonas shigelloides
SBT59720.1 Na(+)dicarboxylate symporter Plesiomonas shigelloides
AIJ07134.1 transporter divalent anion:Na+ symporter family Edwardsiella anguillarum ET080813
ARD20114.1 Anion transporter Edwardsiella piscicida
AQV96274.1 Anion transporter Edwardsiella hoshinae
ATI65019.1 Anion transporter Edwardsiella tarda
AJH16648.1 solute carrier family 13 (sodium-dependent dicarboxylate transporter) member 2/3/5 Myroides profundi
ALU25483.1 transporter Myroides odoratimimus
EHQ42385.1 anion transporter Myroides odoratus DSM 2801
ATA82100.1 transporter Capnocytophaga leadbetteri
SNV12571.1 Na(+)dicarboxylate symporter Capnocytophaga haemolytica
ATA83012.1 transporter Capnocytophaga sputigena
ACU92343.1 anion transporter Capnocytophaga ochracea DSM 7271
ATA86712.1 transporter Capnocytophaga gingivalis
AJC85066.1 DASS family sodium/dicarboxylate symporter Campylobacter peloridis LMG 23910
AJC88067.1 DASS family sodium/dicarboxylate symporter Campylobacter insulaenigrae NCTC 12927
ARE80823.1 DASS family sodium/anion symporter Campylobacter helveticus
ARJ56866.1 DASS family sodium/anion symporter Campylobacter cuniculorum DSM 23162 LMG 24588
AEC17406.1 Sodium:sulfate symporter transmembrane region Gallibacterium anatis UMN179
CBW15017.1 unnamed protein product Haemophilus parainfluenzae T3T1
SNV56205.1 di-and tricarboxylate transporter Haemophilus pittmaniae
CKG77814.1 di-and tricarboxylate transporter Haemophilus influenzae
AKU63210.1 transporter Aggregatibacter aphrophilus
ANU81783.1 transporter Aggregatibacter actinomycetemcomitans
ACA32505.1 anion transporter Histophilus somni 2336
ARU76668.1 transporter Histophilus somni
SNV82450.1 sodium/sulfate symporter Pasteurella multocida
SNV85192.1 sodium/sulfate symporter Pasteurella dagmatis
AAU38522.1 CitT protein Mannheimia succiniciproducens MBEL55E
ABR74922.1 anion transporter Actinobacillus succinogenes 130Z
AIZ78270.1 transporter Actinobacillus equuli subsp. equuli
SNV25394.1 di- and tricarboxylate transporter Actinobacillus suis
ACE61955.1 hypothetical protein APP7_1303 Actinobacillus pleuropneumoniae serovar 7 str. AP76
AMW16249.1 transporter Glaesserella parasuis
AHG76007.1 transporter Mannheimia varigena USDA-ARS-USMARC-1296
ASW70950.1 transporter Mannheimia haemolytica
AHG84472.1 transporter Bibersteinia trehalosi USDA-ARS-USMARC-189
AIZ78271.1 transporter Actinobacillus equuli subsp. equuli
SNV25401.1 CitT protein Actinobacillus suis
AAU37985.1 CitT protein Mannheimia succiniciproducens MBEL55E
AHG84606.1 CitT protein Bibersteinia trehalosi USDA-ARS-USMARC-189
ACX82492.1 transporter Aggregatibacter actinomycetemcomitans D11S-1
ANU82213.1 transporter Aggregatibacter actinomycetemcomitans
CBW15384.1 unnamed protein product Haemophilus parainfluenzae T3T1
SNV89673.1 di-and tricarboxylate transporter Haemophilus pittmaniae
AMD89325.1 sodium:proton antiporter Desulfovibrio fairfieldensis
ASN94365.1 anion:sodium symporter Enterocloster bolteae
CBL25620.1 anion transporter Ruminococcus torques L2-14
ATA67342.1 transporter Capnocytophaga cynodegmi
ATA88418.1 transporter Capnocytophaga stomatis
ATA93084.1 transporter Capnocytophaga canimorsus
ANW97023.1 transporter Wenyinghzhangia fucanilytica
AFL97634.1 anion transporter Ornithobacterium rhinotracheale DSM 15997
EHQ42004.1 anion transporter Myroides odoratus DSM 2801
AJH16204.1 solute carrier family 13 (sodium-dependent dicarboxylate transporter) member 2/3/5 Myroides profundi
ALU25034.1 divalent cation transporter Myroides odoratimimus
AEC16573.1 sodium:sulfate symporter transmembrane region Gallibacterium anatis UMN179
SNV76903.1 sodium-dependent dicarboxylate transporter SdcS Pasteurella multocida
SMD43530.1 solute carrier family 13 (sodium-dependent dicarboxylate transporter) member 2/3/5 Aquiflexum balticum DSM 16537
ABB44423.1 Sodium/sulphate symporter Sulfurimonas denitrificans DSM 1251
ACY15952.1 anion transporter Haliangium ochraceum DSM 14365
AOP33727.1 sodium:sulfate symporter Leptospira hyperaryensis
ASV10779.1 sodium:sulfate symporter Leptospira santarosai
AKU65994.1 transporter Schaalia meyeri
SDT85390.1 solute carrier family 13 (sodium-dependent dicarboxylate transporter) member 2/3/5 Schaalia radingae
SNV89440.1 putative dicarboxylate uptake system Corynebacterium cystitidis
ASB51111.1 transporter Alkalitalea saponilacus
ASE18451.1 dihydroorotate dehydrogenase Prevotella melaninogenica
AUI56059.1 dihydroorotate dehydrogenase Prevotella jejuni
ANR73464.1 dihydroorotate dehydrogenase Prevotella scopos JCM 17725
AEA19996.1 transporter DASS family Prevotella denticola F0289
AKU69887.1 dihydroorotate dehydrogenase Prevotella fusca JCM 17724
BAU17339.1 sodium-dependent transporter Prevotella intermedia
ALQ47954.1 dihydroorotate dehydrogenase Prevotella enoeca
ATA83209.1 dihydroorotate dehydrogenase Capnocytophaga sputigena
ACU93469.1 anion transporter Capnocytophaga ochracea DSM 7271
ATA80889.1 dihydroorotate dehydrogenase Capnocytophaga leadbetteri
AGB29881.1 anion transporter Prevotella dentalis DSM 3688
ADE83476.1 transporter divalent anion:sodium symporter (DASS) family Prevotella ruminicola 23
ANE32595.1 DASS family sodium/dicarboxylate symporter Campylobacter hyointestinalis subsp. hyointestinalis LMG 9260
ADV44987.1 anion transporter Bacteroides helcogenes P 36-108
ALJ44070.1 Sodium-dependent dicarboxylate transporter SdcS Bacteroides thetaiotaomicron
ANL58681.1 dihydroorotate dehydrogenase Bacteroides caecimuris
ASW16375.1 dihydroorotate dehydrogenase Bacteroides caccae
EGJ72076.1 anion transporter Bacteroides coprosuis DSM 18011
AMO54831.1 sodium/di- and tricarboxylate cotransporter Endozoicomonas montiporae CL-33
AMO54833.1 sodium/di- and tricarboxylate cotransporter Endozoicomonas montiporae CL-33
APP05902.1 dihydroorotate dehydrogenase Vibrio harveyi
AQW58312.1 dihydroorotate dehydrogenase Vibrio owensii
ARR43252.1 dihydroorotate dehydrogenase Vibrio campbellii
ATI46928.1 dihydroorotate dehydrogenase Vibrio parahaemolyticus
ACY50412.1 transporter Vibrio antiquarius
ANQ27525.1 dihydroorotate dehydrogenase Vibrio natriegens
ARP48120.1 Sodium-dependent dicarboxylate transporter SdcS Vibrio alginolyticus
ASI95231.1 dihydroorotate dehydrogenase Vibrio rotiferianus
AUL96897.1 Transporter putative Vibrio vulnificus
ASI89680.1 dihydroorotate dehydrogenase Vibrio mediterranei
AIW15234.1 dihydroorotate dehydrogenase Vibrio tubiashii ATCC 19109
KPH28116.1 dihydroorotate dehydrogenase Vibrio coralliilyticus
ANU35239.1 putative transporter Vibrio scophthalmi
CAV17283.1 putative transporter Vibrio atlanticus
ANO33975.1 dihydroorotate dehydrogenase Vibrio breoganii
CCO59240.1 putative SODIUM/DICARBOXYLATE COTRANSPORTER Vibrio nigripulchritudo
CDQ49139.1 putative sodium/sulphate symporter Vibrio anguillarum
ASA55509.1 dihydroorotate dehydrogenase Vibrio gazogenes
AMF95504.1 SLC13 family permease Vibrio fluvialis
AOW83723.1 dihydroorotate dehydrogenase Vibrio mimicus
SNC56963.1 Na(+)/dicarboxylate symporter Vibrio cholerae
AJR06832.1 putative transporter Photobacterium gaetbulicola Gung47
CAG21658.1 putative transporter Photobacterium profundum SS9
AMG30857.1 SLC13 family permease Grimontia hollisae
ATG73044.1 dihydroorotate dehydrogenase Zobellella denitrificans
ADL25270.1 anion transporter Fibrobacter succinogenes subsp. succinogenes S85
ASB35572.1 anion transporter Akkermansia muciniphila
AJQ92389.1 Di- and tricarboxylate transporter Gynuella sunshinyii YC6258
ABM04080.1 anion transporter Psychromonas ingrahamii 37
ABM04082.1 anion transporter Psychromonas ingrahamii 37
ALP42143.1 sodium-dependent transporter Aeromonas schubertii
ATY80542.1 dihydroorotate dehydrogenase Aeromonas veronii
ATU99045.1 dihydroorotate dehydrogenase Aeromonas salmonicida
SBT60994.1 Na(+)dicarboxylate symporter Plesiomonas shigelloides
AMG32071.1 SLC13 family permease Grimontia hollisae
ABV37811.1 transporter divalent anion:sodium symporter family Shewanella sediminis HAW-EB3
ARR49661.1 dihydroorotate dehydrogenase Photobacterium damselae subsp. damselae
CAG19627.1 putative transporter Photobacterium profundum SS9
AJR09251.1 putative transporter Photobacterium gaetbulicola Gung47
ASI90642.1 dihydroorotate dehydrogenase Vibrio mediterranei
CCO57205.1 putative SODIUM/DICARBOXYLATE COTRANSPORTER Vibrio nigripulchritudo


AAW85547.1 transporter divalent anion:sodium symporter family Aliivibrio fischeri ES114
CED71136.1 putative sodium/sulfate symporter Aliivibrio wodanis
ANU36343.1 putative transporter Vibrio scophthalmi
AOW84000.1 dihydroorotate dehydrogenase Vibrio mimicus
SNC56009.1 Na(+)/dicarboxylate symporter Vibrio cholerae
AUI86043.1 dihydroorotate dehydrogenase Vibrio azureus
AUL96038.1 Transporter putative Vibrio vulnificus
AQW59384.1 dihydroorotate dehydrogenase Vibrio owensii
ARR44265.1 dihydroorotate dehydrogenase Vibrio campbellii
APP04352.1 dihydroorotate dehydrogenase Vibrio harveyi
ADT86955.1 hypothetical transporter Vibrio furnissii NCTC 11218
AMF96135.1 dihydroorotate dehydrogenase Vibrio fluvialis
CAV18414.1 putative transporter Vibrio atlanticus
CDQ50296.1 putative transporter Vibrio anguillarum
ASI96633.1 dihydroorotate dehydrogenase Vibrio rotiferianus
ANQ26304.1 dihydroorotate dehydrogenase Vibrio natriegens
ATI48301.1 SLC13 family permease Vibrio parahaemolyticus
ACY51746.1 transporter Vibrio antiquarius
ARP46497.1 Sodium-dependent dicarboxylate transporter SdcS Vibrio alginolyticus
ABZ96373.1 Conserved hypothetical protein putative membrane protein Leptospira biflexa serovar Patoc strain Patoc 1 (Paris)
ACR12291.1 transporter divalent anion:Na+ symporter family Terendinibacter turnerae T7901
ARU55627.1 DASS family transporter Oleiphilus messinensis
ATX78134.1 sodium:sulfate symporter Reinekea forsetii
BAS28709.1 transporter Limnochorda pilosa
ABF91466.1 sodium:sulfate symporter Myxococcus xanthus DK 1622
ATB48520.1 sodium:sulfate symporter Myxococcus macrosporus DSM 14697
CAM69535.1 putative sodium/sulphate symporter Leishmania infantum JPCM5
CBZ35716.1 sodium/sulphate symporter putative Leishmania donovani
CAJ05751.1 putative sodium/sulphate symporter Leishmania major strain Friedlin
CBZ28667.1 putative sodium/sulphate symporter Leishmania mexicana MHOM/GT/2001/U1103
AIN99940.1 sodium/sulfate symporter putative Leishmania panamensis
CAM39844.1 putative sodium/sulphate symporter Leishmania braziliensis MHOM/BR/75/M2904
EAN80069.1 Sodium/sulphate symporter putative Trypanosoma brucei brucei TREU927
BAM81695.1 probable low-affinity phosphate transporter Pho91p Cyanidioschyzon merolae strain 10D
AUB25470.1 hypothetical protein CKF44_001962 Cryptococcus neoformans var. grubii
AUB25470.1 phosphate transporter Cryptococcus neoformans var. grubii
ADV22567.1 Inorganic phosphate transporter putative Cryptococcus gattii WM276
PAV16470.1 Sodium sulfate symporter Phellinus noxius
PAV16470.1 Sodium sulfate symporter Pyrrhoderma noxium
SHO75875.1 Similar to S.cerevisiae protein PHO91 (Low-affinity vacuolar phosphate transporter) Malassezia sympodialis ATCC 42132
SAM85044.1 probable PHO91-similarity to Pho87p and Pho90p Ustilago bromivora
CBQ69716.1 probable PHO91-similarity to Pho87p and Pho90p Sporisorium reilianum SRZ2
KIS66697.1 putative inorganic phosphate transmembrane transporter Ustilago maydis 521
CCT74136.1 related to PHO87 protein Fusarium fujikuroi IMI 58289
EWG52638.1 hypothetical protein FVEG_11326 Fusarium verticillioides 7600
EWG52639.1 hypothetical protein FVEG_11326 Fusarium verticillioides 7600
KNB15281.1 hypothetical protein FOXG_13901 Fusarium oxysporum f. sp. lycopersici 4287
EWG52640.1 hypothetical protein FVEG_11326 Fusarium verticillioides 7600
EKJ69226.1 hypothetical protein FPSE_10595 Fusarium pseudograminearum CS3096
ESU08777.1 hypothetical protein FGSG_12273 Fusarium graminearum PH-1
EYB25023.1 hypothetical protein FGSG_12273 Fusarium graminearum
OAQ71115.1 multidrug resistance protein 1 2 3 (p glycoprotein 1 2 3) Pochonia chlamydosporia 170
KYK58446.1 putative PHO87 protein Drechmeria coniospora
KUI68045.1 hypothetical protein VM1G_03279 Valsa mali
OBR03508.1 SPX domain-containing protein Colletotrichum higginsianum IMI 349063
EHA48305.1 SPX domain-containing protein Magnaporthe oryzae 70-15
EHA48305.1 SPX domain-containing protein Pyricularia oryzae 70-15
EAA27007.3 SPX domain-containing protein Neurospora crassa OR74A
AEO61566.1 hypothetical protein MYCTH_2311845 Thermothelomyces thermophila ATCC 42464
AEO61566.1 hypothetical protein MYCTH_2311845 Thermothelomyces thermophila ATCC 42464
AEO71324.1 hypothetical protein THITE_2123538 Thermothielavioides terrestris NRRL 8126
AEO71324.1 hypothetical protein THITE_2123538 Thielavia terrestris NRRL 8126
APA12492.1 hypothetical protein sscle_09g072620 Sclerotinia sclerotiorum 1980 UF-70
ATZ49047.1 hypothetical protein BCIN_04g02470 Botrytis cinerea B05.10
PIB00761.1 putative transporter Cercospora beticola
SMY23843.1 unnamed protein product Zymoseptoria tritici ST99CH_1A5
KZN93850.1 putative transporter Penicillium chrysogenum
CBF89425.1 TPA phosphate transport (Eurofung) Aspergillus nidulans FGSC A4
XP_001388746.2 plasma membrane phosphate transporter Pho87 Aspergillus niger CBS 513.88
BAE58036.1 unnamed protein product Aspergillus oryzae RIB40
EAL88138.1 plasma membrane phosphate transporter Pho87 putative Aspergillus fumigatus Af293
CAA18293.1 membrane transporter (predicted) Schizosaccharomyces pombe
XP_013020242.1 membrane transporter Schizosaccharomyces octosporus yFS286
AQW06211.1 hypothetical protein YALI1_E36540g Yarrowia lipolytica
CAG80222.1 YALI0E31064p Yarrowia lipolytica CLIB122
ANB15340.1 Pho91p Sugiyamaella lignohabitans
AQZ13778.1 PHO91 (YNR013C) Zygosaccharomyces parabailii
CDH11007.1 probable Low-affinity phosphate transporter PHO91 Zygosaccharomyces bailii ISA1307
CDF91635.1 ZYBA0S12-03444g1.1 Zygosaccharomyces bailii CLIB 213
CDH17257.1 probable Low-affinity phosphate transporter PHO91 Zygosaccharomyces bailii ISA1307
AQZ18194.1 PHO91 (YNR013C) Zygosaccharomyces parabailii
GAV53886.1 hypothetical protein ZYGR_0AK03880 Zygosaccharomyces rouxii NBRC 110957
CAR27892.1 ZYRO0D08998p Zygosaccharomyces rouxii
GAV48974.1 hypothetical protein ZYGR_0N03790 Zygosaccharomyces rouxii NBRC 110957
CCE90054.1 hypothetical protein TDEL_0A07220 Torulaspora delbrueckii
OXB47732.1 hypothetical protein B1J92_I06320g Candida glabrata
BAP73200.1 uncharacterized transporter YNR013C Kluyveromyces marxianus
CAG98067.1 KLLA0F06160p Kluyveromyces lactis
SCU83017.1 LAMI_0C01706g1.1 Lachancea mirantina
SCW00353.1 LAFE_0C02322g1.1 Lachancea fermentati
CAR21276.1 KLTH0A05962p Lachancea thermotolerans CBS 6340
SCV03946.1 LAME_0H14510g1.1 Lachancea meyersii CBS 8951
SCU78165.1 LANO_0A02278g1.1 Lachancea nothofagi CBS 11611
SCU80240.1 LADA_0B05886g1.1 Lachancea dasiensis CBS 10888
CCC67021.1 hypothetical protein NCAS_0A04630 Naumovozyma castellii CBS 4309
CCD27360.1 hypothetical protein NDAI_0K01690 Naumovozyma dairenensis CBS 421
KOG97166.1 PHO91-like protein Saccharomyces eubayanus
EJS41896.1 pho91p Saccharomyces arboricola H-6
KZV08641.1 PHO91 Saccharomyces cerevisiae
CCE64421.1 hypothetical protein TPHA_0H02170 Tetrapisispora phaffii CBS 4417
CCH62109.1 hypothetical protein TBLA_0G01650 Tetrapisispora blattae CBS 6284
CCF58682.1 hypothetical protein KAFR_0F00850 Kazachstania africana CBS 2517
CCK69806.1 hypothetical protein KNAG_0D00540 Kazachstania naganishii CBS 8797
AAS52209.2 ADR289Cp Eremothecium gossypii ATCC 10895
AET37864.1 Hypothetical protein Ecym_2111 Eremothecium cymbalariae DBVPG 7215
AMD21453.1 HER174Cp Eremothecium sinecaudum
CCH42708.1 putative membrane protein Wickerhamomyces ciferrii
ODV76005.1 hypothetical protein CYBJADRAFT_165372 Cyberlindnera jadinii NRRL_Y-1542
AOW25891.1 Pho91p Candida albicans SC5314
CAX44451.1 constitutive low-affinity phosphate transporter putative Candida dubliniensis CD36
XP_002550314.1 hypothetical protein CTRG_04512 Candida tropicalis MYA-3404
CCE40782.1 hypothetical protein CPAR2_108190 Candida parapsilosis
CCG22637.1 Pho91 low-affinity phosphate transporter Candida orthopsilosis Co_90-125
ABN66240.2 low-affinity phosphate transporter Scheffersomyces stipitis CBS 6054
SGZ49180.1 CIC11C00000002907 Candida intermedia
XP_002616076.1 hypothetical protein CLUG_03257 Clavispora lusitaniae ATCC 42720
XP_025339998.1 hypothetical protein CXQ85_001353 Candida haemulonis
XP_028892879.1 uncharacterized protein CJI97_000494 Candida auris
XP_001482569.1 hypothetical protein PGUG_05589 Meyerozyma guilliermondii ATCC 6260
CAG86855.2 DEHA2D05676p Debaryomyces hansenii CBS767
CCE72693.1 Piso0_000284 Millerozyma farinosa CBS 7064
CCE73254.1 Piso0_000284 Millerozyma farinosa CBS 7064
ANZ73761.1 BA75_00526T0 Komagataella pastoris
AOA64048.1 GQ67_03567T0 Komagataella phaffii
XP_029319619.1 uncharacterized protein C5L36_0A07420 Pichia kudriavzevii
ESW97410.1 putative transporter C3B8_04c Ogataea parapolymorpha DL-1
XP_018212262.1 hypothetical protein OGAPODRAFT_15724 Ogataea polymorpha
ESW95885.1 Inorganic phosphate transporter PHO87 Ogataea parapolymorpha DL-1
XP_018213462.1 hypothetical protein OGAPODRAFT_15408 Ogataea polymorpha
AOW26258.1 SPX domain-containing inorganic phosphate transporter Candida albicans SC5314
CAX44818.1 inorganic phosphate transporter PHO87 putative Candida dubliniensis CD36
XP_002550541.1 hypothetical protein CTRG_04839 Candida tropicalis MYA-3404
EAZ63809.2 phosphate permease Scheffersomyces stipitis CBS 6054
CAG98936.2 DEHA2E06842p Debaryomyces hansenii CBS767
CCE83733.1 Piso0_004319 Millerozyma farinosa CBS 7064
CCE84764.1 Piso0_004319 Millerozyma farinosa CBS 7064
SGZ56293.1 CIC11C00000003765 Candida intermedia
XP_025340986.1 hypothetical protein CXQ85_001825 Candida haemulonis
XP_028889757.1 uncharacterized protein CJI97_002493 Candida auris
CCE40693.1 hypothetical protein CPAR2_107280 Candida parapsilosis
CCG22553.1 Pho87 phosphate permease Candida orthopsilosis Co_90-125
CCH42353.1 putative membrane protein Wickerhamomyces ciferrii
ODV72231.1 SPX-domain-containing protein Cyberlindnera jadinii NRRL_Y-1542
AAS53999.2 AFR628Cp Eremothecium gossypii ATCC 10895
AET41432.1 Hypothetical protein Ecym_8141 Eremothecium cymbalariae DBVPG 7215
AMD19585.1 HCL366Wp Eremothecium sinecaudum
BAP69344.1 inorganic phosphate transporter PHO87 Kluyveromyces marxianus
CAH01201.1 KLLA0C03454p Kluyveromyces lactis
SCU86930.1 LADA_0H03532g1.1 Lachancea dasiensis CBS 10888
SCU99492.1 LAME_0G03378g1.1 Lachancea meyersii CBS 8951
SCV03308.1 LANO_0G03356g1.1 Lachancea nothofagi CBS 11611
CAR23931.1 KLTH0F03652p Lachancea thermotolerans CBS 6340
SCV00838.1 LAMI_0G07624g1.1 Lachancea mirantina
SCW03126.1 LAFE_0G03510g1.1 Lachancea fermentati
KZV10076.1 PHO90 Saccharomyces cerevisiae
EJS43199.1 pho90p Saccharomyces arboricola H-6
KOG98428.1 PHO90-like protein Saccharomyces eubayanus
CCF58752.1 hypothetical protein KAFR_0F01550 Kazachstania africana CBS 2517
CCK69328.1 hypothetical protein KNAG_0C02170 Kazachstania naganishii CBS 8797
CCH63059.1 hypothetical protein TBLA_0J00590 Tetrapisispora blattae CBS 6284



Figures S4. Docking studies and pore radius trajectories

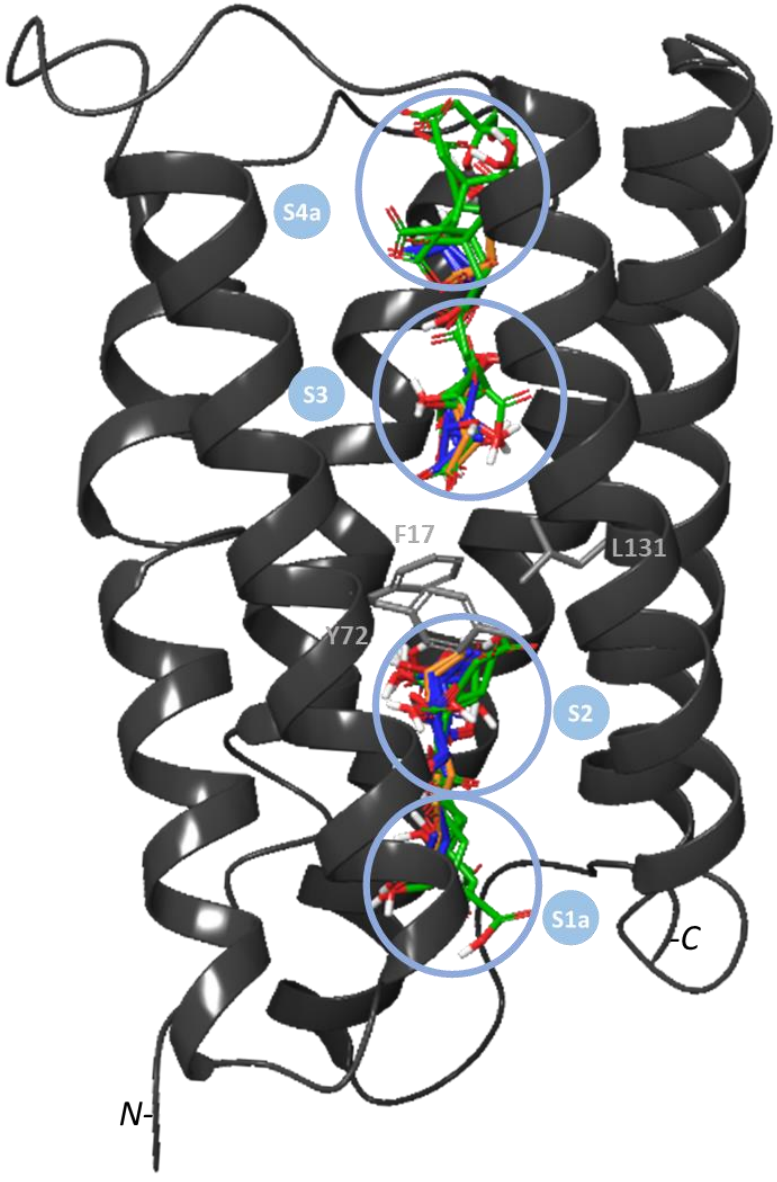


Figure S4.1 Molecular docking of *Escherichia coli* Satp 3D-model, based on SatP_Ck structure, with the substrates lactate (blue ligand), succinate (orange ligand) and citrate (green ligand). The four binding sites are depicted from S1 to S4 sites, including the localization of N- and C-terminal of the protein. The narrowest hydrophobic constriction of the anion pathway formed by F17, Y72 and L131 is also represented.

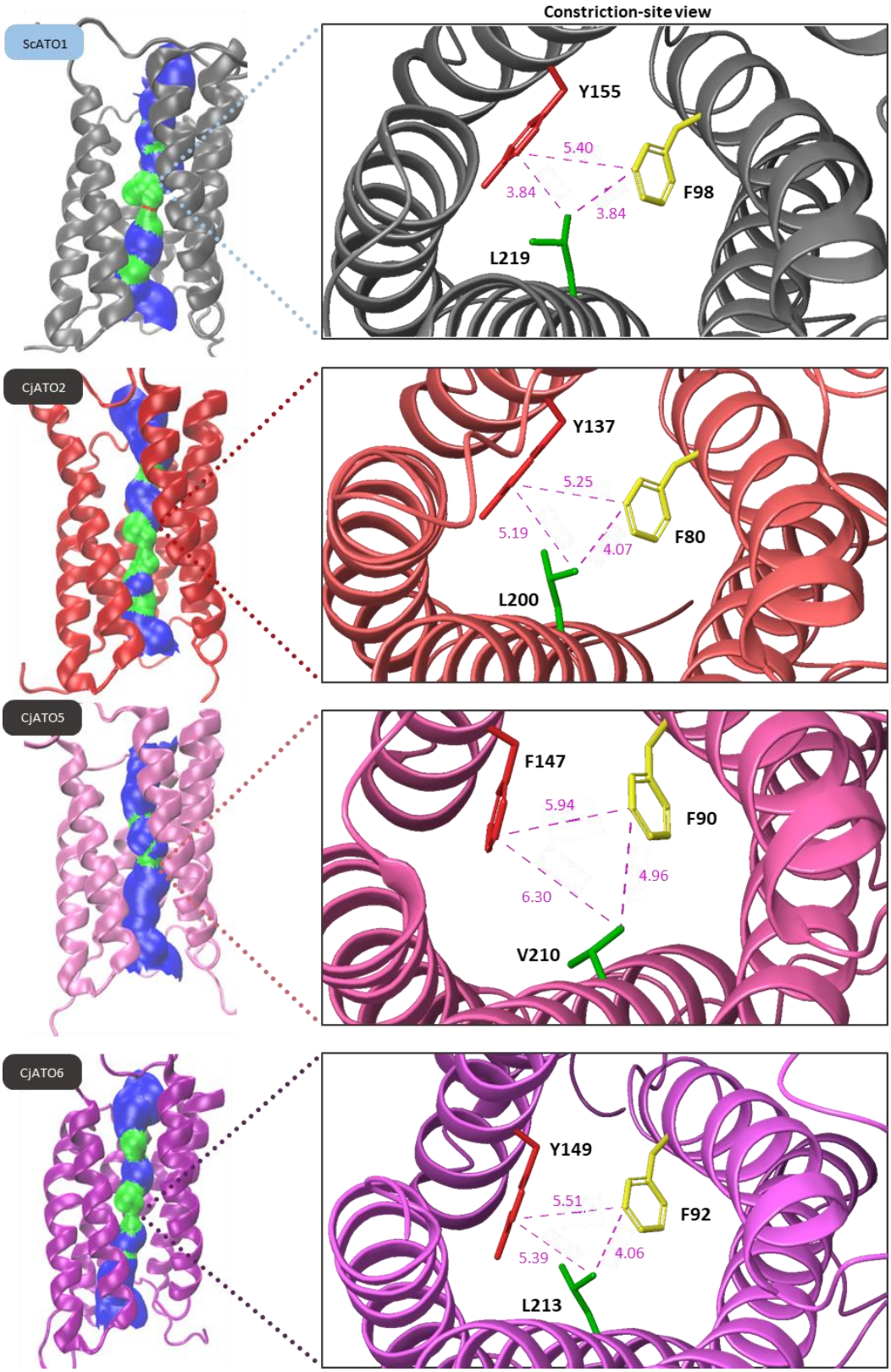


Figure S4.2 Channel structures of Ato proteins are depicted in the left-side in cartoon, having the color scheme represented by blue (larger aperture), green (intermediate pore size) and red (more constricted) colors. In the right-side is depicted the top view of the narrowest hydrophobic constriction site of the anion pathway, including the predicted distances (Å) between represented residues.

CHAPTER VI

Phylogenetical, functional, and structural analysis of the CexA citrate exporter from *Aspergillus niger*

Note:

The work presented in this chapter is in preparation for publication.

CHAPTER VI

Phylogenetical, functional, and structural analysis of the CexA citrate exporter from *Aspergillus niger*

ABSTRACT

The recently identified citrate exporter from *Aspergillus niger* (AnCexA) is a member of the major facilitator superfamily (MFS) and the Drug-H^+ antiporter (DHA1) family (TC 2.A.1.2). Phylogenetic analysis revealed that the DHA1 family is mostly dispersed in eukaryotic genomes. The disruption of AnCexA in *A. niger* completely abolishes citric acid secretion, while its heterologous expression in *Saccharomyces cerevisiae* promoted the secretion of citric acid during growth on glucose. In this study, we overexpressed the AnCexA in a *Saccharomyces cerevisiae* to further characterize its activity and specificity. We found that *A. niger* transporter is highly specific for citric and isocitric acids, and besides its role as a citrate exporter, it is also able to import citrate with low affinity and high capacity at pH 5.5 (K_m=29.4 mM of citric acid and V_{max}=11.40 nmol citric acid s^{-1} mg^{-1} dry wt.). In addition, we explored the structure-function relationship of AnCexA using rational site-directed mutagenesis based on 3D structure prediction, molecular docking analysis, and the identification of conserved amino acid residues within the DHA1 family members. Mutants with altered activity and specificity were obtained and provide the first insights on the structure of the citrate exporter AnCexA, which can ultimately lead to the improvement of citrate bioproduction.

KEYWORDS: Plasma membrane transport; Citrate export; AnCexA transporter; DHA1 transporter family; *Aspergillus niger*; Heterologous expression; *In silico* studies

6.1 INTRODUCTION

Organic acids, such as carboxylic acids, are structurally diverse chemicals, commonly used in many industrial formulations from food, pharmaceutical, and cosmetic to polymer/plastics sectors [1]. Citric acid, a well-known carboxylic acid, is employed in the food and beverage industry due to its antioxidant properties, as a food preservative, or as an acidifier by boosting the flavors and aromas from fruit juices, soft drinks, and other beverages [2]. In 2020, the global citric acid market achieved a volume of 2.39 million tons. By 2026, the citric acid market is estimated to reach 2.91 million tons [3]. A great diversity of microbes from bacteria, fungi, to yeast have been reported to be involved in the production of citric acid [4-6]. *Aspergillus niger* is the oldest industrial workhorse, revealing outstanding properties that turn it ideal for industrial fermentation. This fungus is used for the industrial production of citric acid since 1923 [7, 8]. Its natural ability to secrete metabolites, such as organic acids and proteins, coupled with its robustness to extreme acid environments, are desired traits that turn *A. niger* into an important industrial bioplatform [9]. Microbes are metabolically engineered to produce specific metabolites and to use particular substrates, like industrial wastes and subproducts [10]. Rewiring the plasma membrane transport of such molecules in microbial cell factories is one of the key-processes to optimize cellular metabolism. Transporter engineering approaches have improved the efficiency of microbial cell factories for the production of desired metabolites [11, 12]. Some current limitations for the industrial exploitation of microbial cell factories include the ability to assimilate specific carbon and energy sources and product accumulation and subsequent toxicity with resulting low product extracellular titers [13, 14]. Low-product yield, reactor productivity, and final product concentration can be caused by severe end-product inhibition [15]. To surpass these bottlenecks, researchers outlined strategies that include the expression of importers and/or exporters, that can be improved by genetic manipulation for increased transport capacity or altered specificity [10, 12, 16]. Additionally, the employment of acid-tolerant microorganisms, like *Saccharomyces cerevisiae*, determines one of the crucial steps in effective bioprocess development [17, 18]. The coupled superior predisposition for growing at lower pH values and strengthened resistance mechanisms to weak carboxylic acid stress turns it into an ideal chassis to produce desired fine chemicals [19]. Although several carboxylate importers are known in fungi, few exporters are currently identified [10]. Two itaconate exporters were already identified, Itp1 in the fungal pathogen *Ustilago maydis* [20], and MfsA in *Aspergillus terreus* [21, 22]. The Mae1p from *Schizosaccharomyces pombe* is a proton/dicarboxylate transporter able to export fumarate, succinate, and malate [23, 24]. Despite being characterized as monocarboxylate importers, when expressed in an *S. cerevisiae* strain engineered for

lactate production, Jen1 and Ady2 were able to export lactic acid [25]. Recently two citrate exporters were identified, AnCexA from *A. niger* [11, 26] and Cex1 from *Yarrowia lipolytica* [27]. AnCexA belongs to the Drug-H^+ Antiporter (DHA1) family (TC 2.A.1.2), while Cex1 belongs to the Drug:H^+ Antiporter 2-Family (DHA2) (TC 2.A.1.3). The disruption of AnCexA impaired citric acid production in *A. niger* and its heterologous expression in *S. cerevisiae*, lead to the secretion of citric acid in glucose-grown cells [11, 26]. Previous studies proposed different molecular mechanisms for citrate export in *A. niger*: passive diffusion, citrate/H^+ antiporter, or ΔpH-driven symport with H^+ ions [4, 28, 29]. In the present study, we present the phylogenetic analysis of the citrate exporter AnCexA, its 3D structure prediction as well as its functional characterization through heterologous expression in *S. cerevisiae*.

6.2 Materials and Methods

6.2.1 Phylogenetic reconstruction of AnCexA

About 10000 proteomes were downloaded from the refseq subsection of the NCBI Assembly platform as individual FASTA files and converted into a local database. Only sequences derived from a single genome of a given species were considered to avoid redundancies. The specimen with the higher number of proteins described in the database was selected. A BLAST search, with a cut-off e-value of 10^{-10} and an associated query-cover value higher than 65%, was carried out on this database using the protein XP_025452994.1 from *Aspergillus niger*. Retrieved protein sequences were aligned using the MAFFT online server [30], which incorporates multiple alignment strategies. Sequences that were not aligning extensively across the conserved region of the alignment were further excluded from the phylogenetic analysis. These sequences, in many cases, could represent the lower quality of the stretch of the genome where they are located or incomplete annotation of the full gene and they should not necessarily be regarded as non-functional genes. Phylogenetic reconstruction was performed using Maximum Likelihood, more appropriate for the deeper divergences under analysis here, using MEGA7 [31] and the Jones-Taylor-Thornton substitution model. Bootstrap was performed for 1000 repetitions. Obtained phylogenetic tree was displayed and edited in FigTree v.1.4.4. (http://tree.bio.ed.ac.uk/).

6.2.2 Yeast cultivations, plasmid constructs, and growth conditions

The yeast strains, plasmids, and primers used in this study are listed in tables 1 and 2, and 3 respectively. In this study, two *S. cerevisiae* CEN.PK strains were used: the CEN.PK 113-5D [32] and the IMX1000 [33]. The *S. cerevisiae* CEN.PK 113-5D transformants were maintained on YP agar plates supplemented

with 2% (w/v) glucose and 0.3 mg/mL hygromycin B. The *S. cerevisiae* IMX1000 strains were maintained in minimal media with the required supplements for the growth of the auxotrophic strains [11, 33].

Yeast cells were grown in yeast nitrogen base (YNB, Difco), 0.67%, w/v (YNB medium), enriched with adequate requirements for prototrophic growth. Carbon sources were glucose (2%, w/v), acetic acid (0.5%, v/v; pH 6.0), lactic acid (0.5%, v/v; pH 5.0), pyruvic acid (0.5% w/v; pH 5.0), fumaric acid (1%, w/v, pH 5.0), succinic acid (1%, w/v, pH 5.0), malic acid (1%, w/v; pH 5.0) and citric acid (1%, w/v; pH 5.5). Yeast growth was carried out at 30°C, both in liquid and solid media. Cultures were harvested during the mid-exponential phase of growth. For growth tests, cells were grown on YNB Glu−Ura media, until the mid-exponential phase and adjusted to an OD_{640nm} of 0.1. A set of three 1:10 serial dilutions were performed and 3 μL of each suspension was inoculated in the desired media, using YNB Glu−Ura as a control carbon source. Cells were incubated at 30°C for 2 days and at 18°C for 14-22 days. At 18°C, carboxylic acid uptake by diffusion is drastically reduced so that growth on CA as sole carbon and energy source is directly dependent on a functional transporter, as described by Soares-Silva *et al.* (2007) [34].

6.2.3 TRANSPORT ASSAYS

Transport assays were performed as previously described by Ribas *et al.* (2017) [35] using labelled [1,5-^{14}C] citric acid (Perkin Elmer, Massachusetts, USA) with a specific activity of 300 and 1500 dpm/nmol. Optimal-fitting of the experimental data for initial uptake rate was done through computer-assisted non-linear regression analysis performed by GraphPad Prism (California, USA) version 6.0 for Windows. The kinetic parameters were obtained at a significance level $p<0.05$. The data presented is the mean value of at least three independent experiments, with three replicas each. Inhibition assays were carried out through the simultaneous addition of labelled and nonlabelled substrates. The determination of non-specific ^{14}C-adsorption to the cells, along with the diffusion component, was performed by adding a mixture of labelled acid and unlabeled acid concentrated 1000-fold and measured in a PerkinElmer Tri-Carb 4810TR liquid scintillation spectrophotometer, with dpm correction.

Table 1. Yeast strains used in this study

Strain	Genotype	Source
***S. cerevisiae* CEN.PK113-5D**	*MAT*a *MAL2-8C SUC2 ura3-52*	Entian and Kötter (2007)
IMX1000 (parental strain CEN.PK113-7D)	*MAT*a *ura3-52 trp1-289 leu2-3112 his3Δ can1Δ::cas9-natNT2 mch1Δ mch2Δ mch5Δ aqy1Δ itr1Δ pdr12Δ mch3Δ mch4Δ yil166cΔ hxt1Δ jen1Δ ato1Δ aqr1Δ thi73Δ fps1Δ aqy2Δ yll053cΔ ato2Δ ato3Δ aqy3Δ tpo2Δ yro2Δ azr1Δ yhl008cΔ tpo3Δ*	Mans *et al.* (2017)
***S. cerevisiae* CEN.PK 113-5D pϕ**	*S. cerevisiae* transformed with arscen plasmid, empty vector	Steiger *et al.* (2019)
***S. cerevisiae* CEN.PK 113-5D pTEF1**	*S. cerevisiae* transformed with p*TEF1-cexA-*t*CYC1* cassette	Steiger *et al.* (2019)
***S. cerevisiae* CEN.PK 113-5D pTPI1**	*S. cerevisiae* transformed with p*TPI1-cexA-*t*CYC1* cassette	Steiger *et al.* (2019)
IMX1000 pϕ	IMX1000 transformed with arscen plasmid, empty vector	This work
IMX1000 pTEF1	IMX1000 transformed with pTEF1-CexA	This work
IMX1000 pCexA-S71A	IMX1000 transformed with pCexA-S71A	This work
IMX1000 pCexA-S75A	IMX1000 transformed with pCexA-S75A	This work
IMX1000 pCexA-N76A	IMX1000 transformed with pCexA-N76A	This work
IMX1000 pCexA-P80A	IMX1000 transformed with pCexA-P80A	This work
IMX1000 pCexA-D84A	IMX1000 transformed with pCexA-D84A	This work
IMX1000 pCexA-G122A	IMX1000 transformed with pCexA-G122A	This work
IMX1000 pCexA-R123A	IMX1000 transformed with pCexA-R123A	This work
IMX1000 pCexA-R124A	IMX1000 transformed with pCexA-R124A	This work
IMX1000 pCexA-R154A	IMX1000 transformed with pCexA-R154A	This work
IMX1000 pCexA-F188A	IMX1000 transformed with pCexA-F188A	This work
IMX1000 pCexA-R192A	IMX1000 transformed with pCexA-R192A	This work
IMX1000 pCexA-Q196A	IMX1000 transformed with pCexA-Q196A	This work
IMX1000 pCexA-P200A	IMX1000 transformed with pCexA-P200A	This work
IMX1000 pCexA-T207A	IMX1000 transformed with pCexA-T207A	This work
IMX1000 pCexA-P235A	IMX1000 transformed with pCexA-P235A	This work
IMX1000 pCexA-E236A	IMX1000 transformed with pCexA-E236A	This work
IMX1000 pCexA-T237A	IMX1000 transformed with pCexA-T237A	This work
IMX1000 pCexA-Y307A	IMX1000 transformed with pCexA-Y307A	This work
IMX1000 pCexA-S311A	IMX1000 transformed with pCexA-S311A	This work
IMX1000 pCexA-S315A	IMX1000 transformed with pCexA-S315A	This work
IMX1000 pCexA-R461A	IMX1000 transformed with pCexA-R461A	This work

Table 2. Plasmids used in this study

Plasmid	Characteristics	Reference
pϕ	BB3_arscen_URA_Hygro; empty vector (low copy plasmid)	118077 in Addgene
pTPI1	pMST-1312; cassette p*TPI1-cexA*t*CYC1*, constitutive expression	118075 in Addgene
pTEF1	pMST-1313; cassette p*TEF1-cexA*t*CYC1*, constitutive expression	118076 in Addgene
pCexA-S71A	pTEF1 with the substitution S71A in CexA	This work
pCexA-S75A	pTEF1 with the substitution S75A in CexA	This work
pCexA-N76A	pTEF1 with the substitution N76A in CexA	This work
pCexA-P80A	pTEF1 with the substitution P80A in CexA	This work
pCexA-D84A	pTEF1 with the substitution D84A in CexA	This work
pCexA-G122A	pTEF1 with the substitution G122A in CexA	This work
pCexA-R123A	pTEF1 with the substitution R123A in CexA	This work
pCexA-R124A	pTEF1 with the substitution R154A in CexA	This work
pCexA-R154A	pTEF1 with the substitution R154A in CexA	This work
pCexA-F188A	pTEF1 with the substitution F188A in CexA	This work
pCexA-R192A	pTEF1 with the substitution R192A in CexA	This work
pCexA-Q196A	pTEF1 with the substitution Q196A in CexA	This work
pCexA-P200A	pTEF1 with the substitution P200A in CexA	This work
pCexA-T207A	pTEF1 with the substitution T207A in CexA	This work
pCexA-P235A	pTEF1 with the substitution P235A in CexA	This work
pCexA-E236A	pTEF1 with the substitution E236A in CexA	This work
pCexA-T237A	pTEF1 with the substitution T237A in CexA	This work
pCexA-Y307A	pTEF1 with the substitution Y307A in CexA	This work
pCexA-S311A	pTEF1 with the substitution S311A in CexA	This work
pCexA-S315A	pTEF1 with the substitution S315A in CexA	This work
pCexA-R461A	pTEF1 with the substitution R461A in CexA	This work

6.2.4 AnCexA structural molecular docking studies

The tridimensional modelling was performed for the AnCexA (XP_001398400.1). For 3D structure prediction, the AnCexA amino acid sequence was threaded through the PDB library using HHPred [36] and LOMETS (Local Meta-Threading-Server) [37]. The top-ranked threading template was the L-lactate transporter from *Syntrophobacter fumaroxidans* (PDB 6G9X; 2.54 Å resolution) [38]. Molecular docking simulations were done as previously described [35], using the ligand structure of citric acid downloaded from the Zinc database [39]. Docking prediction was carried out with deprotonated forms of citric acid with the protonation states adjusted to match the desired pH. The results were exported in the .mol2 format. Substrate 3D structures were built by inputting canonical SMILES strings in the UCSF Chimera [40], being minimized before molecular docking simulations in PyRx software [41] using AutoDock Vina. All simulated interactions were analyzed in 2D and 3D pose views using both Chimera and Maestro v11.2. To validate any structural differences that might indicate different substrate preferences for the wild-type AnCexA protein and mutant alleles, we used the HOLE program (2.2.005 Linux) to predict the pore radius [42]. The image for the pore was obtained using the Visual Molecular Dynamics program (VMD, 1.9.3) [43]. The pore radius of the wild-type and the mutants of AnCexA proteins were compared in a graph with the coordinate in the direction of the channel vector serving as the X-axis.

6.2.5 Construction of AnCexAp mutations, cloning, and expression in *S. cerevisiae*

Oligonucleotide-directed mutagenesis was performed as previously described [34] using the pTEF1-CexA plasmid (60 ng) [11]. Twenty-one mutations of the AnCexA were obtained. The oligonucleotides used are listed in table 3. Mutations were confirmed by sequencing. The mutant versions of AnCexA were transformed in a *S. cerevisiae* IMX1000 strain [33] using the High Efficiency Transformation Method [44] and transformants were selected by complementation of uracil auxotrophy.

Table 3. Oligonucleotides used for strain construction, cloning, and expression

Name	Sequence
S71A_fwd	CTCGCTGCCATATTTGCTCCGCTTTCGTCGAAC
S71A_rev	GTTCGACGAAAGCGGAGCAAATATGGCAGCGAG
S75A_fwd	CATATTTTCTCCGCTTTCGGCTAACATTTACTTCCCTG
S75A_rev	CAGGGAAGTAAATGTTAGCCGAAAGCGGAGAAAATATG
N76A_fwd	TCTCCGCTTTCGTCGGCTATTTACTTCCCTGCC
N76A_rev	GGCAGGGAAGTAAATAGCCGACGAAAGCGGAGA
P80A_fwd	GTCGAACATTTACTTCGCTGCCCTGGATGATG
P80A_rev	CATCATCCAGGGCAGCGAAGTAAATGTTCGAC
D84A_fwd	CTTCCCTGCCCTGGATGCTGTCTCGAAATCCCTC
D84A_rev	GAGGGATTTCGAGACAGCATCCAGGGCAGGGAAG
G122A_fwd	CATGTCAGACGCCACAGCTAGACGGCCTGTCTTTATTG
G122A_rev	CAATAAAGACAGGCCGTCTAGCTGTGGCGTCTGACATG
R123A_fwd	GTCAGACGCCACAGGTGCTCGGCCTGTCTTTATTG
R123A_rev	CAATAAAGACAGGCCGAGCACCTGTGGCGTCTGAC
R124A_fwd	GACGCCACAGGTAGAGCTCCTGTCTTTATTGGAAC
R124A_rev	GTTCCAATAAAGACAGGAGCTCTACCTGTGGCGTC
R154A_fwd	GAGCTCATGGCCTTCGCTGCCTTGCAGGCTGCTG
R154A_rev	CAGCAGCCTGCAAGGCAGCGAAGGCCATGAGCTC
F188A_fwd	GTAGCTTGGTGGGTATCGCTGGTGGAGTTCGCATG
F188A_rev	CATGCGAACTCCACCAGCGATACCCACCAAGCTAC
R192A_fwd	GTATCTTCGGTGGAGTTGCTATGCTTGGACAGGGAATC
R192A_rev	GATTCCCTGTCCAAGCATAGCAACTCCACCGAAGATAC
Q196A_fwd	GTTCGCATGCTTGGAGCTGGAATCGGGCCGGTTTTC
Q196A_rev	GAAAACCGGCCCGATTCCAGCTCCAAGCATGCGAAC
P200A_fwd	GGACAGGGAATCGGGGCTGTTTTCGGCGGCATTTTC
P200A_rev	GAAAATGCCGCCGAAAACAGCCCCGATTCCCTGTCC
T207A_fwd	GTTTTCGGCGGCATTTTCGCTCAGTATCTCGGATATC
T207A_rev	GATATCCGAGATACTGAGCGAAAATGCCGCCGAAAAC
P235A_fwd	CATTCTGGTGCTTCTTGCTGAGACATTGAGGCCAATTG
P235A_rev	CAATTGGCCTCAATGTCTCAGCAAGAAGCACCAGAATG
E236A_fwd	CTGGTGCTTCTTCCGGCTACATTGAGGCCAATTG
E236A_rev	CAATTGGCCTCAATGTAGCCGGAAGAAGCACCAG
T237A_fwd	GTGCTTCTTCCGGAGGCTTTGAGGCCAATTGCTG
T237A_rev	CAGCAATTGGCCTCAAAGCCTCCGGAAGAAGCAC
Y307A_fwd	CTTTGGAAGTATCGTGGCTACAGTGTGGAGCATG
Y307A_rev	CATGCTCCACACTGTAGCCACGATACTTCCAAAG
S311A_fwd	GTGTACACAGTGTGGGCTATGGTGACATCCAGTAC
S311A_rev	GTACTGGATGTCACCATAGCCCACACTGTGTACAC
S315A_fwd	GTGGAGCATGGTGACAGCTAGTACCACCGACCTCTTC
S315A_rev	GAAGAGGTCGGTGGTACTAGCTGTCACCATGCTCCAC
R461A_fwd	GTGAACAATCTGATGGCTTGCCTGCTTGGAGCTG
R461A_rev	CAGCTCCAAGCAGGCAAGCCATCAGATTGTTCAC
Primer seq_F	GGTCATGTCCTCAACCACGTCCTCA
Primer seq_R	AAGCCTAGTTGCCGTTGGCTTTGGCCT

6.3 RESULTS

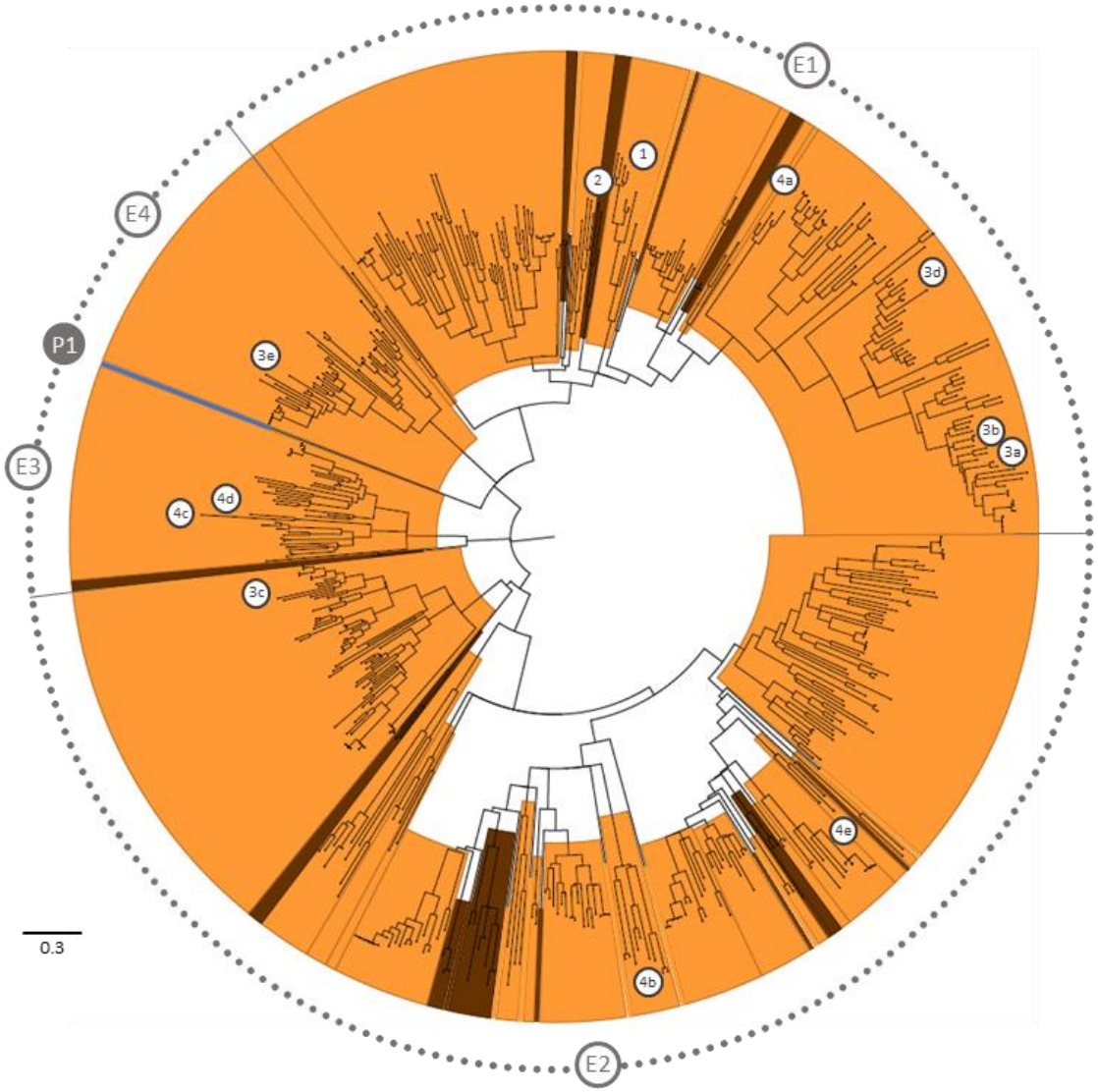
6.3.1 EVOLUTIONARY RELATIONSHIPS OF ANCEXA

A phylogenetic analysis was carried out to study the evolution of the members from the Drug-H^+ antiporter (DHA1) family (TC 2.A.1.2) present in eukaryotic and prokaryotic genomes. The homolog nominations and bootstrap values from the phylogenetic reconstruction are all available in the Supplementary material section (Figure S1). The BLASTP search in the NCBI's assembly database using complete genomes retrieved 581 hits. A total of twenty-one sequences that lacked large conserved regions were excluded resulting in a final dataset of 560 sequences.

Sixteen protein homologs to AnCexA were found in the *A. niger* CBS 513.88 strain in this database (Figure 1). For a further study of the present tree, we have highlighted sections according to the distribution of homologs among eukaryotic and prokaryotic organisms. The prokaryotic clade is formed by the presence of a single homolog from *Bacillus anthracis*. The yeast and filamentous fungi homologs present can be divided into four main eukaryotic clades (E1-E4, in figure1).

Eukaryotic divisions were defined by the presence of previously characterized transporters. The E1 clade includes the CexA protein from *A. niger* [11], the Itp1 protein, an itaconate transporter characterized in *Ustilago maydis* [20], three homologs from *S. cerevisiae*, the multidrug transporters Qdr1p [45], Qdr2p [46] and Aqr1p [47, 48], and one homolog from *Y. lipolytica,* YALI0F03751p [27], and other homologs present in ascomycetes and basidiomycetes. Aqr1p confers resistance to short-chain monocarboxylic acids and quinidine [47, 48], and was implicated in the excretion of excess amino acids in *S. cerevisiae* cells. Qdr1p is involved in the spore wall assembly and provides resistance to drugs like quinidine, barban, ketoconazole, fluconazole [45]. Besides its involvement in drug resistance, Qdr2p is also recognized as an importer of potassium ions and an exporter of copper [49].

The E2 clade is split into two subclades. On the E2-major subsection are the other three homologs detected in *U. maydis*, the remaining homologs from *Y. lipolytica*, including YALI0E21241g and YALI0F25597g characterized as not being involved in citrate secretion [27], some homologs from other basidiomycetes such as *Cryptococcus*, *Sporisorium*, *Phellinus* and *Malassezia* species, the other homologs from *A. niger*, and a large section of homologs from other yeast and fungi. In another subclade is the multidrug transporter Qdr3p from *S. cerevisiae* [50], and other homologs present in ascomycetes and basidiomycetes from *Phellinus* and *Cryptococcus* species.



DHA1 Homologs

1. *Aspergillus niger* AnCexA ; (16 other homologs)
2. *Ustilago maydis* UmItp1 ; (3 other homologs)
3. *Saccharomyces cerevisiae* 3a-Qdr1p; 3b-Qdr2p; 3c-Qdr3p; 3d-Aqr1p; 3e-Dtr1p
4. *Yarrowia lipolytica* 4a-YALI0F03751g; 4b-YALI0F25597g; 4c-YALI0C00847g; 4d-YALI0E10483g; 4e-YALI0E21241g ; (13 other homologs)

Figure 1. Maximum likelihood phylogenetic tree of DHA1 family (TCDB 2.A.1.2) members present in eukaryotic and prokaryotic genomes. Branch lengths are proportional to sequence divergence. Groups indicated as E1, E2, E3, E4, P1 were created to facilitate tree description in the main text and are not meant to provide any type of classification. Major taxonomic groups are indicated in shades of blue – bacteria, orange – ascomycetes, and brown – basidiomycetes. Homologs relevant to the discussion of the manuscript are here highlighted.

Close to this latter subclade is found another major E3 clade that includes two of the homologs from *Y. lipolytica*, YALI0C00847g, and YALI0E10483g, also previously characterized as not being involved in citrate production [27] and other homologs present in several ascomycetes and *Cryptococcus* species. E4 clade contains the putative dityrosine transporter Dtr1p from *S. cerevisiae* [51] and several homologs present in other yeasts species. The AnCexAp characterized in this study is phylogenetically closer to other homologs from *Fusarium* species (*F. fujikuroi* and *F. oxysporum*), *Colletotrichum higginsianum*, *Neurospora crassa*, *Sclerotinia sclerotiorum*, *Botrytis cinerea*, *Penicillium chrysogenum,* and one homolog from *Aspergillus oryzae*. A strong presence of AnCexA homologs in several ascomycetes was evidenced. The few basidiomycetes members are dispersed across the phylogenetic tree.

6.3.2 Functional studies of AnCexA in *Saccharomyces cerevisiae*

The *S. cerevisiae* CEN.PK 113-5D strain was used for heterologous expression of the AnCexA under the control of *TPI1* and *TEF1* constitutive promoters. Cells were grown in media containing mono-, di- and tricarboxylic acids as sole carbon and energy at 18 and 30°C. As observed in Figure 2, all the *S. cerevisiae* CEN.PK 113-5D transformed strains presented growth on all carbon sources, both at 18°C and 30°C, with increased growth in the last temperature. The negative control grew better than the cells expressing the AnCexA protein. Most probably, cells expressing AnCexA are exporting citrate, as previously reported for *S. cerevisiae* cells [11] which can be essential to sustain long-term cellular growth on solid media.

The uptake rate of radiolabelled ^{14}C-citric acid 1.0 mM revealed a higher citrate uptake of the *S. cerevisiae* CEN.PK 113-5D strain harboring pAnCexA-TEF1 compared to cells transformed with pAnCexA-TPI1 (data not shown). Therefore, the effect of the pH on the initial uptake rates of citrate was determined in glucose-grown cells only considering the *S. cerevisiae* CEN.PK 113-5D pAnCexA-TEF1 strain (Figure 3A). The highest transport of citrate occurred at pH 7.0. To assess the transporter activity at different growth stages (see Supplementary material section, Figure S2), citrate uptake (10.0 mM, 30°C, pH 7.0) was evaluated in four time-points during the exponential growth phase (Figure 3B) revealing that citrate uptake capacity increased during the progression of the growth on glucose.

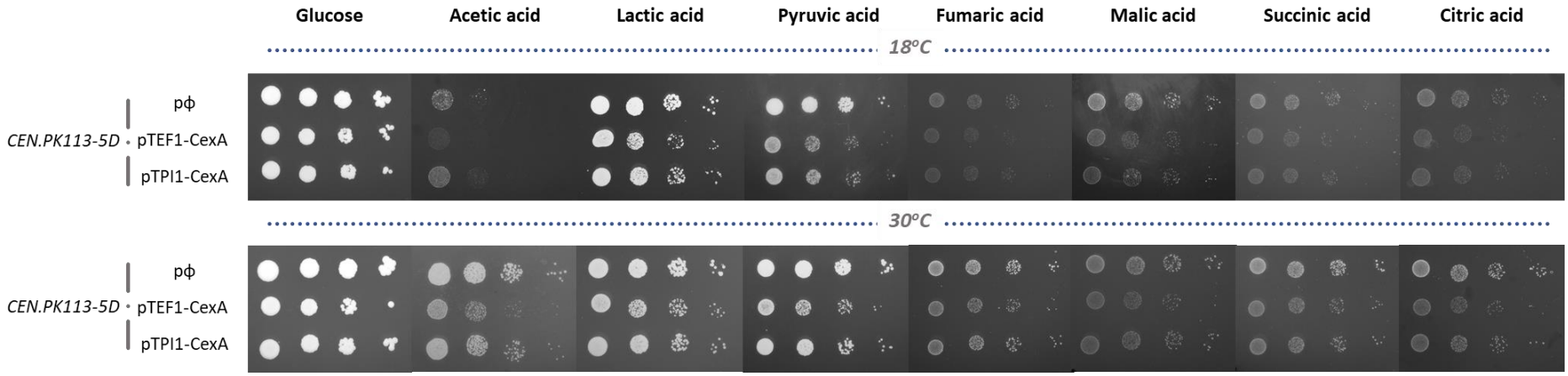
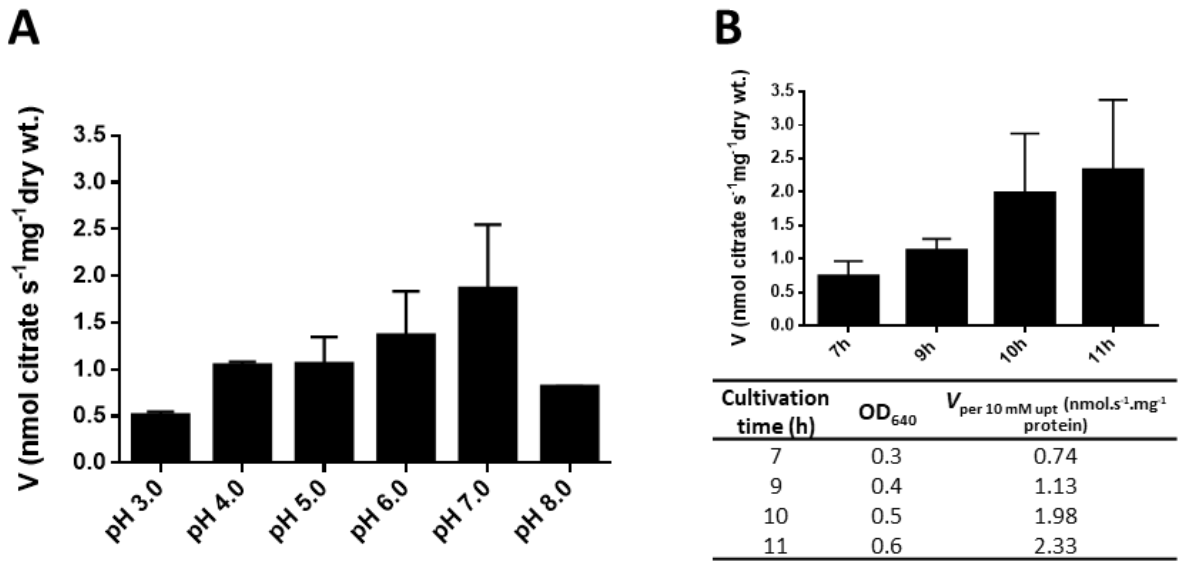


Figure 2. Growth phenotypes of *S. cerevisiae* CEN.PK 113-5D strains expressing *Aspergillus niger* CexA transporter under the control of TEF1 and TPI1 promoters. Different media containing glucose (2% w/v), acetic acid (0.5% v/v; pH 6.0), lactic acid (0.5% v/v; pH 5.0), pyruvic acid (0.5% w/v; pH 5.0), fumaric acid (1% w/v; pH 5.0), malic acid (1% w/v; pH 5.0), succinic acid (1% w/v; pH 5.0) and citric acid (1% w/v; pH 5.5) were tested as sole carbon and energy sources. The negative control corresponds to the strain carrying the corresponding empty vector (pϕ). Cells suspension were 1/10 serially diluted; 3 µL drops of each dilution were spotted in plates and incubated at 18°C for 22 days and 30°C for 2 days.



Cultivation time (h)	OD_{640}	$V_{per\ 10\ mM\ upt}$ (nmol.s^{-1}.mg^{-1} protein)
7	0.3	0.74
9	0.4	1.13
10	0.5	1.98
11	0.6	2.33

Figure 3. Uptake of ^{14}C-citric acid 10.0 mM in glucose-grown cells of *S. cerevisiae* CEN.PK 113-5D transformed with pAnCexA-TEF1. A) Effect of the pH on the uptake of the acid. B) Time-course of the acid uptake during the exponential growth phase. The data shown are mean values of at least three independent experiments and the error bars represent the standard deviation.

The kinetic parameters for the initial uptake rates of radiolabelled citric acid at pH 5.5 and pH 7.0, 30°C, were determined in glucose-grown cells of *S. cerevisiae* CEN.PK 113-5D transformed with AnCexA, collected at the O.D. of 0.5 (Figure 4). At pH 7.0, a linear kinetic was obtained for citric acid uptake (Figure 4A), as for pH 5.5, citrate uptake presented a Michaelis–Menten kinetics with the following K_m 29.4 ± 13.7 mM of citric acid and V_{max} 11.40 ± 3.66 nmol s^{-1} mg^{-1} dry wt. (Figure 4B). These latter findings suggest the involvement of AnCexA as low-affinity transporter for citrate in *S. cerevisiae* and evidence its role as a citrate importer. The specificity and energetics of AnCexA were further assessed at pH 5.5 (Figure 5). None of the mono- and dicarboxylic acids tested had an inhibition effect on citrate uptake (Figure 5A) however it was inhibited by isocitric acid (Figure 5B) suggesting that only tricarboxylates are substrates of AnCexA. Additionally, the protonophore CCCP (carbonyl cyanide m-chlorophenylhydrazone), a recognized metabolic inhibitor that collapses the proton motive force, did not affect citrate transport, revealing that the uptake of citrate is not dependent on the proton motive force (Figure 5A).

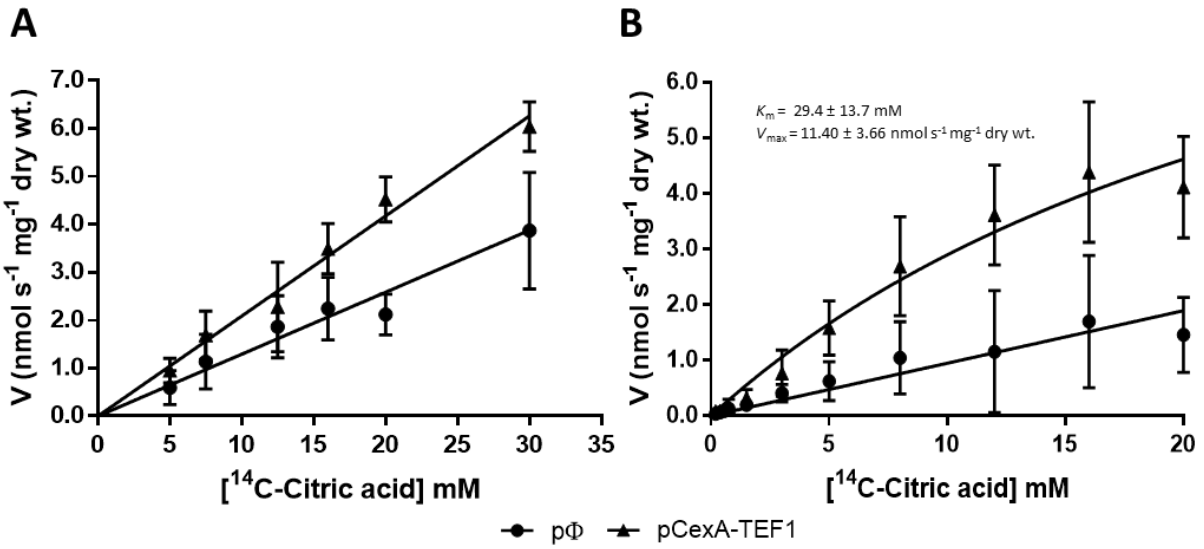


Figure 4. Citric acid transport in *S. cerevisiae* CEN.PK 113-5D cells expressing AnCexA under the control of TEF1. Initial uptake rates of radiolabelled ^{14}C-citric acid as a function of the acid concentration at A) pH 7.0 and B) pH 5.5, at 30°C. Cells were grown on YNB-glucose as the sole carbon and energy source and collected in the mid-exponential growth phase (OD=0.5). The data shown are mean values of at least three independent experiments. Error bars correspond to the standard deviations.

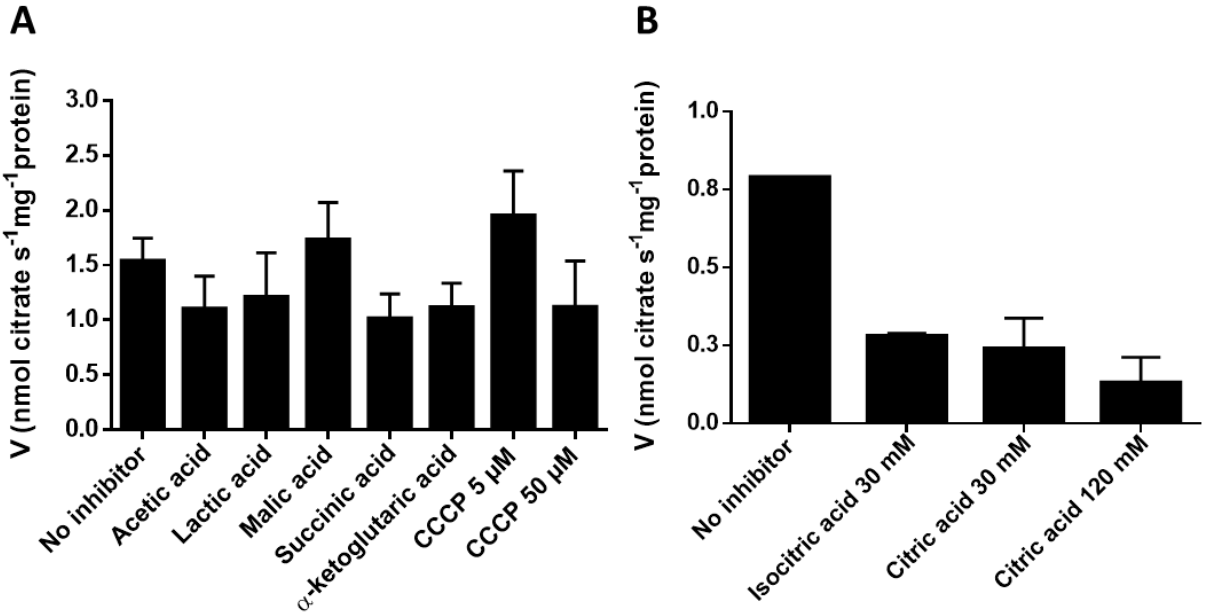


Figure 5. Specificity and energetics of citric acid uptake at pH 5.5, 30°C in *S. cerevisiae* CEN.PK 113-5D cells expressing AnCexA under the control of TEF1. A) ^{14}C-citric acid uptake (12 mM) in the presence of non-labelled acetic, lactic, malic, succinic, α-ketoglutaric acids (120 mM), and CCCP (5 and 50 μM). B) ^{14}C-citric acid (3 mM) uptake in the absence and presence of non-labelled citric (30 mM and 120 mM) and isocitric (30 mM) acids. The data shown are mean values of at least two independent experiments and the error bars represent the SD.

6.3.3 ANCEXA STRUCTURAL CHARACTERIZATION

The 3D-model structure of AnCexA (Figure 6) was predicted based on the crystal structure of the MFS transporter from *Syntrophobacter fumaroxidans*. This transporter presents twelve transmembrane segments with 14% identity with AnCexA and BLASTP analysis presented an E-value of $3.7e^{-35}$. The AnCexA structure contains four predicted binding sites for citrate (1-4, Figure 6A), one located in the extracellular side of the membrane and the other three closer to the central pore of the protein. The putative interactions of citrate with amino acid residues at each site were also predicted by this model (Figure 6B). Tables S1 and S2 describe the amino acid residues with predicted strong intramolecular interactions with deprotonated forms of citric acid and the estimated binding affinity values [kcal/mol] by PyRx software for the docking of AnCexAp with the citrate form adjusted for pH 5.5. In the first binding site, citrate interacts with T207, I215, and T318 residues establishing hydrogen bonds. In the second binding site, citrate also interacts through hydrogen bonds with S75, N76, Q196, Y307, W310, and T314 residues. In the third binding site, C343 is found to interact with citrate through a hydrogen bond. The residue R192 presents a shared salt bridge interaction, that is relevant for the citrate binding located at the third and fourth binding sites. Finally, on the fourth binding site, besides R192, citrate also interacts with S71 through a hydrogen bond. Based on this analysis we have selected these residues as targets for mutagenesis of AnCexA (Figure 7). Further amino acid residues were selected based on the residue conservation of the AnCexA alignment (Figure 7A). Site-directed mutagenesis in CexA was performed by replacing each residue with an alanine residue (Figure 7B). The mutations were tested on media containing mono-, di- and tricarboxylic acids as sole carbon and energy sources at 18°C and 30°C (Figure 8). The *S. cerevisiae* IMX1000 strain was used as a host background for the heterologous expression of *AnCexA* alleles, under the control of the *TEF1* promoter. The growth phenotypes at 30°C revealed three distinct types of mutations (Figure 8A).

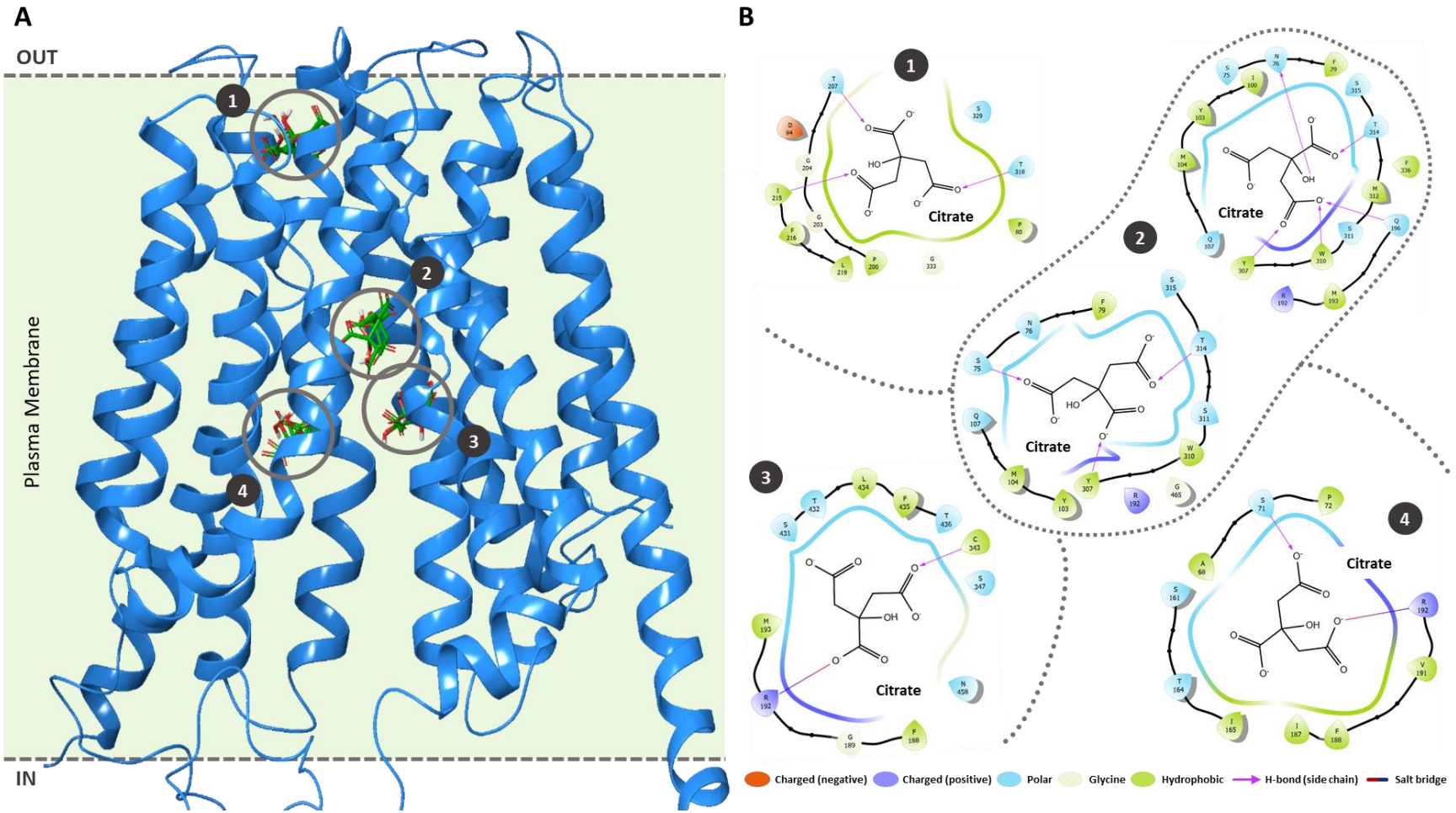


Figure 6. Structure-function analysis of the *Aspergillus niger* CexAp. A) Molecular docking of citrate with the 3D-model of AnCexA, based on the crystal structure of the MFS transporter from *S. fumaroxidans* (PDB 6G9X). Four binding sites (1-4) were found for citrate in the *A. niger* CexA. B) The 2D view represents the interactions between citrate and AnCexA residues.

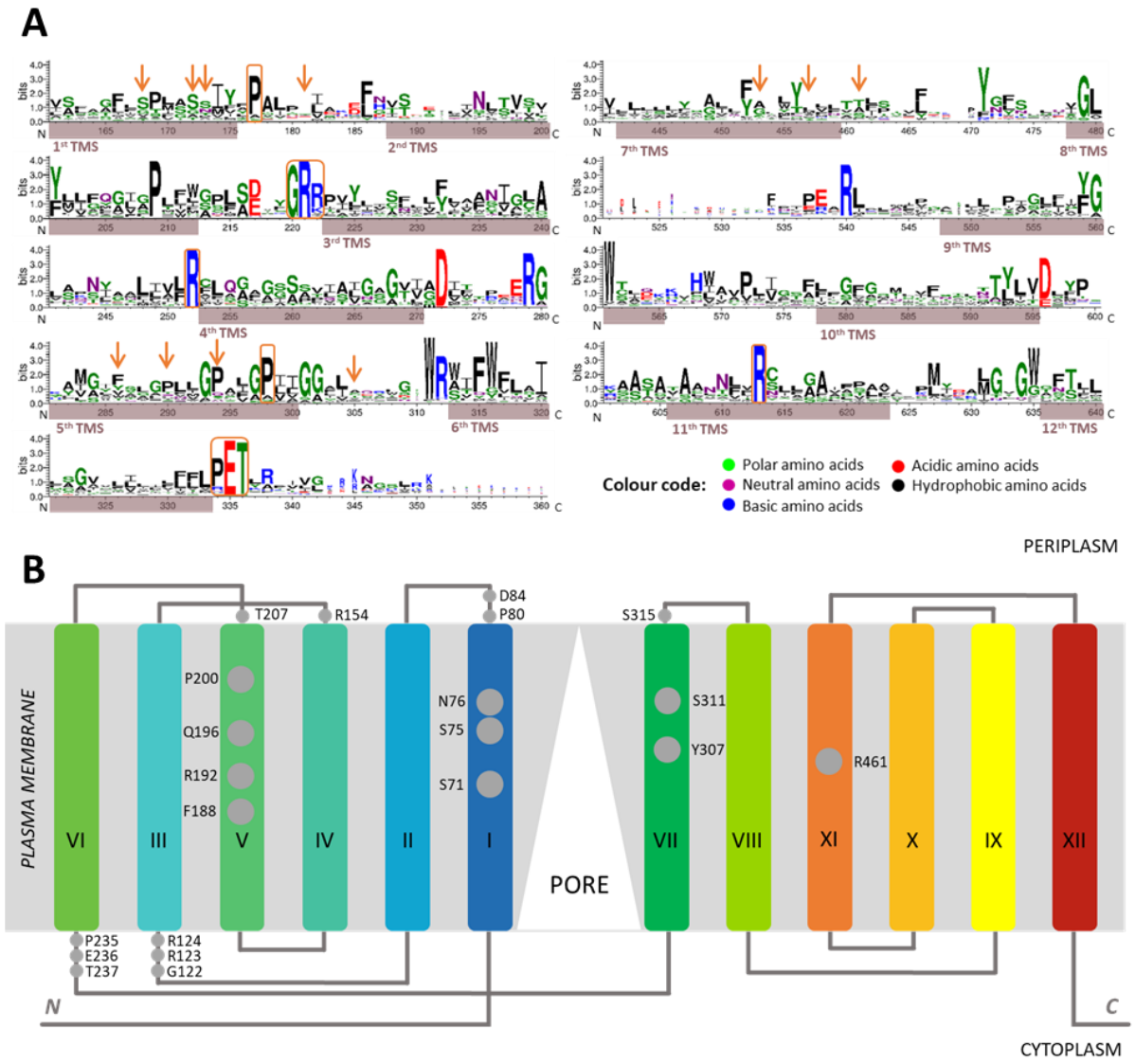


Figure 7. Structure-function study of AnCexA. A) The conservation logo resulted from the alignment of AnCexA homologs visualized with WebLogo (https://weblogo.berkeley.edu). The regions presented correspond to residues with more than 90% of aligned sequences. Transmembrane segments (TMSs) predicted with PSI/TM-Coffee (http://tcoffee.crg.cat/apps/tcoffee/do:tmcoffee) and TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) are highlighted in brown in the logo, and orange arrows and boxes (highly conserved) indicate amino acids mutated in this study. B) Predicted topology of CexA. Transmembrane segments are colored with the rainbow scheme from I to XII TMS. The residues mutated for alanine are shown as grey circles: I TMS (S71, S75, N76, P80, D84), beginning of III TMS (G122, R123, R124), IV TMS (R154), V TMS (F188, R192, Q196, P200, T207), VI TMS (P235, E236, T237), VII TMS (Y307, S311, S315) and XI TMS (R461). The hydrophilic pore is depicted by the white triangle between the I and VII TMSs, pointing to the transport of carboxylate molecules from periplasm to cytoplasm. The N- and C-terminal of the exporter are annotated by N and C, respectively.

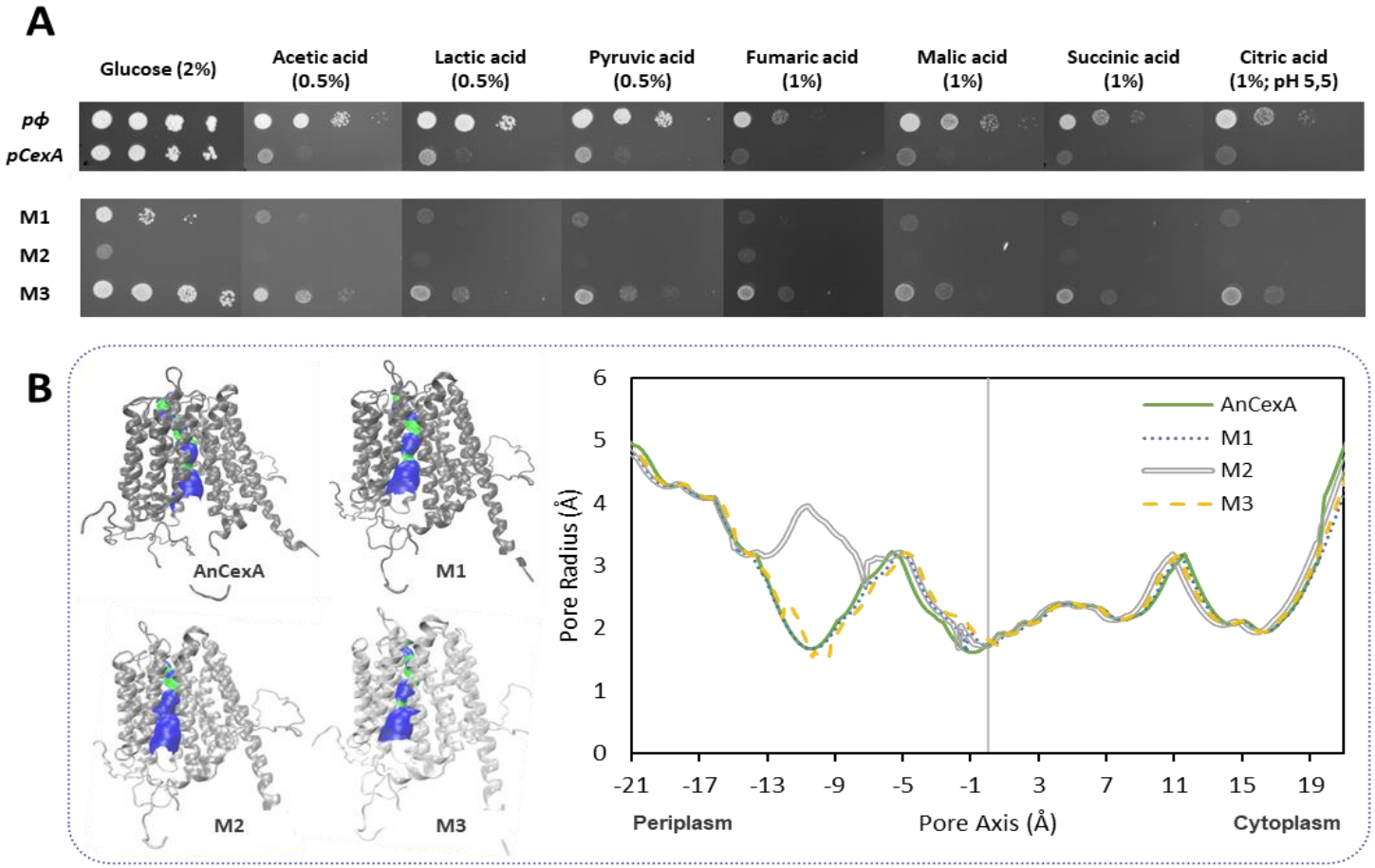


Figure 8. Structural-functional analysis of AnCexA. A) Growth phenotypes of *S. cerevisiae* IMX1000 strains expressing AnCexA transporter, transformed with the empty vector and respective AnCexA mutant versions M1, M2 and M3, at 30°C grown on glucose (2% w/v), acetic acid (0.5% v/v; pH 6.0), lactic acid (0.5% v/v; pH 5.0), pyruvic acid (0.5% w/v; pH 5.0), fumaric acid (1% w/v ; pH 5.0), malic acid (1% w/v; pH 5.0), succinic acid (1% w/v; pH 5.0) and citric acid (1% w/v; pH 5.5) as sole carbon and energy sources. The negative control corresponds to the strain carrying the corresponding empty vector (pϕ). Cells were serially diluted; 3 µL drops of each dilution were spotted in plates and incubated for 4 days. B) Cartoon of the simulations using HOLE software for pore shape prediction of AnCexA and mutant proteins M1, M2, and M3: color scheme is ruled by blue (larger aperture), green, and red (more constricted). The plot represents the simulations for the pore radius profiles along the AnCexA (green line), M1 (blue-dotted line), M2 (grey line), and M3 (yellow dashed-line) proteins.

Mutant M1 presented mild growth in all carbon sources, exhibiting a growth profile similar to cells expressing AnCexA. This indicated that the mutation did not cause significant alterations at the level of the protein transport capacity and specificity. In the M2 mutant, a toxic phenotype was detected which might be coupled to a gain of function of the transporter, potentially favoring the export of citrate and thus decreasing the ability of cells to grow on organic acids. For the mutant M3 improved growth was found on all media with carboxylic acid used as the sole carbon and energy source. This latter mutation might be critical for citrate transport resulting in a loss of function since a similar phenotype was displayed by the negative control.

The prediction of the pore shape for the AnCexA protein and the mutants M1, M2, and M3 was explored with the Hole software (Figure 8B). An increase of about 2.0 Å on the pore radius was observed in the initial section in the M2 mutant when compared to the wild-type AnCexA. As for M1 and M3 mutants' minor differences were observed.

6.4 Discussion

The phylogenetic analysis CexA homologs belonging to the DHA1 transporter family revealed a high prevalence in fungi. The single member of the prokaryotic clade is likely to represent a case of horizontal gene transfer. On the AnCexA clade, four homologs are functionally characterized: the itaconate transporter Itp1p from *U. maydis* [20], and three multidrug transporters from *S. cerevisiae* (Qdr1, Qdr2, Aqr1) [45-48]. From these, Aqr1p is described as a transporter of carboxylic acids, conferring resistance to monocarboxylic acids and quinidine, being also associated with the excretion of amino acids [47, 48]. Qdr1 is involved in the resistance to drugs e.g. quinidine, ketoconazole, and fluconazole, and the in spore wall assembly [45]. Qdr2p presents a wide substrate specificity, including the transport of several drugs and mono- and divalent cations, and the export of copper, also contributing to potassium homeostasis in cells [46]. *S. cerevisiae* has twelve proteins belonging to the DHA1 family, however, the remaining proteins were not included in the CexA-phylogenetic tree, e.g. Flr1, Hol1, Tpo1, Tpo2, Tpo3, Tpo4, and Yhk8 [52, 53].

In *Y. lipolytica,* the five AnCexA homologs are not involved in citrate export (YALI0F03751g, YALI0C00847g, YALI0E10483g, YALI0F25597g, and YALI0E21241g) and recently uncovered Cex1 citrate transporter belongs to another transporter family [27]. The function of this group of genes still needs to be characterized.

Among the AnCexA homologs present in the tree, only AnCexA has been identified as a citrate exporter. The number of AnCexA homologs is also strain dependent. In the genome of the *A. niger* strain ATCC 1015, only six putative homologs (>30% identity) were identified when compared to the seventeen homologs found in the *A. niger* strain CBS 513.88 reference genome (Figure 1). This is in accordance with a high variation reported for this species verified by exo-metabolite profiling and phylogenetic studies [54]. In fact, the genome sequencing of the *A. niger* CBS 513.88 strain revealed an extensive number of MFS transporters, such as the AnCexA homologs [55]. Nonetheless, the functionally characterized AnCexA and the AnCexA from CBS 513.88 vary only in one residue I21T that is localized in the N-terminal. In this work, the heterologous expression of AnCexA in *S. cerevisiae* resulted in a decreased growth on all carbon sources tested. The growth phenotype obtained in YNB-citric acid (1%) as sole carbon and energy source is in accordance with previous results [11], where a decreased growth was observed when the citrate exporter is expressed, under the control of a constitutive system, in *S. cerevisiae* cells. The observed growth phenotypes in carboxylic acids may be associated with the export of citrate, or even of other carboxylic acids, although the transporter seems to be specific for citrate and isocitrate. In this scenario, we hypothesize that the intracellular depletion of citrate and isocitrate affects cell growth performance on carboxylic acids. In the study by Steiger *et al.* (2019), the constitutive expression of CexA in *A. niger* had an impact on cellular traits, namely in the reduction of conidiation pattern verified upon growth in solid media plates [11]. In *S. cerevisiae* it is possible that the expression of AnCexAp also interferes with other cellular processes.

The CexA citrate uptake activity was optimal at pH values between 5.0 to 7.0. At a pH below the pK_a of the acid, the undissociated form predominates, being able to enter the cell by simple diffusion [56, 57]. For a pH value above the pK_a of the acid, the charged anionic form prevails, needing a transporter to cross the biological membrane [56]. In our findings, the increase in citrate uptake occurred at higher pH suggesting that a charged citrate is transported by AnCexA (pK_{a1}=3.13; pK_{a2}=4.76; pK_{a3}=6.39; 6.40) [4]. The reduction in the citrate uptake observed at pH 8.0, is in agreement with previous reports, as this pH value is considered critical and non-physiological for *S. cerevisiae* cells [58]. The optimal pH for citrate transport was pH 7.0. At this pH, citric acid is in the fully deprotonated state (pH > pK_{a3}) and simple diffusion is limited by its cell membrane permeability.

The kinetic parameters for citrate uptake obtained at pH 5.5 correspond to a low-affinity and high capacity transporter with the following kinetic parameters K_m 29.4 ± 13.7 mM of citric acid and V_{max} 11.40 ± 3.66 nmol s^{-1} mg^{-1} dry wt. These values are in accordance with the primary role reported for AnCexA, citrate export. In order, to be an efficient exporter, the transporter has to present a low affinity

for substrate import, thus preventing a futile cycle. Prior studies have associated low-affinity transport systems for carboxylic acids as protein exporters. The low-affinity plasma-membrane monocarboxylate transporter MCT4 (TC 2.A.1.13.6) is reported as being an exporter with a major physiological role in monocarboxylate efflux [59-61].

Inhibition assays suggest that citrate and isocitrate are transported by AnCexAp in *S. cerevisiae.* The AnCexAp transporter was the second plasma membrane citrate transporter identified in fungi [11, 26]. The first was the AnBest1, an anion efflux channel from *Aspergillus nidulans* [62] and more recently the citrate exporter YlCex1 from *Y. lipolytica* [27] and the citrate importer PkJEN2-2 from *P. kudriavzevii* [63] were also identified. However, the kinetic parameters for citrate transport of these three transporters were not determined.

The effect of the protonophore CCCP revealed the non-dependence of this transporter on a proton motive force, although one of the mechanisms hypothesized for citrate export was a citrate/H^+ antiporter [11]. Thus, further assays are needed to clarify the driving force of the citrate transport by AnCexAp, such as the utilization of ionophores valinomycin or monensin, to verify if citrate transport is influenced by the transmembrane electrochemical potential.

The structure-function studies of the AnCexA transporter were based on a 3D model and docking studies. The AnCexAp 3D-model structure was predicted based on the recently reported crystal structure of a MFS transporter belonging to a Solute Carrier (SLC) member 16 family from *Syntrophobacter fumaroxidans* [38]. In this latter family, several family members are carboxylic acid transporters, e.g. the MCT1-4 that encode proton:linked monocarboxylate transporters with affinity for L-lactate and pyruvate substrates [64, 65]. Molecular docking studies of AnCexA suggest four putative binding sites for citrate, one located in the extracellular side of the membrane and the other three closer to the central pore of the protein. Transmembrane segments (TMSs) I, V and VII contain the majority of the residues involved in strong intramolecular interactions with the deprotonated citric acid. Interestingly, the residue R192 seems to interact with citrate through a salt bridge. The nature of this interaction is likely to contribute to the conformational stability of the protein structure and specificity [66, 67]. With the present mutations, we observed three distinct growth patterns (Figure 8A): (i) cell growth similar to native AnCexA; (ii) growth similar to the negative control and, finally (iii) a lower growth than the native AnCexA. The projection of the pore radius in the mutant M2 revealed a considerable increase of the pore radius near the central zone. This suggests a larger aperture of the pore size that could result in an improvement of transport capacity as these mutants present a decreased growth when compared to wild-type AnCexA. Further studies are needed to confirm this hypothesis, namely the evaluation of the transport capacity and citrate

production in these mutants. The site-directed mutagenesis approach enabled the identification of critical residues involved in citrate transport.

6.5 CONCLUSIONS

Tricarboxylic acids, such as citric and isocitric acids, are a group of chemical building-blocks with a high demand in the global market driven by their applicability in many sectors from food to beverage fields, pharmaceutical to personal care products to the industry of cleaning products (detergents and cleaners) [3]. The microbial production of citric acid via fermentation has been strongly related to filamentous fungi, particularly to *A. niger* [68]. Membrane transporter proteins are key players for the optimization of final product export, directly influencing the productivity of cell factories. In this work, we have characterized the newly identified *A. niger* CexA transporter by heterologous expression in *S. cerevisiae*. Our results indicate that, besides its role as an exporter, AnCexA can work as an importer of citric acid, although with low affinity.

Further studies are needed to uncover the nature of the molecular transport mechanism involved in citrate uptake/efflux by AnCexA. Site-directed mutagenesis strategies played an important role in the identification of crucial residues for protein function. The combination of 3D-model refinement, ligand docking, and identification of conserved domains by multiple sequence alignment, has become a powerful tool for the prediction of key functional residues involved in substrate specificity and transporter function. Further experiments are necessary to confirm if the decreased growth in carboxylic acids presented by mutant alleles is due to an improved citrate export capacity. Inhibition assays determine if mutant alleles present altered substrate specificity. In addition, protein localization assays are necessary to determine if the alleles presenting increased growth are the result of AnCexA loss of function or incorrect protein localization, or low expression levels. In the future, an extended set of mutations will further uncover the structure-function relationship of this transporter, helping to elucidate the mechanism for substrate recognition and specificity of AnCexA.

6.6 REFERENCES

1. Agler, M., Wrenn, B., Zinder, S.,Angenent, L., *Waste to bioproduct conversion with undefined mixed cultures: the carboxylate platform.* Trends in biotechnology, **2011**. 29(2): p. 70-78
2. Behera, B., Mishra, R.,Mohapatra, S., *Microbial citric acid: Production, properties, application, and future perspectives.* Food Frontiers, **2021**
3. Analysts, G.I., *Global Citric Acid Industry*, **2020**. p. 1-351,
4. Papagianni, M., *Advances in citric acid fermentation by Aspergillus niger: biochemical aspects, membrane transport and modeling.* Biotechnology advances, **2007**. 25(3): p. 244-263
5. Crolla, A.,Kennedy, K., *Optimization of citric acid production from Candida lipolytica Y-1095 using n-paraffin.* Journal of biotechnology, **2001**. 89(1): p. 27-40
6. Vandenberghe, L., Soccol, C., Pandey, A.,Lebeault, J., *Microbial production of citric acid.* Brazilian Archives of Biology and Technology, **1999**. 42(3): p. 263-276
7. Tong, Z., Zheng, X., Tong, Y., Shi, Y.,Sun, J., *Systems metabolic engineering for citric acid production by Aspergillus niger in the post-genomic era.* Microbial cell factories, **2019**. 18(1): p. 1-15
8. Show, P., Oladele, K., Siew, Q., Aziz Zakry, F., Lan, J.,Ling, T., *Overview of citric acid production from Aspergillus niger.* Frontiers in Life Science, **2015**. 8(3): p. 271-283
9. Parshikov, I., Woodling, K.,Sutherland, J., *Biotransformations of organic compounds mediated by cultures of Aspergillus niger.* Applied microbiology and biotechnology, **2015**. 99(17): p. 6971-6986
10. Soares-Silva, I., Ribas, D., Sousa-Silva, M., Azevedo-Silva, J., Rendulić, T.,Casal, M., *Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications.* FEMS microbiology letters, **2020**
11. Steiger, M.G., Rassinger, A., Mattanovich, D.,Sauer, M., *Engineering of the citrate exporter protein enables high citric acid production in Aspergillus niger.* Metabolic engineering, **2019**. 52: p. 224-231
12. Kell, D., Swainston, N., Pir, P.,Oliver, S., *Membrane transporter engineering in industrial biotechnology and whole cell biocatalysis.* Trends in biotechnology, **2015**. 33(4): p. 237-246
13. Karagiosis, S.,Baker, S., *Fungal cell factories.* Food and industrial bioproducts and bioprocessing., **2012**: p. 205-219
14. Daugulis, A., *Integrated fermentation and recovery processes.* Current opinion in biotechnology, **1994**. 5(2): p. 192-195
15. Yang, S., Huang, H., Tay, A., Qin, W., De Guzman, L.,Nicolas, E., *Chapter 16 - Extractive Fermentation for the Production of Carboxylic Acids*, in *Bioprocessing for Value-Added Products from Renewable Resources*, S.-T. Yang, Editor **2007**, Elsevier: Amsterdam. p. 421-446, https://doi.org/10.1016/B978-044452114-9/50017-7
16. Boyarskiy, S.,Tullman-Ercek, D., *Getting pumped: membrane efflux transporters for enhanced biomolecule production.* Current opinion in chemical biology, **2015**. 28: p. 15-19
17. Alonso, S., Rendueles, M.,Díaz, M., *Microbial production of specialty organic acids from renewable and waste materials.* Critical reviews in biotechnology, **2015**. 35(4): p. 497-513

18. Nielsen, J., *Yeast systems biology: model organism and cell factory.* Biotechnology journal, **2019**. 14(9): p. 1800421
19. Leber, C.,Da Silva, N., *Engineering of Saccharomyces cerevisiae for the synthesis of short chain fatty acids.* Biotechnology and bioengineering, **2014**. 111(2): p. 347-358
20. Geiser, E., Przybilla, S., Friedrich, A., Buckel, W., Wierckx, N., Blank, L.,Bölker, M., *Ustilago maydis produces itaconic acid via the unusual intermediate trans-aconitate.* Microbial biotechnology, **2016**. 9(1): p. 116-126
21. Hossain, A.H., Li, A., Brickwedde, A., Wilms, L., Caspers, M., Overkamp, K.,Punt, P.J., *Rewiring a secondary metabolite pathway towards itaconic acid production in Aspergillus niger.* Microbial cell factories, **2016**. 15(1): p. 130
22. Huang, X., Lu, X., Li, Y., Li, X.,Li, J.-J., *Improving itaconic acid production through genetic engineering of an industrial Aspergillus terreus strain.* Microbial cell factories, **2014**. 13(1): p. 119
23. Camarasa, C., Bidard, F., Bony, M., Barre, P.,Dequin, S., *Characterization of Schizosaccharomyces pombe malate permease by expression in Saccharomyces cerevisiae.* Applied and Environmental Microbiology, **2001**. 67(9): p. 4144-4151, 10.1128/aem.67.9.4144-4151.2001
24. Darbani, B., Stovicek, V., van der Hoek, S.A.,Borodina, I., *Engineering energetically efficient transport of dicarboxylic acids in yeast Saccharomyces cerevisiae.* Proceedings of the National Academy of Sciences, **2019**: p. 201900287
25. Pacheco, A., Talaia, G., Sá-Pessoa, J., Bessa, D., Gonçalves, M.J., Moreira, R., Paiva, S., Casal, M.,Queirós, O., *Lactic acid production in Saccharomyces cerevisiae is modulated by expression of the monocarboxylate transporters Jen1 and Ady2.* FEMS Yeast Res, **2012**. 12(3): p. 375-381
26. Odoni, D., Vazquez-Vilar, M., van Gaal, M., Schonewille, T., Martins dos Santos, V., Tamayo-Ramos, J., Suarez-Diez, M.,Schaap, P., *Aspergillus niger citrate exporter revealed by comparison of two alternative citrate producing conditions.* FEMS microbiology letters, **2019**. 366(7): p. fnz071
27. Erian, A.M., Egermeier, M., Rassinger, A., Marx, H.,Sauer, M., *Identification of the citrate exporter Cex1 of Yarrowia lipolytica.* FEMS Yeast Res, **2020**
28. Burgstaller, W., *Thermodynamic boundary conditions suggest that a passive transport step suffices for citrate excretion in Aspergillus and Penicillium.* Microbiology, **2006**. 152(3): p. 887-893
29. Netik, A., Torres, N., Riol, J.,Kubicek, C., *Uptake and export of citric acid by Aspergillus niger is reciprocally regulated by manganese ions.* Biochimica et Biophysica Acta (BBA)-Biomembranes, **1997**. 1326(2): p. 287-294
30. Katoh, K., Rozewicki, J.,Yamada, K., *MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization.* Briefings in Bioinformatics, **2017**. 20(4): p. 1160-1166, 10.1093/bib/bbx108
31. Kumar, S., Stecher, G.,Tamura, K., *MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets.* Molecular biology and evolution, **2016**. 33(7): p. 1870-1874
32. Entian, K.,Kötter, P., *25 yeast genetic strain and plasmid collections.* Methods in microbiology, **2007**. 36: p. 629-666

33. Mans, R., Hassing, E.-J., Wijsman, M., Giezekamp, A., Pronk, J.T., Daran, J.-M.,van Maris, A.J.A., *A CRISPR/Cas9-based exploration into the elusive mechanism for lactate export in Saccharomyces cerevisiae.* FEMS Yeast Res, **2017**. 17(8), 10.1093/femsyr/fox085
34. Soares-Silva, I., Paiva, S., Diallinas, G.,Casal, M., *The conserved sequence NXX [S/T] HX [S/T] QDXXXT of the lactate/pyruvate: H+ symporter subfamily defines the function of the substrate translocation pathway.* Molecular membrane biology, **2007**. 24(5-6): p. 464-474
35. Ribas, D., Sa-Pessoa, J., Soares-Silva, I., Paiva, S., Nygard, Y., Ruohonen, L., Penttila, M.,Casal, M., *Yeast as a tool to express sugar acid transporters with biotechnological interest.* FEMS Yeast Res, **2017**. 17(2), 10.1093/femsyr/fox005
36. Zimmermann, L., Stephens, A., Nam, S., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A.,Alva, V., *A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core.* Journal of molecular biology, **2018**. 430(15): p. 2237-2243
37. Zheng, W., Zhang, C., Wuyun, Q., Pearce, R., Li, Y.,Zhang, Y., *LOMETS2: improved meta-threading server for fold-recognition and structure-based function annotation for distant-homology proteins.* Nucleic acids research, **2019**. 47(W1): p. W429-W436
38. Bosshart, P., Kalbermatter, D., Bonetti, S.,Fotiadis, D., *Mechanistic basis of L-lactate transport in the SLC16 solute carrier family.* Nat Commun, **2019**. 10(1): p. 1-11
39. Irwin, J.J.,Shoichet, B.K., *ZINC− A free database of commercially available compounds for virtual screening.* Journal of chemical information and modeling, **2005**. 45(1): p. 177-182
40. Pettersen, E., Goddard, T., Huang, C., Couch, G., Greenblatt, D., Meng, E.,Ferrin, T., *UCSF Chimera—a visualization system for exploratory research and analysis.* Journal of computational chemistry, **2004**. 25(13): p. 1605-1612
41. Dallakyan, S.,Olson, A., *Small-molecule library screening by docking with PyRx*, in *Chemical biology***2015**, Springer. p. 243-250
42. Smart, O., Neduvelil, J., Wang, X., Wallace, B.,Sansom, M., *HOLE: a program for the analysis of the pore dimensions of ion channel structural models.* Journal of molecular graphics, **1996**. 14(6): p. 354-360
43. Humphrey, W., Dalke, A.,Schulten, K., *VMD: visual molecular dynamics.* Journal of molecular graphics, **1996**. 14(1): p. 33-38
44. Gietz, R.,Woods, R., *Yeast transformation by the LiAc/SS Carrier DNA/PEG method*, in *Yeast Protocol***2006**, Springer. p. 107-120
45. Nunes, P., Tenreiro, S.,Sá-Correia, I., *Resistance and adaptation to quinidine in Saccharomyces cerevisiae: role of QDR1 (YIL120w), encoding a plasma membrane transporter of the major facilitator superfamily required for multidrug resistance.* Antimicrobial agents and chemotherapy, **2001**. 45(5): p. 1528-1534
46. Vargas, R., Tenreiro, S., Teixeira, M., Fernandes, A.,Sá-Correia, I., *Saccharomyces cerevisiae multidrug transporter Qdr2p (Yil121wp): localization and function as a quinidine resistance determinant.* Antimicrobial agents and chemotherapy, **2004**. 48(7): p. 2531-2537
47. Tenreiro, S., Nunes, P., Viegas, C., Neves, M., Teixeira, M., Cabral, M.,Sá-Correia, I., *AQR1 gene (ORF YNL065w) encodes a plasma membrane transporter of the major facilitator superfamily that confers resistance to short-chain monocarboxylic acids and quinidine in Saccharomyces cerevisiae.* Biochemical and Biophysical Research Communications, **2002**. 292(3): p. 741-748

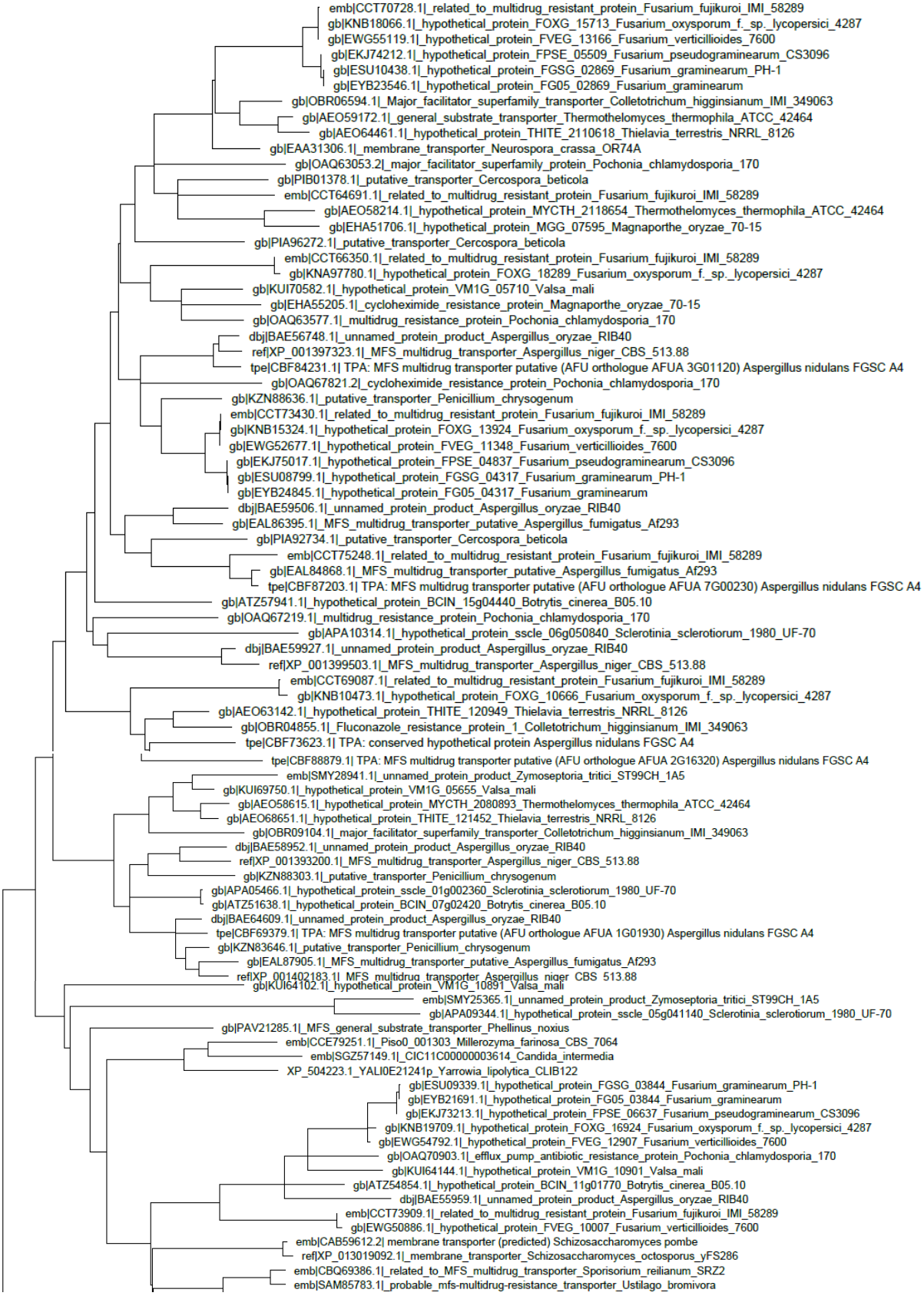
48. Velasco, I., Tenreiro, S., Calderon, I.,André, B., *Saccharomyces cerevisiae Aqr1 is an internal-membrane transporter involved in excretion of amino acids.* Eukaryotic Cell, **2004**. 3(6): p. 1492-1503
49. Vargas, R., García-Salcedo, R., Tenreiro, S., Teixeira, M., Fernandes, A., Ramos, J.,Sá-Correia, I., *Saccharomyces cerevisiae multidrug resistance transporter Qdr2 is implicated in potassium uptake, providing a physiological advantage to quinidine-stressed cells.* Eukaryotic Cell, **2007**. 6(2): p. 134-142
50. Tenreiro, S., Vargas, R., Teixeira, M., Magnani, C.,Sá-Correia, I., *The yeast multidrug transporter Qdr3 (Ybr043c): localization and role as a determinant of resistance to quinidine, barban, cisplatin, and bleomycin.* Biochemical and Biophysical Research Communications, **2005**. 327(3): p. 952-959
51. Felder, T., Bogengruber, E., Tenreiro, S., Ellinger, A., Sá-Correia, I.,Briza, P., *Dtr1p, a multidrug resistance transporter of the major facilitator superfamily, plays an essential role in spore wall maturation in Saccharomyces cerevisiae.* Eukaryotic Cell, **2002**. 1(5): p. 799-810
52. Dos Santos, S., Teixeira, M., Dias, P.,Sá-Correia, I., *MFS transporters required for multidrug/multixenobiotic (MD/MX) resistance in the model yeast: understanding their physiological function through post-genomic approaches.* Frontiers in physiology, **2014**. 5: p. 180
53. Dias, P.J., Seret, M., Goffeau, A., Correia, I.,Baret, P., *Evolution of the 12-spanner drug: H+ antiporter DHA1 family in hemiascomycetous yeasts.* Omics: a journal of integrative biology, **2010**. 14(6): p. 701-710
54. Andersen, M., Salazar, M., Schaap, P., van de Vondervoort, P., Culley, D., Thykaer, J., Frisvad, J., Nielsen, K., Albang, R., Albermann, K., Berka, R., Braus, G., Braus-Stromeyer, S., Corrochano, L., Dai, Z., van Dijck, P., Hofmann, G., Lasure, L., Magnuson, J., Menke, H., Meijer, M., Meijer, S., Nielsen, J., Nielsen, M., van Ooyen, A., Pel, H., Poulsen, L., Samson, R., Stam, H., Tsang, A., van den Brink, J., Atkins, A., Aerts, A., Shapiro, H., Pangilinan, J., Salamov, A., Lou, Y., Lindquist, E., Lucas, S., Grimwood, J., Grigoriev, I., Kubicek, C., Martinez, D., van Peij, N., Roubos, J., Nielsen, J.,Baker, S., *Comparative genomics of citric-acid-producing Aspergillus niger ATCC 1015 versus enzyme-producing CBS 513.88.* Genome research, **2011**. 21(6): p. 885-897, 10.1101/gr.112169.110
55. Pel, H., De Winde, J., Archer, D., Dyer, P., Hofmann, G., Schaap, P., Turner, G., De Vries, R., Albang, R.,Albermann, K., *Genome sequencing and analysis of the versatile cell factory Aspergillus niger CBS 513.88.* Nature biotechnology, **2007**. 25(2): p. 221-231
56. Casal, M., Paiva, S., Queirós, O.,Soares-Silva, I., *Transport of carboxylic acids in yeasts.* FEMS microbiology reviews, **2008**. 32(6): p. 974-994
57. Casal, M., Queirós, O., Talaia, G., Ribas, D.,Paiva, S., *Carboxylic acids plasma membrane transporters in Saccharomyces cerevisiae*, in *Yeast membrane transport***2016**, Springer. p. 229-251
58. Peña, A., Sánchez, N., Álvarez, H., Calahorra, M.,Ramírez, J., *Effects of high medium pH on growth, metabolism and transport in Saccharomyces cerevisiae.* FEMS Yeast Res, **2015**. 15(2)
59. Wilson, M., Jackson, V., Heddle, C., Price, N., Pilegaard, H., Juel, C., Bonen, A., Montgomery, I., Hutter, O.,Halestrap, A., *Lactic acid efflux from white skeletal muscle is catalyzed by the*

monocarboxylate transporter isoform MCT3. Journal of Biological Chemistry, **1998**. 273(26): p. 15920-15926

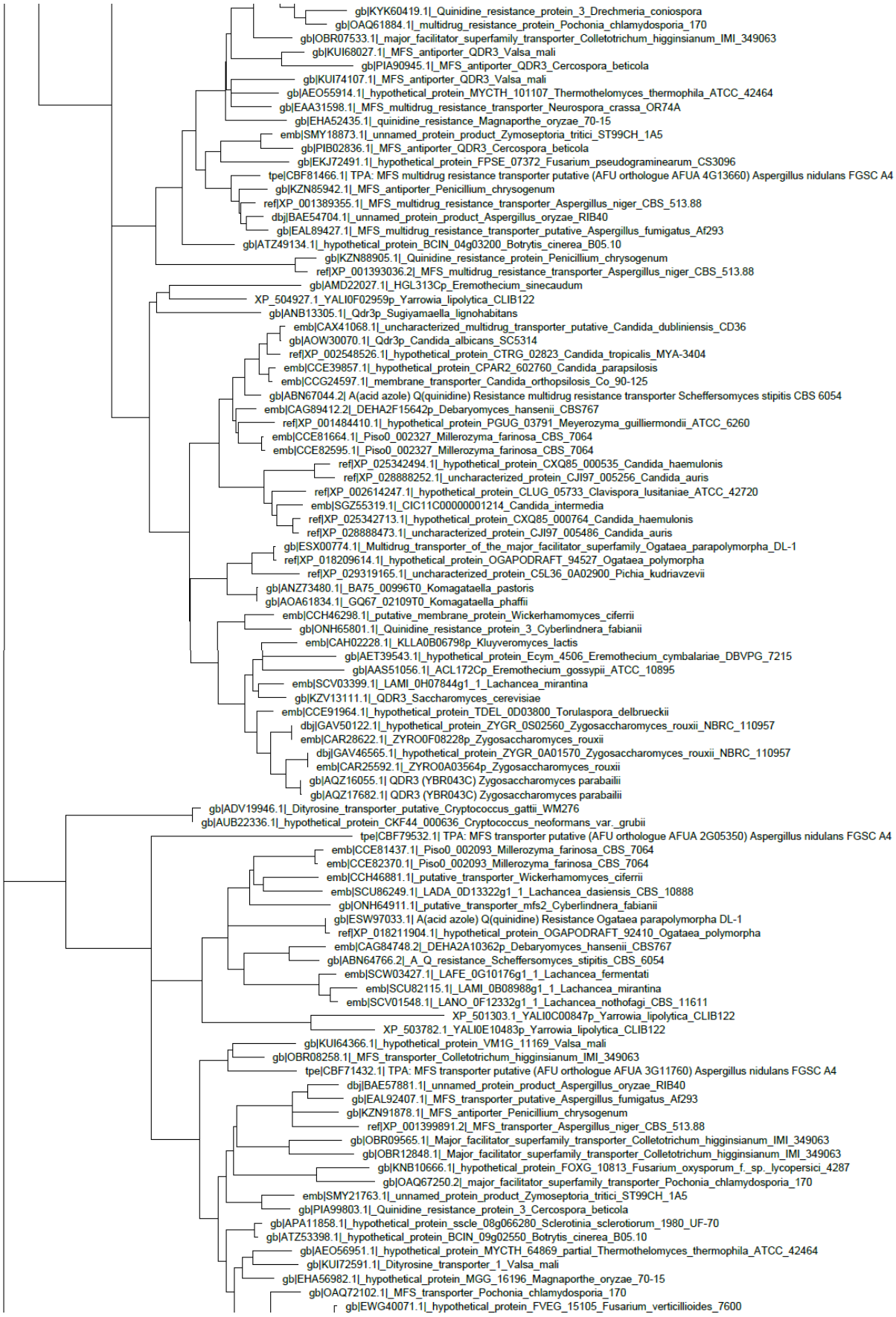
60. Halestrap, A., *Monocarboxylic acid transport.* Comprehensive Physiology, **2011**. 3(4): p. 1611-1643
61. Dimmer, K., Friedrich, B., Lang, F., Deitmer, J.,Bröer, S., *The low-affinity monocarboxylate transporter MCT4 is adapted to the export of lactate in highly glycolytic cells.* Biochemical journal, **2000**. 350(1): p. 219-227
62. Roberts, S.K., Milnes, J.,Caddick, M., *Characterisation of AnBEST1, a functional anion channel in the plasma membrane of the filamentous fungus, Aspergillus nidulans.* Fungal Genetics and Biology, **2011**. 48(9): p. 928-938, https://doi.org/10.1016/j.fgb.2011.05.004
63. Xi, Y., Zhan, T., Xu, H., Chen, J., Bi, C., Fan, F.,Zhang, X., *Characterization of JEN family carboxylate transporters from the acid-tolerant yeast Pichia kudriavzevii and their applications in succinic acid production.* Microbial biotechnology, **2021**
64. Halestrap, A., *The SLC16 gene family–structure, role and regulation in health and disease.* Molecular aspects of medicine, **2013**. 34(2-3): p. 337-349
65. Halestrap, A.,Meredith, D., *The SLC16 gene family—from monocarboxylate transporters (MCTs) to aromatic amino acid transporters and beyond.* Pflügers Archiv, **2004**. 447(5): p. 619-628
66. Kurczab, R., Śliwa, P., Rataj, K., Kafel, R.,Bojarski, A., *Salt bridge in ligand–protein complexes—systematic theoretical and statistical investigations.* Journal of chemical information and modeling, **2018**. 58(11): p. 2224-2238
67. Bosshard, H., Marti, D.,Jelesarov, I., *Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings.* Journal of Molecular Recognition, **2004**. 17(1): p. 1-16
68. Berovic, M.,Legisa, M., *Citric acid production.* Biotechnology annual review, **2007**. 13: p. 303-343

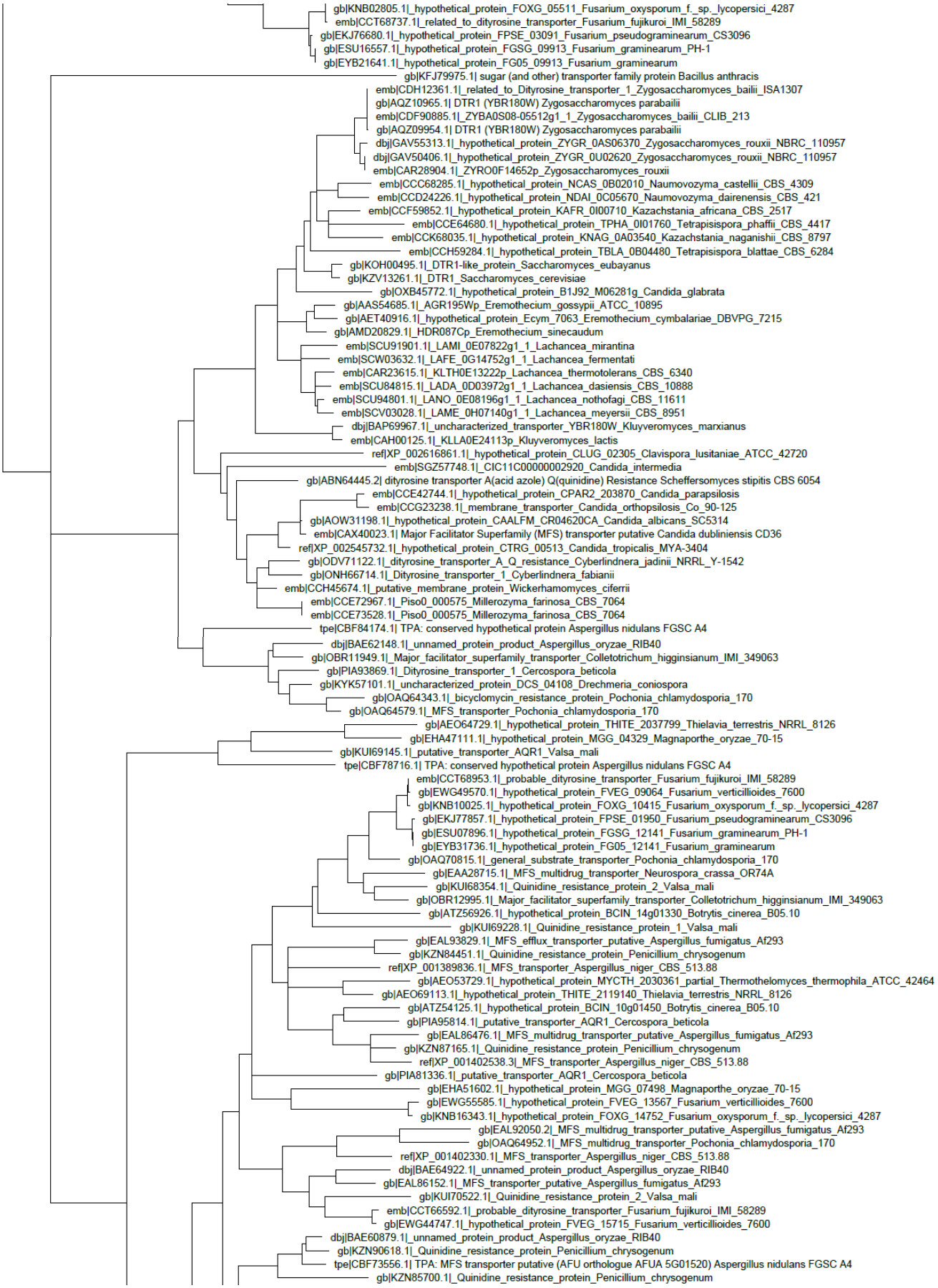
SUPPLEMENTARY MATERIAL

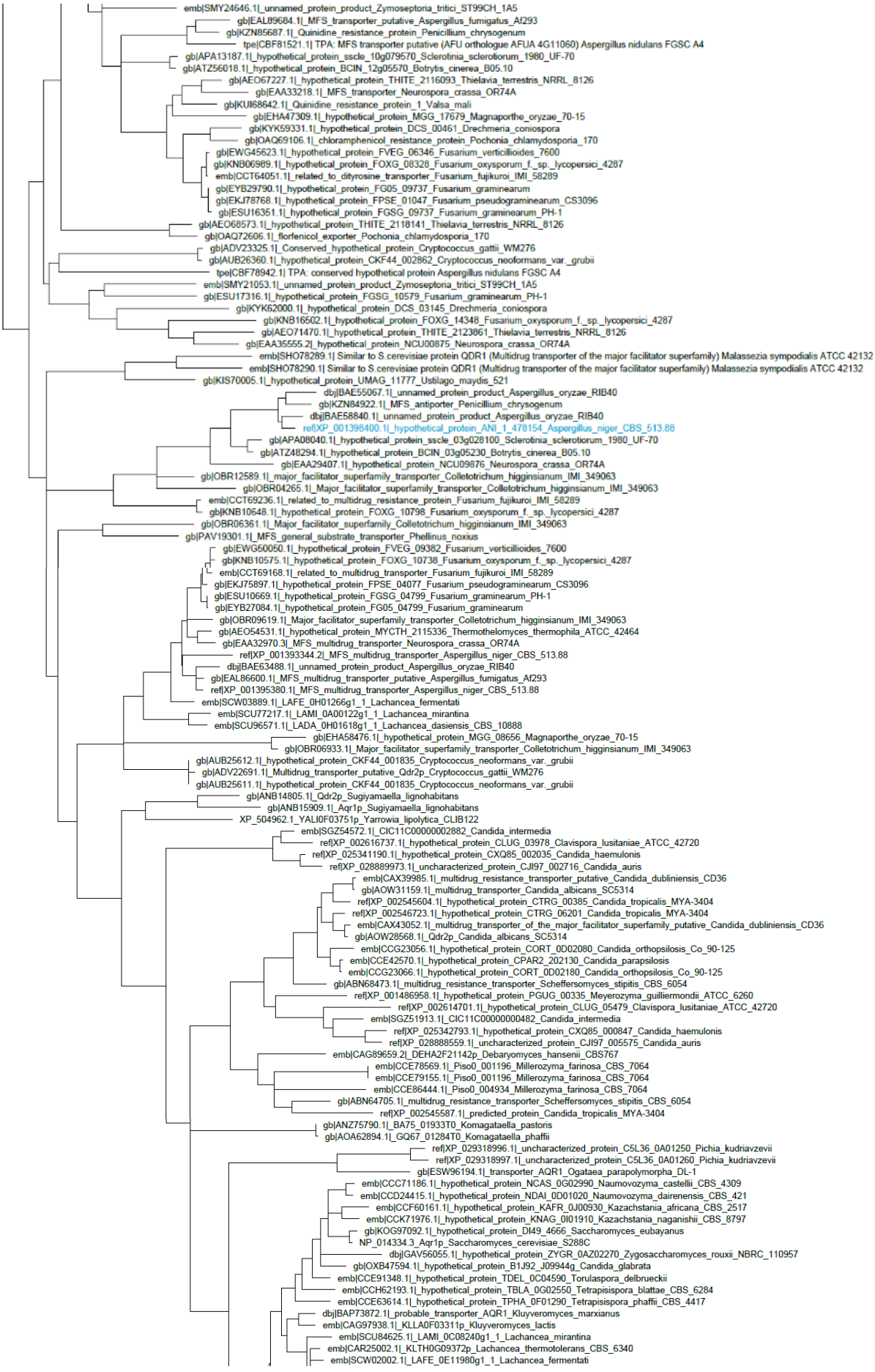
S1. Detailed AnCexA-phylogenetic tree with all homolog nominations.

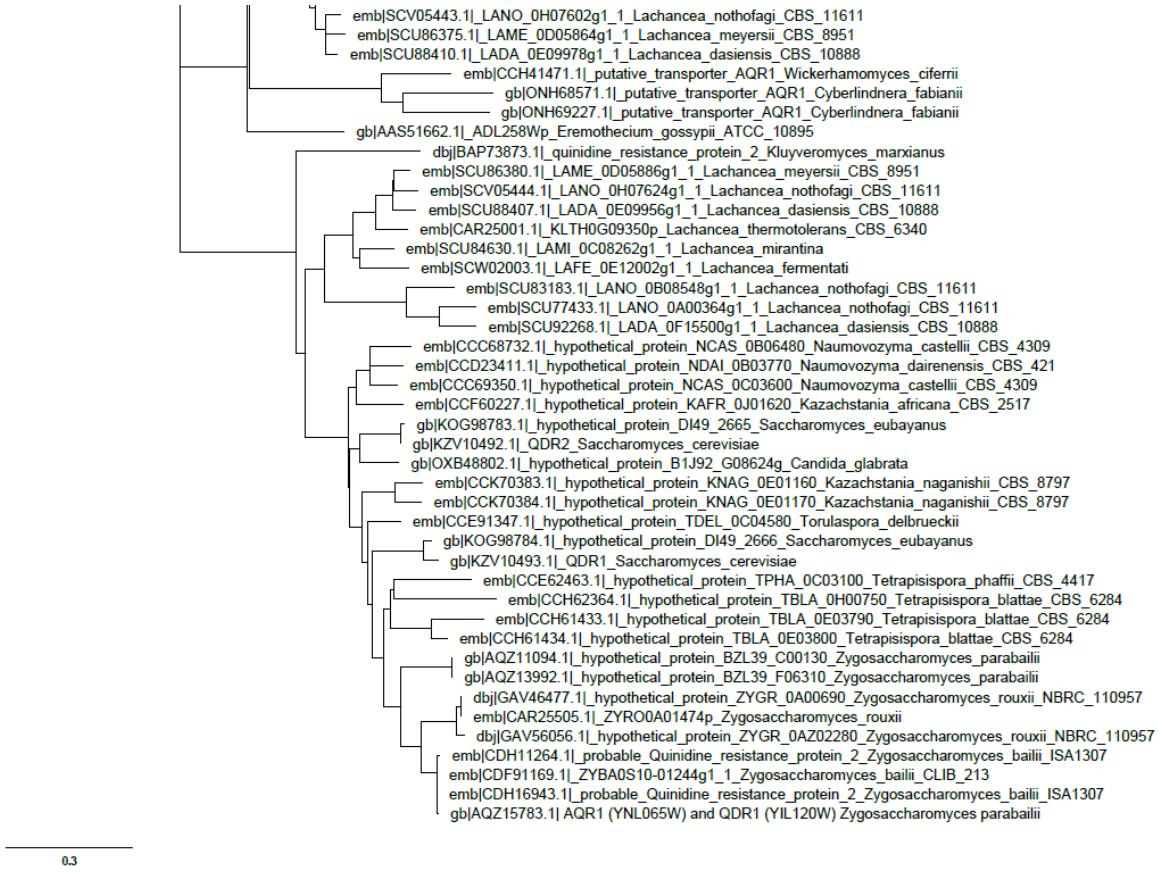



emb|SHO79823.1| Similar to S.cerevisiae protein FLR1 (Plasma membrane transporter of the major facilitator superfamily) Malassezia sympodialis ATCC 42132
dbj|BAE57729.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
ref|XP_001400436.1|_MFS_multidrug_transporter_Aspergillus_niger_CBS_513.88
gb|ANB15002.1|_Tpo1p_Sugiyamaella_lignohabitans
gb|PAV16754.1|_MFS_polyamine_transporter_Phellinus_noxius
gb|APA13108.1|_hypothetical_protein_sscle_10g078780_Sclerotinia_sclerotiorum_1980_UF-70
gb|ATZ55900.1|_hypothetical_protein_BCIN_12g04490_Botrytis_cinerea_B05.10
gb|OBR02488.1|_Major_facilitator_superfamily_transporter_Colletotrichum_higginsianum_IMI_349063
gb|KNB10734.1|_hypothetical_protein_FOXG_10866_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|OAQ71223.1|_multidrug_resistance_protein_Pochonia_chlamydosporia_170
gb|KUI68182.1|_hypothetical_protein_VM1G_04063_Valsa_mali
dbj|BAE62217.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
tpe|CBF78056.1| TPA: MFS multidrug transporter putative (AFU orthologue AFUA 6G03040) Aspergillus nidulans FGSC A4
dbj|BAE60182.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|EAL85734.1|_MFS_multidrug_transporter_putative_Aspergillus_fumigatus_Af293
gb|OAQ62371.1|_MFS_transporter_Pochonia_chlamydosporia_170
ref|XP_001396212.1|_MFS_multidrug_transporter_Aspergillus_niger_CBS_513.88
ref|XP_001394704.2|_MFS_multidrug_transporter_Aspergillus_niger_CBS_513.88
gb|PIB02553.1|_putative_transporter_Cercospora_beticola
gb|APA09951.1|_hypothetical_protein_sscle_05g047210_Sclerotinia_sclerotiorum_1980_UF-70
emb|SMY18972.1|_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
gb|PIA88696.1|_putative_transporter_Cercospora_beticola
gb|KZN89571.1|_putative_transporter_Penicillium_chrysogenum
tpe|CBF71792.1| TPA: conserved hypothetical protein Aspergillus nidulans FGSC A4
gb|OBR12977.1|_MFS_multidrug_transporter_Colletotrichum_higginsianum_IMI_349063
emb|CCT65000.1|_related_to_multidrug_resistant_protein_Fusarium_fujikuroi_IMI_58289
gb|KNB02984.1|_hypothetical_protein_FOXG_05602_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|EWG41299.1|_hypothetical_protein_FVEG_03436_Fusarium_verticillioides_7600
gb|ESU08294.1|_hypothetical_protein_FGSG_11525_Fusarium_graminearum_PH-1
gb|EYB21869.1|_hypothetical_protein_FG05_11525_Fusarium_graminearum
gb|OBR06822.1|_MFS_multidrug_transporter_Colletotrichum_higginsianum_IMI_349063
gb|APA06336.1|_hypothetical_protein_sscle_01g011060_Sclerotinia_sclerotiorum_1980_UF-70
gb|ATZ55329.1|_hypothetical_protein_BCIN_11g05940_Botrytis_cinerea_B05.10
gb|OBR16457.1|_Multidrug_resistance_protein_Colletotrichum_higginsianum_IMI_349063
XP_501444.1_YALI0C04499p_Yarrowia_lipolytica_CLIB122
XP_501568.1_YALI0C07722p_Yarrowia_lipolytica_CLIB122
XP_505874.1_YALI0F25597p_Yarrowia_lipolytica_CLIB122
XP_501443.1_YALI0C04477p_Yarrowia_lipolytica_CLIB122
gb|ANZ76993.1|_BA75_03977T0_Komagataella_pastoris
gb|AOA61076.1|_GQ67_01728T0_Komagataella_phaffii
dbj|BAE63710.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
emb|CCT63091.1|_related_to_multidrug_resistant_protein_Fusarium_fujikuroi_IMI_58289
gb|KNA95865.1|_hypothetical_protein_FOXG_01267_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|EWG36065.1|_hypothetical_protein_FVEG_00229_Fusarium_verticillioides_7600
gb|EHA50726.1|_fluconazole_resistance_protein_1_Magnaporthe_oryzae_70-15
gb|AEO64097.1|_hypothetical_protein_THITE_2109873_Thielavia_terrestris_NRRL_8126
gb|EDO65246.2|_hypothetical_protein_NCU10314_Neurospora_crassa_OR74A
gb|KZN88954.1|_putative_drug/proton_antiporter_Penicillium_chrysogenum
gb|OAQ73873.1|_multidrug_resistance_protein_Pochonia_chlamydosporia_170
emb|CCT70182.1|_related_to_multidrug_resistant_protein_Fusarium_fujikuroi_IMI_58289
gb|EWG39033.1|_hypothetical_protein_FVEG_02044_Fusarium_verticillioides_7600
gb|KZN85529.1|_putative_drug/proton_antiporter_Penicillium_chrysogenum
ref|XP_001400577.2|_MFS_transporter_Aspergillus_niger_CBS_513.88
dbj|BAE59728.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|PIA91736.1|_putative_drug/proton_antiporter_Cercospora_beticola
gb|APA08577.1|_hypothetical_protein_sscle_04g033470_Sclerotinia_sclerotiorum_1980_UF-70
gb|ATZ58201.1|_hypothetical_protein_BCIN_16g00530_Botrytis_cinerea_B05.10
emb|CBQ72989.1|_related_to_multidrug_resistant_protein_Sporisorium_reilianum_SRZ2
XP_500222.1_YALI0A18920p_Yarrowia_lipolytica_CLIB122
XP_505386.1_YALI0F13739p_Yarrowia_lipolytica_CLIB122
gb|AEO63931.1|_hypothetical_protein_THITE_2141914_Thielavia_terrestris_NRRL_8126
XP_500099.1_YALI0A15576p_Yarrowia_lipolytica_CLIB122
XP_505105.1_YALI0F07062p_Yarrowia_lipolytica_CLIB122
XP_505146.1_YALI0F08063p_Yarrowia_lipolytica_CLIB122
gb|EWG53172.1|_hypothetical_protein_FVEG_11656_Fusarium_verticillioides_7600
dbj|BAE64493.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|ADV23623.1|_Polyamine_transport-related_protein_putative_Cryptococcus_gattii_WM276
gb|AUB26660.1|_hypothetical_protein_CKF44_003112_Cryptococcus_neoformans_var._grubii
gb|PAV21995.1|_MFS_general_substrate_transporter_Phellinus_noxius
emb|SAM81401.1|_probable_mfs-multidrug-resistance_transporter_Ustilago_bromivora
gb|KIS70849.1|_putative_mfs-multidrug-resistance_transporter_Ustilago_maydis_521
gb|KIS66248.1|_hypothetical_protein_UMAG_05248_Ustilago_maydis_521
emb|CBQ73118.1|_probable_mfs-multidrug-resistance_transporter_Sporisorium_reilianum_SRZ2
gb|KIS69389.1|_putative_mfs-multidrug-resistance_transporter_Ustilago_maydis_521
gb|AUB22127.1|_hypothetical_protein_CKF44_000465_Cryptococcus_neoformans_var._grubii
gb|ADV24402.1|_spermine_transporter_putative_Cryptococcus_gattii_WM276
gb|AUB27442.1|_hypothetical_protein_CKF44_004298_Cryptococcus_neoformans_var._grubii
gb|AUB28478.1|_hypothetical_protein_CKF44_005478_Cryptococcus_neoformans_var._grubii
gb|KZN93571.1|_Polyamine_transporter_Penicillium_chrysogenum
XP_501168.1_YALI0B21142p_Yarrowia_lipolytica_CLIB122
emb|CCE72443.1|_Piso0_000011_Millerozyma_farinosa_CBS_7064
emb|CCE73002.1|_Piso0_000011_Millerozyma_farinosa_CBS_7064
gb|ONH68432.1|_Polyamine_transporter_3_Cyberlindnera_fabianii
gb|PIB02543.1|_putative_MFS-type_transporter_Cercospora_beticola
gb|APA14254.1|_hypothetical_protein_sscle_12g090240_Sclerotinia_sclerotiorum_1980_UF-70
gb|ATZ47468.1|_hypothetical_protein_BCIN_02g07440_Botrytis_cinerea_B05.10
gb|EAL86034.1|_MFS_multidrug_transporter_putative_Aspergillus_fumigatus_Af293
gb|AEO70680.1|_hypothetical_protein_THITE_2122357_Thielavia_terrestris_NRRL_8126
gb|OAQ68731.2|_MFS_multidrug_transporter_Pochonia_chlamydosporia_170
gb|EKJ75527.1|_hypothetical_protein_FPSE_04302_Fusarium_pseudograminearum_CS3096
emb|CCT64016.1|_probable_multidrug_resistance_protein_Fusarium_fujikuroi_IMI_58289
gb|KNB07069.1|_hypothetical_protein_FOXG_08371_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|KNB07071.1|_hypothetical_protein_FOXG_08371_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|KNB07067.1|_hypothetical_protein_FOXG_08371_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|KNB07068.1|_hypothetical_protein_FOXG_08371_Fusarium_oxysporum_f._sp._lycopersici_4287
emb|SMY21311.1|_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
tpe|CBF73862.1| TPA: conserved hypothetical protein Aspergillus nidulans FGSC A4
gb|PIA97565.1|_putative_transporter_Cercospora_beticola
gb|PIA99416.1|_putative_transporter_mfs2_Cercospora_beticola
gb|APA06215.1|_hypothetical_protein_sscle_01g009850_Sclerotinia_sclerotiorum_1980_UF-70
gb|ATZ55282.1|_hypothetical_protein_BCIN_11g05520_Botrytis_cinerea_B05.10
XP_503088.1_YALI0D20834p_Yarrowia_lipolytica_CLIB122
XP_505392.1_YALI0F13981p_Yarrowia_lipolytica_CLIB122
XP_504964.1_YALI0F03795p_Yarrowia_lipolytica_CLIB122
dbj|BAE55142.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|APA13051.1|_hypothetical_protein_sscle_10g078210_Sclerotinia_sclerotiorum_1980_UF-70
dbj|BAE55871.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
tpe|CBF89705.1| TPA: conserved hypothetical protein Aspergillus nidulans FGSC A4
gb|KZN90505.1|_putative_transporter_Penicillium_chrysogenum
gb|EAL93187.1|_MFS_transporter_putative_Aspergillus_fumigatus_Af293
tpe|CBF70261.1| TPA: MFS transporter putative (AFU orthologue AFUA 2G08910) Aspergillus nidulans FGSC A4
gb|PAV17244.1|_MFS_general_substrate_transporter_Phellinus_noxius
gb|PAV21643.1|_MFS_general_substrate_transporter_Phellinus_noxius
gb|AUB24047.1|_hypothetical_protein_CKF44_004599_Cryptococcus_neoformans_var._grubii
emb|CCT75627.1|_related_to_dityrosine_transporter_Fusarium_fujikuroi_IMI_58289
gb|EWG51978.1|_hypothetical_protein_FVEG_10820_Fusarium_verticillioides_7600
gb|KNB12810.1|_hypothetical_protein_FOXG_12306_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|KNB12809.1|_hypothetical_protein_FOXG_12306_Fusarium_oxysporum_f._sp._lycopersici_4287
emb|CCT68543.1|_related_to_dityrosine_transporter_Fusarium_fujikuroi_IMI_58289
gb|KNB02382.1|_hypothetical_protein_FOXG_05286_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|EWG40462.1|_hypothetical_protein_FVEG_02843_Fusarium_verticillioides_7600
gb|ESU16796.1|_hypothetical_protein_FGSG_10120_Fusarium_graminearum_PH-1
gb|EKJ77162.1|_hypothetical_protein_FPSE_02612_Fusarium_pseudograminearum_CS3096
gb|EYB25751.1|_hypothetical_protein_FG05_10120_Fusarium_graminearum
gb|OAQ65507.1|_multidrug_resistance_protein_Pochonia_chlamydosporia_170


gb|KYK60419.1|_Quinidine_resistance_protein_3_Drechmeria_coniospora
gb|OAQ61884.1|_multidrug_resistance_protein_Pochonia_chlamydosporia_170
gb|OBR07533.1|_major_facilitator_superfamily_transporter_Colletotrichum_higginsianum_IMI_349063
gb|KUI68027.1|_MFS_antiporter_QDR3_Valsa_mali
gb|PIA90945.1|_MFS_antiporter_QDR3_Cercospora_beticola
gb|KUI74107.1|_MFS_antiporter_QDR3_Valsa_mali
gb|AEO55914.1|_hypothetical_protein_MYCTH_101107_Thermothelomyces_thermophila_ATCC_42464
gb|EAA31598.1|_MFS_multidrug_resistance_transporter_Neurospora_crassa_OR74A
gb|EHA52435.1|_quinidine_resistance_Magnaporthe_oryzae_70-15
emb|SMY18873.1|_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
gb|PIB02836.1|_MFS_antiporter_QDR3_Cercospora_beticola
gb|EKJ72491.1|_hypothetical_protein_FPSE_07372_Fusarium_pseudograminearum_CS3096
tpe|CBF81466.1|_TPA:_MFS_multidrug_resistance_transporter_putative_(AFU_orthologue_AFUA_4G13660)_Aspergillus_nidulans_FGSC_A4
gb|KZN85942.1|_MFS_antiporter_Penicillium_chrysogenum
ref|XP_001389355.1|_MFS_multidrug_resistance_transporter_Aspergillus_niger_CBS_513.88
dbj|BAE54704.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|EAL89427.1|_MFS_multidrug_resistance_transporter_putative_Aspergillus_fumigatus_Af293
gb|ATZ49134.1|_hypothetical_protein_BCIN_04g03200_Botrytis_cinerea_B05.10
gb|KZN88905.1|_Quinidine_resistance_protein_Penicillium_chrysogenum
ref|XP_001393036.2|_MFS_multidrug_resistance_transporter_Aspergillus_niger_CBS_513.88
gb|AMD22027.1|_HGL313Cp_Eremothecium_sinecaudum
XP_504927.1_YALI0F02959p_Yarrowia_lipolytica_CLIB122
gb|ANB13305.1|_Qdr3p_Sugiyamaella_lignohabitans
emb|CAX41068.1|_uncharacterized_multidrug_transporter_putative_Candida_dubliniensis_CD36
gb|AOW30070.1|_Qdr3p_Candida_albicans_SC5314
ref|XP_002548526.1|_hypothetical_protein_CTRG_02823_Candida_tropicalis_MYA-3404
emb|CCE39857.1|_hypothetical_protein_CPAR2_602760_Candida_parapsilosis
emb|CCG24597.1|_membrane_transporter_Candida_orthopsilosis_Co_90-125
gb|ABN67044.2|_A(acid azole) Q(quinidine) Resistance multidrug resistance transporter Scheffersomyces stipitis CBS 6054
emb|CAG89412.2|_DEHA2F15642p_Debaryomyces_hansenii_CBS767
ref|XP_001484410.1|_hypothetical_protein_PGUG_03791_Meyerozyma_guilliermondii_ATCC_6260
emb|CCE81664.1|_Piso0_002327_Millerozyma_farinosa_CBS_7064
emb|CCE82595.1|_Piso0_002327_Millerozyma_farinosa_CBS_7064
ref|XP_025342494.1|_hypothetical_protein_CXQ85_000535_Candida_haemulonis
ref|XP_028888252.1|_uncharacterized_protein_CJI97_005256_Candida_auris
ref|XP_002614247.1|_hypothetical_protein_CLUG_05733_Clavispora_lusitaniae_ATCC_42720
emb|SGZ55319.1|_CIC11C00000001214_Candida_intermedia
ref|XP_025342713.1|_hypothetical_protein_CXQ85_000764_Candida_haemulonis
ref|XP_028888473.1|_uncharacterized_protein_CJI97_005486_Candida_auris
gb|ESX00774.1|_Multidrug_transporter_of_the_major_facilitator_superfamily_Ogataea_parapolymorpha_DL-1
ref|XP_018209614.1|_hypothetical_protein_OGAPODRAFT_94527_Ogataea_polymorpha
ref|XP_029319165.1|_uncharacterized_protein_C5L36_0A02900_Pichia_kudriavzevii
gb|ANZ73480.1|_BA75_00996T0_Komagataella_pastoris
gb|AOA61834.1|_GQ67_02109T0_Komagataella_phaffii
emb|CCH46298.1|_putative_membrane_protein_Wickerhamomyces_ciferrii
gb|ONH65801.1|_Quinidine_resistance_protein_3_Cyberlindnera_fabianii
emb|CAH02228.1|_KLLA0B06798p_Kluyveromyces_lactis
gb|AET39543.1|_hypothetical_protein_Ecym_4506_Eremothecium_cymbalariae_DBVPG_7215
gb|AAS51056.1|_ACL172Cp_Eremothecium_gossypii_ATCC_10895
emb|SCV03399.1|_LAMI_0H07844g1_1_Lachancea_mirantina
gb|KZV13111.1|_QDR3_Saccharomyces_cerevisiae
emb|CCE91964.1|_hypothetical_protein_TDEL_0D03800_Torulaspora_delbrueckii
dbj|GAV50122.1|_hypothetical_protein_ZYGR_0S02560_Zygosaccharomyces_rouxii_NBRC_110957
emb|CAR28622.1|_ZYRO0F08228p_Zygosaccharomyces_rouxii
dbj|GAV46565.1|_hypothetical_protein_ZYGR_0A01570_Zygosaccharomyces_rouxii_NBRC_110957
emb|CAR25592.1|_ZYRO0A03564p_Zygosaccharomyces_rouxii
gb|AQZ16055.1|_QDR3_(YBR043C)_Zygosaccharomyces_parabailii
gb|AQZ17682.1|_QDR3_(YBR043C)_Zygosaccharomyces_parabailii
gb|ADV19946.1|_Dityrosine_transporter_putative_Cryptococcus_gattii_WM276
gb|AUB22336.1|_hypothetical_protein_CKF44_000636_Cryptococcus_neoformans_var._grubii
tpe|CBF79532.1|_TPA:_MFS_transporter_putative_(AFU_orthologue_AFUA_2G05350)_Aspergillus_nidulans_FGSC_A4
emb|CCE81437.1|_Piso0_002093_Millerozyma_farinosa_CBS_7064
emb|CCE82370.1|_Piso0_002093_Millerozyma_farinosa_CBS_7064
emb|CCH46881.1|_putative_transporter_Wickerhamomyces_ciferrii
emb|SCU86249.1|_LADA_0D13322g1_1_Lachancea_dasiensis_CBS_10888
gb|ONH64911.1|_putative_transporter_mfs2_Cyberlindnera_fabianii
gb|ESW97033.1|_A(acid azole) Q(quinidine) Resistance Ogataea parapolymorpha DL-1
ref|XP_018211904.1|_hypothetical_protein_OGAPODRAFT_92410_Ogataea_polymorpha
emb|CAG84748.2|_DEHA2A10362p_Debaryomyces_hansenii_CBS767
gb|ABN64766.2|_A_Q_resistance_Scheffersomyces_stipitis_CBS_6054
emb|SCW03427.1|_LAFE_0G10176g1_1_Lachancea_fermentati
emb|SCU82115.1|_LAMI_0B08988g1_1_Lachancea_mirantina
emb|SCV01548.1|_LANO_0F12332g1_1_Lachancea_nothofagi_CBS_11611
XP_501303.1_YALI0C00847p_Yarrowia_lipolytica_CLIB122
XP_503782.1_YALI0E10483p_Yarrowia_lipolytica_CLIB122
gb|KUI64366.1|_hypothetical_protein_VM1G_11169_Valsa_mali
gb|OBR08258.1|_MFS_transporter_Colletotrichum_higginsianum_IMI_349063
tpe|CBF71432.1|_TPA:_MFS_transporter_putative_(AFU_orthologue_AFUA_3G11760)_Aspergillus_nidulans_FGSC_A4
dbj|BAE57881.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|EAL92407.1|_MFS_transporter_putative_Aspergillus_fumigatus_Af293
gb|KZN91878.1|_MFS_antiporter_Penicillium_chrysogenum
ref|XP_001399891.2|_MFS_transporter_Aspergillus_niger_CBS_513.88
gb|OBR09565.1|_Major_facilitator_superfamily_transporter_Colletotrichum_higginsianum_IMI_349063
gb|OBR12848.1|_Major_facilitator_superfamily_transporter_Colletotrichum_higginsianum_IMI_349063
gb|KNB10666.1|_hypothetical_protein_FOXG_10813_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|OAQ67250.2|_major_facilitator_superfamily_transporter_Pochonia_chlamydosporia_170
emb|SMY21763.1|_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
gb|PIA99803.1|_Quinidine_resistance_protein_3_Cercospora_beticola
gb|APA11858.1|_hypothetical_protein_sscle_08g066280_Sclerotinia_sclerotiorum_1980_UF-70
gb|ATZ53398.1|_hypothetical_protein_BCIN_09g02550_Botrytis_cinerea_B05.10
gb|AEO56951.1|_hypothetical_protein_MYCTH_64869_partial_Thermothelomyces_thermophila_ATCC_42464
gb|KUI72591.1|_Dityrosine_transporter_1_Valsa_mali
gb|EHA56982.1|_hypothetical_protein_MGG_16196_Magnaporthe_oryzae_70-15
gb|OAQ72102.1|_MFS_transporter_Pochonia_chlamydosporia_170
gb|EWG40071.1|_hypothetical_protein_FVEG_15105_Fusarium_verticillioides_7600


gb|KNB02805.1|_hypothetical_protein_FOXG_05511_Fusarium_oxysporum_f._sp._lycopersici_4287
emb|CCT68737.1|_related_to_dityrosine_transporter_Fusarium_fujikuroi_IMI_58289
gb|EKJ76680.1|_hypothetical_protein_FPSE_03091_Fusarium_pseudograminearum_CS3096
gb|ESU16557.1|_hypothetical_protein_FGSG_09913_Fusarium_graminearum_PH-1
gb|EYB21641.1|_hypothetical_protein_FG05_09913_Fusarium_graminearum
gb|KFJ79975.1| sugar (and other) transporter family protein Bacillus anthracis
emb|CDH12361.1|_related_to_Dityrosine_transporter_1_Zygosaccharomyces_bailii_ISA1307
gb|AQZ10965.1| DTR1 (YBR180W) Zygosaccharomyces parabailii
emb|CDF90885.1|_ZYBA0S08-05512g1_1_Zygosaccharomyces_bailii_CLIB_213
gb|AQZ09954.1| DTR1 (YBR180W) Zygosaccharomyces parabailii
dbj|GAV55313.1|_hypothetical_protein_ZYGR_0AS06370_Zygosaccharomyces_rouxii_NBRC_110957
dbj|GAV50406.1|_hypothetical_protein_ZYGR_0U02620_Zygosaccharomyces_rouxii_NBRC_110957
emb|CAR28904.1|_ZYRO0F14652p_Zygosaccharomyces_rouxii
emb|CCC68285.1|_hypothetical_protein_NCAS_0B02010_Naumovozyma_castellii_CBS_4309
emb|CCD24226.1|_hypothetical_protein_NDAI_0C05670_Naumovozyma_dairenensis_CBS_421
emb|CCF59852.1|_hypothetical_protein_KAFR_0I00710_Kazachstania_africana_CBS_2517
emb|CCE64680.1|_hypothetical_protein_TPHA_0I01760_Tetrapisispora_phaffii_CBS_4417
emb|CCK68035.1|_hypothetical_protein_KNAG_0A03540_Kazachstania_naganishii_CBS_8797
emb|CCH59284.1|_hypothetical_protein_TBLA_0B04480_Tetrapisispora_blattae_CBS_6284
gb|KOH00495.1|_DTR1-like_protein_Saccharomyces_eubayanus
gb|KZV13261.1|_DTR1_Saccharomyces_cerevisiae
gb|OXB45772.1|_hypothetical_protein_B1J92_M06281g_Candida_glabrata
gb|AAS54685.1|_AGR195Wp_Eremothecium_gossypii_ATCC_10895
gb|AET40916.1|_hypothetical_protein_Ecym_7063_Eremothecium_cymbalariae_DBVPG_7215
gb|AMD20829.1|_HDR087Cp_Eremothecium_sinecaudum
emb|SCU91901.1|_LAMI_0E07822g1_1_Lachancea_mirantina
emb|SCW03632.1|_LAFE_0G14752g1_1_Lachancea_fermentati
emb|CAR23615.1|_KLTH0E13222p_Lachancea_thermotolerans_CBS_6340
emb|SCU84815.1|_LADA_0D03972g1_1_Lachancea_dasiensis_CBS_10888
emb|SCU94801.1|_LANO_0E08196g1_1_Lachancea_nothofagi_CBS_11611
emb|SCV03028.1|_LAME_0H07140g1_1_Lachancea_meyersii_CBS_8951
dbj|BAP69967.1|_uncharacterized_transporter_YBR180W_Kluyveromyces_marxianus
emb|CAH00125.1|_KLLA0E24113p_Kluyveromyces_lactis
ref|XP_002616861.1|_hypothetical_protein_CLUG_02305_Clavispora_lusitaniae_ATCC_42720
emb|SGZ57748.1|_CIC11C00000002920_Candida_intermedia
gb|ABN64445.2| dityrosine transporter A(acid azole) Q(quinidine) Resistance Scheffersomyces stipitis CBS 6054
emb|CCE42744.1|_hypothetical_protein_CPAR2_203870_Candida_parapsilosis
emb|CCG23238.1|_membrane_transporter_Candida_orthopsilosis_Co_90-125
gb|AOW31198.1|_hypothetical_protein_CAALFM_CR04620CA_Candida_albicans_SC5314
emb|CAX40023.1| Major Facilitator Superfamily (MFS) transporter putative Candida dubliniensis CD36
ref|XP_002545732.1|_hypothetical_protein_CTRG_00513_Candida_tropicalis_MYA-3404
gb|ODV71122.1|_dityrosine_transporter_A_Q_resistance_Cyberlindnera_jadinii_NRRL_Y-1542
gb|ONH66714.1|_Dityrosine_transporter_1_Cyberlindnera_fabianii
emb|CCH45674.1|_putative_membrane_protein_Wickerhamomyces_ciferrii
emb|CCE72967.1|_Piso0_000575_Millerozyma_farinosa_CBS_7064
emb|CCE73528.1|_Piso0_000575_Millerozyma_farinosa_CBS_7064
tpe|CBF84174.1| TPA: conserved hypothetical protein Aspergillus nidulans FGSC A4
dbj|BAE62148.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|OBR11949.1|_Major_facilitator_superfamily_transporter_Colletotrichum_higginsianum_IMI_349063
gb|PIA93869.1|_Dityrosine_transporter_1_Cercospora_beticola
gb|KYK57101.1|_uncharacterized_protein_DCS_04108_Drechmeria_coniospora
gb|OAQ64343.1|_bicyclomycin_resistance_protein_Pochonia_chlamydosporia_170
gb|OAQ64579.1|_MFS_transporter_Pochonia_chlamydosporia_170
gb|AEO64729.1|_hypothetical_protein_THITE_2037799_Thielavia_terrestris_NRRL_8126
gb|EHA47111.1|_hypothetical_protein_MGG_04329_Magnaporthe_oryzae_70-15
gb|KUI69145.1|_putative_transporter_AQR1_Valsa_mali
tpe|CBF78716.1| TPA: conserved hypothetical protein Aspergillus nidulans FGSC A4
emb|CCT68953.1|_probable_dityrosine_transporter_Fusarium_fujikuroi_IMI_58289
gb|EWG49570.1|_hypothetical_protein_FVEG_09064_Fusarium_verticillioides_7600
gb|KNB10025.1|_hypothetical_protein_FOXG_10415_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|EKJ77857.1|_hypothetical_protein_FPSE_01950_Fusarium_pseudograminearum_CS3096
gb|ESU07896.1|_hypothetical_protein_FGSG_12141_Fusarium_graminearum_PH-1
gb|EYB31736.1|_hypothetical_protein_FG05_12141_Fusarium_graminearum
gb|OAQ70815.1|_general_substrate_transporter_Pochonia_chlamydosporia_170
gb|EAA28715.1|_MFS_multidrug_transporter_Neurospora_crassa_OR74A
gb|KUI68354.1|_Quinidine_resistance_protein_2_Valsa_mali
gb|OBR12995.1|_Major_facilitator_superfamily_transporter_Colletotrichum_higginsianum_IMI_349063
gb|ATZ56926.1|_hypothetical_protein_BCIN_14g01330_Botrytis_cinerea_B05.10
gb|KUI69228.1|_Quinidine_resistance_protein_1_Valsa_mali
gb|EAL93829.1|_MFS_efflux_transporter_putative_Aspergillus_fumigatus_Af293
gb|KZN84451.1|_Quinidine_resistance_protein_Penicillium_chrysogenum
ref|XP_001389836.1|_MFS_transporter_Aspergillus_niger_CBS_513.88
gb|AEO53729.1|_hypothetical_protein_MYCTH_2030361_partial_Thermothelomyces_thermophila_ATCC_42464
gb|AEO69113.1|_hypothetical_protein_THITE_2119140_Thielavia_terrestris_NRRL_8126
gb|ATZ54125.1|_hypothetical_protein_BCIN_10g01450_Botrytis_cinerea_B05.10
gb|PIA95814.1|_putative_transporter_AQR1_Cercospora_beticola
gb|EAL86476.1|_MFS_multidrug_transporter_putative_Aspergillus_fumigatus_Af293
gb|KZN87165.1|_Quinidine_resistance_protein_Penicillium_chrysogenum
ref|XP_001402538.3|_MFS_transporter_Aspergillus_niger_CBS_513.88
gb|PIA81336.1|_putative_transporter_AQR1_Cercospora_beticola
gb|EHA51602.1|_hypothetical_protein_MGG_07498_Magnaporthe_oryzae_70-15
gb|EWG55585.1|_hypothetical_protein_FVEG_13567_Fusarium_verticillioides_7600
gb|KNB16343.1|_hypothetical_protein_FOXG_14752_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|EAL92050.2|_MFS_multidrug_transporter_putative_Aspergillus_fumigatus_Af293
gb|OAQ64952.1|_MFS_multidrug_transporter_Pochonia_chlamydosporia_170
ref|XP_001402330.1|_MFS_transporter_Aspergillus_niger_CBS_513.88
dbj|BAE64922.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|EAL86152.1|_MFS_transporter_putative_Aspergillus_fumigatus_Af293
gb|KUI70522.1|_Quinidine_resistance_protein_2_Valsa_mali
emb|CCT66592.1|_probable_dityrosine_transporter_Fusarium_fujikuroi_IMI_58289
gb|EWG44747.1|_hypothetical_protein_FVEG_15715_Fusarium_verticillioides_7600
dbj|BAE60879.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|KZN90618.1|_Quinidine_resistance_protein_Penicillium_chrysogenum
tpe|CBF73556.1| TPA: MFS transporter putative (AFU orthologue AFUA 5G01520) Aspergillus nidulans FGSC A4
gb|KZN85700.1|_Quinidine_resistance_protein_Penicillium_chrysogenum


emb|SMY24646.1|_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
gb|EAL89684.1|_MFS_transporter_putative_Aspergillus_fumigatus_Af293
gb|KZN85687.1|_Quinidine_resistance_protein_Penicillium_chrysogenum
tpe|CBF81521.1|_TPA: MFS transporter putative (AFU orthologue AFUA 4G11060) Aspergillus nidulans FGSC A4
gb|APA13187.1|_hypothetical_protein_sscle_10g079570_Sclerotinia_sclerotiorum_1980_UF-70
gb|ATZ56018.1|_hypothetical_protein_BCIN_12g05570_Botrytis_cinerea_B05.10
gb|AEO67227.1|_hypothetical_protein_THITE_2116093_Thielavia_terrestris_NRRL_8126
gb|EAA33218.1|_MFS_transporter_Neurospora_crassa_OR74A
gb|KUI68642.1|_Quinidine_resistance_protein_1_Valsa_mali
gb|EHA47309.1|_hypothetical_protein_MGG_17679_Magnaporthe_oryzae_70-15
gb|KYK59331.1|_hypothetical_protein_DCS_00461_Drechmeria_coniospora
gb|OAQ69106.1|_chloramphenicol_resistance_protein_Pochonia_chlamydosporia_170
gb|EWG45623.1|_hypothetical_protein_FVEG_06346_Fusarium_verticillioides_7600
gb|KNB06989.1|_hypothetical_protein_FOXG_08328_Fusarium_oxysporum_f._sp._lycopersici_4287
emb|CCT64051.1|_related_to_dityrosine_transporter_Fusarium_fujikuroi_IMI_58289
gb|EYB29790.1|_hypothetical_protein_FG05_09737_Fusarium_graminearum
gb|EKJ78768.1|_hypothetical_protein_FPSE_01047_Fusarium_pseudograminearum_CS3096
gb|ESU16351.1|_hypothetical_protein_FGSG_09737_Fusarium_graminearum_PH-1
gb|AEO68573.1|_hypothetical_protein_THITE_2118141_Thielavia_terrestris_NRRL_8126
gb|OAQ72606.1|_florfenicol_exporter_Pochonia_chlamydosporia_170
gb|ADV23325.1|_Conserved_hypothetical_protein_Cryptococcus_gattii_WM276
gb|AUB26360.1|_hypothetical_protein_CKF44_002862_Cryptococcus_neoformans_var._grubii
tpe|CBF78942.1|_TPA: conserved hypothetical protein Aspergillus nidulans FGSC A4
emb|SMY21053.1|_unnamed_protein_product_Zymoseptoria_tritici_ST99CH_1A5
gb|ESU17316.1|_hypothetical_protein_FGSG_10579_Fusarium_graminearum_PH-1
gb|KYK62000.1|_hypothetical_protein_DCS_03145_Drechmeria_coniospora
gb|KNB16502.1|_hypothetical_protein_FOXG_14348_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|AEO71470.1|_hypothetical_protein_THITE_2123861_Thielavia_terrestris_NRRL_8126
gb|EAA35555.2|_hypothetical_protein_NCU00875_Neurospora_crassa_OR74A
emb|SHO78289.1|_Similar to S.cerevisiae protein QDR1 (Multidrug transporter of the major facilitator superfamily) Malassezia sympodialis ATCC 42132
emb|SHO78290.1|_Similar to S.cerevisiae protein QDR1 (Multidrug transporter of the major facilitator superfamily) Malassezia sympodialis ATCC 42132
gb|KIS70005.1|_hypothetical_protein_UMAG_11777_Ustilago_maydis_521
dbj|BAE55067.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|KZN84922.1|_MFS_antiporter_Penicillium_chrysogenum
dbj|BAE58840.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
ref|XP_001398400.1|_hypothetical_protein_ANI_1_478154_Aspergillus_niger_CBS_513.88
gb|APA08040.1|_hypothetical_protein_sscle_03g028100_Sclerotinia_sclerotiorum_1980_UF-70
gb|ATZ48294.1|_hypothetical_protein_BCIN_03g05230_Botrytis_cinerea_B05.10
gb|EAA29407.1|_hypothetical_protein_NCU09876_Neurospora_crassa_OR74A
gb|OBR12589.1|_major_facilitator_superfamily_transporter_Colletotrichum_higginsianum_IMI_349063
gb|OBR04265.1|_Major_facilitator_superfamily_transporter_Colletotrichum_higginsianum_IMI_349063
emb|CCT69236.1|_related_to_multidrug_resistance_protein_Fusarium_fujikuroi_IMI_58289
gb|KNB10648.1|_hypothetical_protein_FOXG_10798_Fusarium_oxysporum_f._sp._lycopersici_4287
gb|OBR06361.1|_Major_facilitator_superfamily_Colletotrichum_higginsianum_IMI_349063
gb|PAV19301.1|_MFS_general_substrate_transporter_Phellinus_noxius
gb|EWG50050.1|_hypothetical_protein_FVEG_09382_Fusarium_verticillioides_7600
gb|KNB10575.1|_hypothetical_protein_FOXG_10738_Fusarium_oxysporum_f._sp._lycopersici_4287
emb|CCT69168.1|_related_to_multidrug_transporter_Fusarium_fujikuroi_IMI_58289
gb|EKJ75897.1|_hypothetical_protein_FPSE_04077_Fusarium_pseudograminearum_CS3096
gb|ESU10669.1|_hypothetical_protein_FGSG_04799_Fusarium_graminearum_PH-1
gb|EYB27084.1|_hypothetical_protein_FG05_04799_Fusarium_graminearum
gb|OBR09619.1|_Major_facilitator_superfamily_transporter_Colletotrichum_higginsianum_IMI_349063
gb|AEO54531.1|_hypothetical_protein_MYCTH_2115336_Thermothelomyces_thermophila_ATCC_42464
gb|EAA32970.3|_MFS_multidrug_transporter_Neurospora_crassa_OR74A
ref|XP_001393344.2|_MFS_multidrug_transporter_Aspergillus_niger_CBS_513.88
dbj|BAE63488.1|_unnamed_protein_product_Aspergillus_oryzae_RIB40
gb|EAL86600.1|_MFS_multidrug_transporter_putative_Aspergillus_fumigatus_Af293
ref|XP_001395380.1|_MFS_multidrug_transporter_Aspergillus_niger_CBS_513.88
emb|SCW03889.1|_LAFE_0H01266g1_1_Lachancea_fermentati
emb|SCU77217.1|_LAMI_0A00122g1_1_Lachancea_mirantina
emb|SCU96571.1|_LADA_0H01618g1_1_Lachancea_dasiensis_CBS_10888
gb|EHA58476.1|_hypothetical_protein_MGG_08656_Magnaporthe_oryzae_70-15
gb|OBR06933.1|_Major_facilitator_superfamily_transporter_Colletotrichum_higginsianum_IMI_349063
gb|AUB25612.1|_hypothetical_protein_CKF44_001835_Cryptococcus_neoformans_var._grubii
gb|ADV22691.1|_Multidrug_transporter_putative_Qdr2p_Cryptococcus_gattii_WM276
gb|AUB25611.1|_hypothetical_protein_CKF44_001835_Cryptococcus_neoformans_var._grubii
gb|ANB14805.1|_Qdr2p_Sugiyamaella_lignohabitans
gb|ANB15909.1|_Aqr1p_Sugiyamaella_lignohabitans
XP_504962.1_YALI0F03751p_Yarrowia_lipolytica_CLIB122
emb|SGZ54572.1|_CIC11C00000002882_Candida_intermedia
ref|XP_002616737.1|_hypothetical_protein_CLUG_03978_Clavispora_lusitaniae_ATCC_42720
ref|XP_025341190.1|_hypothetical_protein_CXQ85_002035_Candida_haemulonis
ref|XP_028889973.1|_uncharacterized_protein_CJI97_002716_Candida_auris
emb|CAX39985.1|_multidrug_resistance_transporter_putative_Candida_dubliniensis_CD36
gb|AOW31159.1|_multidrug_transporter_Candida_albicans_SC5314
ref|XP_002545604.1|_hypothetical_protein_CTRG_00385_Candida_tropicalis_MYA-3404
ref|XP_002546723.1|_hypothetical_protein_CTRG_06201_Candida_tropicalis_MYA-3404
emb|CAX43052.1|_multidrug_transporter_of_the_major_facilitator_superfamily_putative_Candida_dubliniensis_CD36
gb|AOW28568.1|_Qdr2p_Candida_albicans_SC5314
emb|CCG23056.1|_hypothetical_protein_CORT_0D02080_Candida_orthopsilosis_Co_90-125
emb|CCE42570.1|_hypothetical_protein_CPAR2_202130_Candida_parapsilosis
emb|CCG23066.1|_hypothetical_protein_CORT_0D02180_Candida_orthopsilosis_Co_90-125
gb|ABN68473.1|_multidrug_resistance_transporter_Scheffersomyces_stipitis_CBS_6054
ref|XP_001486958.1|_hypothetical_protein_PGUG_00335_Meyerozyma_guilliermondii_ATCC_6260
ref|XP_002614701.1|_hypothetical_protein_CLUG_05479_Clavispora_lusitaniae_ATCC_42720
emb|SGZ51913.1|_CIC11C00000000482_Candida_intermedia
ref|XP_025342793.1|_hypothetical_protein_CXQ85_000847_Candida_haemulonis
ref|XP_028888559.1|_uncharacterized_protein_CJI97_005575_Candida_auris
emb|CAG89659.2|_DEHA2F21142p_Debaryomyces_hansenii_CBS767
emb|CCE78569.1|_Piso0_001196_Millerozyma_farinosa_CBS_7064
emb|CCE79155.1|_Piso0_001196_Millerozyma_farinosa_CBS_7064
emb|CCE86444.1|_Piso0_004934_Millerozyma_farinosa_CBS_7064
gb|ABN64705.1|_multidrug_resistance_transporter_Scheffersomyces_stipitis_CBS_6054
ref|XP_002545587.1|_predicted_protein_Candida_tropicalis_MYA-3404
gb|ANZ75790.1|_BA75_01933T0_Komagataella_pastoris
gb|AOA62894.1|_GQ67_01284T0_Komagataella_phaffii
ref|XP_029318996.1|_uncharacterized_protein_C5L36_0A01250_Pichia_kudriavzevii
ref|XP_029318997.1|_uncharacterized_protein_C5L36_0A01260_Pichia_kudriavzevii
gb|ESW96194.1|_transporter_AQR1_Ogataea_parapolymorpha_DL-1
emb|CCC71186.1|_hypothetical_protein_NCAS_0G02990_Naumovozyma_castellii_CBS_4309
emb|CCD24415.1|_hypothetical_protein_NDAI_0D01020_Naumovozyma_dairenensis_CBS_421
emb|CCF60161.1|_hypothetical_protein_KAFR_0J00930_Kazachstania_africana_CBS_2517
emb|CCK71976.1|_hypothetical_protein_KNAG_0I01910_Kazachstania_naganishii_CBS_8797
gb|KOG97092.1|_hypothetical_protein_DI49_4666_Saccharomyces_eubayanus
NP_014334.3_Aqr1p_Saccharomyces_cerevisiae_S288C
dbj|GAV56055.1|_hypothetical_protein_ZYGR_0AZ02270_Zygosaccharomyces_rouxii_NBRC_110957
gb|OXB47594.1|_hypothetical_protein_B1J92_J09944g_Candida_glabrata
emb|CCE91348.1|_hypothetical_protein_TDEL_0C04590_Torulaspora_delbrueckii
emb|CCH62193.1|_hypothetical_protein_TBLA_0G02550_Tetrapisispora_blattae_CBS_6284
emb|CCE63614.1|_hypothetical_protein_TPHA_0F01290_Tetrapisispora_phaffii_CBS_4417
dbj|BAP73872.1|_probable_transporter_AQR1_Kluyveromyces_marxianus
emb|CAG97938.1|_KLLA0F03311p_Kluyveromyces_lactis
emb|SCU84625.1|_LAMI_0C08240g1_1_Lachancea_mirantina
emb|CAR25002.1|_KLTH0G09372p_Lachancea_thermotolerans_CBS_6340
emb|SCW02002.1|_LAFE_0E11980g1_1_Lachancea_fermentati



S2. Growth curve of *S. cerevisiae* CEN.PK 113-5D strain expressing pAnCexA-TEF. Cells were grown in 2% glucose media, at 30°C. Time points where the uptake of radiolabeled citric acid (10 mM) was measured are highlighted in the blue square.

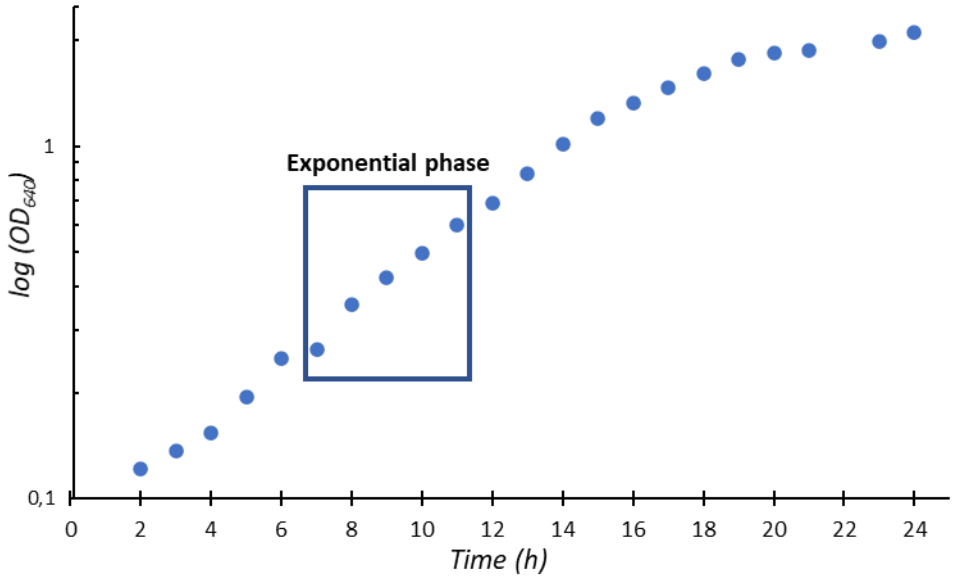


Table S1. AnCexA residues predicted to present strong intramolecular interactions with citrate by molecular docking studies. Grey initials IL/H stands for ionic ligation and hydrogen bond.

3D-Protein template	Citrate		
	(-1)	*(-2)*	*(-3)*
AnCexA	A68; S75; N76; P80; D84; S161; G189; R192(*IL/H*); Q196; P200; T207; I215; Y307; S311; T314; S315; T318; S329; C343; S431; T432	P80; D84; R192(*IL/H*); Q196; P200; T207; I215; Y307; W310; T314; T318; S329; C343; T432	S71; N76; R192(*IL/H*); Q196; T207; Y307; W310; T314; I215; T318; C343

Table S2. Average of the binding affinity values [kcal/mol] calculated with PyRx software for the docking of AnCexA with citrate.

3D-Protein template	Average of binding affinities (kcal/mol) at different binding sites			
	Citrate			
	S1	**S2**	**S3**	**S4**
AnCexA	-5,7	-5,83	-5,3	-4,6

CHAPTER VII

The Acetate Uptake Transporter Family motif "NPAPLGL(M/S)" is Essential for Substrate Uptake

Adapted from:

Ribas, D., Soares-Silva, I., Vieira, D., Sousa-Silva, M., Sá-Pessoa, J., Azevedo-Silva, J., Viegas, S., Arraiano, S., Diallinas, G., Paiva, S., Soares, P., Casal, M. (2019). The Acetate Uptake Transporter family motif "NPAPLGL(M/S)" is essential for substrate uptake and binding. *Fungal Genetics and Biology*, 122, 1-10. doi.org/10.1016/j.fgb.2018.10.001

[1] Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, Portugal

[2] Centre of Molecular and Environmental Biology (CBMA), Department of Biology, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal

Personal contribution: I collaborated in the heterologous expression of GPR1 and ATO2 (FUN34), the transport assays and epifluorescent microscopy, in addition to the analysis of data and manuscript writing.

Note: This chapter was previously reported in the Doctoral thesis of David Ribas, which is entitled "Characterization of carboxylate transporters: functional and structural studies for the improvement of cell factories" and is available in the "Repositorium" of the University of Minho.

CHAPTER VII

The acetate uptake transporter family motif "NPAPLGL(M/S)" is essential for substrate uptake

ABSTRACT

Organic acids are recognized as one of the most prevalent compounds in ecosystems, thus the transport and assimilation of these molecules represent an adaptive advantage for organisms. The AceTr family members are associated with the active transport of organic acids, namely acetate and succinate. The phylogenetic analysis shows this family is dispersed in the tree of life. However, in eukaryotes, it is almost limited to microbes, though reaching a prevalence close to 100% in fungi, with an essential role in spore development. Aiming at deepening the knowledge in this family, we studied the acetate permease AceP from *Methanosarcina acetivorans*, as the first functionally characterized archaeal member of this family. Furthermore, we demonstrate that the yeast Gpr1 from *Yarrowia lipolytica* is an acetate permease, whereas the Ady2 closest homologue in *Saccharomyces cerevisiae*, Fun34, has no role in acetate uptake. In this work, we describe the functional role of the AceTr conserved motif NPAPLGL(M/S). We further unveiled the role of the amino acid residues R122 and Q125 of SatP as essential for protein activity.

KEYWORDS: AceTr family phylogeny; Acetate transport; Succinate transport; Carboxylate transporters; Cell membrane

7.1 INTRODUCTION

Organic acids are one of the most prevalent and abundant organic compounds on earth's surface and they are vital intermediates for cell metabolism [1]. The cell homeostasis of every organism in an ecosystem is highly dependent on the transport of specific molecules across the cell membrane. The transport of organic acids is influenced by abiotic factors, such as pH and temperature. When the environmental pH is below the p*Ka* of the acid, the protonated form predominates and crosses the cell membrane by passive diffusion according to its diffusion coefficient. However, when the pH is above the acid p*Ka* the anionic form of the acid predominates, requiring a transporter protein to cross the membrane [2, 3].

The AceTr family was recently reassigned as The Acetate Uptake Transporter Family (TCDB 2.A.96) due to the functional characterization of several of its members as acetate transporters. AceTr proteins have six predicted transmembrane segments (TMS), sharing the conserved motif NPAPLGL(M/S) located at the beginning of the first TMS [3, 4]. The *Yarrowia lipolytica* Gpr1 protein was the first to be characterized. Its transcription was found to be regulated by the presence of acetic acid or ethanol [4], being involved in acetic acid sensitivity, cell and colony morphology, yeast-to-hyphae transition and cell lifespan [4, 5]. The AceTr member of *Saccharomyces cerevisiae*, Ady2, was initially described as critical for proper ascus formation in the presence of acetate [6]. The *ADY2* gene was later found to be strongly overexpressed in acetate growth conditions, and acetate uptake assays confirmed the role of the Ady2 protein in acetate uptake [7]. S*. cerevisiae* has two other AceTr homologues, Ato3 and Fun34 (Ato2), which were reported to be involved in ammonium export, along with Ady2 [8]. In the filamentous fungi *Aspergillus nidulans,* four AceTr homologues are described, AcpA, AcpB, AcpC, and AlcS, [9-11]. The AcpA protein was reported to be essential for the uptake and use of acetate as a sole carbon source and it is thought to contribute to spore maintenance and homeostasis [9, 11]. AcpB was found to be responsible for residual acetate transport in mycelia, whereas no function related to acetate transport was found for AcpC [11]. The AlcS protein is not involved in the transport of carboxylic acids [10,12]. In *Escherichia coli* only one AceTr homolog reported is the SatP protein. Unlike other AceTr members, it has specificity for both mono- and dicarboxylic acids being a succinate/acetate proton symporter, which is mostly active during exponential phase [13]. In archaea, the *aceP* gene from *Methanosarcina acetivorans* was found to be overexpressed in the presence of acetate, however, this high level of expression is only reached when acetate is used as a sole carbon source [14]. Recently two AceTr homologues, of the human pathogen *Cryptococcus neoformans*, referred as Ady2 and Ato2, were

reported to play a role in acetate transport, which was demonstrated to be critical for the survival of this species in stress and starvation conditions [15]. In this study, we carried out a phylogenetic analysis of AceTr family and further characterized two of its members. We heterologously expressed the *GPR1* from *Y. lipolytica* in *S. cerevisiae* and *aceP* from *M. acetivorans* in *E. coli*. We also analyzed the role of each residue of the conserved motif NPAPLGL(M/F) in substrate transport of both membrane proteins SatP and Ady2. Complementary molecular docking analyzes were used to uncover putative binding sites in the vicinity of this conserved motif.

7.2 MATERIALS AND METHODS

7.2.1 HOMOLOGY SEARCH

We downloaded over 10000 proteomes from NCBI Assembly platform as individual FASTA files. A BLAST search, with a cut-off e-value 10^{-5}, was performed on this database using three queries, a bacterial, an archaeal and an eukaryotic member: Satp from the organism *E. coli*, AceP from *M. acetivorans* and Ady2 from *S. cerevisiae*. To avoid redundancies, only sequences from a single genome of a given species were considered.

7.2.2 ALIGNMENT AND PHYLOGENETIC RECONSTRUCTION

Retrieved protein sequences were aligned with PROMALS3D [16], a multiple-alignment algorithm that incorporates secondary structure prediction. Sequences that were not aligning extensively across the conserved region of the alignment were further excluded from the phylogenetic analysis. A phylogenetic reconstruction was performed using Maximum Likelihood, more appropriate for the deeper divergences under analysis here, using MEGA6 [17] and the Jones-Taylor-Thornton (JTT) substitution model. Bootstrap was performed for 1000 repetitions. Obtained phylogenetic tree was displayed and edited in FigTree v.1.3.1. (http://tree.bio.ed.ac.uk/).

7.2.3 STRAINS, PLASMIDS AND GROWTH CONDITIONS

The strains and plasmids used in this work are listed in Tables 1 and 2 respectively. The *S. cerevisiae* strain W303-1A *jen1*Δ *ady2*Δ, lacking monocarboxylate uptake capacity, was used to express the *ADY2* alleles and the *GPR1* from *Y. lipolytica*. The cultures were maintained on slants of yeast extract (1%, w/v), peptone (1%, w/v), glucose (2%, w/v) and agar (2%, w/v) or yeast nitrogen base (Difco), 0.67%, w/v (YNB medium), supplemented with adequate requirements for prototrophic growth.

For growth tests, cells were grown as previously described [18]. The *E. coli* strain *actP*Δ *lldp*Δ *satP*Δ (assigned in this study as *E. coli* 3Δ) was used to express the *SatP* alleles and the *aceP* gene from *M. acetivorans*. Bacterial strains were grown as previously described [13].

Table 1. List of plasmids used in this study

Plasmid	Characteristics	Reference
pUC18	High-copy plasmid, constitutive expression	(Norrander *et al.* 1983) [19]
PSatP	pUC18 derivative; constitutive expression of *SatP*	(Sá-Pessoa et al, 2013) [13]
pSatP-N8A	pSatP with the substitution N8A in SatP	This study
pSatP-P9A	pSatP with the substitution P9A in SatP	This study
pSatP-A10T	pSatP with the substitution A10T in SatP	This study
pSatP-P11A	pSatP with the substitution P11A in SatP	This study
pSatP-L12A	pSatP with the substitution L12A in SatP	This study
pSatP-G13A	pSatP with the substitution G13A in SatP	This study
pSatP-L14A	pSatP with the substitution L14A in SatP	This study
pSatP-M15A	pSatP with the substitution M15A in SatP	This study
p416GPD	Glyceraldehyde-3-phosphate dehydrogenase (GPD) promoter	(Mumberg *et al.*, 1995) [20]
pDS1::GFP	p416 derivative, constitutive expression of JEN1::GFP	(Soares-Silva *et al.*, 2007) [18]
pAdy2::GFP	p416 derivative, constitutive expression of ADY2::GFP	This study
pAdy2::GFP-N89A	pAdy2::GFP with substitution N89A in Ady2	This study
pAdy2::GFP-P90A	pAdy2::GFP with substitution P90A in Ady2	This study
pAdy2::GFP-A91T	pAdy2::GFP with substitution A91T in Ady2	This study
pAdy2::GFP-P92A	pAdy2::GFP with substitution P92A in Ady2	This study
pAdy2::GFP-L93A	pAdy2::GFP with substitution L93A in Ady2	This study
pAdy2::GFP-G94A	pAdy2::GFP with substitution G94A in Ady2	This study
pAdy2::GFP-L95A	pAdy2::GFP with substitution L95A in Ady2	This study
pAdy2::GFP-S92A	pAdy2::GFP with substitution S92A in Ady2	This study
PAceP	pUC18 derivative; constitutive expression of *aceP*	This study
p416::Gpr1	p416GPD with constitutive expression of *GPR1*	This study
p416::Gpr1-GFP	p416GPD with constitutive expression of *GPR1::GFP*	This study

Table 2. List of strains used in this study

Strain	Genotype	Reference
E. coli MG1693	thyA715	(Arraiano *et al.* 1988) [21]
E. coli lldP	*MG1693 lldP (lldP::Cm^R)*	This study
E. coli 3Δ	*MG1693 yaaH actP lldP (ΔyaaH*	This study
***E. coli 3Δ* pUC18**	*E. coli* 3Δ transformed with pUC18	This study
***E. coli 3Δ* pSatP**	*E. coli* 3Δ transformed with pSatP	This study
***E. coli 3Δ* pSatP-N8A**	*E. coli* 3Δ transformed with pSatP-N8A	This study
***E. coli 3Δ* pSatP-P9A**	*E. coli* 3Δ transformed with pSatP-P9A	This study
***E. coli 3Δ* pSatP-A10T**	*E. coli* 3Δ transformed with pSatP-A10T	This study
***E. coli 3Δ* pSatP-P11A**	*E. coli* 3Δ transformed with pSatP-P11A	This study
***E. coli 3Δ* pSatP-L12A**	*E. coli* 3Δ transformed with pSatP-L12A	This study
***E. coli 3Δ* pSatP-G13A**	*E. coli* 3Δ transformed with pSatP-G13A	This study
***E. coli 3Δ* pSatP-L14A**	*E. coli* 3Δ transformed with pSatP-L14A	This study
***E. coli 3Δ* pSatP-M15A**	*E. coli* 3Δ transformed with pSatP-M15A	This study
E. coli 3Δ pAceP	*E. coli* 3Δ transformed with pAceP	This study
***S. cerevisiae* W303-1A**	MATα *ade2 leu2 his3 trp1 ura3*	(Thomas and Rothstein 1989) [22]
S. cerevisiae jen1Δ ady2Δ	W303-1A; *JEN1::KanMX4*	(Soares-Silva *et al.*, 2007)
S. cerevisiae jen1Δ ady2Δ	*jen1Δ ady2Δ* transformed with	(Soares-Silva *et al.*, 2007)
***S. cerevisiae jen1Δ ady2Δ* pAdy2**	*jen1Δ ady2Δ* transformed with	This study
S. cerevisiae jen1Δ ady2Δ	*jen1Δ ady2Δ transformed with*	This study
S. cerevisiae jen1Δ ady2Δ	*jen1Δ ady2Δ transformed with*	This study
S. cerevisiae jen1Δ ady2Δ	*jen1Δ ady2Δ transformed with*	This study
S. cerevisiae jen1Δ ady2Δ	*jen1Δ ady2Δ* transformed with	This study
S. cerevisiae jen1Δ ady2Δ	*jen1Δ ady2Δ* transformed with	This study
S. cerevisiae jen1Δ ady2Δ	*jen1Δ ady2Δ* transformed with	This study
S. cerevisiae jen1Δ ady2Δ	*jen1Δ ady2Δ* transformed with	This study
S. cerevisiae jen1Δ ady2Δ	*jen1Δ ady2Δ* transformed with	This study
***Yarrowia lipolytica* PYCC 4811**	Wild-type	Collection
***S. cerevisiae* S288c**	Wild-type	Collection
***S. cerevisiae jen1Δ ady2Δ* pGpr1**	*jen1Δ ady2Δ* transformed with	This study
***S. cerevisiae jen1Δ ady2Δ* pGpr1-**	*jen1Δ ady2Δ* transformed with	This study

7.2.4 CONSTRUCTION OF *E. COLI ACTPΔ LLDPΔ SATPΔ*

The *lldP* null-mutant was constructed using the primer pairs dlldP1/dlldP2 and following the λ-red recombinase method [23] with a few modifications, as described previously [13, 24]. The chloramphenicol-resistance cassette of plasmid pKD3 replaces nucleotides +50 to +1587 of the *lldP* gene. The gene deletion was verified by colony PCR using the primer pair P1lldP/P2lldP and the chromosomal mutation subsequently transferred to a fresh genetic background (MG1693 strain) by P1 phage transduction, to generate *E. coli* MG1693 *lldP* (Δ*lldP*::Cm^R) strain. To construct the triple *yaaH/ actP/ lldP* mutant, the antibiotic resistance genes of the double *yaaH/ actP* mutant (BBC234) (Sá-Pessoa et al., 2013) were eliminated using a helper plasmid encoding the FLP recombinase (pCP20) and following the procedures previously described [23]. The loss of the kanamycin and chloramphenicol resistance was confirmed, and the strain used as receptor, in the P1 phage transduction of *lldP* mutation from the single *lldP* null-mutant. The chloramphenicol resistance was used for selection of the *lldP* mutation and the presence of the three, *yaaH*, *actP* and *lldP*, gene deletions were confirmed by PCR using specific primers: P1yaaH/P2yaaH for *yaaH*; P1ActP/P4actP for *actP* and P1lldP/P2lldP for *lldP* (primers listed in Table 3).

7.2.5 CONSTRUCTION OF ADY2 AND SATP MUTANTS

Site-directed mutagenesis was performed as previously described [18]. For *GFP* fusion proteins with the *ADY2*, GAP repair technique was performed as previously described [18], using the primers listed in Table 3.

7.2.6 HETEROLOGOUS EXPRESSION OF *GPR1*

The pGPR1 plasmid was constructed by restriction enzyme cloning and the pGPR1::GFP plasmid through GAP repair technique [25]. The GPR1 gene was amplified by PCR, from genomic DNA extracted from *Yarrowia lipolytica* PYCC 4811 [26], using the primers Gpr1_FWD and Gpr1_REV (Table 3). The amplified GPR1 was digested using *Hind*III and *Sal*I enzymes (*ThermoFisher Scientific*; USA). The final product was inserted into the vector p416GPD, previously digested with the same restriction enzymes using the same conditions. The resultant plasmid pGPR1 was cloned into *S. cerevisiae* W303-1A *jen1Δ ady2Δ* strain. To obtain the plasmid pGPR1-GFP, the GFP gene was inserted as described by Bessa et al., 2012, using the primers Gpr1-gfp_FWD and the Gpr1- gfp_REV (Table 3).

Table 3. Oligonucleotides used for strain construction, cloning, expression, site-directed mutagenesis and GFP tagging.

Name	Sequence
N89A_FWD	GTGCACAAATTTGCTGCTCCTGCGCCCTTAGG
N89A_REV	CCTAAGGGCGCAGGAGCAGCAAATTTGTGCAC
P90A_FWD	CACAAATTTGCTAATGCTGCGCCCTTAGGTC
P90A_REV	GACCTAAGGGCGCAGCATTAGCAAATTTGTG
A91T_FWD	CAAATTTGCTAATCCTACGCCCTTAGGTCTTTC
A91T_REV	GAAAGACCTAAGGGCGTAGGATTAGCAAATTTG
P92A_FWD	GCTAATCCTGCGGCCTTAGGTCTTTCAGCC
P92A_REV	GGCTGAAAGACCTAAGGCCGCAGGATTAGC
L93A_FWD	GCTAATCCTGCGCCCGCAGGTCTTTCAGCCTTC
L93A_REV	GAAGGCTGAAAGACCTGCGGGCGCAGGATTAGC
G94A_FWD	GCTAATCCTGCGCCCTTAGCTCTTTCAGCCTTCGCG
G94A_REV	CGCGAAGGCTGAAAGAGCTAAGGGCGCAGGATTAGC
L95A_FWD	CCTGCGCCCTTAGGTGCTTCAGCCTTCGCGTTG
L95A_REV	CAACGCGAAGGCTGAAGCACCTAAGGGCGCAGG
N8A_FWD	CAACACTAAGTTGGCTGCTCCGGCACCGCTGGGC
N8A_REV	GCCCAGCGGTGCCGGAGCAGCCAACTTAGTGTTG
P9A_FWD	CACTAAGTTGGCTAATGCGGCACCGCTGGGCCTG
P9A_REV	CAGGCCCAGCGGTGCCGCATTAGCCAACTTAGTG
A10T_FWD	CTAAGTTGGCTAATCCGACACCGCTGGGCCTGATGG
A10T_REV	CCATCAGGCCCAGCGGTGTCGGATTAGCCAACTTAG
P11A_FWD	GTTGGCTAATCCGGCAGCGCTGGGCCTGATGGGC
P11A_REV	GCCCATCAGGCCCAGCGCTGCCGGATTAGCCAAC
L12A_FWD	GCTAATCCGGCACCGGCGGGCCTGATGGGCTTC
L12A_REV	GAAGCCCATCAGGCCCGCCGGTGCCGGATTAGC
G13A_FWD	CTAATCCGGCACCGCTGGCCCTGATGGGCTTCGGC
G13A_REV	GCCGAAGCCCATCAGGGCCAGCGGTGCCGGATTAG
L14A_FWD	CCGGCACCGCTGGGCGCGATGGGCTTCGGCATG
L14A_REV	CATGCCGAAGCCCATCGCGCCCAGCGGTGCCGG
M15A_FWD	GGCACCGCTGGGCCTGGCGGGCTTCGGCATGACC
M15A_FREV	GGTCATGCCGAAGCCCGCCAGGCCCAGCGGTGCC
ady2_gfp_fwd	CCCATTACCATCTACTGAAAGGGTAATCTTTAGTAAAGGAGAAGAAC
ady2_gfp_rev	CATAACTAATTACATGACTCGAGCTATTTGTATAGTTCATCCATGCC
dlldP1	GTTAAGACATAAGCCTGAAGCGTGGTGATCACGCCCACTATACAGGTGAAGAGTGTAGGCTGGAGCTGCTTC
dlldP2	ATGAATCTCTGGCAACAAAACTACGATCCCGCCGGGAATATCTGGCTTTCGGTCCATATGAATATCCTCCTTAG
P1lldP	ACGATGTGCGTGGACTCCAG
P2lldP	ATCAATCAGCGCCCGCAC
P1yaaH	ATGCCGCGCCCTGAAAACACTAC
P2yaaH	AGTGCAAGACGCGACGTTAGCGAAT
P1actP	TCTACATCTGGCGGGCGAAC
P4actP	ACAGAGTGGTTATCGTTAATCAG
Gpr1_FWD	GGCAAGCTTATGAACACCGAAATCCCC
Gpr1_REV	GCGGTCGACTTAGTCCTTCTTGACGAA
Gpr1-gfp_FWD	TTTCCTCAAAGAGATTAAATACTGCTACTGAAAATATGAACACCGAAATCCCCGA
Gpr1-gfp_REV	ACTCGAGGTCGACGGTATCGATAAGCTTTACTATTTGTATAGTTCATCCA

7.2.7 Cloning strategy of *aceP* gene

The *aceP* gene was chemically synthesized with an optimal codon usage for expression in *E. coli* since no expression was obtained with the original gene cloned into the pUC18 vector. The synthetic *aceP* gene was codon optimized by the OptimumGene software tool (GenScript, Piscataway, NJ, USA Inc.). The synthetic version of *aceP* flanked by *KpnI* and *NdeI* restriction sites was cloned in the puc18 vector, originating the pAceP vector.

7.2.8 Transport assays

Measurement of transport activity in yeast and bacterial strains was performed as previously described by Soares-Silva et al. (2007) and Sá-Pessoa et al. (2013), respectively. The radiolabeled substrates used were the following: [1-^{14}C] acetic acid, sodium salt (GE Healthcare, London, UK) and [2,3-^{14}C] succinic acid (Moravek Biochemicals, California, USA).

7.2.9 Epifluorescent microscopy

Microscopy analysis was performed as described in Soares-Silva et al. (2007) on a Leica DM5000B epifluorescent microscope with a Leica DFC 350FX R2 digital camera using the LAS AF V1.4.1 software.

7.2.10 Structural modeling and molecular docking simulations

To obtain the predicted 3D structure of SatP transporter, the amino residue sequence was threaded through The Robetta server (http://robetta.bakerlab.org) [27]. The molecular docking of substrates with SatP transporter was performed as previously described [28].

7.3 Results

7.3.1 Phylogenetic analysis of AceTr family

A total of 818 hits were obtained for the BLAST search in the NCBI's Assembly database, using complete genomes only, with the 3 proteins of the AceTr family: Ady2, SatP, *and* AceP. Homologues were present across a wide range of organisms: most of them were detected in bacteria, archaea, and fungi. Hypothetical AceTr homologues were also detected in the genus *Leishmania* of eukaryotic euglenids, usually as multiple copies per organism, but also in three Viridiplantae of the family Bathycoccaceae and the Rhodophyta *Cyanidioschyzon merolae*. In order to maximize phylogenetic accuracy 12 sequences

that lacked large conserved regions were excluded from a final dataset of 806 sequences in the phylogenetic tree. The phylogenetic tree is displayed in Figure 1.

Relevant homologues to the present study are indicated. For a clearer description of the tree, some sections were labelled as E1, E2, and P1 to P4. The Phylogenetic tree suggests a basal split between eukaryotic and prokaryotic organisms. The bootstrap value is low as expected for such a deep divergence in a protein sequence however, this deep divergence suggests a reliable reconstruction. Eukaryotes and prokaryotes form almost two monophyletic clades with an exception from each subclade. The exception in the eukaryotic clade is the protein of bacteria *Gordonia polyisoprenivorans*. The relatively long branch of the tree and the undefined position in relation to the neighbouring eukaryotic counterpartssuggests that the inclusion of the bacteria *G. polyisoprenivorans* might be dubious. The exception in prokaryotic clade is a clade containing one set of homologues from the pathogenic trypanosome *Leishmania* genus. The clade is quite deeply embroiled in various prokaryotic sub-branches suggesting that it likely represents a case of horizontal gene transfer.

The prokaryotic subclade shows bacterial and archaeal homologues intertwined within the tree. The tree splits deeply into two branches, labelled P1-3 and P4 in Figure 1. The less common of these two branches (P4) contains mostly homologues present in the bacterial phylum Firmicutes and there is no functional characterization described on any of the members of this clade. The second clade (P1-3) also contains another split between a clade containing a high percentage of homologues present in Actinobacteria (P3), including a monophyletic branch from this group, a monophyletic archaea sub-branch and homologues from various other bacteria (including from Phyla Firmicutes, Proteobacteria and others), and a second more frequent clade (P1-2) where bacterial homologues from phylum Proteobacteria are the most common (forming the totality of minor clade P2). The large subclade P1 also contains other bacteria, several archaea and the mentioned-above subclade of eukaryotic *Leishmania* genus. The P1 subclade contains the SatP from bacteria *E. coli* [13] and AceP (this study) from archaea *M. acetivorans*, two transporters functionally characterized as acetate transporters as well as two other homologues for the latter organism not functionally characterized (MA_0103 and MA_4393). It is important to note that while prokaryotic homologues dominate the tree, only 15% of the genomes in the database contained at least one member of the AceTr family (14% in bacteria and 31% in archaea), reflecting mostly the large numbers of sequenced prokaryotic genomes.

The eukaryotic branch has a deep divergence between homologues present in fungi and other eukaryotes. This latter clade only contains 9 representative sequences, four from *Leishmania sp., Thalassiosira pseudonana* (a marine diatom), *Cyanidioschyzon merolae* (red algae) photoautotrophic organism that

inhabits sulfate-rich hot springs and three in green algae [29]. The presence of this deep clade suggest a probable more ancient presence in a wider range of eukaryotes, although the AceTr family was only found at 10% of the genomes other than fungi analyzed, a value that is likely overestimated due to the substantial number of *Leishmania* genus genomes that contain homologues. The main eukaryotic clade contains homologues present in fungi. Homologues belonging to the AceTr family were strikingly detected in 97% of the genomes analyzed and the fungi without AceTr members belonged to the *Encephalitozoon* genus. This clade splits also into two clades. The minor one (E2) contains mostly homologues present in filamentous ascomycetes, including two homologues present in *Aspergillus nidulans*.

A second one (E1) includes homologues from a wider taxonomic range of fungi, including Basidiomycota and a few filamentous Ascomycetes, namely experimentally tested homologues in *Aspergillus nidulans AcpA, AcpB, and AcpC*. The AcpC homologue in *Aspergillus nidulans* that is not associated with an acetate transport activity [11], is located in an earlier branching of this clade together with the homologues from *Cryptococcus neoformans* involved in acetate transport [15]. In the remaining clade, most sequences were mainly detected in Saccharomycetes, including a large sub-clade formed by Saccharomycetales only, although one subclade includes the AcpA and AcpB and the homologue in *Schizosaccharomyces pombe*. The large clade containing only Saccharomycetales has the three homologues of *S. cerevisiae*. Ady2 and Ato2 are very similar phylogenetically with Ato3, suggesting a duplication and evolution of the gene dating to an earlier divergence within the Saccharomycetales evolution. This duplication probably occurred before the whole genome duplication event in yeast, since the yeasts *Kluyveromyces marxianus* and *Kluyveromyces lactis* have homologues in a cluster together with both Ady2/Ato2 and Ato3 from *S. cerevisiae*. This major clade of Saccharomycetales also contains the homologues of *Yarrowia lipolytica.* Apart from the homologue Gpr1, experimentally tested here, the organism contains a set of homologues (five) forming a cluster (probably the result of recent duplications of the gene) whose function is yet undetermined.

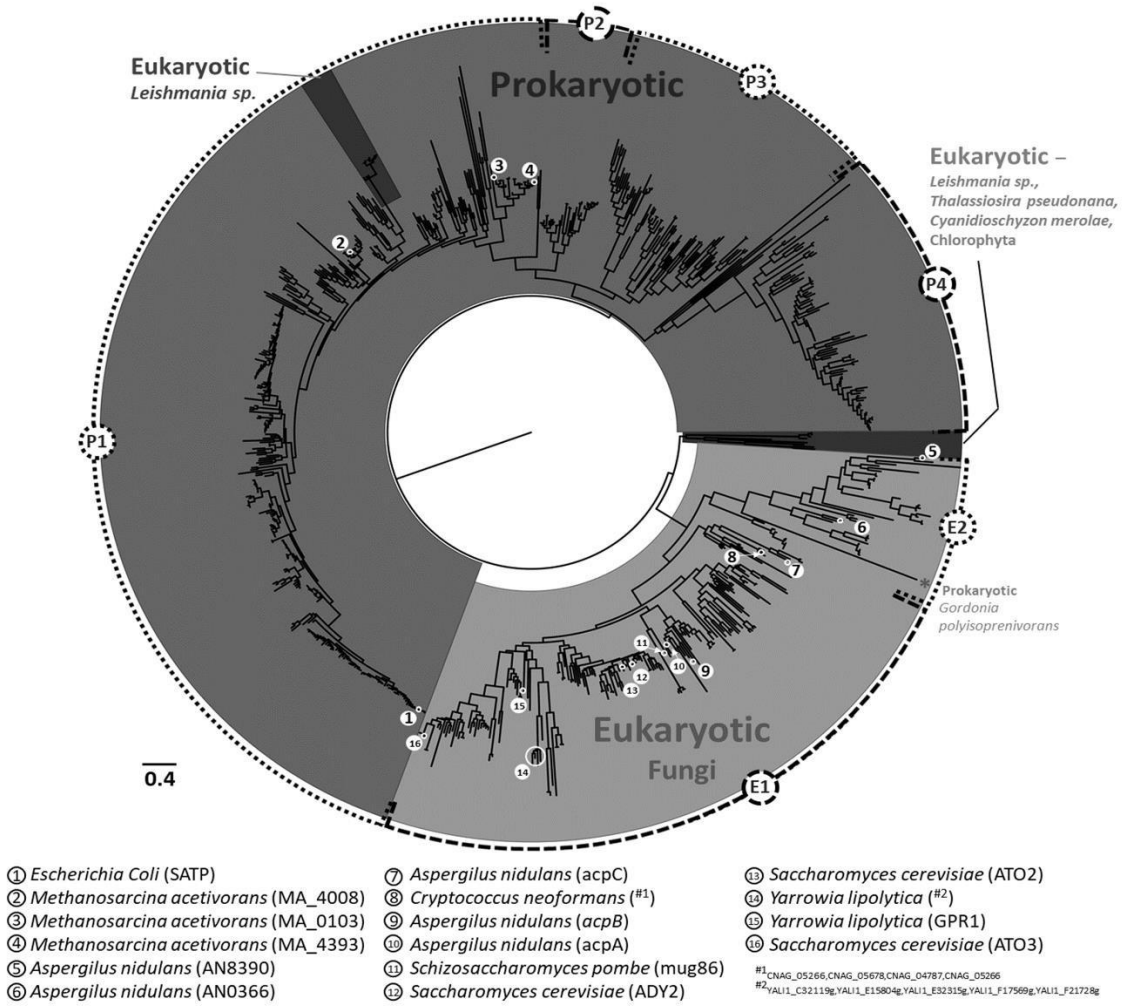
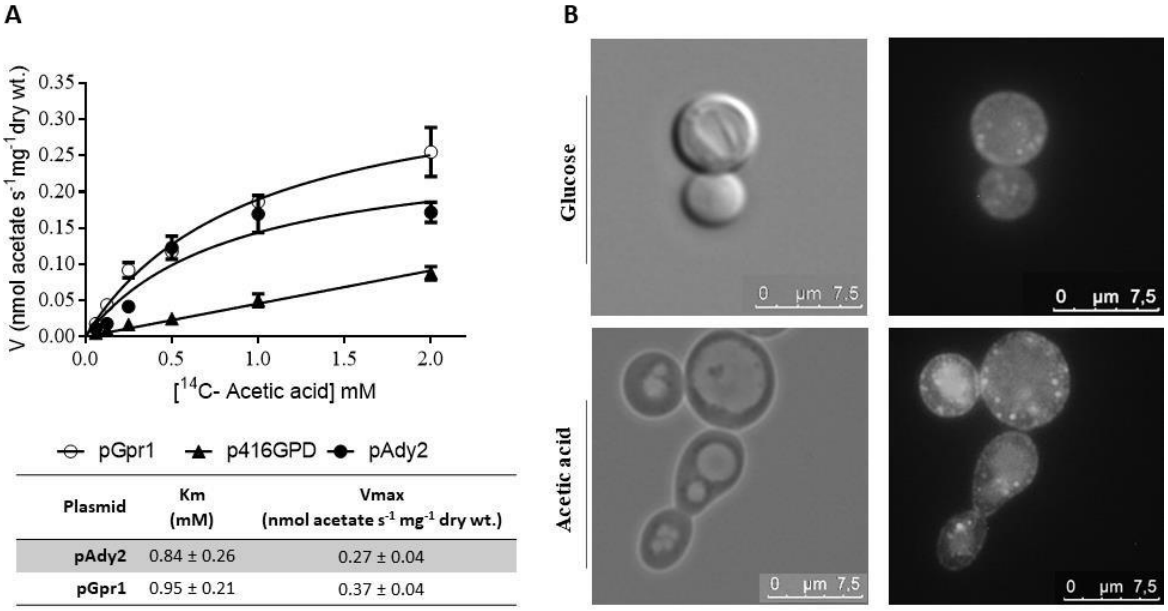


Figure 1. Maximum likelihood phylogenetic tree of AceTr family (TCDB 2.A.96) present in eukaryotic and prokaryotic genomes. Branch lengths are proportional to sequence divergence. Major taxonomic groups are indicated in different shades of grey. Homologues relevant for the discussion through the manuscript are highlighted. Groups indicated as E1, E2, P1, P2, P3, and P4 were created to facilitate the following of the tree description in the main text and are not meant to provide any type of classification.

7.3.2 GPR1 MEDIATES ACETATE TRANSPORT IN *S. CEREVISIAE*

The functionality of the Gpr1 membrane protein from *Y. lipolytica* was accessed through heterologous expression in a *S. cerevisiae* W303-1A *jen1*Δ *ady2*Δstrain, which does not have activity for plasma membrane carboxylate transport under the conditions tested [18, 28, 30, 31]. The control strains and *S. cerevisiae* strain expressing *GPR1* were grown in minimal medium YNB-Glu and then transferred to minimal media containing 0.5% acetic acid during 6 h. The expression of Gpr1 was able to restore the acetate-mediated transport in *S. cerevisiae jen1*Δ *ady2*Δ strain with a K_m of 0.95 ± 0.21 mM and V_{max} of 0.37 ± 0.04 nmol acetate $s^{-1}mg^{-1}$dry weight, at pH6.0 (Figure 2A). A GFP-tagged version of Gpr1p was analyzed by fluorescence microscopy revealing that the fusion protein was localized at the plasma membrane (Figure 2B).



Plasmid	Km (mM)	Vmax (nmol acetate s^{-1} mg^{-1} dry wt.)
pAdy2	0.84 ± 0.26	0.27 ± 0.04
pGpr1	0.95 ± 0.21	0.37 ± 0.04

Figure 2. A) Initial uptake rates of radiolabeled ^{14}C-acetic acid at different concentrations by *S. cerevisiae* W303-1A *jen1*Δ *ady2*Δ cells heterologously expressing p*Gpr1*, p*Ady2* and p416GPD as a control at pH 6.0, 30°C and respective kinetic parameters. Cells were cultivated in glucose until exponential growth phase, washed and transferred to YNB supplemented with acetic acid (0.5%) during 6 hours. B) Epifluorescence and contrast phase microscopy of a GFP-tagged version of *S. cerevisiae* W303-1A *jen1*Δ *ady2*Δ cells expressing *GPR1*, before and after 4h of exposure of cells to 0.5% acetic acid, at pH6, 30°C. Scale bars = 7.5 µm.

7.3.3 THE *ACEP* GENE CODES FOR A HIGHLY SPECIFIC ACETATE-PROTON SYMPORTER

Methanosarcina acetivorans uses acetate as a source of carbon and as a source of energy by breaking down acetate to produce carbon dioxide and methane. It was previously shown that *aceP* gene from *M. acetirovans* is overexpressed in the presence of acetate (Rohlin and Gunsalus, 2010).

The *M. acetivorans* species has an optimal growth temperature between 35 ºC and 45 ºC and optimal pH between 6.5 and 7 [32], close to the model organism *E. coli,* which led us to attempt the heterologous expression of *aceP* in this last organism. In addition, the tools for genetic manipulation of *M. acetivorans* are scarce, and culture conditions are difficult to achieve. The heterologous expression of *aceP* was performed in the *E. coli 3*Δ strain, deleted for three monocarboxylate transporters. An optimized version of *aceP* was created using the GeneScript codon usage adaptation algorithm. The vector pAceP was transformed in the *E. coli 3*Δ strain, and the acetate uptake measured after growth in the minimal medium during exponential phase. The *E. coli 3*Δ pAceP was able to recover the mediated transport of acetate, with a Michaelis-Menten kinetics, displaying a K_m of 0.49 ± 0.07 mM of acetic acid and a V_{max} of 46.4 ± 2.9 nmol min^{-1} mg^{-1} protein (Figure 3A).

To further characterize the transport energetics of the *E. coli 3*Δ pAceP strain, the acetate uptake was accessed in the presence of several relevant inhibitors. The protonophore CCCP (carbonyl cyanide *m*-chlorophenylhydrazone), which collapses the proton motive force, lowered acetate transport to almost negligible values at pH 6.0 in cells expressing AceP (Figure 3B). The potassium ionophore valinomycin and the sodium ionophore monensin, which disrupt the membrane electrical potential ($\Delta\psi$), had no significant effect on acetate uptake. These data suggest that acetate uptake occurs by a proton symporter mechanism, similar to other AceTr members.

The measurement of labelled acetate uptake in the presence of non-labelled acids allowed the identification of potential inhibitors of acetic acid uptake, being indicative of an interaction of these acids with the transporter protein. None of the carboxylic acids tested inhibited the acetate uptake in *E. coli 3*Δ cells expressing *aceP* (Figure 3C). Thus, this acetate transporter is not able to bind the carboxylic acids here tested, namely the monocarboxylates lactate, pyruvate and formate, and the dicarboxylates malate and succinate.

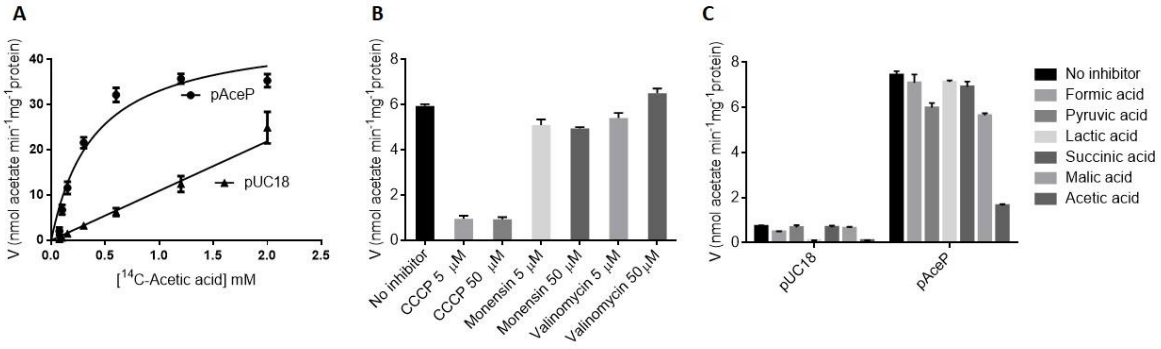


Figure 3. Acetic acid kinetics and transport energetics of *E. coli 3*Δ heterologously expressing pAceP. A) Initial uptake rates of radiolabelled ^{14}C-acetic acid at different concentration by *E. coli 3*Δ heterologously expressing pAceP or pUC18, at pH 6.0, 30◦C. Cells were cultivated in glucose until exponential growth phase. B) Transport energetics: effect of pH and of CCCP, valinomycin, and monensin on the uptake of 0.1 mM ^{14}C-acetic acid in cells of *E. coli 3*Δ pAceP. Cells were pre-incubated with the compounds mentioned, at the concentration indicated and pH 6.0 for 1 min before adding the radiolabelled substrate. C) transport specificity: the uptake of 0.1 mM ^{14}C-acetic acid in in *E. coli 3*Δ heterologously expressing pAceP or pUC18 measured in the absence and presence of non-labelled formic, pyruvic, lactic, succinic and malic acid (10 mM), pH 6.0 and 30◦C. Cells were pre-incubated with the compounds mentioned, at the concentration indicated and pH 6.0 for 1 min before adding the radiolabelled substrate. Each data point represents the mean± S.D. for three independent experiments.

7.3.4 The conserved motif NPAPLGL(M/S) plays a role in substrate uptake

A multiple alignment of several AceTr family members described in the literature as acetate transporters, as well as the AceP protein above characterized, was performed. Among the conserved regions identified, the motif ^{89}NPAPLGL(M/S)^{96} (numbers refer to Ady2 sequence) exhibited strong conservation in AceTr family members. Most of the conserved residues are polar and hydrophobic and thus have the potential to interact with charged substrates, e.g. acetate, succinate. The Ady2 topology prediction using the PROTTER web server suggests that this motif is found at the beginning of the first transmembrane segment (TMS I) in the transition between the cytoplasm and the plasma membrane (data not shown). A similar output was also obtained for the topology prediction of SatP transporter. Site-directed mutagenesis was performed in the conserved motif NPAPLGL(M/S) of Ady2 and SatP, replacing each conserved residue by an alanine residue, excluding the alanine residues 89 (Ady2) and 10 (SatP), which were replaced each by a tyrosine residue. All the mutants from both Ady2 or SatP presented an abolished acetate transport profile (Figure 4 A, B). Since SatP is known to mediate the transport of

succinate, we also measured succinate uptake in the SatP mutant alleles (Figure 4B). Similar to acetate uptake, the SatP mutant alleles displayed a succinate uptake capacity significantly reduced when compared to the wild-type allele. The Ady2 mutant alleles were also grown in YNB acetic acid (0.5%) as sole carbon and energy source. All Ady2 mutant alleles displayed poor growth when compared with the native allele.

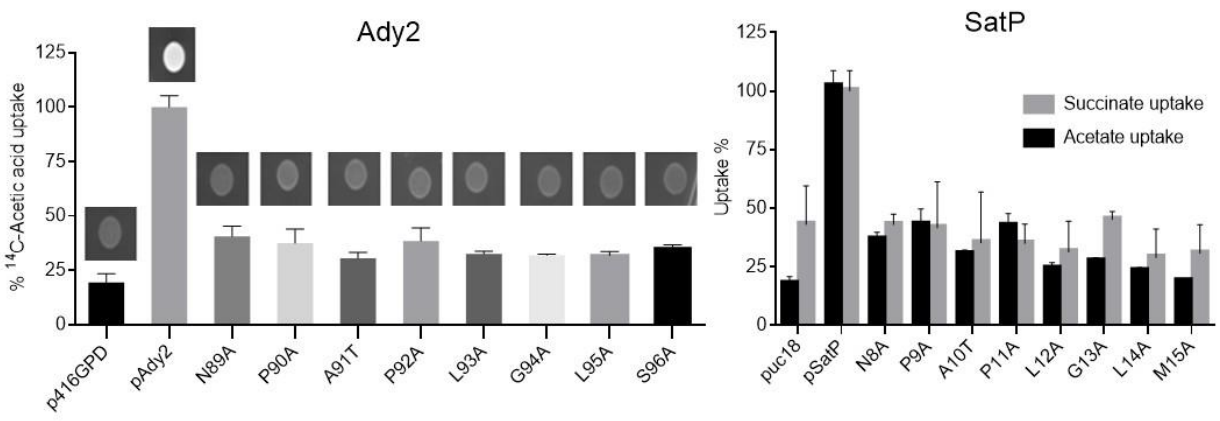


Figure 4. A) The percentage of 0.1 mM ^{14}C-acetic acid uptake, at pH 6.0, considering the velocity of transport found for the *S. cerevisiae ady2*Δ *jen1*Δ strains expressing pAdy2 as 100%. Cells were collected at mid-exponential growth phase from YNB glucose 1 % (w/v), Cells were cultivated in glucose until mid- exponential growth phase, washed and transferred during 6 h to YNB supplement with the acetic acid (0.5%) and solid YNB acetic acid (0.5 %) for growth tests shown in the images above each bar of Ady2 alleles in p416GPD. B) The percentage of 0.5 mM ^{14}C-acetic acid uptake, at pH 6.0, considering the velocity of transport found for the *E. coli 3*Δ strains expressing pSatP as 100%. Cells were collected at mid-exponential growth phase from minimal medium with glucose 1% (w/v). The data shown are mean values of at least three independent experiments and the error bars represent the standard deviation.

7.3.5 THE MOTIF NPAPLGL(M/S) IS NOT CRITICAL FOR MEMBRANE TARGETING OF ADY2

Single mutations in the conserved motif NPAPLGL(M/S) residues studied led to lack of Ady2 mediated acetate transport activity. In order to formally distinguish whether this is due to an incorrect protein location or lack of transporter function, we used mutant alleles expressing an Ady2::GFP chimeric transporter. The Ady2::GFP localization was determined by fluorescence microscopy. In all cases Ady2::GFP mutant alleles were localized, similar to the wild-type version, at the plasma membrane level after 6h induction in YNB acetic acid as it was found for the wild- type (Figure 5). These results strongly suggest that in these mutants presenting low or no transport capacity this is not a consequence of stability or trafficking, but rather due to an effect on the mechanism of substrate transport *per se*.

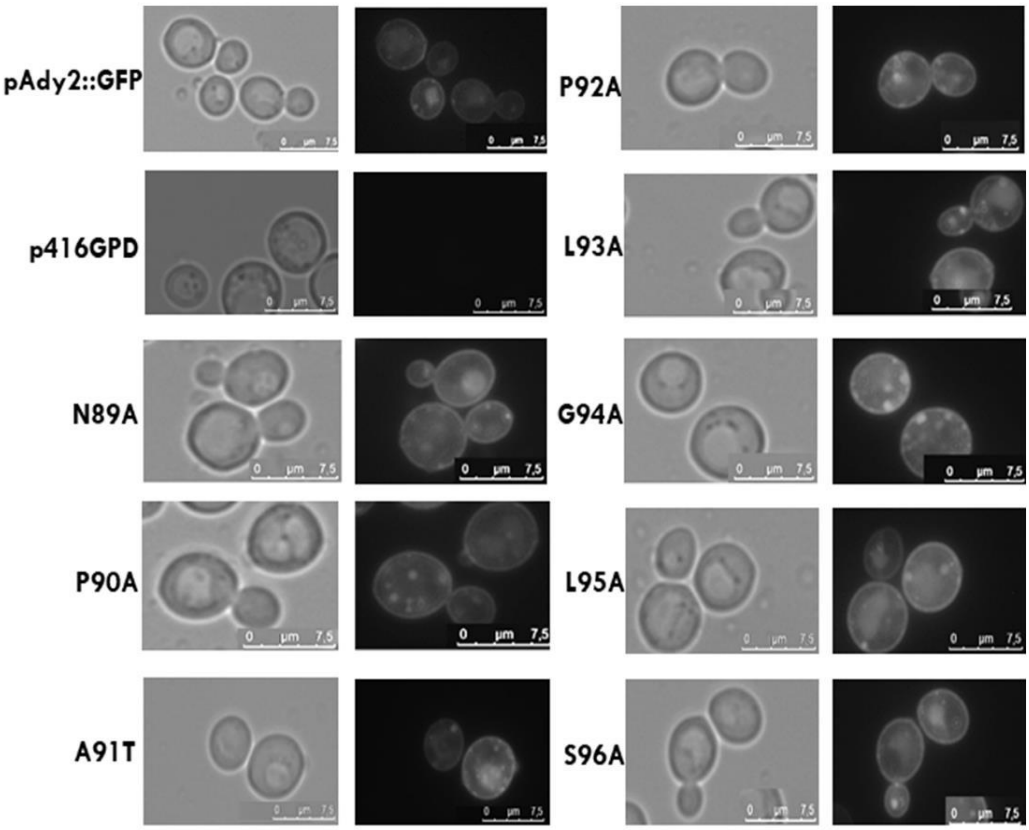


Figure 5. **Epifluorescence microscopy localization of Ady2::GFP expressed in *S. cerevisiae ady2Δ jen1Δ*.** Cells were collected at mid-exponential growth phase from YNB glucose 1% (w/v), washed and transferred during 6 h to YNB, supplement with the acetic acid (0.5%), and observed by epifluorescence microscopy.

7.3.6 The conserved motif NPAPLGL(M/S) is involved in the formation of a putative substrate binding site

To further understand the specific role of each residue in protein activity and specificity, molecular docking analysis was performed to uncover putative interactions between the transporter and substrate ligand molecules at the molecular level (Figure 6). Only the 3D predicted structure from SatP displayed a level of confidence required for further molecular docking analysis, and thus docking analysis was not possible with the Ady2. Molecular docking results suggest that the residues of the conserved motif ^{8}NPAPLGL(M/S)^{15} of SatP may be involved in the formation of a putative binding site (Figure 6). It is noticeable that acetate and succinate molecules share a similar spatial location, in which the carboxyl group of acetate is superimposed with one of the carboxyl groups of succinate. The residues A10, P11, L12, G13, L14, and M15 form a hydrophobic binding pocket for substrates, quite close to the cytoplasmic pore entrance. In addition, L12, G13, L14 and M15 residues interact covalently with the carboxyl group of both acetate and succinate by hydrogen bonds.

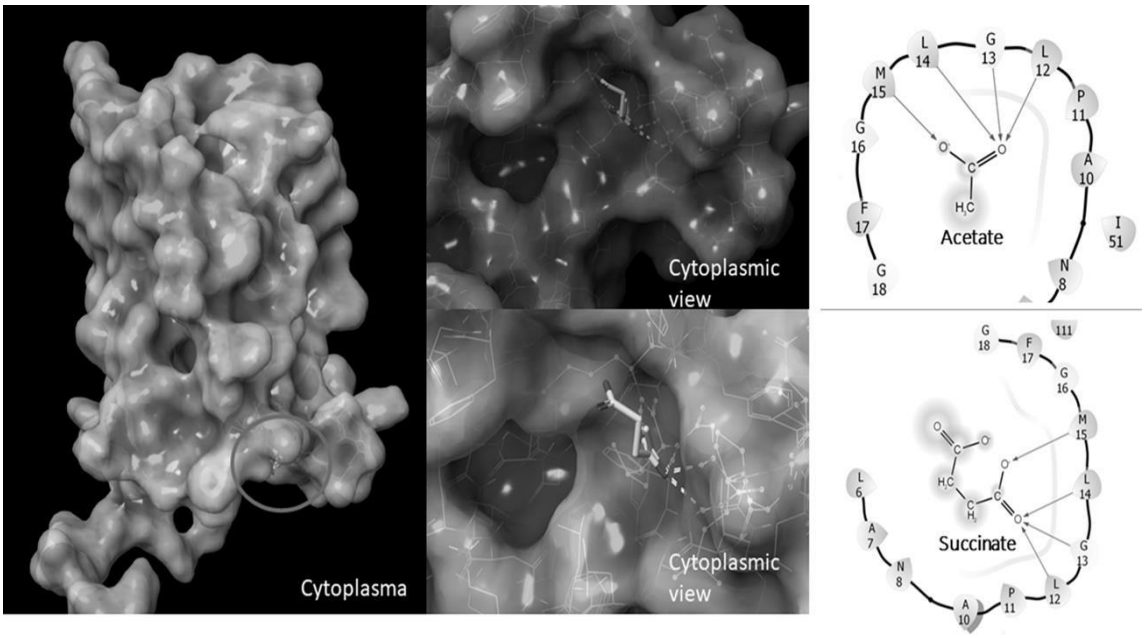


Figure 6. Molecular docking of SatP with acetate and succinate. A) Transversal and cytoplasmic view of SatP protein surface with the identification of the putative binding site surrounded by a grey circle. Ligand molecules are represented in their stick molecular structure. B) 2D view of the interactions between ligands (acetate and succinate) and SatP residues. Black arrows represent hydrogen bonds between ligands atoms and residue atoms.

7.4. DISCUSSION

The AceTr-Acetate Transporters are found in all domains of life. The phylogenetic analysis of all the homologues so far sequenced, suggests basal split between eukaryotic and prokaryotic organisms that form almost two monophyletic clades, with an exception from each subclade. Overall, the phylogenetic tree suggests that an ancestral form of this family was present in the primordial organisms in the root of the tree of life. However, although present in a taxonomically diverse range of organisms, the AceTr homologues are limited to very specific groups of prokaryotic and eukaryotic organisms. This fact strongly appoints to their loss in most organisms during evolution. While this trend is obvious for most eukaryotic organisms outside fungi (homologues detected in only 10% of analyzed genomes), the picture is similar for prokaryotic organisms, with members of the AceTr family detected in only 15% of the genomes in the database. On a drastic shift, homologues were detected in 97% of the fungal genomes analyzed. These data suggest that these transporters play an important role in fungi. In fact, the Ady2p from *S. cerevisiae* was described as essential for proper ascus formation [6] and later AcpA from *A. nidulans* as the major acetate transporte during germination of conidiospores [11]. In addition, the Ady2 homologue of *Cryptococcus neoformans* was also associated with acetate transport and described as indispensable for survival during physiologically relevant starvation conditions [15]. The deletion of *C. neoformans* Ady2 and Ato2 was also found to impair virulence factors-ability to grow at human body temperature, produce melanine and the capsule structure, which are major immune modulators [15]. Based on this evidence, it is tempting to postulate that this family of transporters is critical for spore germination process in fungi, and thus greatly conserved throughout evolution. According to the phylogenetic analysis, the only fungal genomes without AceTr members were the *Encephalitozoon* that belongs to phylum Microsporidia. These are atypical fungi, first considered a deeply branching protist lineage, lacking mitochondria and peroxisomes, that are obligate intracellular parasites infesting many animal groups [33].

In this work, we also extended the number of AceTr members functionally characterized as acetate transporters, and for the first time to an archaealmember, the *aceP* gene from *M. acetivorans*. The *Methanosarcineae* are the most metabolically diverse methanogens, thrive in a broad range of environments, and are unique among the Archaea in forming complex multicellular structures. Species, like *M. acetivorans* display high methanogenesis metabolic activity. In nature, the methane produced is mostly derived from acetate with a strong impact on global warming [34]. Here we demonstrate that the *aceP* from *M. acetivorans* is mediating the transport of acetic acid, by an acetate-proton symport mechanism, highly specific for acetate.

Additionally, we characterized the acetate transport capacity of Gpr1 from *Y. lipolytica*, an AceTr family member, until now only described as being involved in cells´ sensitivity to acetic acid[4]. Gpr1p was reported to have an important role over stress cue, triggered by the exposure of cells to acetic acid, and to be involved in the repression of genes that encode glyoxylate cycle enzymes [4,5]. Fluorescence analysis of *S. cerevisiae* cells harbouring Gpr1-GFP revealed that the fusion protein was detected at the plasma membrane, as expected [4, 35]. The yeast *Y. lipolytica* is reported to display specific physiological, metabolic and genomic characteristics, which differentiates it from the model yeast *S. cerevisiae* [36]. Although the whole genome duplication event did not occur in *Y. lipolytica*, its genome size is almost de double of *S. cerevisiae.*

Nevertheless, Gpr1 and Ady2, two of the AceTr members of these distinct yeast species, maintained their activity as acetate transporters over evolution. To gain insights into the structure of AceTr family members, the most conserved motif NPAPLGL(M/S) of Ady2 and SatP proteins, was analyzed by site-directed mutagenesis. Until now, no protein structure has been determined for members of this family. Given the results presented herein the motif, NPAPLGL(M/S) can be considered as a 'signature' motif of this group of transporters, as substitutions in the residues of this domain abolished the transport activity and resulted in poor growth in acetate, as sole carbon and energy source. Nevertheless, the membrane targeting of Ady2 was not affected by mutations in residues of Ady2 motif NPAPLGL(M/S). In a similar mutagenesis approach made by Gentsch and colleagues (2008), several residues of Ady2 were randomly mutated and characterized regarding acetic acid sensitivity in *S. cerevisiae*; the mutant alleles A70V, F71V, T74I, A88V, E144G N145D, M211K, and F212S were found to be critical for acetic acid sensitivity. Interestingly, the A88 residue was placed right before the NPAPLGL(M/S) motif. In addition, it was also reported that the deletion of NPAPLGL(M/S) motif in the Gpr1 protein heterologously expressed in *S. cerevisiae* strain, deleted for all AceTr homologues, displayed alterations in acetic acid sensitivity [37]. Thus, the NPAPLGL(M/S) is critical for transport activity but the targeting and membrane stability in not driven by residues of this motif, at least for Ady2. Molecular docking analysis led to the identification of a putative binding site formed by the "^{8}NPAPLGL(M/S)^{15}" motif in a SatP 3D predicted structure. Although at the moment, there is no direct experimental evidence that the AceTr protein is an acetate exporter, several reports propose the role of these proteins in acetate export [4, 13, 37]. The predicted location of this critical conserved motif in the vicinity of the protein pore entrance at the cytoplasm side reinforces the hypothesis that SatP plays a role in acetate export.

According to this study, the members of AceTr family were lost during the evolution of both eukaryotic and prokaryotic organisms, excluding few exceptions. Interestingly this evolutionary pattern is also

displayed by other carboxylate transporter family, designated as the Jen1 family (TCDB 2.A.1.12). The Jen1 family, which includes members characterized as carboxylate transporters, is present in fungi and bacteria, but there are no homologues in higher eukaryotes, where other membrane proteins such as the members of the Monocarboxylate Transporter Family (TCDB 2.A.1.13), seem responsible for the transport of carboxylic acids [2, 3].

It is biologically reasonable to consider that evolutionary pressures led to alternate physiological processes to evolve from a catalogue of alternatives, present in ancestral organisms (especially in prokaryotic organisms characterized by their evolutionary plasticity). However, we must consider the presence of AceTr family members in almost all analyzed fungi organisms. This suggests the existence of a specific and essential role in fungal evolution, possibly in sexual reproduction through sporulation.

7.5. REFERENCES

1. Xiao M, Wu F: A review of environmental characteristics and effects of low-molecular weight organic acids in the surface ecosystem. *Journal of Environmental Sciences* 2014, 26(5):935-954.
2. Casal M, Paiva S, Queiros O, Soares-Silva I: Transport of carboxylic acids in yeasts. *FEMS microbiology reviews* 2008, 32(6):974-994.
3. Casal M, Queiros O, Talaia G, Ribas D, Paiva S: Carboxylic Acids Plasma Membrane Transporters in Saccharomyces cerevisiae. *Advances in experimental medicine and biology* 2016, 892:229-251.
4. Augstein A, Barth K, Gentsch M, Kohlwein SD, Barth G: Characterization, localization and functional analysis of Gpr1p, a protein affecting sensitivity to acetic acid in the yeast Yarrowia lipolytica. *Microbiology* 2003, 149(3):589-600.
5. Tzschoppe K, Augstein A, Bauer R, Kohlwein SD, Barth G: trans-dominant mutations in the GPR1 gene cause high sensitivity to acetic acid and ethanol in the yeast Yarrowia lipolytica. *Yeast* 1999, 15(15):1645-1656.
6. Rabitsch KP, Toth A, Galova M, Schleiffer A, Schaffner G, Aigner E, Rupp C, Penkner AM, Moreno-Borchart AC, Primig M *et al*: A screen for genes required for meiosis and spore formation based on whole-genome expression. *Current biology : CB* 2001, 11(13):1001-1009.
7. Paiva S, Devaux F, Barbosa S, Jacq C, Casal M: Ady2p is essential for the acetate permease activity in the yeast Saccharomyces cerevisiae. *Yeast* 2004, 21(3):201-210.
8. Palková Z, Devaux F, Icicova M, Minarikova L, Le Crom S, Jacq C: Ammonia pulses and metabolic oscillations guide yeast colony development. *Molecular biology of the cell* 2002, 13(11):3901-3914.

9. Robellet X, Flipphi M, Pegot S, Maccabe AP, Velot C: AcpA, a member of the GPR1/FUN34/YaaH membrane protein family, is essential for acetate permease activity in the hyphal fungus Aspergillus nidulans. *The Biochemical journal* 2008, 412(3):485-493.
10. Fillinger S, Felenbok B: A newly identified gene cluster in Aspergillus nidulans comprises five novel genes localized in the alc region that are controlled both by the specific transactivator AlcR and the general carbon-catabolite repressor CreA. *Molecular microbiology* 1996, 20(3):475-488.
11. Sá-Pessoa J, Amillis S, Casal M, Diallinas G: Expression and specificity profile of the major acetate transporter AcpA in Aspergillus nidulans. *Fungal genetics and biology : FG & B* 2015, 76:93-103.
12. Flipphi M, Robellet X, Dequier E, Leschelle X, Felenbok B, Velot C: Functional analysis of alcS, a gene of the alc cluster in Aspergillus nidulans. *Fungal genetics and biology : FG & B* 2006, 43(4):247-260.
13. Sá-Pessoa J, Paiva S, Ribas D, Silva IJ, Viegas SC, Arraiano CM, Casal M: SATP (YaaH), a succinate-acetate transporter protein in Escherichia coli. *The Biochemical journal* 2013, 454(3):585-595.
14. Rohlin L, Gunsalus RP: Carbon-dependent control of electron transfer and central carbon pathway genes for methane biosynthesis in the Archaean, Methanosarcina acetivorans strain C2A. *BMC microbiology* 2010, 10:62.
15. Kisirkoi G: Acetate Transport Is Essential for Survival and Virulence of *Cryptococcus neoformans*. *All Dissertations* 2017, 1915.
16. Pei J, Kim B-H, Grishin NV: PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic acids research* 2008, 36(7):2295-2300.
17. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S: MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular biology and evolution* 2013, 30(12):2725-2729.
18. Soares-Silva I, Paiva S, Diallinas G, Casal M: The conserved sequence NXX[S/T]HX[S/T]QDXXXT of the lactate/pyruvate:H(+) symporter subfamily defines the function of the substrate translocation pathway. *Molecular membrane biology* 2007, 24(5-6):464-474.
19. Norrander J, Kempe T, Messing J: Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* 1983, 26(1):101-106.
20. Mumberg D, Müller R, Funk M: Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 1995, 156(1):119-122.
21. Arraiano CM, Yancey SD, Kushner SR: Stabilization of discrete mRNA breakdown products in ams pnp rnb multiple mutants of Escherichia coli K-12. *Journal of Bacteriology* 1988, 170(10):4625-4633.
22. Thomas BJ, Rothstein R: The genetic control of direct-repeat recombination in Saccharomyces: The effect of rad52 and rad1 on mitotic recombination at GAL10, a transcriptionally regulated gene. *Genetics* 1989.
23. Datsenko KA, Wanner BL: One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* 2000, 97(12):6640-6645.

24. Viegas SC, Pfeiffer V, Sittka A, Silva IJ, Vogel J, Arraiano CM: Characterization of the role of ribonucleases in Salmonella small RNA decay. *Nucleic acids research* 2007, 35(22):7651-7664.
25. Bessa D, Pereira F, Moreira R, Johansson B, Queirós O: Improved gap repair cloning in yeast: treatment of the gapped vector with Taq DNA polymerase avoids vector self-ligation. *Yeast* 2012, 29(10):419-423.
26. Lõoke M, Kristjuhan K, Kristjuhan A: EXTRACTION OF GENOMIC DNA FROM YEASTS FOR PCR-BASED APPLICATIONS. *BioTechniques* 2011, 50(5):325-328.
27. Kim DE, Chivian D, Baker D: Protein structure prediction and analysis using the Robetta server. *Nucleic acids research* 2004, 32(Web Server issue):W526-W531.
28. Ribas D, Sá-Pessoa J, Soares-Silva I, Paiva S, Nygård Y, Ruohonen L, Penttilä M, M. C: Yeast as a tool to express sugar acid transporters with biotechnological interest. *FEMS Yeast Res* 2017, 17(2 (fox005)).
29. Fujiwara T, Ohnuma M, Yoshida M, Kuroiwa T, Hirano T: Gene Targeting in the Red Alga Cyanidioschyzon merolae: Single- and Multi-Copy Insertion Using Authentic and Chimeric Selection Markers. *PloS one* 2013, 8(9):e73608.
30. Soares-Silva I, Ribas D, Foskolou IP, Barata B, Bessa D, Paiva S, Queiros O, Casal M: The Debaryomyces hansenii carboxylate transporters Jen1 homologues are functional in Saccharomyces cerevisiae. *FEMS yeast research* 2015, 15(8).
31. Queirós O, Pereira L, Paiva S, Moradas-Ferreira P, Casal M: Functional analysis of Kluyveromyces lactis carboxylic acids permeases: heterologous expression of KlJEN1 and KlJEN2 genes. *Current genetics* 2007, 51(3):161-169.
32. Sowers KR, Baron SF, Ferry JG: Methanosarcina acetivorans sp. nov., an Acetotrophic Methane-Producing Bacterium Isolated from Marine Sediments. *Applied and environmental microbiology* 1984, 47(5):971-978.
33. Katinka MD, Duprat S, Cornillot E, Méténier G, Thomarat F, Prensier G, Barbe V, Peyretaillade E, Brottier P, Wincker P *et al*: Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi. *Nature* 2001, 414:450.
34. Galagan JE, Nusbaum C, Roy A, Endrizzi MG, Macdonald P, FitzHugh W, al. e: The Genome of M. acetivorans Reveals Extensive Metabolic and Physiological Diversity. *Genome Research* 2002, 12(4):532-542.
35. Matthäus F, Barth G: The Gpr1/Fun34/YaaH Protein Family in the Nonconventional Yeast Yarrowia lipolytica and the Conventional Yeast Saccharomyces cerevisiae. In: *Yarrowia lipolytica: Genetics, Genomics, and Physiology.* Edited by Barth G. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013: 145-163.
36. Nicaud J-M: Yarrowia lipolytica. *Yeast* 2012, 29(10):409-418.
37. Gentsch M, Kuschel M, Schlegel S, Barth G: Mutations at different sites in members of the Gpr1/Fun34/YaaH protein family cause hypersensitivity to acetic acid in Saccharomyces cerevisiae as well as in Yarrowia lipolytica. *FEMS yeast research* 2007, 7(3):380-390.

CHAPTER VIII

Evolutionary engineering reveals amino acid substitutions in Ato2 and Ato3 that allow improved growth of *Saccharomyces cerevisiae* on lactic acid

Adapted from:

Baldi, N.[1*], De Valk, S.[1*], Sousa-Silva, M.[2,3], Casal, M.[2,3], Soares-Silva, I.[2,3], Mans, R.[1] (2021).
Evolutionary engineering reveals amino acid substitutions in Ato2 and Ato3 that allow improved growth of *Saccharomyces cerevisiae* on lactic acid

[1] Department of Biotechnology, Delft University of Technology, Delft, The Netherlands
[2] Institute of Science and Innovation for Bio-Sustainability (IB-S), University of Minho, Portugal
[3] Centre of Molecular and Environmental Biology (CBMA), Department of Biology, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal

[*] Shared-first authorship.
Submitted to FEMS Yeast Research.

Personal contribution: I performed the *in silico* structural analysis presented in this work, namely the 3D modelling and molecular docking of the transporters Ady2, Ato2 and Ato3 and mutant alelles. I also participated in the manuscript writing.
Note: Part of this chapter is also reported in the Doctoral thesis of Nicolo Baldi (TU Delft, NL).

CHAPTER VIII

Evolutionary engineering reveals amino acid substitutions in Ato2 and Ato3 that allow improved growth of Saccharomyces cerevisiae on lactic acid

ABSTRACT

In *Saccharomyces cerevisiae*, the complete set of proteins involved in transport of lactic acid across the cell membrane has not been identified, which limits its potential as industrial production host for this bulk chemical. In this study we aimed to identify transport proteins not previously described to be involved in lactic acid transport via a combination of directed evolution, whole-genome resequencing and reverse engineering. Evolution of a strain lacking all known lactic acid transporters on lactate led to the identification of mutated Ato2 and Ato3 as two novel lactic acid transport proteins. When compared to previously identified *S. cerevisiae* genes involved in lactic acid transport, *ATO3*^{T284C} was able to facilitate the highest growth rate (0.15 ± 0.01 h^{-1}) on this carbon source. A comparison between (evolved) sequences and 3D models of the transport proteins showed that many mutations resulted in a widening of the narrowest hydrophobic constriction of the anion channel. We hypothesize that this observation, sometimes in combination with an increased binding affinity of lactic acid in the binding sites adjacent to this restriction site, are responsible for the improved lactic acid transport in the evolved proteins.

KEYWORDS: Adaptive laboratory evolution; *Saccharomyces cerevisiae*; Ato transporters; Monocarboxylate transporters; Molecular docking studies

8.1 INTRODUCTION

Carboxylic acids are widely applied as platform molecules in the chemical, pharmaceutical, food and beverage industries. Some organis acids, like succinic and lactic acid, are currently produced on industrial scale using biotechnological processes [1-3]. Advantages of bioprocesses are their potential sustainability compared to their petrochemical counterparts and, in the case of chiral compounds, their ability to selectively produce one of the enantiomers. The latter is of importance for the production of biodegradable polymers, which requires the precursors to be optically pure [4]. One of these optically active molecules is lactic acid, which is used as a preservative in the dairy industry and as precursor for bioplastic formation. The demand for lactic acid was 1.220.000 tons in 2016 and is expected to further increase by 16.2% by 2025 [5]. At present, lactic acid is mostly produced by prokaryotic hosts (reviewed by Es, Mousavi Khaneghah [6] and McKinlay, Vieille [7]), with titers reaching upwards of 182 g L^{-1} in fed-batch cultures [8]. The limited tolerance of these hosts to low pH implicates that neutralizing agents must be added during microbial fermentation under industrial conditions. Subsequent retrieval of the undissociated acid requires acidification of the broth, with concomitant production of gypsum, resulting in additional costs in downstream processing [9].

Despite exhibiting the highest growth rate at pH 5.0-5.5 [10], the yeast *Saccharomyces cerevisiae* is known for its tolerance to low pH, and could therefore be an interesting production host for carboxylic acids [11-13]. Early attempts to engineer this organisms for industrial production of lactic acid have been successful [14, 15] and, more recently, fed-batch cultures of engineered *S. cerevisiae* strains reached titers of 82 g L^{-1} and 142 g L^{-1} for the D and L isomer, respectively [16, 17]. Despite our ability to engineer the production several organic acids in *S. cerevisiae*, little is known about their transport across the yeast membrane [11, 18]. Transport of carboxylic acids forms an important optimization target for metabolic engineering, to stimulate product export (increasing productivity and titers) and to increase tolerance to weak organic acid stress [11, 19, 20]. Diffusion of undissociated lactic acid across yeast cell membranes is unlikely to be the main mechanism for lactic acid export under physiological conditions [21-23]. Therefore, further investigation of the genes involved in lactic acid export in yeast remains important to enable the design and construction of future industrial cell factories.

Two genes encoding transporters for monocarboxylic acids have been identified in *S. cerevisiae*: *JEN1* and *ADY2*. Jen1 is a member of the Major Facilitator Superfamily (MFS) which enables uptake of lactic, acetic and pyruvic acid, while Ady2, a member of the AceTr family, has been described to function predominantly as an acetate transporter [24-26]. While deletion of both genes abolished growth on lactic

acid as the sole carbon source, export of lactic acid in an engineered *S. cerevisiae* strain remained unaffected by combined deletion of *JEN1* and *ADY2* [27, 28]. The difficulty of identifying the export system(s) for lactic acid was further illustrated by a recent study, in which the combined deletion of 25 (putative) transporter-encoding genes in a single yeast strain did not affect the rate of lactic acid export [29].

A powerful strategy to identify genes involved in a specific physiological function is the use of adaptive laboratory evolution. Application of a selective pressure is used to enrich for mutants with the phenotype of interest, which can subsequently be analysed by whole genome sequencing to identify mutated genes related to the evolved phenotype [30]. Whereas the use of this strategy to select for mutants with altered lactic acid export remains a challenge, adaptive evolution can be directly applied to select for improved lactic acid uptake. This concept was demonstrated in a previous study, in which laboratory evolution of a *jen1Δ* strain in culture medium with lactic acid as sole carbon source led to the identification of mutated *ADY2* alleles that had an increased uptake capacity for lactic acid [31].

In this study, we use adaptive laboratory evolution to identify additional transporters, which upon mutation can efficiently catalyze lactic acid uptake in *S. cerevisiae*. Subsequently, we overexpress the complete suite of native and evolved lactic acid transporters and characterize the ability of the resulting strains to grow on a variety of organic acids. Finally, we identify specific amino acid residues in the evolved transporters which play a key role in determining their ability to transport lactic acid and provide a mechanistic explanation using three dimensional structure predictions combined with molecular docking analysis.

8.2 Materials and Methods

8.2.1 Strains and maintenance

The *S. cerevisiae* strains used in this study (Table) share the CEN.PK113-7D or the CEN.PK2-1C genetic backgrounds [32]. Stock cultures of *S. cerevisiae* were grown aerobically in 500 mL round-bottom shake flasks containing 100 mL synthetic medium (SM) [33] or YP medium (10 g L^{-1} Bacto yeast extract, 20 g L^{-1} Bacto peptone) supplemented with 20 g L^{-1} glucose. When needed, auxotrophic requirements were complemented via addition of 150 mg L^{-1} uracil, 100 mg L^{-1} histidine, 500 mg L^{-1} leucine and/or 75 mg L^{-1} tryptophan [34]. For plate cultivation, 2% (w/v) agar was added to the medium prior to heat sterilization. Stock cultures of *E. coli* XL1-Blue Subcloning Grade Competent Cells (Agilent, Santa Clara, CA, USA) that were used for plasmid propagation were grown in LB medium (5 g L^{-1} Bacto yeast extract, 10 g L^{-1} Bacto

tryptone, 10 g L^{-1} NaCl) supplemented with 100 mg L^{-1} ampicillin. Media were autoclaved at 121°C for 20 min and supplements and antibiotics were filter sterilized and added to the media prior to use. Frozen culture stocks were prepared by addition of sterile glycerol (to a final concentration of 30% v/v) to exponentially growing shake flask cultures of *S. cerevisiae* or overnight cultures of *E. coli* and 1 mL aliquots were stored at -80°C.

8.2.2 MOLECULAR BIOLOGY TECHNIQUES

Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA) was used for PCR amplification for cloning purposes. Diagnostic PCRs were performed using DreamTaq PCR Master Mix (2X) (Thermo Fisher Scientific). In both cases, the manufacturer's protocol was followed, with the exception of the use of lower primer concentrations (0.2 μM each). Desalted (DST) oligonucleotide primers were used, except for primers binding to coding regions, which were PAGE purified. Primers were purchased from Sigma Aldrich (Saint Louis, MO, USA), with the exception of primers 17453 and 17453, which were purchased from Ella Biotech (Planegg, Germany). For diagnostic PCR, yeast genomic DNA was isolated as described by Lõoke, Kristjuhan [35]. Commercial kits for DNA extraction and purification were used for small-scale DNA isolation (Sigma Aldrich), PCR cleanup (Sigma Aldrich), and gel extraction (Zymo Research, Irvine, CA, USA). Restriction analysis of constructed plasmids was performed using FastDigest restriction enzymes (Thermo Scientific). Gibson assembly of linear DNA fragments was performed using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs, Ipswich, MA, USA) in a total reaction volume of 5 μL. Transformation of chemically competent *E. coli* XL1-Blue (Agilent) was performed according to the manufacturer's protocol.

8.2.3 PLASMID CONSTRUCTION

The plasmids and oligonucleotide primers used in this study are listed in Table 1 and Supplementary Table 1, respectively. All plasmids were constructed by Gibson assembly of two linear fragments. With the exception of the fragments used for the construction of plasmid pUDR420, all fragments were PCR-amplified from either a template plasmid or from genomic DNA. A detailed description of the parts used to make each plasmid can be found in Supplementary Table 2.

Plasmid pUDR405 was constructed by Gibson assembly of two linear fragments, both obtained via PCR amplification of plasmid pROS13 using primers 8664 and 6262 (for the *JEN1*-gRNA_2μ_*ADY2*-gRNA insert) and 6005 (for the plasmid backbone), as previously described by Mans, van Rossum [36]. Plasmid pUDR420 was constructed by Gibson assembly of a double-stranded DNA fragment, obtained by

annealing the complementary single-stranded oligonucleotides 8691 and 13552, and a vector backbone amplified from plasmid pMEL13 using primers 6005 and 6006. For construction of pUDE813, the linear p426-TEF plasmid backbone was amplified from plasmid p426-TEF using primers 5921 and 10547 and the *ATO3* open reading frame (ORF) was amplified from yeast strain CEN.PK113-7D genomic DNA using primers 13513 and 13514. Subsequently, Gibson assembly of the linear p426-TEF plasmid backbone and the *ATO3* insert yielded pUDE813. pUDE814, pUDE1001, pUDE1002, pUDE1003, pUDE1004, pUDE1021 and pUDE1022 were constructed similar to pUDE813, using primers 5921 and 10547 to amplify the linear p427-TEF plasmid backbone. The inserts were amplified from genomic DNA of strain CEN.PK113-7D (for wildtype genes) or from genomic DNA of the corresponding evolved strain (for mutated genes) using primers 13513 and 13514 (pUDE814), 17170 and 17171 (pUDE1001), 17168 and 17169 (pUDE1002, pUDE1003 and pUDE1004) or 17452 and 17453 (pUDE1021 and pUDE1022). For construction of pUDC319, plasmid p426-TEF was amplified using primers 2949 and 17741 and the CEN6 origin of replication was amplified from pUDC156 using primers 17742 and 17743. Subsequently, Gibson assembly of the linear p426-TEF plasmid fragment and the CEN6 fragment yielded pUDC319. pUDC320, pUDC321, pUDC322, pUDC323, pUDC324, pUDC325, pUDC326 and pUDC327 were constructed in a similar way using the same primers, but the linear plasmid fragment was amplified from pUDE813, pUDE814, pUDE1001, pUDE1002, pUDE1003, pUDE1004, pUDE1021 and pUDE1022, respectively.

Table 1. Plasmids used in this study

Name	Relevant characteristic	Origin
pROS13	2μm ampR *kanMX* gRNA-*CAN1* gRNA-*ADE2*	[36]
pMEL13	2μm ampR *kanMX* gRNA-*CAN1*	[36]
pUDR405	2μm ampR *kan*MX gRNA-*JEN1* gRNA-*ADY2*	This study
pUDR420	2μm ampR *kanMX* gRNA-*ATO3*	This study
p426-TEF	2μm *URA3 pTEF1-tCYC1*	[37]
pUDE813	2μm *URA3 pTEF1-ATO3-tCYC1*	This study
pUDE814	2μm *URA3 pTEF1-ATO3*^{T284C}*-tCYC1*	This study
pUDE1001	2μm *URA3 pTEF1-JEN1-tCYC1*	This study
pUDE1002	2μm *URA3 pTEF1-ADY2-tCYC1*	This study
pUDE1003	2μm *URA3 pTEF1- ADY2*^{C755G} *-tCYC1*	This study
pUDE1004	2μm *URA3 pTEF1- ADY2*^{C655G}*-tCYC1*	This study
pUDE1021	2μm *URA3 pTEF1-ATO2-tCYC1*	This study

pUDE1022	2μm *URA3 pTEF1-ATO2^{T653C}-tCYC1*	This study
pUDC156	*CEN6 URA3 pTEF-CAS9-tCYC1*	[38]
pUDC319	*CEN6 URA3 pTEF-tCYC1*	This study
pUDC320	*CEN6 URA3 pTEF1-ATO3-tCYC1*	This study
pUDC321	*CEN6 URA3 pTEF1-ATO3^{T284C}-tCYC1*	This study
pUDC322	*CEN6 URA3 pTEF1-JEN1-tCYC1*	This study
pUDC323	*CEN6 URA3 pTEF1-ADY2-tCYC1*	This study
pUDC324	*CEN6 URA3 pTEF1- ADY2^{C755G} -tCYC1*	This study
pUDC325	*CEN6 URA3 pTEF1- ADY2^{C655G}-tCYC1*	This study
pUDC326	*CEN6 URA3 pTEF1-ATO2-tCYC1*	This study
pUDC327	*CEN6 URA3 pTEF1-ATO2^{T653C}-tCYC1*	This study

8.2.4 STRAIN CONSTRUCTION

S. cerevisiae strains were transformed with the LiAc/ssDNA method [39]. For transformations with a dominant marker, the transformation mixture was plated on YP plates, containing glucose (20 g L^{-1}) as carbon source, and supplemented with 200 mg L^{-1} G418 (Invitrogen, Carlsbad, CA, USA). Gene deletions were performed as previously described [36]. For transformation of plasmids harboring an auxotrophic marker, transformed cells were plated on SM medium with glucose (20 g L^{-1}) as a carbon source and when needed, appropriate auxotrophic requirements were supplemented. The tryptophan auxotrophy of IMX1000 was the result of a single point mutation in the *TRP1* gene (*trp1-289*) [40] and was spontaneously reverted by plating the strain on SM medium supplemented with uracil, histidine and leucine, and picking a tryptophan prototrophic colony, yielding strain IMX2486. Strain IMX2487 was constructed by transforming IMX2486 with a linear fragment, obtained by PCR amplification of the *LEU2* gene from CEN.PK113-7D, using primers 1742 and 1743. Strain IMX2488 was constructed by transforming IMX2487 with a linear fragment, obtained by PCR amplification of the *HIS3* gene from CEN.PK113-7D, using primers 1738 and 3755. Strain IMK875 was constructed by transforming the Cas9-expressing strain IMX585 with plasmid pUDR405 and two double stranded repair oligonucleotides obtained by annealing oligonucleotides 8597 to 8598 and 8665 to 8666. Strain IMK876 was constructed by transforming the Cas9-expressing strain IMX581 with plasmid pUDR405 and two double stranded repair oligonucleotides obtained by annealing oligonucleotides 8597 to 8598 and 8665 to 8666. Strains IMK882 and IMK883 were obtained by transforming strains IMK875 and IMK876, respectively, with plasmid pUDR420 and a double stranded repair oligonucleotide obtained by annealing oligonucleotides 14120 and 14121. Plasmids p426-TEF, pUDE813, pUDE814, pUDE1001, pUDE1002, pUDE1003,

pUDE1004, pUDE1021, pUDE1022, pUDC319, pUDC320, pUDC321, pUDC322, pUDC323, pUDC324, pUDC325, pUDC326 and pUDC327 were transformed in strain IMX2488, yielding IME581, IME582, IME583, IME584, IME585, IME586, IME587, IME588, IME589, IMC164, IMC165, IMC166, IMC167, IMC168, IMC169, IMC170, IMC171 and IMC172, respectively. Single colony isolates from evolution cultures ('IMS'-strains) were obtained by plating the evolved culture on solid medium and restreaking a grown colony to a fresh plate three consecutive times.

Table 2. *Saccharomyces cerevisiae* strains used in this study

Strain name	**Relevant genotype** [a]	**Origin**
CEN.PK113-7D	Prototrophic reference, *MAT*a	[32]
IMX581	*MAT*a *ura3-52 can1::cas9-natNT2*	[36]
IMX585	*MAT*a *can1::cas9-natNT2*	[36]
IMK341	*MAT*a *ura3::loxP ady2::loxP-hphNT1-loxP jen1::loxP*	[27]
IMW004	*MAT*a *URA3 ADY2*^{C755G} *jen1::loxP-KanMX4-loxP*	[27]
IMW005	*MAT*a *URA3 ADY2*^{C655G} *jen1::loxP-KanMX4-loxP*	[27]
IMX1000	*MAT*a *ura3-52 trp1-289 leu2-3112 his3Δ can1Δ::cas9-natNT2 mch1Δ mch2Δ mch5Δ aqy1Δ itr1Δ pdr12Δ mch3Δ mch4Δ yil166cΔ hxt1Δ jen1Δ ady2Δ aqr1Δ thi73Δ fps1Δ aqy2Δ yll053cΔ ato2Δ ato3Δ aqy3Δ tpo2Δ yro2Δ azr1Δ yhl008cΔ tpo3Δ*	[41]
IMK875	*MAT*a *can1::cas9-natNT2 jen1Δ ady2Δ*	This study
IMK876	*MAT*a *can1::cas9-natNT2 ura3-52 jen1Δ ady2Δ*	This study
IMK882	*MAT*a *can1::cas9-natNT2 jen1Δ ady2Δ ato3Δ*	This study
IMK883	*MAT*a *can1::cas9-natNT2 ura3-52 jen1Δ ady2Δ ato3Δ*	This study
IMS807	IMK341 evolved for growth on lactate, evolution line A	This study
IMS808	IMK341 evolved for growth on lactate, evolution line A	This study
IMS809	IMK341 evolved for growth on lactate, evolution line A	This study
IMS810	IMK341 evolved for growth on lactate, evolution line B	This study
IMS811	IMK341 evolved for growth on lactate, evolution line B	This study
IMS1122	IMK882 evolved for growth on lactate	This study
IMS1123	IMK882 evolved for growth on lactate	This study
IMS1130	IMK882 evolved for growth on lactate	This study
IMX2486	IMX1000 *ura3-52 TRP1, leu2-3112, his3Δ*	This study
IMX2487	IMX1000 *ura3-52 TRP1*, *LEU2*, *his3Δ*	This study
IMX2488	IMX1000 *ura3-52 TRP1*, *LEU2*, *HIS3*	This study

IME581	IMX2488 p426-TEF (2μm)	This study
IME582	IMX2488 pUDE813 (2μm *ATO3*)	This study
IME583	IMX2488 pUDE814 (2μm *ATO3*^{T284C})	This study
IME584	IMX2488 pUDE1001 (2μm *JEN1*)	This study
IME585	IMX2488 pUDE1002 (2μm *ADY2*)	This study
IME586	IMX2488 pUDE1003 (2μm *ADY2*^{C755G})	This study
IME587	IMX2488 pUDE1004 (2μm *ADY2*^{C655G})	This study
IME588	IMX2488 pUDE1021 (2μm *ATO2*)	This study
IME589	IMX2488 pUDE1022 (2μm *ATO2*^{T653C})	This study
IMC164	IMX2488 pUDC319 (*CEN6*)	This study
IMC165	IMX2488 pUDC320 (*CEN6 ATO3*)	This study
IMC166	IMX2488 pUDC321 (*CEN6 ATO3*^{T284C})	This study
IMC167	IMX2488 pUDC322 (*CEN6 JEN1*)	This study
IMC168	IMX2488 pUDC323 (*CEN6 ADY2*)	This study
IMC169	IMX2488 pUDC324 (*CEN6 ADY2*^{C755G})	This study
IMC170	IMX2488 pUDC325 (*CEN6 ADY2*^{C655G})	This study
IMC171	IMX2488 pUDC326 (*CEN6 ATO2*)	This study
IMC172	IMX2488 pUDC327 (*CEN6 ATO2*^{T653C})	This study

8.2.5 MEDIA AND CULTIVATION

Evolution experiments were performed in 500 mL shake-flask cultures containing 100 mL synthetic medium [33] with 84 mM lactic acid as sole carbon source. The pH of the medium was set at 5.0 and the cultures were incubated at 30°C in an Innova incubator shaker (New Brunswick Scientific, Edison, NJ, USA) set at 200 rpm. Auxotrophic requirements were supplemented as needed.

Strains were characterized in SM supplemented with different carbon sources. To achieve an initial carbon concentration of 250 mM, the culture media contained either 42 mM glucose, 83 mM lactic acid, 125 mM acetic acid or 83 mM pyruvic acid. The characterization was performed in a Growth-Profiler system (EnzyScreen, Heemstede, The Netherlands) equipped with 96-well plates in a culture volume of 250 µL, set at 250 rpm and 30°C. The measurement interval was set at 30 minutes. Raw green values were corrected for well-to-well variation using measurements of a 96-well plate containing a culture with an externally determined optical density of 3.75 in all wells. Optical densities were calculated by converting green values (corrected for well-to-well variation) using a calibration curve that was determined by fitting a third-degree polynomial through 22 measurements of cultures with known OD values between 0.1 and 20. Growth rates were calculated using the calculated optical densities of at least 15 points in the

exponential phase. Exponential growth was assumed when an exponential curve could be fitted with an R^2 of at least 0.985.

8.2.6 Analytical methods

Culture optical density at 660 nm was measured with a Libra S11 spectrophotometer (Biochrom, Cambridge, United Kingdom). In order to measure within the linear range of the instrument (OD between 0.1 and 0.3), cultures were diluted in an appropriate amount of demineralized water. Metabolite concentrations in culture supernatants and media were analyzed using an Agilent 1260 Infinity HPLC system equipped with a Bio-rad Aminex HPX-87H ion exchange column, operated at 60°C with 5 mM H_2SO_4 as mobile phase at a flow rate of 0.600 mL min^{-1}.

8.2.7 DNA extraction and whole genome sequencing

Strain IMK341 and the evolved single colony isolates (IMS-strains) were grown in 500 mL shake flasks containing 100 mL YP medium supplemented with glucose (20 g L^{-1}) as a carbon source. The cultures were incubated at 30°C until the strains reached stationary phase and genomic DNA was isolated using the Qiagen 100/G kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions and quantified using a Qubit® Fluorometer 2.0 (Thermo Fisher Scientific). The isolated DNA was sequenced in-house on a MiSeq (Illumina, San Diego, CA, USA) with 300 bp paired-end reads using TruSeq PCR-free library preparation (Illumina). For all the strains, the reads were mapped onto the *S. cerevisiae* CEN.PK113-7D genome [42] using the Burrows–Wheeler Alignment tool (BWA) and further processed using SAMtools and Pilon for variant calling [43-45].

8.2.8 3D Modelling and Molecular docking of Ady2, Ato2 and Ato3

The three-dimensional modelling analysis was performed for the protein sequences of Ato1, $Ady2^{L219V}$, $Ady2^{A252G}$, Ato2, $Ato2^{L218S}$, Ato3 and $Ato3^{F95S}$. The amino acid sequences were retrieved from the *Saccharomyces* Genome Database [46]. To determine the predicted transporter 3D structures, the amino acid sequences were threaded through the PDB library using LOMETS (Local Meta-Threading-Server) [47]. The *Citrobacter koseri* succinate acetate permease (CkSatP, PDB 5YS3) was the top ranked template threading identified in LOMETS for Ady2, Ato2 and Ato3 [48]. Since the CkSatP three-dimensional modelling obtained the best score for protein structure prediction, it was further considered for molecular docking analysis. Molecular docking simulations were performed as described by Ribas *et al.,* 2017 [49]. Ligand structures of acetic, lactic and pyruvic acids for all target proteins in

the study were downloaded from Zinc database [50]. Structures used for docking were all confirmed in Maestro v11.2 before ligand-protein simulations using AutoDock Vina in PyRx software [51]. The docking studies were performed with dissociated forms of each carboxylic acid. The protonation states were adjusted to match a pH of 5.0-6.0 and exported in the mol2 format. Docking was performed with four docking-boxes for each protein complex, containing top, bottom and middle-structure parts for a more robust use of Autodock Vina program. Exhaustiveness parameter was set up for 1000 calculations for each one of grid-zones defined for docking. The generated docking scores and 2-3D pose views were evaluated for the establishment of molecular interactions and ligand binding affinities.

8.3. Results

8.3.1 Laboratory evolution on lactic acid leads to point mutations in Ato2 or Ato3

In an attempt to identify additional transporters able to catalyze uptake of lactic acid transport after gaining point mutations, we incubated strains IMK341 and IMX1000 in duplicate shake flasks cultures containing synthetic medium with lactic acid as the sole carbon source. In IMK341 the known carboxylic acid transporters *JEN1* and *ADY2* are knocked out (*jen1Δ*, *ady2Δ*), whereas IMX1000 contains a further 23 deletions in putative lactic acid transporter-encoding genes [41]. After 9 weeks, growth was observed in both cultures of IMK341 whereas no growth was observed after 12 weeks of incubation of IMX1000. Whole-genome sequencing of evolved IMK341 (*jen1Δ*, *ady2Δ*) cell lines, isolated after transfer to fresh medium, revealed three to seven non-synonymous SNPs in each mutant and no chromosomal duplications or rearrangements (Table 3). Strikingly, all evolved isolates shared an identical mutation in *ATO3* (*ATO3*^{T284C}), which encodes a protein described to be involved in ammonium transport [52]. To investigate the role of *ATO3* in lactic acid uptake, we overexpressed both the native and evolved *ATO3* in IMK883 (*jen1Δ*, *ady2Δ*, *ato3Δ*) and tested the strains for growth on SM lactic acid plates. After 5 days, only the reference strain CEN.PK113-7D and the strain carrying the *ATO3*^{T284C} allele were able to grow (Supplementary Figure 1), indicating that the T284C mutation in *ATO3* was responsible for the evolved phenotype. We then combined the deletion of *JEN1*, *ADY2* and *ATO3* in strain IMK882 (*jen1Δ, ady2Δ, ato3Δ*) and repeated the evolution. After 5 and 12 days, growth was observed in two of the cultures from which strains IMS1122 and IMS1123 were isolated after transfer to a flask with fresh medium. In both single colony isolates, five SNPs were present (Table 3), including a common mutation in *ATO2*, (*ATO2*^{T653C}), which has also been described as an ammonium transporter together with *ATO3* and *ADY2*

[52]. Finally, the evolution was repeated with IMK982 (*jen1*Δ, *ady2*Δ, at*o3*Δ, *ato2*Δ), but no growth was observed after 12 weeks of incubation.

Table 3. Amino acid changes identified by whole-genome sequencing of single colony isolates evolved for growth in medium containing lactic acid as sole carbon source. Isolates IMS807 to IMS811 are derived from IMK341 (jen1Δ, ady2Δ) and IMS1122 and IMS1123 are derived from IMK882 (jen1Δ, ady2Δ, ato3Δ)). IMS807, IM808 and IMS809 are isolates from evolution line #1 and IMS810 and IMS811 are isolates from evolution line #2. The mutation Sip5^{*490Q} indicates the loss of the stop codon.

IMK341 evolution #1			**IMK341 evolution #2**		**IMK822 evolution #1**	**IMK822 evolution #2**
IMS807	IMS808	IMS809	IMS810	IMS811	IMS1122	IMS1123
Ato3^{F95S}	**Ato3^{F95S}**	**Ato3^{F95S}**	**Ato3^{F95S}**	**Ato3^{F95S}**	**Ato2^{L218S}**	**Ato2^{L218S}**
Mms2^{Y58C}	Mms2^{Y58C}	Mms2^{Y58C}	Sip5^{*490Q}	Sip5^{*490Q}	Lrg1^{H979N}	Whi2^{E119*}
Pih1^{D147Y}	Pih1^{D147Y}	Pih1^{D147Y}	Ssn2^{M1280R}	Lip5^{R4L}	Ykr051w^{Y285H}	Ykr051w^{Y285H}
Uba1^{L952F}		Drn1^{P213L}			Jjj1^{H356Q}	Jjj1^{H356Q}
Stv1^{L275F}					Trm10^{A49V}	Trm10^{A49V}
Whi2^{E187*}						
Vba4^{P198L}						

8.3.2 Overexpression of mutated transporters enables rapid growth in liquid medium with lactic acid as sole carbon source

Strikingly, the evolved transporters able to catalyze uptake of lactic acid (*ATO2* and *ATO3* in this study, and *ADY2* in work by de Kok et al., 2012) represent all members of the Acetate Uptake Transporter Family (TCDB 2.A.96). To characterize impact of the mutations on the ability of the mutated transport proteins to catalyze uptake of organic acids, we individually overexpressed *JEN1, ADY2, ATO2* and *ATO3* and their mutated alleles via centromeric vectors in IMX2488, a strain background in which 25 (putative) organic acid transporters were deleted (Table 2). Whereas no viable cultures could be obtained with strains overexpressing wildtype *ATO2 ,* all other IMX2488-derived transporter expressing strains had similar growth rates in liquid medium with 42 mM glucose as carbon source compared to the empty vector control (IMC164), indicating no major physiological adaptations to the overexpression of the transporters when grown on glucose (Figure 1, top left panel). Overexpression of the transporter variants from multicopy vectors resulted in a growth rate reduction of up to 66% compared to the empty vector reference when grown on glucose and were therefore not tested further (Supplementary Figure 2). In accordance with previous research, strains overexpressing *ADY2*, *ADY2*^{c755G} and *ADY2*^{c655G} showed a

maximum specific growth rate of 0.02 ± 0.01 h^{-1}, 0.08 ± 0.01 h^{-1} and 0.10 ± 0.01 h^{-1} when grown in medium containing 83 mM lactic acid as carbon source, respectively (de Kok et al., 2012). Surprisingly, strains expressing the evolved *ATO2*^{T653C} and *ATO3*^{T284C} alleles outperformed all the other tested strains, with a maximum specific growth rate of 0.11 ± 0.01 h^{-1} and 0.15 ± 0.01 h^{-1}, respectively (Figure 1, top right panel and Supplementary Figure 5). These represent the highest reported growth rates reported for *S. cerevisiae* on this carbon source and indicate that, similar to the role of evolved Ato3 in IMS807-811, the mutations in Ato2 are responsible for the evolved phenotypes observed in IMS1122 and IMS1123 (Table 3).

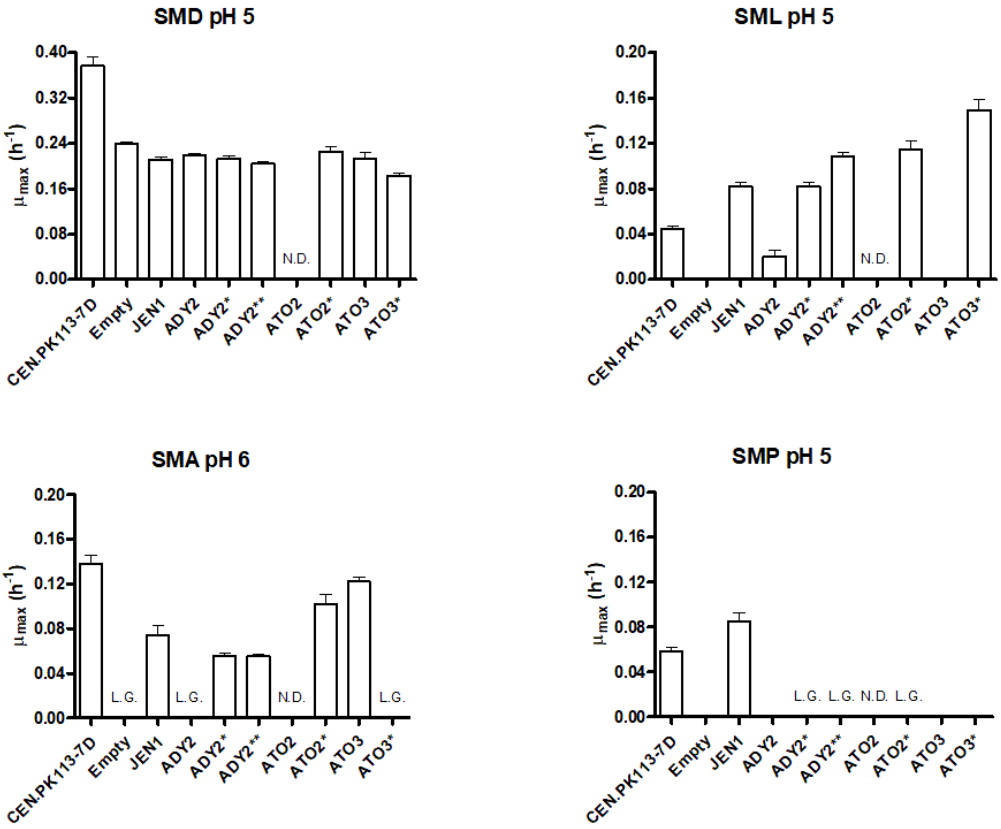


Figure 1. Growth rates on different carbon sources of *S. cerevisiae* reference strain CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty vector or a vector carrying the indicated organic acid transporter. Bars and error bars represent the average and standard deviation of three independent experiments. SMD: synthetic medium with 42 mM glucose. SML: synthetic medium with 83 mM lactic acid. SMA: synthetic medium with 125 mM acetic acid. SMP: synthetic medium with 83 mM pyruvic acid. Empty: empty plasmid. ADY2*: *ADY2*^{C755G} allele. ADY2**: *ADY2*^{C655G} allele. ATO2*: *ATO2*^{T653C} allele. ATO3*: *ATO3*^{T284C} allele. For some experiments, a linear increase in optical density was observed, which impeded the determination of an exponential growth rate (indicated by L.G.). N.D.: not determined.

8.3.3 MUTATIONS IN ATO2 AND ATO3 CHANGE THE UPTAKE CAPACITY OF ACETATE AND PYRUVATE

After demonstrating that the point mutations increased the catalytic activity of Ato2, Ato3 and Ady2 for lactic acid transport, we also investigated their ability to transport acetic and pyruvic acids (Figure 1, bottom panels and Supplementary Figures 6 and 7). In liquid medium at pH 5.0 with 125 mM acetic acid (pK_a of 4.75), no growth was observed for any of the strains with the 25-deletion background, likely caused by acetic acid toxicity due to the absence of essential acetic acid exporters (Supplementary Figure 3). However, at pH 6.0 different growth characteristics were observed in strains expressing *ADY2* and *ATO3* compared to their evolved counterparts. Whereas expression of native *ADY2* resulted in slow non-exponential growth, expression of *ADY2*^{C755G} or *ADY2*^{C655G} improved growth performance. Surprisingly, native ATO3 was apparently able to catalyze acetate transport, a function not previously described for this gene, whereas only slow, non-exponential growth was observed for the evolved *ATO3*^{T284C} variant. In medium containing 83 mM pyruvic acid, no exponential growth was observed for any of the strains expressing Ato2, Ato3 or Ady2 variants. However, slow, non-exponential growth was observed for strains expressing *ATO2*^{T653C} or any variant of ADY2 which could indicate a minor change in affinity for this substrate caused by the point mutations.

8.3.4 PROTEIN MODELLING REVEALS MUTATIONS IN THE CENTRAL HYDROPHOBIC CONSTRICTION SITE AS IMPORTANT FACTOR IN DETERMINING SUBSTRATE SPECIFICITY

In order to establish a link between the observed phenotypes and the structure alterations of transporters carrying the mutated amino acid residues, the 3D protein structures of Ady2, Ato2 and Ato3 were predicted based on the crystal structure of the *Citrobacter koseri* acetate anion channel SatP (PDB 5YS3), a bacterial member of the AceTr family [48]. When combined with sequence alignment of Ady2p, Ato2p and Ato3p, it showed that the Leu219Val mutation in Ady2, the Leu218Ser mutation in Ato2 and the Phe95Ser mutation in Ato3 are amongst three amino acid residues that were previously identified to be essential for the formation of the central narrowest hydrophobic constriction of the anion pathway in *C. koseri* SatP [48] (Figure 2 and Figure 3). Specifically, these changes result in the substitution of the amino acid side group with a smaller (and in the case of Ato2 and Ato3 a more hydrophilic) alternative (Ato3 is shown in Figure 3 and the models for Ady2 and Ato2 can be found in Supplementary Figures 9 and 10). Based on these models, we estimated the distance between these three hydrophobic residues and found an increased distance in the *ADY2*^{C655G}, *ATO3*^{T284C} and *ATO2*^{T653C} encoded mutants compared to

their corresponding wildtype protein, leading to a larger aperture in the pore of the channel (table 4). We hypothesize that this increased size of the hydrophobic constriction may allow larger substrates to pass through, thus increasing substrate specificity and transport capacity.

To investigate if the mutations affected the presence and affinity of binding sites for acetate, lactate and pyruvate, docking of ligands in the protein structures was simulated using AutoDock Vina (Supplementary Figure 10, Supplementary table 3 and 4). In all proteins, both wildtype and mutated models, four binding sites were identified for acetate, which is in accordance with what has previously been reported for the homolog *Ck*SatP [48]. Of these four binding sites, two, which are located closest to the hydrophobic constriction, also consistently bind lactate and pyruvate. Strikingly, mutations in Ady2, Ato2 and Ato3 led to an increased lactate affinity of at least one of these two sites closest to the hydrophobic constriction, which might have contributed to the increased lactate transport capacity. No clear correlation was found between the physiology observed for strains overexpressing the different protein variants when grown on acetate and pyruvate and the corresponding binding affinities of these two ligands (Supplementary table 3).

Figure 2. Multiple sequence alignment of *Citrobacter koseri* SatP and *Saccharomyces cerevisiae* Ady2, Ato2 and Ato3. The multiple sequence alignment was built with ClustalOmega (https://www.ebi.ac.uk/Tools/msa/clustalo/). Localization of transmembrane segments (TMSs) was predicted with PSI/TM-Coffee (http://tcoffee.crg.cat/apps/tcoffee/do:tmcoffee). Blue rectangles indicate residues of the narrowest constriction site F98-Y155-L219 (amino acid numbers reports to Ady2) (Qiu *et al.*, 2018). Bold, underlined letters indicate the mutated residue.

Table 4. Average distances (in Å) between different amino acids in the constriction pore of Ady2, Ato2 Ato3 and mutated alleles, calculated using the corresponding protein model. Bold values in the table indicate distances which are at least 1 Å larger than the one calculated in the reference structure.

Protein	Amino acid residues			Protein	Amino acid residues			Protein	Amino acid residues		
	219&98	98&155	155&219		218&97	97&154	154&218		215&95	95&152	152&215
Ady2	4,57	6,85	3,96	Ato2	4,16	5,74	4,37	Ato3	4,04	4,60	4,11
Ady2 L219V	4,37	6,94	**5,34**	Ato2 L218S	**5,94**	5,64	**5,57**	Ato3 F95S	**7,69**	**8,50**	4,51
Ady2 A252G	4,49	6,74	3,89								

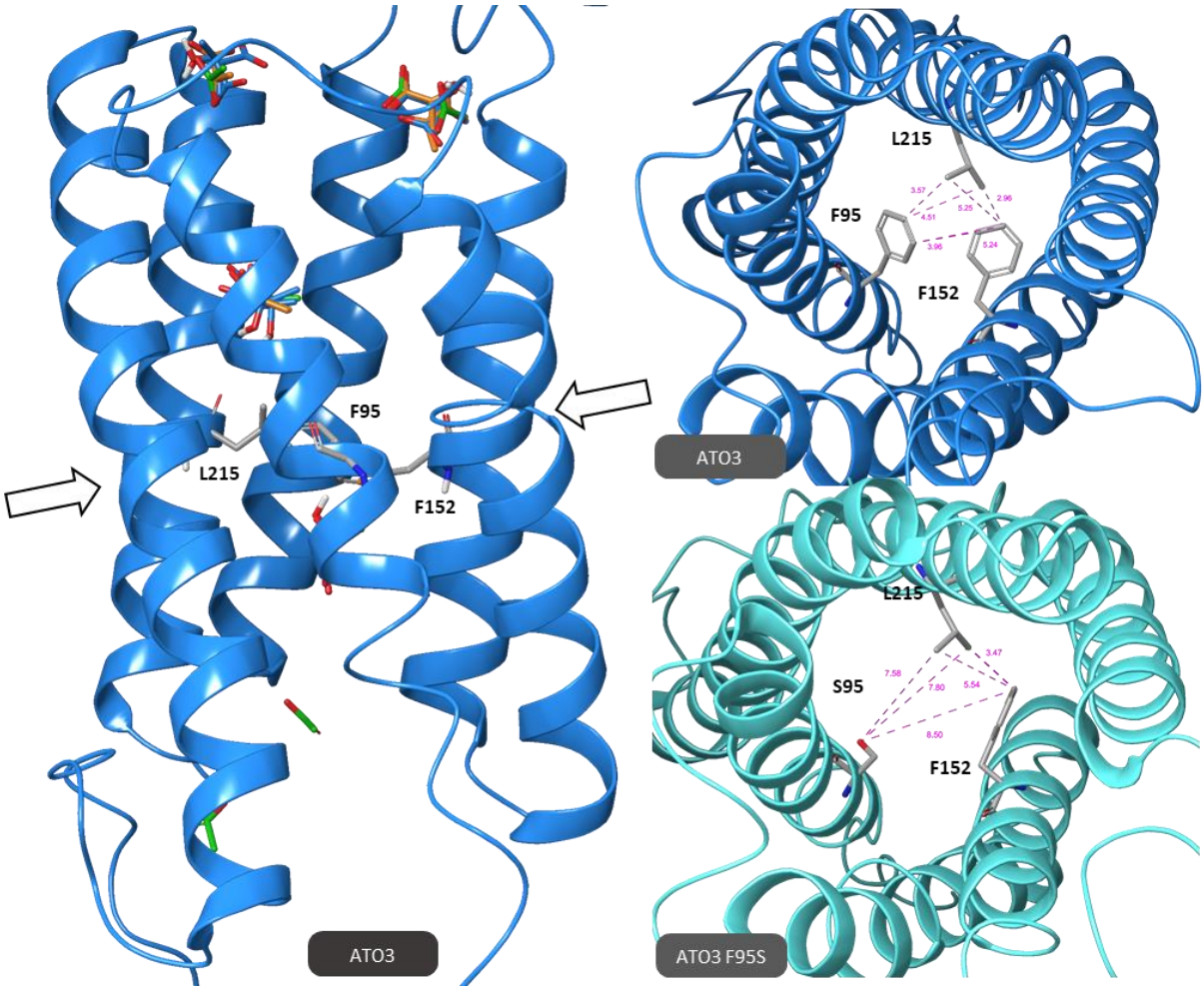


Figure 3. 3D-Models of the transporters Ato3 (dark blue) and Ato3 F95S (light blue). Left: side view of Ato3. Arrows indicate the hydrophobic constriction site, consisting of F95, L215 and F152. Binding sites for acetate (green ligand), lactate (blue ligand) and pyruvate (orange ligand) are here presented. Right, top view of either Ato3 (top) or Ato3^{F95S}(bottom). The amino acids involved in the constriction site are shown. Purple lines and values indicate distances (in Å) between different anchor points of amino acids.

8.4 DISCUSSION

In this study, we report the identification and characterization of a family of transporter genes which, upon mutation, are able to efficiently catalyze the import of lactic acid in *S. cerevisiae*. As rational engineering to identify lactic acid transporters remains elusive [18, 41], we used adaptive laboratory evolution to select for mutants capable of consuming lactic acid, which led to the identification of mutations in *ATO3* (*ATO3*^{T284C}) and *ATO2* (*ATO2*^{T653C}). Together with *ADY2*, *ATO2* and *ATO3* were previously described to code for ammonium transporters (Ammonium Transport Outwards) based on two observations: the high expression levels of these genes when *S. cerevisiae* exports ammonium, and the presence of a motif associated with ammonium transport in the encoded proteins [52]. However, the function of *ADY2* has previously been assessed by Rabitsch, Toth [53], who identified it as a gene required for correct spore formation, and thus named it as *ADY2* (Accumulation of DYads). In view of the observations in our study, where *ADY2*, *ATO2* and *ATO3* and their evolved variants catalyzed uptake of lactic and acetic acid, and the absence of mechanistic studies aimed at illustrating the phenomenon of ammonium export, we support the recent proposition by Alves, Sousa-Silva [54] to rename these genes, present in *S. cerevisiae* and other yeasts, as "Acetate Transporter Ortholog". Our finding that besides Ady2, native Ato3 is able to catalyze acetate transport is in accordance with the observation that expression of both genes encoding these proteins are induced in yeast cells shifted from glucose to acetic acid as carbon source [26].

For physiological studies focused on organic acid substrate uptake, a platform strain devoid of organic acid importers is a useful tool as it enables characterization based on growth rate. The prototrophic strain IMC164 (25 deletions and empty vector) used in this study, exhibited a ~36% decreased growth rate on glucose (0.24 ± 0.00 h^{-1}, Figure 1 and Supplementary Figure 4) compared to the reference laboratory strain CEN.PK113-7D (0.38 ± 0.02 h^{-1}), which is in accordance with previous observations in bioreactor cultures [41].

No growth was observed for IMC164 on medium containing either lactic acid or pyruvate as sole carbon source (Figure 1), demonstrating that this is a suitable strain background to test pyruvic and lactic acid transport capacity of transporter variants. In contrast, when grown on acetic acid at pH 6.0, IMC164 exhibited non-exponential linear growth (Supplementary Figure 6), suggesting acetic acid diffusion, or the presence of at least one gene involved in acetate transport in this strain background. It was reported by Kok, Nijkamp [27] that the overexpression of *ADY2*, under the control of the strong glycolytic promoter *TEF1*, was sufficient to enable slow growth (μ_{max} ~0.02 h^{-1}) in medium containing lactic acid as sole carbon source.

While the native alleles of *ATO3* and likely *ATO2* were not able to sustain growth on lactic acid medium, their mutated versions (*ATO2*^{T653C} and *ATO3*^{T284C}) enabled high growth rates, with the highest growth rate determined at 0.15 ± 0.01 h^{-1} for the strain harboring *ATO3*^{T284C}.

To the best of our knowledge, this growth rate represents the highest reported growth rate of *S. cerevisiae* expressing a single transport protein on lactic acid and is close to the growth rate observed by de Kok et al. (2012) of 0.14 h^{-1} by a strain expressing *ADY2*^{C655G}. This 3-fold increase in growth rate of the engineered strain compared to the reference strain CEN.PK113-7D indicates that, in non-engineered *S. cerevisiae* strains, growth on lactic acid is likely limited by its transport into the cell, and not the capacity to be further metabolized. Therefore, for future work that requires fast consumption of lactic acid, overexpression of *ATO3*^{T284C} can be considered.

To infer the function of the newly identified mutations, the 3D structures of Ady2 (Ato1), Ato2 and Ato3 were modeled using the known structure of a homologous transporter CkSatP as a scaffold [48]. Out of the four identified mutations in the evolved transporters, three modified one of the three amino acids in the narrowest hydrophobic constriction of the protein, allowing for a larger aperture to be formed. This observation is in line with previous research, in which changes in these hydrophobic residues in the *Escherichia coli* SatP homolog have been associated with a change of substrate specificity [55][56]. Simulation of ligand docking in the predicted Ato protein structures showed that the identified mutations led to an increased binding affinity of lactate in the core of the protein, whereas a similar consistent change in binding affinity for acetate and pyruvate was not observed. We postulate that an increased binding affinity upon mutation may contribute to increased transport capacity by facilitating passage of the ligand through the hydrophobic constriction, although the increased size of the hydrophobic constriction is probably the main contributor to the evolved phenotype. Other mechanisms may also contribute to an improved transport capacity, as observed for the A252G mutation which is located outside the constriction pore. These may include an improved transition between the closed to open state of the transporter or increased stability in the plasma membrane.

In this study, we show that laboratory evolution is a powerful tool for the identification of genes involved in substrate transport and resulted in the identification of Ato3^{F95S}, which enables the highest growth rate on lactic acid by *S. cerevisiae* reported in strains expressing a single transport protein thus far. In addition, the presented data on transporter structure and function led to the identification of important amino-acid residues that dictate substrate specificity of *S. cerevisiae* carboxylic acid transporters, which could potentially aid in future rational engineering and annotation of additional proteins involved in organic acid transport.

8.5 REFERENCES

1. Verwaal, R., Wu, L., Damveld, R.A.,Sagt, C.M.J., *Succinic acid production in a eukaryotic cell* WIPO, Editor **2007**, Dsm Ip Assets B.V.,
2. Jansen, M.L.A., Heijen, J.J.,Verwaal, R., *Process for preparing dicarboxylic acids employing fungal cells*, **2017**, DSM IP Assets BV,
3. Abdel-Rahman, M.A., Tashiro, Y.,Sonomoto, K., *Recent advances in lactic acid production by microbial fermentation processes.* Biotechnol Adv, **2013**. 31(6): p. 877-902, https://www.doi.org/10.1016/j.biotechadv.2013.04.002
4. Borodina, I.,Nielsen, J., *Advances in metabolic engineering of yeast Saccharomyces cerevisiae for production of chemicals.* Biotechnology Journal, **2014**. 9(5): p. 609-620, 10.1002/biot.201300445
5. Singhvi, M., Zendo, T.,Sonomoto, K., *Free lactic acid production under acidic conditions by lactic acid bacteria strains: challenges and future prospects.* Appl Microbiol Biotechnol, **2018**. 102(14): p. 5911-5924, https://www.doi.org/10.1007/s00253-018-9092-4
6. Es, I., Mousavi Khaneghah, A., Barba, F.J., Saraiva, J.A., Sant'Ana, A.S.,Hashemi, S.M.B., *Recent advancements in lactic acid production - a review.* Food Res Int, **2018**. 107: p. 763-770, https://www.doi.org/10.1016/j.foodres.2018.01.001
7. McKinlay, J.B., Vieille, C.,Zeikus, J.G., *Prospects for a bio-based succinate industry.* Appl Microbiol Biotechnol, **2007**. 76(4): p. 727-40, https://www.doi.org/10.1007/s00253-007-1057-y
8. Subramanian, M.R., Talluri, S.,Christopher, L.P., *Production of lactic acid using a new homofermentative Enterococcus faecalis isolate.* Microb Biotechnol, **2015**. 8(2): p. 221-9, https://www.doi.org/10.1111/1751-7915.12133
9. Lopez-Garzon, C.S.,Straathof, A.J., *Recovery of carboxylic acids produced by fermentation.* Biotechnol Adv, **2014**. 32(5): p. 873-904, https://www.doi.org/10.1016/j.biotechadv.2014.04.002
10. Verduyn, C., Postma, E., Scheffers, W.A.,van Dijken, J.P., *Physiology of Saccharomyces cerevisiae in anaerobic glucose-limited chemostat culturesx.* Microbiology, **1990**. 136(3): p. 395-403, https://www.doi.org/10.1099/00221287-136-3-395
11. Abbott, D.A., Zelle, R.M., Pronk, J.T.,van Maris, A.J.A., *Metabolic engineering of Saccharomyces cerevisiae for production of carboxylic acids: current status and challenges.* FEMS Yeast Research, **2009**. 9: p. 1123-1136
12. Della-Bianca, B.E.,Gombert, A.K., *Stress tolerance and growth physiology of yeast strains from the Brazilian fuel ethanol industry.* Antonie Van Leeuwenhoek, **2013**. 104(6): p. 1083-95, https://www.doi.org/10.1007/s10482-013-0030-2
13. Carmelo, V., Bogaerts, P.,Sa-Correia, I., *Activity of plasma membrane H+-ATPase and expression of PMA1 and PMA2 genes in Saccharomyces cerevisiae cells grown at optimal and low pH.* Archives of microbiology, **1996**. 166(5): p. 315-320, https://doi.org/10.1007/s002030050389

14. Dequin, S.,Barre, P., *Mixed lactic acid–alcoholic fermentation by Saccharomyes cerevisiae expressing the Lactobacillus casei L (+)–LDH.* Bio/Technology, **1994**. 12(2): p. 173-177, https://doi.org/10.1038/nbt0294-173
15. Porro, D., Brambilla, L., Ranzi, B.M., Martegani, E.,Alberghina, L., *Development of metabolically engineered Saccharomyces cerevisiae cells for the production of lactic acid.* Biotechnol Prog, **1995**. 11(3): p. 294-8, https://www.doi.org/10.1021/bp00033a009
16. Baek, S.H., Kwon, E.Y., Bae, S.J., Cho, B.R., Kim, S.Y.,Hahn, J.S., *Improvement of d-Lactic Acid Production in Saccharomyces cerevisiae Under Acidic Conditions by Evolutionary and Rational Metabolic Engineering.* Biotechnol J, **2017**. 12(10), https://www.doi.org/10.1002/biot.201700015
17. Song, J.Y., Park, J.S., Kang, C.D., Cho, H.Y., Yang, D., Lee, S.,Cho, K.M., *Introduction of a bacterial acetyl-CoA synthesis pathway improves lactic acid production in Saccharomyces cerevisiae.* Metab Eng, **2016**. 35: p. 38-45, https://www.doi.org/10.1016/j.ymben.2015.09.006
18. Borodina, I., *Understanding metabolite transport gives an upper hand in strain development.* Microb Biotechnol, **2019**. 12(1): p. 69-70, https://www.doi.org/10.1111/1751-7915.13347
19. Casal, M., Paiva, S., Queirós, O.,Soarez-Silva, I., *Transport of carboxylic acids in yeasts.* FEMS Microbiology Reviews, **2008**. 32: p. 974–994
20. Soares-Silva, I., Ribas, D., Sousa-Silva, M., Azevedo-Silva, J., Rendulić, T.,Casal, M., *Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications.* FEMS microbiology letters, **2020**
21. Gabba, M., Frallicciardi, J., van 't Klooster, J., Henderson, R., Syga, L., Mans, R., van Maris, A.J.A.,Poolman, B., *Weak Acid Permeation in Synthetic Lipid Vesicles and Across the Yeast Plasma Membrane.* Biophys J, **2020**. 118(2): p. 422-434, https://www.doi.org/10.1016/j.bpj.2019.11.3384
22. Cassio, F., Leao, C.,van Uden, N., *Transport of lactate and other short-chain monocarboxylates in the yeast Saccharomyces cerevisiae.* Appl Environ Microbiol, **1987**. 53(3): p. 509-13, https://doi.org/10.1128/AEM.53.3.509-513.1987
23. van Maris, A.J., Winkler, A.A., Porro, D., van Dijken, J.P.,Pronk, J.T., *Homofermentative lactate production cannot sustain anaerobic growth of engineered Saccharomyces cerevisiae: possible consequence of energy-dependent lactate export.* Appl Environ Microbiol, **2004**. 70(5): p. 2898-905, https://www.doi.org/10.1128/aem.70.5.2898-2905.2004
24. Casal, M., Paiva, S., Andrade, R.P., Gancedo, C.,Leão, C., *The Lactate-Proton Symport of Saccharomyces cerevisiae Is Encoded by JEN1.* Journal of Bacteriology, **1999**. 181(8): p. 2620-2623
25. Akita, O., Nishimori, C., Shimamoto, T., Fujii, T.,Iefuji, H., *Transport of Pyruvate in Saccharomyces cerevisiae and Cloning of the Gene Encoded Pyruvate Permease.* Bioscience, Biotechnology, and Biochemistry, **2000**. 64(5): p. 980-984
26. Paiva, S., Devaux, F., Barbosa, S., Jacq, C.,Casal, M., *Ady2p is essential for the acetate permease activity in the yeast Saccharomyces cerevisiae.* Yeast, **2004**. 21: p. 201-210
27. Kok, S., Nijkamp, J.F., Oud, B., Roque, F.C., Ridder, D., Daran, J.-M., Pronk, J.T.,Maris, A.J.A., *Laboratory evolution of new lactate transporter genes in a jen1Δ mutant of Saccharomyces*

cerevisiae and their identification as ADY2 alleles by whole-genome resequencing and transcriptome analysis. FEMS Yeast Research, **2012**. 12(3): p. 359-374, https://www.doi.org/10.1111/j.1567-1364.2011.00787.x

28. Pacheco, A., Talaia, G., Sa-Pessoa, J., Bessa, D., Goncalves, M.J., Moreira, R., Paiva, S., Casal, M.,Queiros, O., *Lactic acid production in Saccharomyces cerevisiae is modulated by expression of the monocarboxylate transporters Jen1 and Ady2.* FEMS Yeast Res, **2012**. 12(3): p. 375-81, https://www.doi.org/10.1111/j.1567-1364.2012.00790.x
29. Mans, R., Hassing, J.-E., Wijsman, M., Giezekamp, A., Pronk, J.T., Daran, J.-M.G.,Van Maris, A.J.A., *A CRISPR/Cas9-based exploration into the elusive mechanism for lactate export in Saccharomyces cerevisiae.* FEMS Yeast Research, **2017**. 17: p. 1-12
30. Mans, R., Daran, J.-M.G.,Pronk, J.T., *Under pressure: evolutionary engineering of yeast strains for improved performance in fuels and chemicals production.* Current Opinion in Biotechnology, **2018**. 50: p. 47-56, 10.1016/j.copbio.2017.10.011
31. de Kok, S., Nijkamp, J.F., Oud, B., Roque, F.C., de Ridder, D., Daran, J.-M.G., Pronk, J.T.,van Maris, A.J.A., *Laboratory evolution of new lactate transporter genes in a jen1Δ mutant of Saccharomyces cerevisiae and their identification as ADY2 alleles by whole-genome resequencing and transcriptome analysis.* FEMS Yeast Research, **2012**. 12: p. 359-374
32. Entian, K.-D.,Kötter, P., *25 Yeast genetic strain and plasmid collections.* Methods in Microbiology, **2007**. 36: p. 629-666, https://doi.org/10.1016/S0580-9517(06)36025-4
33. Verduyn, C., Postma, E., Scheffers, W.A.,Van Dijken, J.P., *Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation.* Yeast, **1992**. 8(7): p. 501-17, https://doi.org/10.1002/yea.320080703
34. Pronk, J.T., *Auxotrophic yeast strains in fundamental and applied research.* Appl Environ Microbiol, **2002**. 68(5): p. 2095-100, https://www.doi.org/10.1128/aem.68.5.2095-2100.2002
35. Lõoke, M., Kristjuhan, K.,Kristjuhan, A., *Extraction of genomic DNA from yeasts for PCR-based applications.* Biotechniques, **2011**. 50(5): p. 325-328, 10.2144/000113672
36. Mans, R., van Rossum, H.M., Wijsman, M., Backx, A., Kuijpers, N.G., van den Broek, M., Daran-Lapujade, P., Pronk, J.T., van Maris, A.J.,Daran, J.M., *CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae.* FEMS Yeast Res, **2015**. 15(2), https://doi.org/10.1093/femsyr/fov004
37. Mumberg, D., Muller, R.,Funk, M., *Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds.* Gene, **1995**. 156(1): p. 119-22, https://www.doi.org/10.1016/0378-1119(95)00037-7
38. Marques, W.L., Mans, R., Marella, E.R., Cordeiro, R.L., van den Broek, M., Daran, J.G., Pronk, J.T., Gombert, A.K.,van Maris, A.J., *Elimination of sucrose transport and hydrolysis in Saccharomyces cerevisiae: a platform strain for engineering sucrose metabolism.* FEMS Yeast Res, **2017**. 17(1): p. fox006, https://www.doi.org/10.1093/femsyr/fox006
39. Gietz, R.D.,Woods, R.A., *Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method.* Methods Enzymol, **2002**. 350: p. 87-96, https://doi.org/10.1016/s0076-6879(02)50957-5

40. Botstein, D., Falco, S.C., Stewart, S.E., Brennan, M., Scherer, S., Stinchcomb, D.T., Struhl, K.,Davis, R.W., *Sterile host yeasts (SHY): a eukaryotic system of biological containment for recombinant DNA experiments.* Gene, **1979**. 8(1): p. 17-24, https://www.doi.org/10.1016/0378-1119(79)90004-0
41. Mans, R., Hassing, E.-J., Wijsman, M., Giezekamp, A., Pronk, J.T., Daran, J.-M.,van Maris, A.J.J.F.y.r., *A CRISPR/Cas9-based exploration into the elusive mechanism for lactate export in Saccharomyces cerevisiae.* **2017**. 17(8): p. fox085, https://doi.org/10.1093/femsyr/fox085
42. Salazar, A.N., Gorter de Vries, A.R., van den Broek, M., Wijsman, M., de la Torre Cortes, P., Brickwedde, A., Brouwers, N., Daran, J.G.,Abeel, T., *Nanopore sequencing enables near-complete de novo assembly of Saccharomyces cerevisiae reference strain CEN.PK113-7D.* FEMS Yeast Res, **2017**. 17(7), https://doi.org/10.1093/femsyr/fox074
43. Li, H.,Durbin, R.J.B., *Fast and accurate long-read alignment with Burrows–Wheeler transform.* **2010**. 26(5): p. 589-595, https://doi.org/10.1093/bioinformatics/btp698
44. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R.,Genome Project Data Processing, S., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, **2009**. 25(16): p. 2078-9, https://doi.org/10.1093/bioinformatics/btp352
45. Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K.,Earl, A.M., *Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement.* PLoS One, **2014**. 9(11): p. e112963, https://doi.org/10.1371/journal.pone.0112963
46. Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S.,Engel, S.R., *Saccharomyces Genome Database: the genomics resource of budding yeast.* Nucleic acids research, **2012**. 40(D1): p. D700-D705, https://doi.org/10.1093/nar/gkr1029
47. Zheng, W., Zhang, C., Wuyun, Q., Pearce, R., Li, Y.,Zhang, Y., *LOMETS2: improved meta-threading server for fold-recognition and structure-based function annotation for distant-homology proteins.* Nucleic acids research, **2019**. 47(W1): p. W429-W436, https://doi.org/10.1093/nar/gkz384
48. Qiu, B., Xia, B., Zhou, Q., Lu, Y., He, M., Hasegawa, K., Ma, Z., Zhang, F., Gu, L.,Mao, Q., *Succinate-acetate permease from Citrobacter koseri is an anion channel that unidirectionally translocates acetate.* Cell research, **2018**. 28(6): p. 644-654, https://doi.org/10.1038/s41422-018-0032-8
49. Ribas, D., Sa-Pessoa, J., Soares-Silva, I., Paiva, S., Nygard, Y., Ruohonen, L., Penttila, M.,Casal, M., *Yeast as a tool to express sugar acid transporters with biotechnological interest.* FEMS Yeast Res, **2017**. 17(2), 10.1093/femsyr/fox005
50. Sterling, T.,Irwin, J., *ZINC 15 – Ligand Discovery for Everyone.* Journal of chemical information and modeling, **2015**. 55(11): p. 2324-2337, 10.1021/acs.jcim.5b00559
51. Trott, O.,Olson, A., *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.* Journal of computational chemistry, **2010**. 31(2): p. 455-461, 10.1002/jcc.21334

52. Palkova, Z., Devaux, F., Icicova, M., Minarikova, L., Le Crom, S.,Jacq, C., *Ammonia pulses and metabolic oscillations guide yeast colony development.* Mol Biol Cell, **2002**. 13(11): p. 3901-14, https://www.doi.org/10.1091/mbc.e01-12-0149
53. Rabitsch, K.P., Toth, A., Galova, M., Schleiffer, A., Schaffner, G., Aigner, E., Rupp, C., Penkner, A.M., Moreno-Borchart, A.C., Primig, M., Esposito, R.E., Klein, F., Knop, M.,Nasmyth, K., *A screen for genes required for meiosis and spore formation based on whole-genome expression.* Curr Biol, **2001**. 11(13): p. 1001-9, https://www.doi.org/10.1016/s0960-9822(01)00274-3
54. Alves, R., Sousa-Silva, M., Vieira, D., Soares, P., Chebaro, Y., Lorenz, M.C., Casal, M., Soares-Silva, I.,Paiva, S., *Carboxylic Acid Transporters in Candida Pathogenesis.* mBio, **2020**. 11(3): p. e00156-20, https://www.doi.org/10.1128/mBio.00156-20
55. Sá-Pessoa, J., Paiva, S., Ribas, D., Silva, I.J., Viegas, S.C., Arraiano, C.M.,Casal, M., *SATP (YaaH), a succinate–acetate transporter protein in Escherichia coli.* Biochemical journal, **2013**. 454(3): p. 585-595
56. Sun, P., Li, J., Zhang, X., Guan, Z., Xiao, Q., Zhao, C., Song, M., Zhou, Y., Mou, L.,Ke, M., *Crystal structure of the bacterial acetate transporter SatP reveals that it forms a hexameric channel.* Journal of Biological Chemistry, **2018**. 293(50): p. 19492-19500, https://www.doi.org/10.1074/jbc.RA118.003876

SUPPLEMENTARY DATA

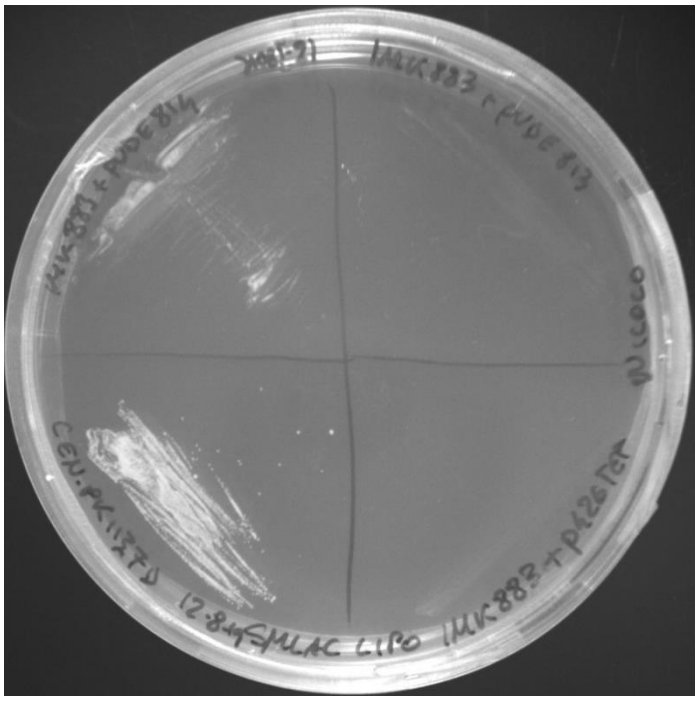
Supplementary Table 1. Primers used in this study

Number	Name	Sequence (5′ -> 3′)	Purpose
8664	JEN1_targetRNA FW SspI	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAG ATAAATGATCGTCACTCAATATTAATTTACGTTTTAGAGCTA GAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC	Construction of pUDR405
6262	CrRNA insert ADY2 fw	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAG ATAAATGATCCCCACCGTAAGAACATAATGGTTTTAGAGCT AGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC	Construction of pUDR405
6005	p426 CRISP rv	GATCATTTATCTTTCACTGCGGAGAAG	Construction of pUDR405 – pUDR420
8691	ATO3_targetRNA FW SspI	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAG ATAAATGATCGAGTATATCTCTTGAATATTGTTTTAGAGCTA GAAATAGCAAGTTAAAATAAG	Construction of pUDR420
13552	ATO3_targetRNA_R V_SspI	CTTATTTTAACTTGCTATTTCTAGCTCTAAAACAATATTCAA GAGATATACTCGATCATTTATCTTTCACTGCGGAGAAGTTT CGAACGCCGAAACATGCGCA	Construction of pUDR420
6006	p426 CRISP fw	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC	Construction of pUDR420
5921	Primer_pTEF1_rv	AAAACTTAGATTAGATTGCTATGCTTTCTTTCTAATGAGC	Linear p426-TEF backbone amplification
10547	p426-GPD backbone rv	TCATGTAATTAGTTATGTCACGC	Linear p426-TEF backbone amplification
13513	pTEF1_ATO3	TACAACTTTTTTTACTTCTTGCTCATTAGAAAGAAAGCATAG CAATCTAATCTAAGTTTTATGACATCGTCTGCTTCTTC	Construction of pUDE813 and pUDE814
13514	tCYC1_ATO3	CGGTTAGAGCGGATGTGGGGGGAGGGCGTGAATGTAAGCG TGACATAACTAATTACATGATTAAGGAGCATTTGGCATTG	Construction of pUDE813 and pUDE814
17168	pTEF1_ADY2_fw	TACAACTTTTTTTACTTCTTGCTCATTAGAAAGAAAGCATAG CAATCTAATCTAAGTTTTATGTCTGACAAGGAACAAACG	Construction of pUDE1002, pUDE1003 and pUDE1004
17169	tCYC1_ADY2_rv	CGGTTAGAGCGGATGTGGGGGGAGGGCGTGAATGTAAGCG TGACATAACTAATTACATGATTAAAAGATTACCCTTTCAGTA G	Construction of pUDE1002, pUDE1003 and pUDE1004
17170	pTEF1_JEN1_fw	TACAACTTTTTTTACTTCTTGCTCATTAGAAAGAAAGCATAG CAATCTAATCTAAGTTTTATGTCGTCGTCAATTACAGATG	Construction of pUDE1001
17171	tCYC1_JEN1_rv	CGGTTAGAGCGGATGTGGGGGGAGGGCGTGAATGTAAGCG TGACATAACTAATTACATGATTAAACGGTCTCAATATGCTCC	Construction of pUDE1001
17452	pTEF1_ATO2_fw	TACAACTTTTTTTACTTCTTGCTCATTAGAAAGAAAGCATAG CAATCTAATCTAAGTTTTATGTCTGACAGAGAACAAAGC	Construction of pUDE1021 and pUDE1022

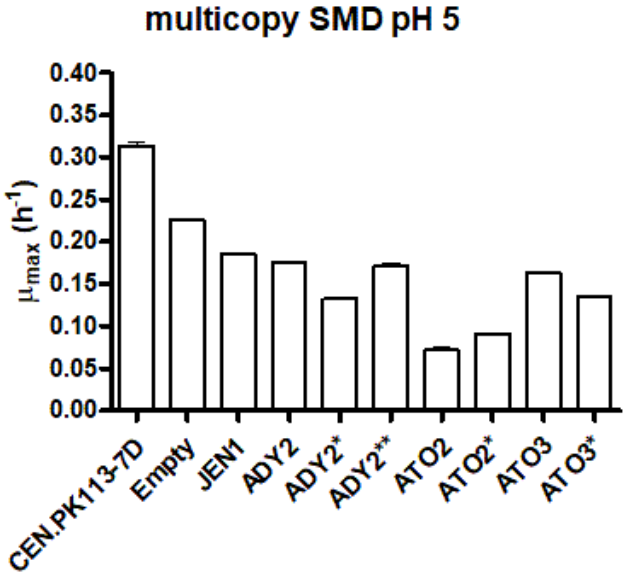
17453	tCYC1_ATO2_rv	CGGTTAGAGCGGATGTGGGGGGAGGGCGTGAATGTAAGCG TGACATAACTAATTACATGATTAGAAGAACACCTTATCATTG C	Construction of pUDE1021 and pUDE1022
17742	p426_CENARS_fw	TAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCG AAAAGTGCCACCTGAACGAACGGATCGCTTGCCTGTAAC	Amplification of CEN6 from pUDC156
17743	p426_CENARS_rv	GATAATATCACAGGAGGTACTAGACTACCTTTCATCCTACA TAAATAGACGCATATAAGTTCCCCGAAAAGTGCCACCTG	Amplification of CEN6 from pUDC156
2949	F Tag Episomal Rev	CGTTCAGGTGGCACTTTTCG	Construction of pUDC319-pUDC327
17741	p426_originremoval	ACTTATATGCGTCTATTTATGTAGGATG	Construction of pUDC319-pUDC327
1742	LEU2 check fw	GGTCGCCTGACGCATATACC	*LEU2* amplification
1743	LEU2 check rv	TAAGGCCGTTTCTGACAGAG	*LEU2* amplification
1738	HIS3 check fw	GCAGGCAAGATAAACGAAGG	*HIS3* amplification
3755	his3 outside rv (B)	CACTTGTTCGCTCAGTTCAG	*HIS3* amplification
8597	JEN1_repair oligo fw	AAGAAGAGTAACAGTTTCAAAAGTTTTTCCTCAAAGAGATTA AATACTGCTACTGAAAATTCACTTTTCATTGCTCTCTAGGG CGTGTTCGCTTCTCTATGTAACTGCATTTCACATATA	Deletion of *JEN1*
8598	JEN1_repair oligo rv	TATATGTGAAATGCAGTTACATAGAGAAGCGAACACGCCCT AGAGAGCAATGAAAAGTGAATTTTCAGTAGCAGTATTTAATC TCTTTGAGGAAAAACTTTTGAAACTGTTACTCTTCTT	Deletion of *JEN1*
8665	ADY2_repair oligo fw	CGACAGCTAACACAGATATAACTAAACAACCACAAAACAAC TCATATACAAACAAATAATGAGCACGACCTACTAATAACGA GAACTATTGAAATAAAAAAGAGTAGTTTTTTATTTTTC	Deletion of *ADY2*
8666	ADY2_repair oligo rv	GAAAAATAAAAAACTACTCTTTTTTATTTCAATAGTTCTCGT TATTAGTAGGTCGTGCTCATTATTTGTTTGTATATGAGTTGT TTTGTGGTTGTTTAGTTATATCTGTGTTAGCTGTCG	Deletion of *ADY2*
14120	ATO3_repair_syn_fw	ATTGAGACGCTCCCCCAGCAGGGTTCGATTGCAGGCGTTT CGCAGGGCAGTAGAATTTCACCTAGACGTGGCCTTCTTGA TGTTGATGTGTACATTGAAGAGCACGTGGGGTTTGTTCT	Deletion of *ATO3*
14121	ATO3_repair_syn_fw	AGAACAAACCCCCACGTGCTCTTCAATGTACACATCAACATC AAGAAGGCCACGTCTAGGTGAAATTCTACTGCCCTGCGAA ACGCCTGCAATCGAACCCTGCTGGGGGAGCGTCTCAAT	Deletion of *ATO3*

Supplementary Table 2. Plasmid construction. Each plasmid was constructed via Gibson assembly of two linear DNA fragments. The template and the primers used to generate those fragments are indicated for each constructed plasmid.

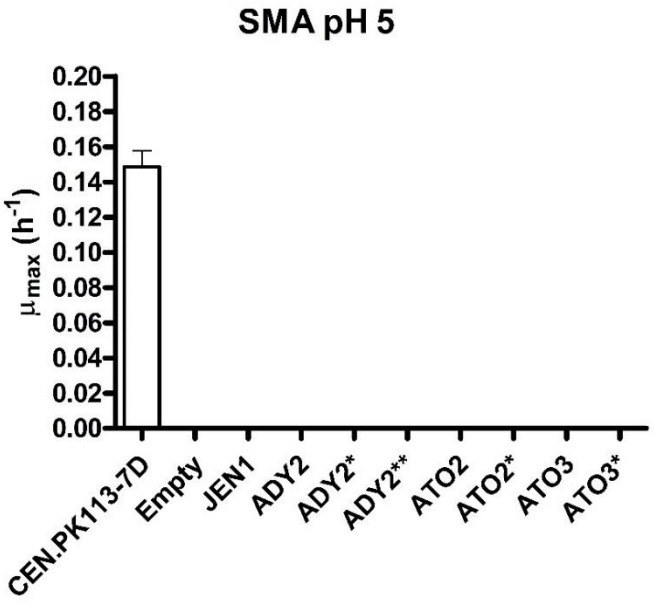
Name	Relevant characteristic	First fragment	Second fragment	Origin
pROS13	2µm ampR *kanMX* gRNA-*CAN1* gRNA-*ADE2*			[36]
pMEL13	2µm ampR *kanMX* gRNA-*CAN1*			[36]
pUDR405	2µm ampR *KAN-MX* gRNA-*JEN1* gRNA-*ADY2*	pROS13 8664+6262	pROS13 6005 (binds twice)	This study
pUDR420	2µm ampR *kanMX* gRNA-*ATO3*	pMEL13 6005+6006	dsDNA oligo formed by: 8691+13552	This study
p426-TEF	*2µ URA3 pTEF1-tCYC1*			[37]
pUDE813	*2µ URA3 pTEF1-ATO3-tCYC1*	p426-TEF 5921+10547	CEN.PK113-7D 13513+13514	This study
pUDE814	*2µ URA3 pTEF1-ATO3^{T284C}-tCYC1*	p426-TEF 5921+10547	IMS807 13513+13514	This study
pUDE1001	*2µ URA3 pTEF1-JEN1-tCYC1*	p426-TEF 5921+10547	CEN.PK113-7D 17170+17171	This study
pUDE1002	*2µ URA3 pTEF1-ADY2-tCYC1*	p426-TEF 5921+10547	CEN.PK113-7D 17168+17169	This study
pUDE1003	*2µ URA3 pTEF1- ADY2^{C755G} -tCYC1*	p426-TEF 5921+10547	IMW004 17168+17169	This study
pUDE1004	*2µ URA3 pTEF1- ADY2^{C655G}-tCYC1*	p426-TEF 5921+10547	IMW005 17168+17169	This study
pUDE1021	*2µ URA3 pTEF1-ATO2-tCYC1*	p426-TEF 5921+10547	CEN.PK113-7D 17452+17453	This study
pUDE1022	*2µ URA3 pTEF1-ATO2^{T653C}-tCYC1*	p426-TEF 5921+10547	IMS1122 17452+17453	This study
pUDC156	*CEN6 URA3 pTEF-CAS9-tCYC1*			Marques *et al.* 2017
pUDC319	*CEN6 URA3 pTEF-tCYC1*	p426-TEF 2949+17741	pUDC156 17742+17743	This study
pUDC320	*CEN6 URA3 pTEF1-ATO3-tCYC1*	pUDE813 2949+17741	pUDC156 17742+17743	This study
pUDC321	*CEN6 URA3 pTEF1-ATO3^{T284C}-tCYC1*	pUDE814 2949+17741	pUDC156 17742+17743	This study
pUDC322	*CEN6 URA3 pTEF1-JEN1-tCYC1*	pUDE1001 2949+17741	pUDC156 17742+17743	This study
pUDC323	*CEN6 URA3 pTEF1-ADY2-tCYC1*	pUDE1002 2949+17741	pUDC156 17742+17743	This study
pUDC324	*CEN6 URA3 pTEF1- ADY2^{C755G} -tCYC1*	pUDE1003 2949+17741	pUDC156 17742+17743	This study
pUDC325	*CEN6 URA3 pTEF1- ADY2^{C655G}-tCYC1*	pUDE1004 2949+17741	pUDC156 17742+17743	This study
pUDC326	*CEN6 URA3 pTEF1-ATO2-tCYC1*	pUDE1021 2949+17741	pUDC156 17742+17743	This study
pUDC327	*CEN6 URA3 pTEF1-ATO2^{T653C}-tCYC1*	pUDE1022 2949+17741	pUDC156 17742+17743	This study



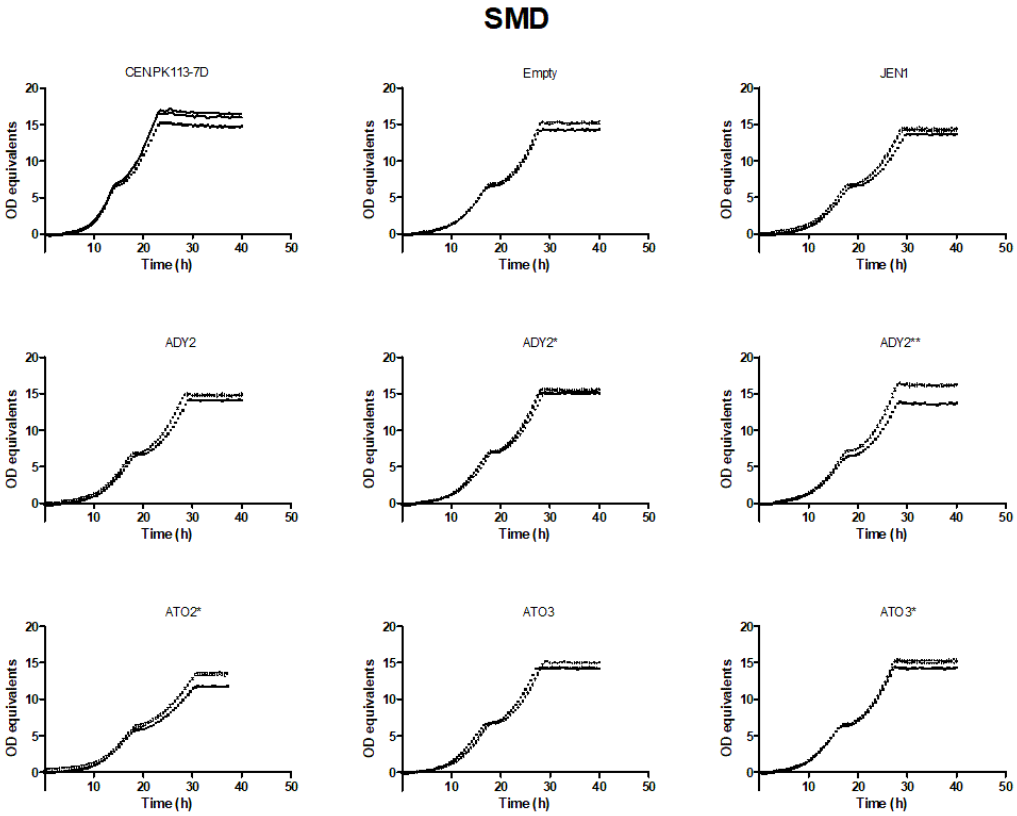
Supplementary Figure 1. Growth of different strains on SM media with lactic acid as the sole carbon source. Bottom left quadrant: prototrophic strain CEN.PK113-7D. Bottom right quadrant: IMK883 (ura3-52, jen1Δ, ady2Δ, ato3Δ) carrying an empty p426-pTEF plasmid. Top right quadrant: IMK883 carrying pUDE813 (p426-pTEF-ATO3). Top left quadrant: IMK883 carrying pUDE814 (p426-pTEF-ATO3^{T284C}). Cells were streaked from a single colony and the plate was incubated at 30 °C for 5 days.



Supplementary Figure 2. Growth rates on SMD of S. cerevisiae reference strain CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty multicopy vector or a multicopy vector containing the indicated organic acid transporter gene. Bars and error bars represent the average and standard deviation of three independent experiments. Empty: empty plasmid. ADY2*: ADY2^{C755G} allele. ADY2**: ADY2^{C655G} allele. ATO2*: ATO2^{T653C} allele. ATO3*: ATO3^{T284C} allele.

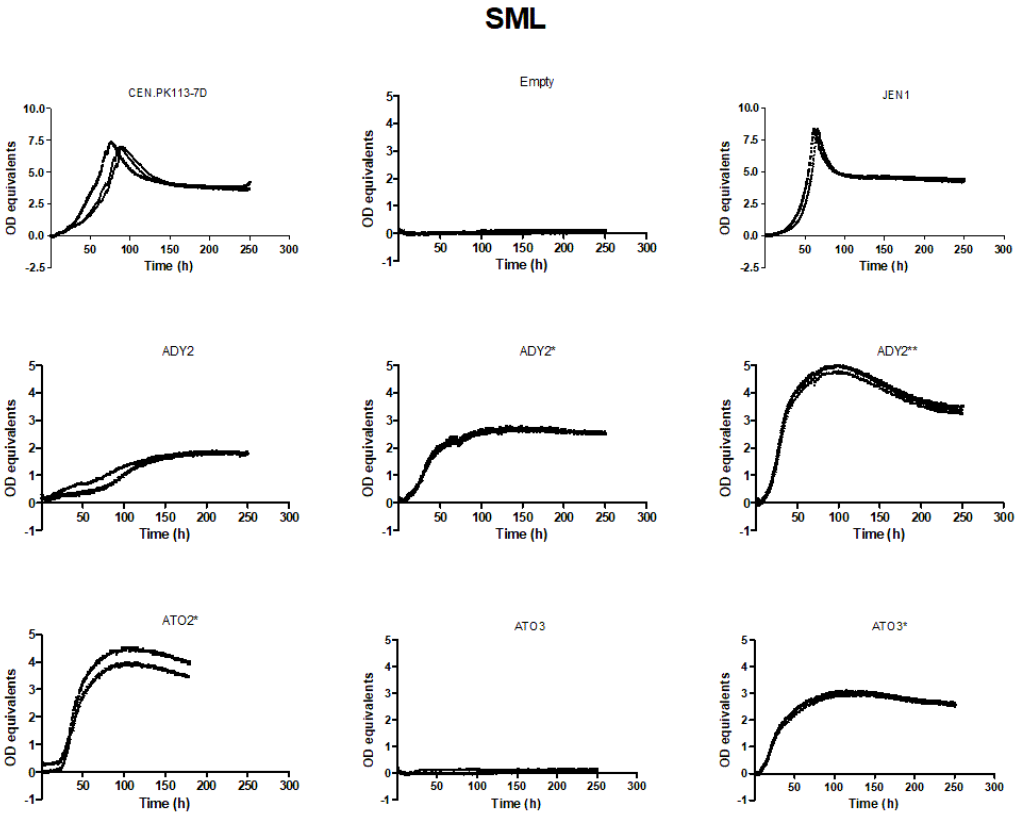


Supplementary Figure 3. Growth rates of S. cerevisiae reference strain CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty centromeric vector or a centromeric vectors containing the indicated organic acid transporter gene. Growth on SMA medium set at pH5. Bars and error bars represent the average and standard deviation of three independent experiments. Empty: empty plasmid. ADY2*: ADY2^{C755G} allele. ADY2**: ADY2^{C655G} allele. ATO2*: ATO2^{T653C} allele. ATO3*: ATO3^{T284C} allele.

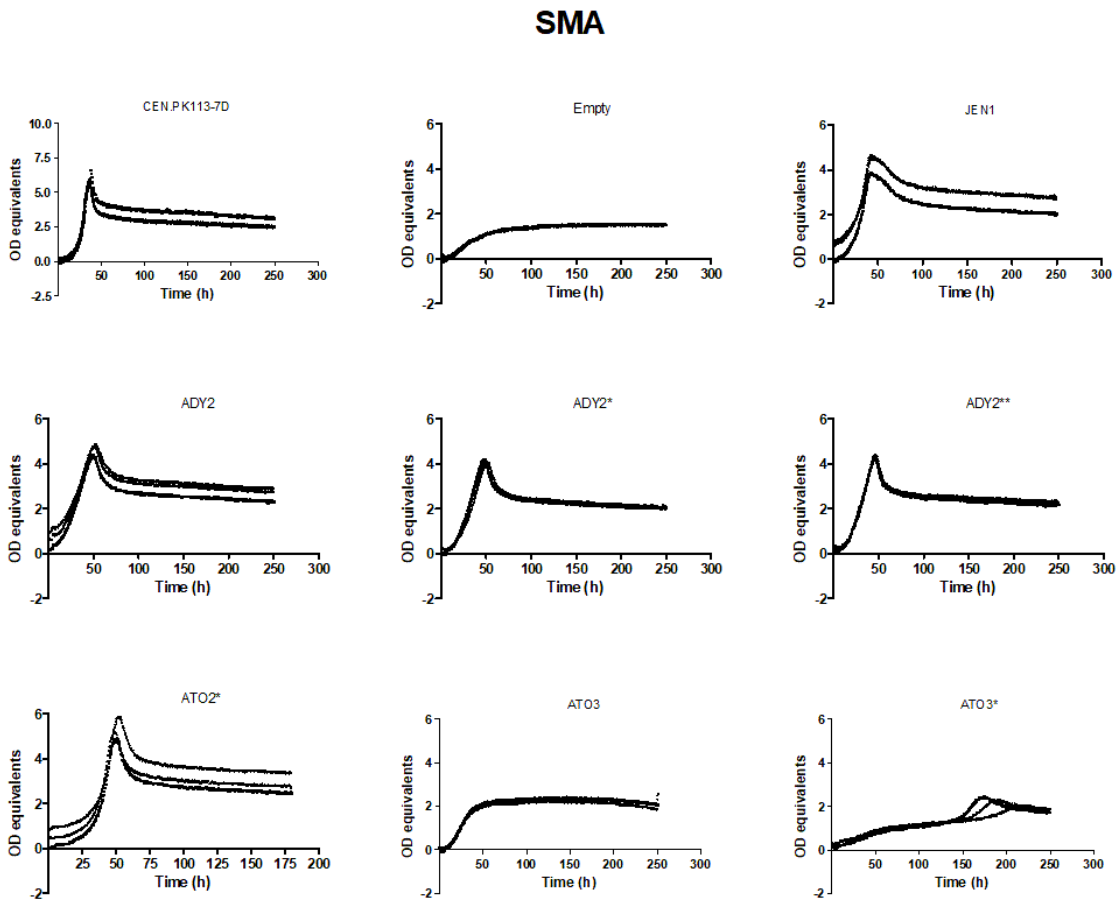


Supplementary Figure 4. Growth profiles in synthetic medium (pH 5.0) with glucose as the sole carbon source of CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty multicopy

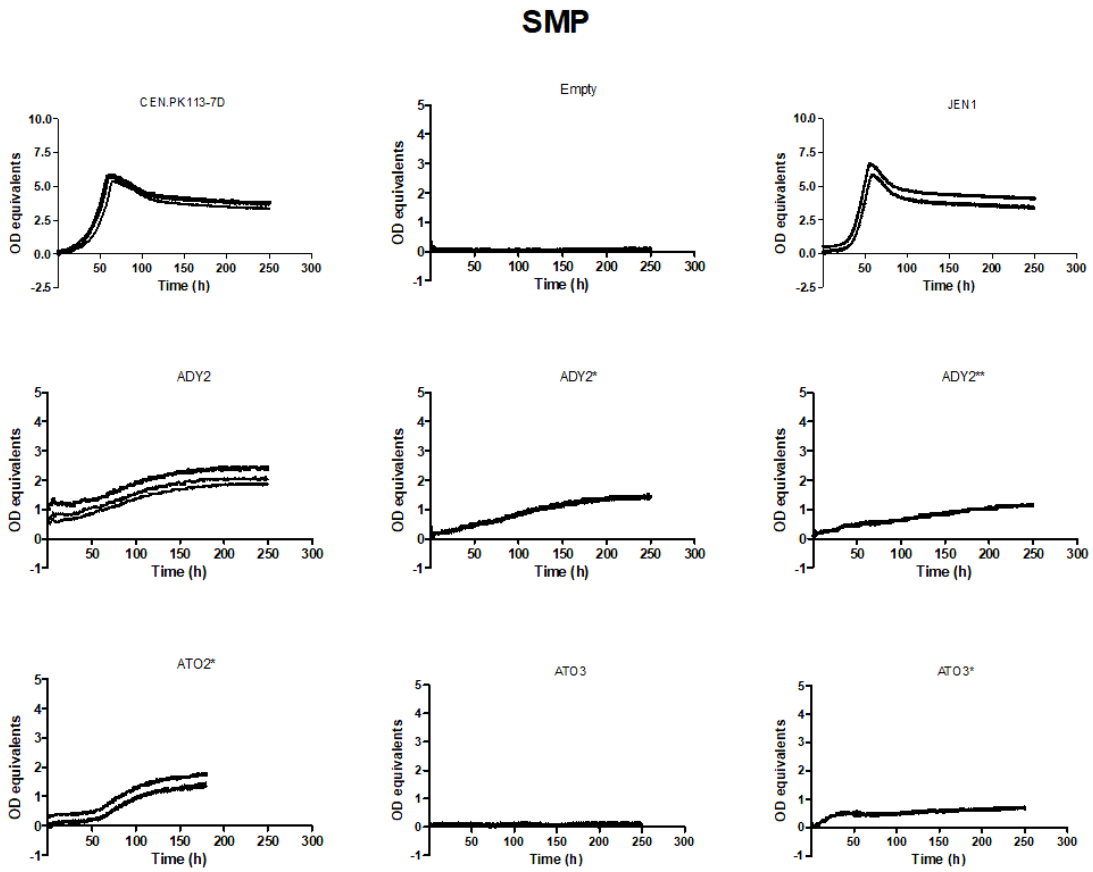
vector or a multicopy vector containing the indicated organic acid transporter gene. Empty: empty plasmid. ADY2*: *ADY2*$^{c.755G}$ allele. ADY2**: *ADY2*$^{c.655G}$ allele. ATO2*: *ATO2*$^{c.653C}$ allele. ATO3*: *ATO3*$^{c.284C}$ allele.



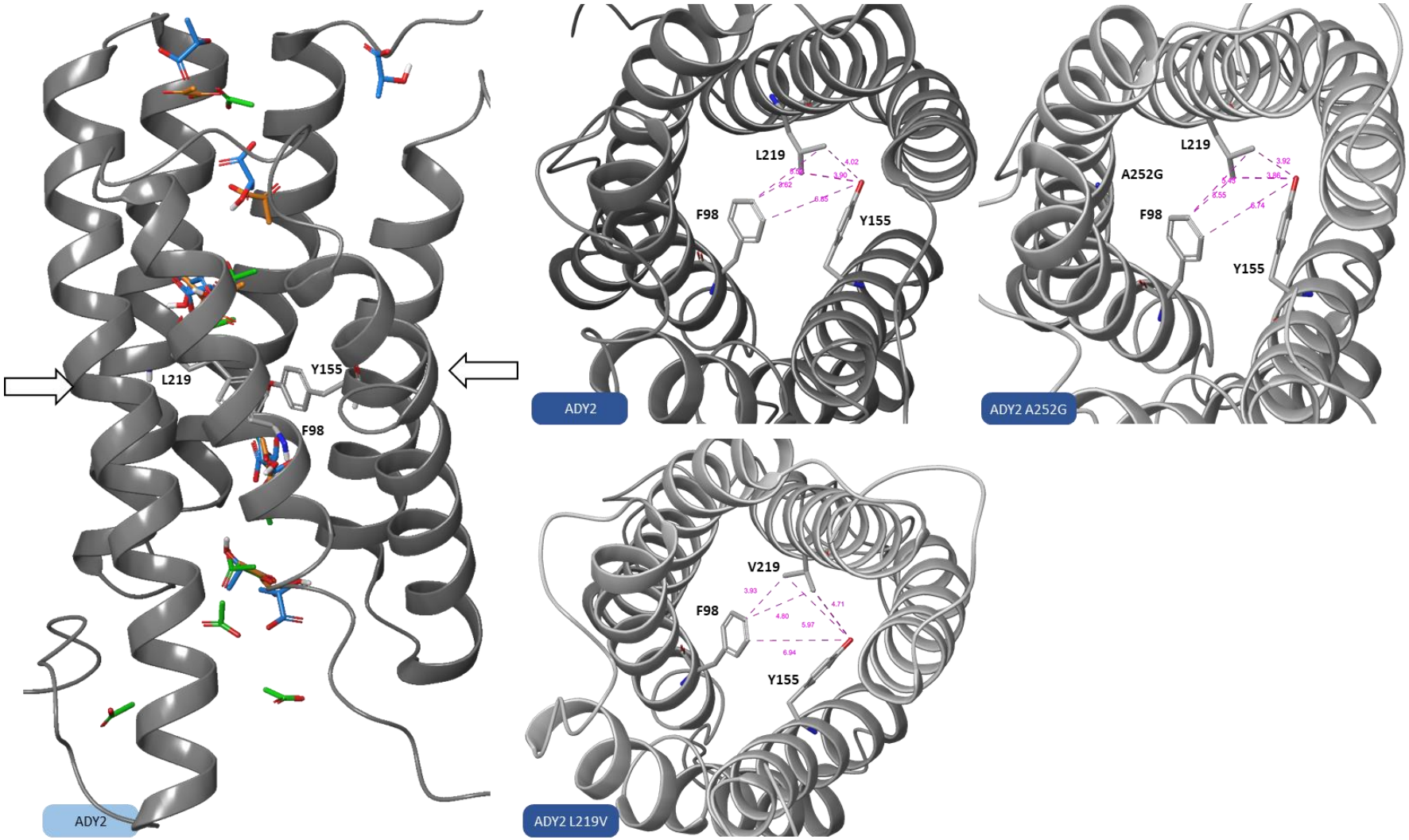
Supplementary Figure 5. Growth profiles in synthetic medium (pH 5.0) with lactate as the sole carbon source of CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty multicopy vector or a multicopy vector containing the indicated organic acid transporter gene. Empty: empty plasmid. ADY2*: *ADY2*$^{c.755G}$ allele. ADY2**: *ADY2*$^{c.655G}$ allele. ATO2*: *ATO2*$^{c.653C}$ allele. ATO3*: *ATO3*$^{c.284C}$ allele.



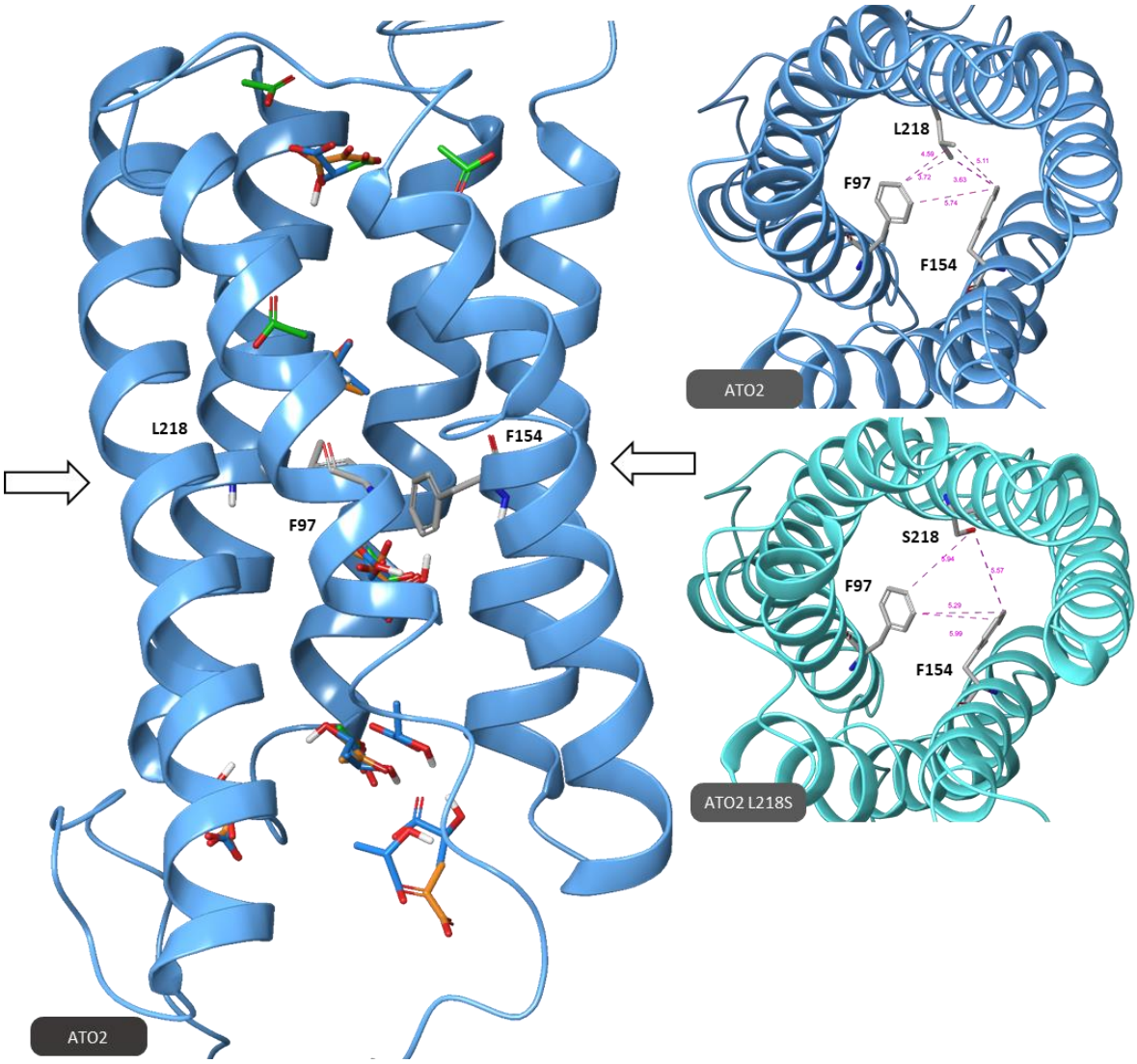
Supplementary Figure 6. Growth profiles in synthetic medium (pH 6.0) with acetate as the sole carbon source of CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty multicopy vector or a multicopy vector containing the indicated organic acid transporter gene. Empty: empty plasmid. ADY2*: *ADY2*^{c755G} allele. ADY2**: *ADY2*^{c655G} allele. ATO2*: *ATO2*^{t653C} allele. ATO3*: *ATO3*^{t284C} allele.



Supplementary Figure 7. Growth profiles in synthetic medium (pH 5) with pyruvate as the sole carbon source of CEN.PK113-7D and the 25-transporter deletion strain IMX2488 expressing an empty multicopy vector or a multicopy vector containing the indicated organic acid transporter gene. Empty: empty plasmid. ADY2*: *ADY2*^{c755G} allele. ADY2**: *ADY2*^{c655G} allele. ATO2*: *ATO2*^{t653C} allele. ATO3*: *ATO3*^{t284C} allele.



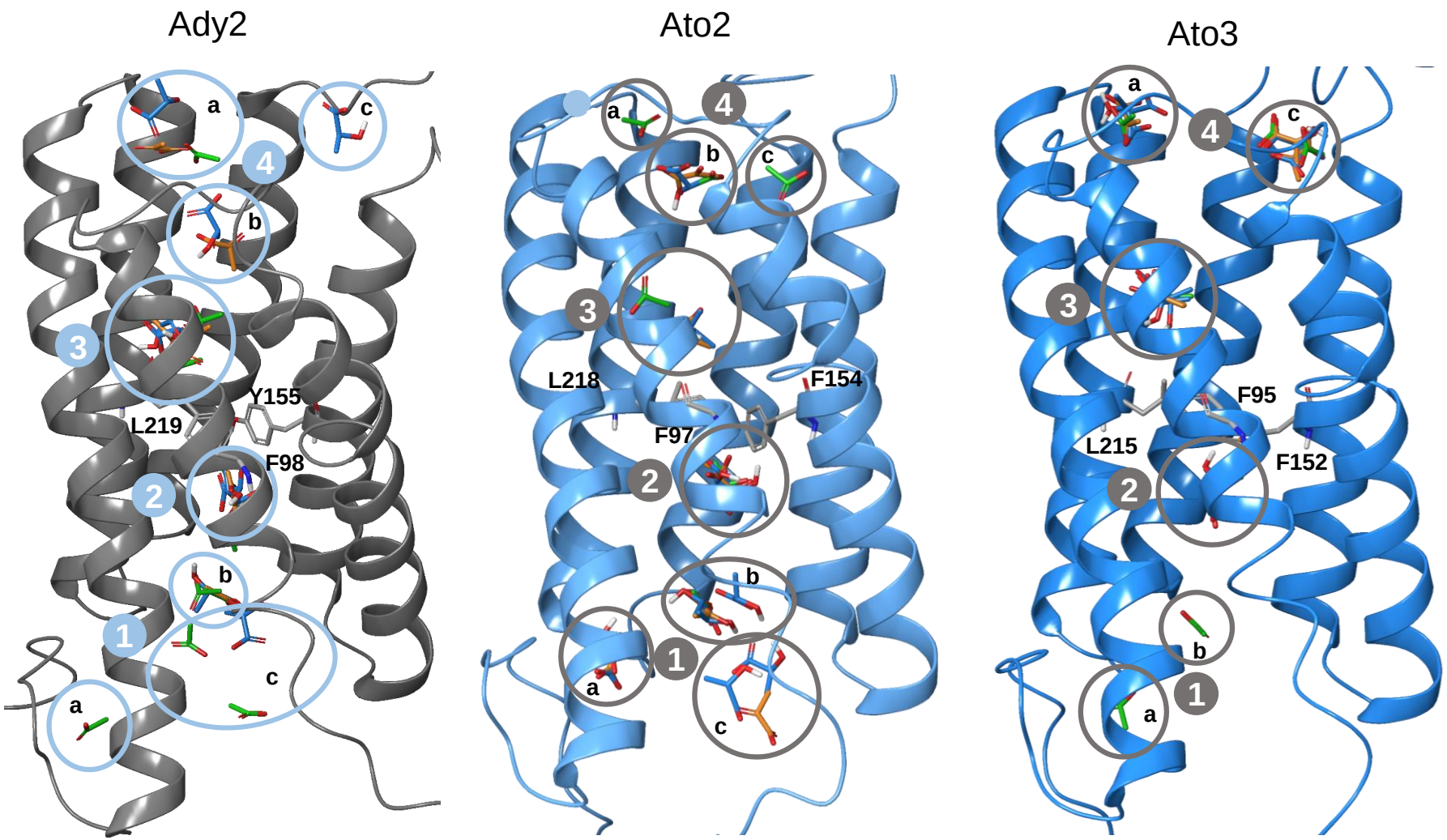
Supplementary Figure 8. 3D model of Ady2, Ady2 C755G and Ady2 C655G alleles. Left, side view of Ady2 where arrows indicate the hydrophobic constriction site. Ady2 binding sites for acetate (green ligand), lactate (blue ligand) and pyruvate (orange ligand) are presented. Right, top view of Ady2, Ady2 C755G and Ady2 C655G alleles. The amino acids involved in the hydrophobic constriction site are shown. Purple lines and values indicate distances (in Å) between different anchor points of amino acids.



Supplementary Figure 9. 3D model of Ato2 and Ato2 L218S allele. Left, side view of Ato2 where the arrows indicate the constriction site. Ato2 binding sites for acetate (green ligand), lactate (blue ligand) and pyruvate (orange ligand) are presented. Right, top view of either Ato2 or Ato2 L218S allele. The amino acids involved in the constriction site are shown. Purple lines and values indicate distances (in Å) between different anchor points of different amino acids.

Supplementary Table 3. Average of the binding affinity values [kcal/mol] calculated with PyRx software for the docking of ligand in the predicted structures of wildtype and mutated Ady2, Ato2 and Ato3.

3D-Protein templates	Average of binding affinities (kcal/mol) at different binding sites								
	Acetate								
	1			2	3	4			
	a	b	c			a	b	c	
Ady2	-2,1	-2,4	-2,3	-2,7	-3,1	-2,7	-	-	
Ady2 A252G	-2,3	-2,2	-2,1	-3,1	-3,1	-2,6	-2,7	-	
Ady2 L219V	-2,4	-	-	-3,0	-3,0	-2,5	-2,7	-	
Ato2	-2,5	-2,6	-	-3,1	-2,9	-2,8	-2,9	-3	
Ato2 L218S	-2,7	-	-	-2,7	-3,1	-3,3	-2,6	-2,5	
Ato3	-2,3	-2,2	-	-2,9	-3,0	-2,2	-	-2,4	
Ato3 F95S	-2,6	-2,4	-	-3,0	-2,9	-2,4	-2,4	-2,4	
	Lactate								
	1			2	3	4			
	a	b	c			a	b	c	d
Ady2	-	-3,1	-3,0	-3,6	-4,3	-3,5	-3,1	-3,1	-
Ady2 A252G	-2,9	-2,8	-2,7	-3,7	-4,4	-3,3	-3,4	-	-2,1
Ady2 L219V	-3,4	-	-	-3,9	-4,4	-	-3,8	-	-
Ato2	-3,2	-3,1	-3,2	-3,9	-3,8	-	-2,9	-	-
Ato2 L218S	-3,5	-	-	-3,8	-4,2	-3,6	-	-3,4	-
Ato3	-	-	-	-3,4	-3,8	-3,1	-	-3,2	-
Ato3 F95S	-3,3	-	-	-4,2	-4,0	-3,2	-3,2	-3,2	-
	Pyruvate								
	1			2	3	4			
	a	b	c			a	b	c	
Ady2	-	-3,1	-	-3,7	-4,2	-3,2	-3,3	-	
Ady2 A252G	-3,0	-	-2,7	-3,9	-4,3	-3,3	-3,3	-	
Ady2 L219V	-3,2	-	-	-4,0	-4,2	-	-3,5	-	
Ato2	-3,3	-3,3	-3,1	-3,9	-4,1	-	-3,6	-	
Ato2 L218S	-3,5	-	-	-3,9	-4,3	-	-	-3,3	
Ato3	-	-	-	-3,6	-4,0	-3,0	-	-3,2	
Ato3 F95S	-3,4	-	-	-4,2	-3,9	-	-3,4	-	



Supplementary figure 10. Molecular docking sites of acetate (green ligand), lactate (blue ligand) and pyruvate (orange ligand) in the predicted structure of Ady2, Ato2 and Ato3, identified using Autodock Vina.

Supplementary Table 4. Residues identified by molecular docking analysis as being involved in the establishment of strong interactions with the indicated ligand. Initials IL stands for ionic ligation.

3D-Protein templates	Acetate	Lactate	Pyruvate
Ady2	N89, N145, S208, T222, H230 (IL), T262	N89, T102, S106, Q133, E140, N145, D182, S208, T222, G229, H230, T238	N145, S208, T222, H230, T238
Ady2 L219V	Y176, T209, T222, H230, T238, N255	T102, N109, Y176, T209, T222, H230, N255	N109, Y166, Y176, T222, T209, H230, N255
Ady2 A252G	N109, R111(IL), Q133, N145, Y176, T209, T222, H230 (IL), T238, N255,	K86, T102, N109, R111(IL), Q133, N145, Y176, T209, T222, H230(IL), T238, N255, G259, T262	Y176, R111(IL), T209, T238, T222, H230, N255, T262
Ato2	Q132, N229, K176(IL), Y175, T237, T221, S207	T101, Q132, N144, S207, T221, N229, T237, N254, T261, R262(IL)	H84, K85(IL), Q132, S207, T221, N229, T237, T261, N254
Ato2 L218S	K176, S180, T208, S218, T221, N229, T237	S105, Q132, D173, A174, K176, D177, S180, L182, T208, T221	D173, K176, S180, L182, T208, T221
Ato3	N106, Y173, T205, T218, M271	C99, N106, G162, Y173, T218, N230	N106, Y173, T218
Ato3 F95S	S95, N106, T218, K234, S258	S95, N106, C158, Y159, T218, D229, N231, K234, S244, S251	S95, N106, T218, K234, S258

CHAPTER IX

General discussion & Future perspectives

GENERAL DISCUSSION

As highlighted throughout this project, organic acids play a central role in cell metabolism regulating several cellular mechanisms and overall cell homeostasis. In addition, compounds like carboxylic acids have a wide versatility, as chemical-building blocks, and applicability in several industrial segments particularly in the food, pharmaceutical and chemical sectors [1]. Currently there is a great demand to obtain such products in a more sustainable way, as an alternative to the traditional petrochemical-based methods. Several processes are being implemented at the industrial level relying on their production via microbial fermentation using yeasts and bacteria [2]. Nevertheless, several constraints still have to be overcome, as the majority of these molecules are reported to be toxic to cell factories leading to cytosol acidification, affecting negatively the productivity and titer of bioprocesses, or even conducting to membrane disruption [3-5]. Thus, it becomes necessary to optimize existing cell factories and explore novel expression hosts, as more efficient and robust microbial cell factories are needed achieve an economically viable industrial production of carboxylic acids.

One of the microorganisms emerging as a promising biorefinery is the yeast *Cyberlindnera jadinii,* as reviewed in Chapter II. Its increased tolerance level to adverse conditions present in industrial fermentation [6], turn this yeast an ideal platform for biotechnological processes. In addition, *C. jadinii* has been widely explored as a source of single-cell protein while being able to synthetize several important compounds, e.g. vitamins, organic acids and proteins [6-9]. In this review, we highlighted ecology, morphology and physiology, taxonomy, life cycle, and genome variation of *C. jadinii* yeast strains. We also present the molecular tools already developed for its genetic manipulation and highlight the need for more efficient genomic manipulation systems. The emerging biotechnological potential of this yeast includes therapeutic applications, the production of food supplements using cost-effective carbon sources, improvement of cosmetic and health care products, and applications in agriculture and wine making industry. Finally, an overview of membrane transporters characterized in this yeast is also presented since they play a pivotal role in cell homeostasis, regulating substrates consumption, metabolic fluxes between organelles and metabolite export [10, 11].

In Chapter III we have reviewed membrane transporters used for the bioproduction of organic acids. The expression of organic acid exporters is reported to improve the microorganism's tolerance to organic acids and increase extracellular titers [12, 13]. The most relevant and emerging cell factories for the production of organic acids are presented in this review, along with the engineering strategies applied to turn them into efficient producers of this family of compounds. The transporter engineering approaches were also highlighted, especially in *Saccharomyces cerevisiae*. Albeit the progresses made on this field,

the redesigning and engineering of optimized membrane transporters for industrial organic acid production remains in an early stage as the functional and structural characterization of transporters is still a laborious process [14].

The experimental work of this PhD thesis was mainly we focused in the identification of novel carboxylic acid transporters and its functional and structural characterization. Next, the most prominent conclusions of the work, are presented.

- **Non-*Saccharomyces* yeasts may become attractive platforms for the bioproduction of carboxylic acids. *C. jadinii* strains display distinct physiological and genomic traits.**

In Chapter IV we explored ten wild yeast species isolated from biowastes. Morphological and physiological traits of the wild species were studied and the yeasts *Pichia kudriavzevii*, *Candida tropicalis* and *Cyberlindnera jadinii* yeasts were identified as promising expression hosts for bioproduction of organic acids. The yeast *C. jadinii* was selected for further analysis where the physiological and genetic diversity of distinct wild isolates and laboratorial strains was explored. These results highlight a set of promising yeasts to be further investigated considering their robust metabolic capabilities.

- ***C. jadinii* transporter proteins from the SHS, AceTr, DASS and SSS families mediate the uptake of acetate, lactate, succinate and citrate in *S. cerevisiae* IMX1000.**

In this work, we made an extensive characterization of the predicted *C. jadinii* transportome, and the most promising carboxylate transporters were functionally characterized by heterologous expression in the *S. cerevisiae* IMX1000 strain. A total of sixteen transporters were uncovered belonging to six transporter families, AceTr (TC 2.A.96), SHS (TC 2.A.1.12), SSS (TC 2.A.21; SLC member 5), TDT (TC 2.A.16), DASS (TC 2.A.47; SLC member 13) and MCT (TC 2.A.1.13; SLC member 16). The carboxylate transporters in *S. cerevisiae* strain encode sixteen acetate transporters, fourteen lactate transporters (CjSlc16; CjSlc5; CjSlc13-1; all CjAto and CjJen), five succinate transporters (CjAto2; CjAto5; CjJen6; CjSlc5 and CjSlc13-1) and four citrate transporters (CjAto5; CjJen6; CjSlc5 and CjSlc13-1). Some of these transporters have a distinct substrate specificity than previously identified homologs. For instance, CjAto2 and CjAto5 are the only AceTr yeast members able to transport succinate so far described, and for the first time a citrate transporter, CjAto5, was functionally characterized in this family. In addition, functional characterized CjSlc5 homologs are described as amino acid transporters.

The modifications detected in specific amino acid residues may be responsible for gain of function for succinate and citrate uptake observed in the CjAto2p, CjAto5p, CjJen6p and CjSlc5p transporters. Since

Jen1 and Ato1 transporters are involved in the efflux of lactate in a *S. cerevisiae* lactic acid producing strain [15], a similar export activity might be present in *C. jadinii* homolog proteins. In view of the already uncovered protein-transporter systems previously reposted in *C. jadinii* [16, 17], we postulate that CjAto1, CjAto3-4 and CjJen1-5 can encode the proton-symporter systems accepting several monocarboxylates. The transporters CjAto5p, CjJen6p, CjSlc13-1p and CjSlc5p may correspond to the facilitated-diffusion systems working as general organic permeases able to accept mono, di and tricarboxylic acids. However, further studies are necessary to reveal the identity of the transporters functionally characterized in *C. jadinii*, i.e. expression patterns, energetics of transporter systems.

The phylogenetic reconstructions of SHS, SSS and DASS family members revealed that SHS and DASS homologs are present in prokaryotic and eukaryotic organisms, whereas SSS members were only found in eukaryotic organisms. The presence of SHS and DASS members across the microbial diversity suggests the occurrence of a specific and essential role of theses members in microbial evolution.

- **The AnCexAp membrane protein transports both citrate and isocitrate in *S. cerevisiae*. Structure function studies reveal critical residues involved in citrate transport.**

The functional and structural characterization of the AnCexA by heterologous expression in *S. cerevisiae* is reported in Chapter VI. In this study, we characterized the import of citrate as a low-affinity system, being highly specific for citric and isocitric acids at pH 5.5. AnCexA is a member of DHA1 transporter family [18]. The phylogenetic analysis of AnCexA homologs revealed their distribution across the microbial diversity, evidencing a high prevalence in basidiomycetes and ascomycetes. A site-directed mutagenesis approach, based on the existence of conserved domains, as well as 3D model prediction and docking studies, identified critical residues for citrate transport. These findings can provide leads on how to engineer AnCexA transporter for improved bioproduction of citrate.

- **The signature motif NPAPLG(M/S) of AceTr family is essential for substrate uptake. The evolutionary analysis of AceTr family members reveals a ubiquitous occurrence in fungi. MaAcep and YlGpr1p are acetate permeases.**

AceTr family members are present in eukaryotic and prokaryotic organisms with a strong incidence in fungi, suggesting the relevance of AceTr members in fungal evolution.

The functionality of several AceTr family members was explored in this work. Ato2p from *S. cerevisiae* and Gpr1p from *Yarrowia lipolytica* were characterized by heterologous expression in *S. cerevisiae jen1*Δ

*ady2*Δ strain and Acep from *Methanosarcina acetivorans* was expressed in *Escherichia coli* triple yaaH/actP/ lldP mutant strain. YlGpr1p and MaAcePp were characterized as acetate transporters.
A structure function study revealed that the NPAPLGL(M/S) conserved motif of AceTr members, EcSatP and ScAto1, is essential for substrate transport, once mutations in this domain suppressed transport activity. Membrane targeting and stability was found not to be affected in these Ato1 mutant alleles.

- **Adaptive laboratory evolution lead to the selection of key amino acid residues from *Sc*Ato2 and *Sc*Ato3 carboxylic acid transporters involved in substrate specificity.**

Chapter VIII presents the characterization of mutated versions of Ato transporters resulting from a directed evolution strategy on lactate, which were able to efficiently catalyze the import of lactic acid in *S. cerevisiae*. Whole-genome resequencing and reverse engineering identified the transporter alleles Ato2 L218S and Ato3 F95S, as being involved in lactic acid transport. The Ato2 L218S mutation also led to the gain of function for acetic acid transport, as the native ScAto2 was not able to transport this acid [19]. The growth rate displayed by Ato3 F95S on media containing lactic acid as sole carbon source was the highest growth rate ($0.15 \pm 0.01\ h^{-1}$) reported so far.
The *in silico* analysis of transporter protein 3D models showed that several mutations resulted in the widening of the narrowest hydrophobic constriction of the pore. This observation, in combination with an increased binding affinity predicted for lactic acid in the binding sites adjacent to this constriction site, can be responsible for the improved lactic acid transport in the evolved proteins
Overall, the presented data on transporter structure and function led to the identification of relevant residues that dictate substrate specificity of *S. cerevisiae* carboxylic acid transporters that could potentially aid in future rational engineering and annotation of additional proteins involved in organic acid transport.

FINAL REMARKS AND FUTURE STEPS

Our studies support the relevance of carboxylate transporters in the biotech industry, expand the list of microbial carboxylate transporters characterized and provide novel phylogenetic, functional and structural insights on members from different transporter families. Altogether, we studied members from seven transporter families, namely AceTr (in Chapters I, V, VII and VIII), SHS, SSS, TDT, DASS, MCT (all in Chapter V) and DHA1 (in Chapter VI), and present a set of studies that demonstrate their role as functional carboxylate transporters.

The future steps will include the functional expression of these carboxylate transporters in microbial strains engineered for the carboxylic acid production. The overexpression of promising carboxylate transporters will contribute to the improvement of biorefineries by increasing carboxylic acid titers, decreasing production costs, turning carboxylic acid bioproduction a competitive alternative to petrochemical-derived products.

The inexistence of an extensive knowledge on the metabolism, regulatory networks, and transport mechanisms of *C. jadinii* hampers the utilization of this yeast in a wider range of biotechnological applications. The genomic characterization of several *C. jadinii* strains will certainly reveal the genetic features underlying the existing interspecies variability, in particular between its teleomorph (*C. jadinii*) and anamorph (*C. utilis*) states. It is also important to determine the ploidy level of the different strains, as they are predicted to vary between 2n-5n. Another important aspect will be the determination of interspecies variation regarding carboxylate transporters, and if present to correlate them with the existing phenotypic differences for the utilization of carboxylic acids as sole carbon and energy sources.

Also, the evaluation of gene expression patterns by qRT-PCR of the different transporters here uncovered, or even a RNAseq approach to determine the expression profile of transporter proteins when cells are grown in different carboxylic acids, will help to characterize the *C. jadinii* transportome. Ultimately, the disruption of genes in *C. jadinii* can unveil the overall contribution of each transporter protein for carboxylate transport. One of the procedures to efficiently perform such deletions is the CRISPR/Cas system that was recently reported for this yeast. To deepen the knowledge on structure-function, site-directed mutagenesis of the residues highlighted throughout the study is also necessary to determine their role in the substrate recognition and protein activity. Further studies are also needed to clarify the energetics associated to the identified transporters. To evaluate the role of these transporters as exporters, carboxylate-producer strains will be engineered to express selected proteins. The bioreactor cultivations coupled with HPLC measurements will allow the assessment of the produced metabolites e.g. lactate, succinate and citrate.

Moreover, across the other uncovered non-*Saccharomyces* yeasts, *C. tropicalis* and *P. kudriavzevii*, it is still missing a deep knowledge over their morphological, physiological and genomic traits for a comparative analysis between the wild yeast isolates against correspond collection strains. The genome sequencing of these latter species will allow the identification of genomic variations occurred due to their adaptation to acidic environments. According to the results here obtained the two yeast species can become promising hosts for the production of organic acids.

Meanwhile, the molecular transport mechanism involved in AnCexAp citrate uptake/efflux it is still unconfirmed. The use of ionophores, such as valinomycin or monensin, can clarify the citrate transport dependency on the electrochemical potential. We also propose a detailed characterization of AnCexA mutants by measuring the transport activity, protein localization and citrate production. These will be of particular interest in mutants that present a growth lower than the wild-type CexA, that could result from an increased export capacity. Inhibition assays can confirm modifications in substrate specificity. Furthermore, to elucidate whether the growth phenotypes associated to a loss of function are a result of an incorrect protein localization or lack of transporter function, protein localization will be evaluated in GFP-tagged mutant alleles.

Future work on AceTr family members, e.g ScAto1, EcSatp and YlGpr1, should be extended to other conserved residues to determine their contribution in substrate specificity, protein localization stability and energetics. This approach should also be applied to evolved lactic acid transporters ScAto2 L218S and ScAto3 F95S.

REFERENCES

1. Alonso, S., Rendueles, M.,Díaz, M., *Microbial production of specialty organic acids from renewable and waste materials.* Critical reviews in biotechnology, **2015**. 35(4): p. 497-513
2. Chen, Y.,Nielsen, J., *Biobased organic acids production by metabolically engineered microorganisms.* Current opinion in biotechnology, **2016**. 37: p. 165-172
3. Legras, J., Erny, C., Le Jeune, C., Lollier, M., Adolphe, Y., Demuyter, C., Delobel, P., Blondin, B.,Karst, F., *Activation of two different resistance mechanisms in Saccharomyces cerevisiae upon exposure to octanoic and decanoic acids.* Applied and Environmental Microbiology, **2010**. 76(22): p. 7526-7535
4. Abbott, D., Knijnenburg, T., De Poorter, L., Reinders, M., Pronk, J.,Van Maris, A., *Generic and specific transcriptional responses to different weak organic acids in anaerobic chemostat cultures of Saccharomyces cerevisiae.* FEMS Yeast Res, **2007**. 7(6): p. 819-833
5. Guan, N.,Liu, L., *Microbial response to acid stress: mechanisms and applications.* Applied microbiology and biotechnology, **2020**. 104(1): p. 51-65
6. Buerth, C., Heilmann, C., Klis, F., de Koster, C., Ernst, J.,Tielker, D., *Growth-dependent secretome of Candida utilis.* Microbiology, **2011**. 157(9): p. 2493-2503, 10.1099/mic.0.049320-0
7. Bekatorou, A., Psarianos, C.,Koutinas, A., *Production of food grade yeasts.* Food Technology and Biotechnology, **2006**. 44(3): p. 407-415

8. Buerth, C., Tielker, D.,Ernst, J., *Candida utilis and Cyberlindnera (Pichia) jadinii: yeast relatives with expanding applications.* Applied microbiology and biotechnology, **2016**. 100(16): p. 6981-6990, https://doi.org/10.1007/s00253-016-7700-8
9. Miura, Y., Kettoku, M., Kato, M., Kobayashi, K.,Kondo, K., *High level production of thermostable alpha-amylase from Sulfolobus solfataricus in high-cell density culture of the food yeast Candida utilis.* Journal of molecular microbiology and biotechnology, **1999**. 1(1): p. 129-134
10. Kell, D.B., Swainston, N., Pir, P.,Oliver, S.G., *Membrane transporter engineering in industrial biotechnology and whole cell biocatalysis.* Trends in Biotechnology, **2015**. 33(4): p. 237-246, 10.1016/j.tibtech.2015.02.001
11. Van der Hoek, S.,Borodina, I., *Transporter engineering in microbial cell factories: the ins, the outs, and the in-betweens.* Current opinion in biotechnology, **2020**. 66: p. 186
12. Abbott, D., Zelle, R., Pronk, J.,Van Maris, A., *Metabolic engineering of Saccharomyces cerevisiae for production of carboxylic acids: current status and challenges.* FEMS Yeast Res, **2009**. 9(8): p. 1123-1136
13. Liu, P.,Jarboe, L., *Metabolic engineering of biocatalysts for carboxylic acids production.* Computational and structural biotechnology journal, **2012**. 3(4): p. e201210011
14. Soares-Silva, I., Ribas, D., Sousa-Silva, M., Azevedo-Silva, J., Rendulić, T.,Casal, M., *Membrane transporters in the bioproduction of organic acids: state of the art and future perspectives for industrial applications.* FEMS microbiology letters, **2020**
15. Pacheco, A., Talaia, G., Sá-Pessoa, J., Bessa, D., Gonçalves, M.J., Moreira, R., Paiva, S., Casal, M.,Queirós, O., *Lactic acid production in Saccharomyces cerevisiae is modulated by expression of the monocarboxylate transporters Jen1 and Ady2.* FEMS Yeast Res, **2012**. 12(3): p. 375-381
16. Cássio, F.,Leão, C., *Low-and high-affinity transport systems for citric acid in the yeast Candida utilis.* Appl. Environ. Microbiol., **1991**. 57(12): p. 3623-3628
17. Cássio, F.,Leão, C., *A comparative study on the transport of L (-) malic acid and other short-chain carboxylic acids in the yeast Candida utilis: Evidence for a general organic acid permease.* Yeast, **1993**. 9(7): p. 743-752
18. Steiger, M.G., Rassinger, A., Mattanovich, D.,Sauer, M., *Engineering of the citrate exporter protein enables high citric acid production in Aspergillus niger.* Metabolic engineering, **2019**. 52: p. 224-231
19. Ribas, D., Soares-Silva, I., Vieira, D., Sousa-Silva, M., Sá-Pessoa, J., Azevedo-Silva, J., Viegas, S., Arraiano, C., Diallinas, G., Paiva, S., Soares, P.,Casal, M., *The acetate uptake transporter family motif "NPAPLGL (M/S)" is essential for substrate uptake.* Fungal Genetics and Biology, **2019**. 122: p. 1-10