



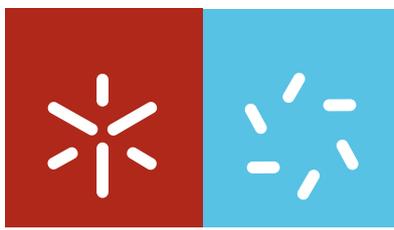
Universidade do Minho

Escola de Ciências

Sara Catarina Costa Laranjeira

**Stress-related genes in *Arabidopsis thaliana*:
investigating the role of squalene epoxidases
in sterol biosynthesis and *EGY3* as a putative
plastidial heat stress determinant**





Universidade do Minho

Escola de Ciências

Sara Catarina Costa Laranjeira

**Stress-related genes in *Arabidopsis thaliana*:
investigating the role of squalene epoxidases
in sterol biosynthesis and *EGY3* as a putative
plastidial heat stress determinant**

Tese de Doutoramento
Doutoramento em Ciências
Especialidade em Biologia

Trabalho realizado sob a orientação de:

Prof. Doutor Rui Manuel Tavares
Prof. Doutor Miguel Botella Mesa
Prof. Doutor Herlânder Azevedo

DECLARAÇÃO

Nome: Sara Catarina Costa Laranjeira

Endereço electrónico: saranjeira@bio.uminho.pt

Telefone: 253604310

Título da tese de doutoramento:

STRESS-RELATED GENES IN *ARABIDOPSIS THALIANA*: INVESTIGATING THE ROLE OF SQUALENE EPOXIDASES IN STEROL BIOSYNTHESIS AND *EGY3* AS A PUTATIVE PLASTIDIAL HEAT STRESS DETERMINANT

Orientadores:

Prof. Doutor Rui Manuel Tavares (Departamento de Biologia, Escola de Ciências, Universidade do Minho, Portugal)

Prof. Doutor Miguel Botella Mesa (Departamento de Biología Molecular y Bioquímica, Universidad de Málaga, Espanha)

Prof. Doutor Herlânder Azevedo (Departamento de Biologia, Escola de Ciências, Universidade do Minho, Portugal)

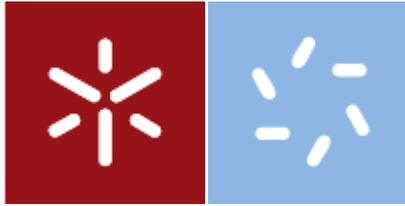
Ano de conclusão: **2011**

Designação do Ramo de Conhecimento do Doutoramento: **Ciências**, especialidade em **Biologia**.

É AUTORIZADA A REPRODUÇÃO PARCIAL DESTA TESE/TRABALHO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Universidade do Minho, ___/___/_____

Assinatura: _____



Universidade do Minho
Escola de Ciências
Departamento de Biologia

Sara Catarina Costa Laranjeira

STRESS-RELATED GENES IN *ARABIDOPSIS THALIANA*: INVESTIGATING
THE ROLE OF SQUALENE EPOXIDASES IN STEROL BIOSYNTHESIS AND
EGY3 AS A PUTATIVE PLASTIDIAL HEAT STRESS DETERMINANT

Tese de Doutoramento
Doutoramento em Ciências
Especialidade em Biologia

Trabalho realizado sob a orientação de:
Prof. Doutor Rui Manuel Tavares
Prof. Doutor Miguel Botella Mesa
Prof. Doutor Herlânder Azevedo

Braga, Setembro de 2011

DECLARAÇÃO

Nome: Sara Catarina Costa Laranjeira

Endereço electrónico: saranjeira@bio.uminho.pt

Telefone: 253604310

Número do Bilhete de Identidade: 11771094

Título da tese de doutoramento:

STRESS-RELATED GENES IN *ARABIDOPSIS THALIANA*: INVESTIGATING THE ROLE OF SQUALENE EPOXIDASES IN STEROL BIOSYNTHESIS AND *EGY3* AS A PUTATIVE PLASTIDIAL HEAT STRESS DETERMINANT

Orientadores:

Prof. Doutor Rui Manuel Tavares (Departamento de Biologia, Escola de Ciências, Universidade do Minho, Portugal)

Prof. Doutor Miguel Botella Mesa (Departamento de Biología Molecular y Bioquímica, Universidad de Málaga, Espanha)

Prof. Doutor Herlânder Azevedo (Departamento de Biologia, Escola de Ciências, Universidade do Minho, Portugal)

Ano de conclusão: **2011**

Designação do Ramo de Conhecimento do Doutoramento: **Ciências**, especialidade em **Biologia**.

É AUTORIZADA A REPRODUÇÃO PARCIAL DESTA TESE/TRABALHO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Universidade do Minho, ____/____/____

Assinatura: _____

Acknowledgments

In a period of tribulation, a mist of joy and fear arise, but as always a man can always conquer his soul and make his own history. Writing a thesis is more than compiling four years of study, of experiments, of joys and sorrows, is going back in time and realizing that no matter what end it takes, it ends in accomplishment. Maybe four years are not enough to accomplish what we wish in the beginning, but learning is the key and if you'll advance in knowledge, scientific thought, people's interaction, dynamic and most of all on accepting the bad results and trying always for the better ones, you know you've made your path.

In this path I must thank so many people, four years are a lot of time, in which so many people came and went into my life. First of all I must thank my supervisors: Rui Tavares, Miguel Botella and Herlânder Azevedo for giving me the opportunity of applying to a fellowship for a PhD with them, giving me a place in their work groups, and making me feel at home. Without them these four years had not existed. I've had a mixed PhD that permitted me to go abroad, broadening even more my perspective and learning.

In Braga, with Rui and Herlânder, they gave me the haven for my ship, it was there that I learned the few steps of a baby in a lab, where I learned almost all the techniques I know today, so for that thank you for all your time spent with me. Thank you for inspiring me hope, clinging in the expectations we all had at the beginning, accepting that things are not just black and white, there are definitely grey areas. Thank you for implementing so much in the Lab and giving me the opportunity to see it all grow and be educated by it. Thank you for the support at all levels, even when time is impossible to stretch and there was a tiny window where I and my work could be. Thank you for all the help even in the darkest hours.

In Málaga, I learned to move in a different way, with a different mind, thank you Miguel for your amazing acceptance, guidance, interest, sharper mind, trying to make me at home, to challenging me to go further, to learn the best I could with the tools and methods you knew and your lab is good at. Thank you for all the skype messaging, good orientation and defining pathways to follow as long as we're discussing possibilities, never as a dictator, but as someone who knows best but don't pride on it. Coordination permitted for what a real orientation should be. I learned much from you on how to prioritise, and give attention to what is most necessary. Thank you for making me move into bigger steps, growing myself as an individual and as a student in this world of science.

A special thanks to all my supervisors for helping me to turn this thesis into a real thing, for all the help in the structure, in the English, in always improving it and making me learn how to do it, it meant a lot for my education. To all my supervisors thank you for believing in me!

In four years many people passed through our lab (BFMP/BVII/BIOFIG - Braga) and I want to thank them all, for the good ambiance, for the camaraderie, for the laughs and even the hard times. Hoping I won't forget anyone, thank you Juliana, Vítor, Humberto, Rómulo, Cátia, Joana, Conde, Manú, Manuela, Mafalda, Marta, Luís, Daniel, João Alves, Rute, Sara F, Patrícia, Ana, Paula, Cristiana, Natacha, Paulinho, Óscar,

Alice, Teresinha, Daniela, Eva, professor Teresa, professor Ana Cunha, João, Eduarda, Inês, António, Franklin, Artur, Richard, Noronha, and Viviane. I won't talk particularly to all but I must say something more about some.

Jules, even with your lately crazy sports fanaticism I guessed it put you saner than me; I enjoyed very well these almost four years of coexistence, many of those in a deserted space, but we coped very well. Thank you for all your help and availability in all work-related things and for listening to me babbling so much when I get a little bit "overwhelmed". It's was great to see you growing into an empowered woman. I'll miss you!

For my smaller work group:

Vítor, my partner in crime since 2006, my bench partner half of those four years, and a good friend, thank you for bringing always an optimistic point of view, for helping me in whatever you knew, but I still resent you being always away so long, but well I know it made wonders for you, so I'm very happy for you, and you're at the final too so the best of luck. You're the luckiest guy I know so you probably won't need it, but luck is never too much. Thank you for being an advisor sometimes, as a friend, annulling any type of competition and being my friend at first, seeking my best interest.

Humberto, my friend, I envy you terribly, you super-sumo guy, you are so organised and meticulous, and updated, I wish I could be more like you, but it's hard to change. I'll try a bit, but I know I can't keep that pace you have, my brain needs definitely more rest. Anyway, thanks for everything, for the critics, even the bad ones, for keeping all tidy (it's a dream for every girl at a lab), though it can be a little annoying too, when you look with your eyes to my bench...Thanks for the support when we coincided in Málaga, for the inside and outside lab time. It was indeed great having someone to talk Portuguese, to share problems and anxieties. I'll never forget the afternoon's snacks when you putted the cookie into the coffee with the same meticulous way you do everything else. And above all else thank you for your friendship!

Daniel, you're a crazy dude, but you're growing and that's nice to see. You refresh the lab and you just don't lament things, you take the task, that's very important. Thank you for your company, help, hearing me talk or yelling at you anyway. I guess I'll miss pulling your ears or pinching you.

João Alves, we worked together and I was uncertain at first but it was a good experience and you made me proud. I wish the best for your future.

Eva, I know you too little but you are so funny that it's impossible not to like you. Hope you'll achieve what you want.

Sara F, I know you too little but you may have the grit necessary, so best of luck, and thanks for the help!!!!

I would like to thank all the other people I interacted in the Biology Department, I won't say them all, but thanks Rita, Andreia, Sofia, João, Raul, Su, Geninha, JP, for embracing the new people so well that friendship flourished. A special thanks to the entire lab teach and investigation techs for saving my life so many times.

There is another great lab I must thank, my lab in Málaga, I appropriated as “mine”, because everyone made me felt at home there, everyone was very friendly, cooperative, helpful, even when I was a pain in the ass and didn't know where the things were, so sorry for being annoying at some times. Though you all knew I missed home terribly you've been more than kind to me, so thanks a lot to the *Laboratorio de Bioquímica y Biotecnología Vegetal* and to its occupants: Vero, Vivi, Karen, Irene, Camilla, Paqui, Ali, Carmen, Itziar, Edu, David, Arni, Cristina, Fabi, Naoufal, Irene Nevado, Jasmin, and of course professor Vitoriano Valpuesta. As time got by the group of genetics was also in the floor so they were “family” too: Adela, Alberto, Rosa, Manolo, Ana, Tábata, Zaira, Edgar, professor Eduardo and of course the people we every day got lunch together as was the case of Ian and so many other friendly people.

Karen, my dear you were my anchor there, my friend, we bonded so fast that going away was more difficult than I thought, I cannot have words enough to tell you how much you mean to me and how much I appreciate your help and friendship there, now and forever. Thanks for everything, and mainly for allowing me into your home, into your beautiful family and into your heart. Except that I won't forget the loss of Portugal against Spain in your home...so sad...Fortunately I have better moments to hold on!!!

Vivi, I can never forget you arriving from Israel and immediately inviting me to eat your amazing gnocchi's, and you hardly even knew me, you're a great cooker!!! From that you've been nothing but kind to me and albeit your problems at one time or the other you stayed brave!!! Thank you for giving me great memories, great laughs in the lab. I can't say how much I enjoyed seeing the World Football Cup 2010 with you and all the people in the lab. It was not really seeing because we were working but you gave me a new sense of that experience and you got so excited and angry and you knew so much about the plays, it was great, I loved it!!!!!!! You're definitely not a common person and that's what interested me most, you are a conqueror albeit complaining a lot, but I guess it comes with the baggage. Wish you the best for your new path now.

David, when I first came to Málaga, you alongside Miguel came to pick me up at the airport, I didn't know what to expect but it went so well, you were a calm person, who taught me what you knew and I'm very grateful for that, thank you so much for wasting time with me.

Ali, my dear Ali, I have so much to thank you for, I would write so many pages, thanks for your technical, personal, and friendly support. You're very special to me and you were always there for me, and I'd love to see the perfection you made on things, I learned a lot just by staring. Thank you for accompanying me on so many places I needed and didn't know how to get there. Thanks for continuing the work and make what I couldn't and can't anymore and bring it to life.

Vero, oh Vero, you are almost as stressed as I am, but you are so good at what you do, I wish I was like that, you inspire people, really. So hard working!!! Thank you for all you taught me and I enjoyed it very much, the learning, the lunches, even being at your bench when you were in Sweden....Hehehe really you are an experiment master!!!

Edu, thank you a lot for all the insight you gave me on Agro, tobacco and all that stuff, in my last experiments you helped me a lot, and you were always ready for helping me, even at a click distance.

Alberto, thank you so much for teaching me some new experiments that were great for my work, I was really excited, and you're an excellent teacher. It was a pleasure to meet you and be guided by you.

Four years went on a blink of an eye but my friends, my true friends helped me so much. Beside the friends that were also colleagues and I've mention them before, there are others who one way or another helped me to cope and endure the road and the final battle. Thank you my friends for being there, comprehensive and allowing me to be none but myself. Thank you Jorge, Isabel, Gi, Marisa, Alberto, Xica, Susana, João, Cristiana, Luís. Thanks to my old buddies, that although some cannot quite understand the stress and commitment to this, I thank you for knowing that they could still be there for me (I hope at least), so thank you Anabela, Nela, Ida, Cris, Bruno, Aires, Patta, Beto and Raky. Thanks to all my boyfriend friends, from university until masters, some of which I had the pleasure to contact more at work as was the case of Jéssica and Nuno. A special thanks to Carla (and Pedro) for the online support over the writing period.

I also want to thank Ana for the amazing help through these last years, you've been more than a friend than most people "are", and without judging your words were ever inspiring. I look to you and I see a good role model, and I wish I could be so self assured as you are, and enlighten the places the way you do. You're very gifted with people, so thanks for giving me a little of your gift.

I must say some special words for two great people; it's rare to have the opportunity of knowing such extraordinary persons, with a heart so pure and thoughtful, with a soul bigger than life. Jorge and Isabel you were sacred to me in those buildings and out of it, you were my temple, my escape, my happiness, I enjoyed seeing you grow along side me through this years, telling our frustrations, our achievements, it brought us close and even in the darkest hour you stood by me no matter what, and that meant the world to me. Even though you knew I had little time to spend with you, you were always comprehensive, and I guess real friends are that, are the ones that respect you no matter what, that do not make demands but wait until you're ready, and no matter how much time it passes everything is the same, the good old same. Someone once said that true friends are the ones that we can be in silence with, and that's how we are, in silence or in laughter, in tears or in excitement, we are forever bond to be friends, the good ones and the ones that endure a lifetime.

My family was so very important to me, my roots are what I am, the leaves I've developed were nurtured by them, every day of my life...They gave me all a child can wish for, they gave me life-models, a very nice childhood, they gave me what was necessary to build me in character and in sensibility to survive in this world. To my extended family I thank all of them, my aunts, uncles, and cousins that even far, kept me on tracks with small words, but encouraging ones. I would like to thank my godparents for all the support and being such open people who demonstrate a different way of life, way ahead of everyone else. Thank you for remembering me at the Galapagos (and every other piece of foreign land) and placing in me the bug of travelling. Also, a great family deserves a place here, for there is one also in my heart, thank you for accepting my particularities and welcoming so well, thanks Artur, Manela and Rodolfo. To my dear parents, mummy and daddy, none of this would've been possible without you, because I am in part what you've

made of me, and even not understanding most of what this all is, I knew you were my safe place, my lifebuoy, fortunately you'll never let me drown, I love you so much!!! I must thank also my sister for whatever guidance she gave me through only 16 years, it helped me enormously. She's the light I look into my heart, the star I look into the sky, the flame that makes me believe because in her own way she was a fighter, and I am a fighter too. Thank you big sis for giving me all the nice memories I can remember, although I'm angry you're not here. Life wrecked me but I like to think it also transformed me into a better and stronger person. A special thanks to all the people that departed life through all those years since I can remember, including all my grandparents. For each one of you I give you my love, appreciation, and a promise that my heart will never be empty as long as you're with me.

One final word goes to my half soul, my love, my passion, my future. You helped me more than you can imagine, especially in the end of this, you were so patient, caring, loving, encouraging. I know you did the best you could for me, so thank you for being unconditionally there. My life would be senseless without you!!! Every bit in my life is important to me and you know I have a hunger for more, I have a hunger that will never stop, an energy that must be spent, and the wants of changing the world. You gave me balance, you always tell me to stop in the right moment. We are two sides of the same coin, one doesn't exist without the other, and so we are in balance. You balance my life and make it all worth it. Thank you for being there, thank you for being you and letting me be me. It's great to be just me. I love you and I can't wait for the future to come!!!

“The fortunate man, in my opinion, is he to whom the gods have granted the power either to do something which is worth recording or to write what is worth reading, and most fortunate of all is the man who can do both.”

Pliny the Younger

Letter to the historian Cornelius Tacitus AD 106

O presente trabalho,
incluindo a sua publicação beneficiou do seguinte apoio
da Fundação para a Ciência e a Tecnologia:
Bolsa de Doutoramento - SFRH/BD/29778/2006,
financiada pelo POPH - QREN - Tipologia 4.1 - Formação Avançada,
comparticipado pelo Fundo Social Europeu e por fundos nacionais do MCTES.



STRESS-RELATED GENES IN *ARABIDOPSIS THALIANA*: INVESTIGATING THE ROLE OF SQUALENE EPOXIDASES IN STEROL BIOSYNTHESIS AND *EGY3* AS A PUTATIVE PLASTIDIAL HEAT STRESS DETERMINANT

Abstract

A worldwide effort in plant science research has been carried out over the last years as a response to the aggravated impact that environmental stresses are having on crop production and yield. The main problem is arguably the lack of water availability in the soil, as a consequence of heat, drought, salt and osmotic stresses. This research includes the fundamental aspects that help to understand the mechanisms behind plant growth and development, as well as their response to the external challenges that determine tolerance leading to increase production and finally survival. Using the model plant *Arabidopsis thaliana*, three genes were studied within the scope of the present thesis. By investigating their role in plant development and particularly abiotic stress, the aim is to contribute to the undergoing effort of improving knowledge-based plant production.

SQE1 encodes the main squalene epoxidase (SQE) in the sterol biosynthetic pathway and has been shown to be an abiotic stress determinant gene required for the regulation of reactive oxygen species and drought tolerance. The *SQE1* homologues, *SQE2* and *SQE3*, have been shown to possess squalene epoxidase activity based on yeast complementation, but their roles in plant development or as potential abiotic stress determinants have not been elucidated. The current study of *SQE* genes in *Arabidopsis* has revealed a complex regulation of sterol biosynthesis. Present analysis shows a low expression of *SQE2* relative to *SQE3* using promoter-*GUS* histochemical analysis, with *SQE3* being highly expressed in seedlings and in reproductive tissues, corroborating microarray expression data available in public databases. Isolation of loss-of-function mutants for both genes together with terbinafine analysis, an inhibitor of SQEs, have shown an *in planta* role for *SQE3* in squalene epoxidation but not for *SQE2*, which suggests a role for this gene different to the biosynthesis of bulk sterols. Interestingly, and based on promoter swap fusions it was also shown that *SQE2* and *SQE3* were unable to complement *SQE1* function in the *dry2/sqe1-5* mutant, excluding redundancy within this gene family. The lack of complementation of *SQE3* cannot be explained by a different subcellular localisation, since translational fusions showed that both, *SQE1* and *SQE3*, are present in the endoplasmic reticulum. Sterol profiling revealed a deregulation in the sterol content of *sqe3-1* shoots, while *dry2/sqe1-5* was already reported to have an important deregulation of sterol content in roots, which indicate different roles depending on the tissue. The double mutant *dry2/sqe3-1* was infertile, indicating that *SQE3* has an important role in plants with reduced *SQE1* activity. Research on the heterozygous *dry2/dry2 SQE3/sqe3-1* showed seed/embryo impairment resulting in a 21% abortion ratio, against the 6% of *dry2/dry2 SQE3/SQE3*, which highlights the importance of sterols produced by *SQE3* during embryo/seed development. A series of phenotype characterisations were conducted

involving the potential role of *SQE2* and *SQE3* in abiotic stresses. Altered responses to heat shock for *sqe2-1* as well as salt, osmotic and ABA treatments for *sqe3-1* were investigated based on gene expression data gathered from public microarray information. However, none displayed differential responses compared to wild-type plants. Biotic stress challenge with the bacterial pathogen *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*) was also analysed for *SQE3*, based on the predicted expression pattern and the putative hypothesis that *SQE3* could provide 2,3-oxidosqualene to the lanosterol pathway. Infection with *Pto* was performed in *sqe3-1* plants, since *LAS1*, the next enzyme in the pathway, also showed induction in these conditions. However, no clear differences were found and therefore our results were inconclusive.

Following a web-based data-mining strategy to search for uncharacterised genes putatively involved in abiotic stress responses, *EGY3*, an EGY-like putative metalloprotease, was identified, for being an unresolved plastid-targeted protein. *EGY3* presents high expression during heat stress and in response to other abiotic stresses. Results of GUS histochemical analysis demonstrated high expression in seedlings, stele and pericycle cells of the root, lateral root primordia, and in several flower organs. Gene expression analysis showed that *EGY3* was induced by heat stress, and loss-of-function *egy3-1* mutants presented a heat-related germination phenotype, with mutant seeds showing more tolerance to high temperatures. Due to the high *EGY3* expression in the root and the presence of a heat-related phenotype, analysis of *egy3-1* root architecture was also performed in both normal and heat stress conditions. Differences between different genotypes were not found under the same conditions; however, significant differences were encountered within the same genotype in response to heat stress, displaying a heat inducible formation of lateral roots in the wild-type, and an increase in total root length in the mutant *egy3-1*. Analysis of developmental and growth phenotypes of *egy3-1* mutants revealed delayed flowering, late growing, bushier plants and prolonged life cycle. These results not only implicate this gene in the heat stress response but also in circadian rhythm mechanisms and the onset of flowering. This was further supported by analysis of the promoter for the identification of *cis*-element presence, with the finding of binding sites of putative TFs related to seed development, abiotic stresses, lateral root, rosette development, and senescence induction. Moreover, *in silico* analysis suggested that *EGY3* is regulated by HsfA2, and is co-expressed with other HSF and HSP, reinforcing its role in heat stress. *EGY3* expression in microarrays of known mutants revealed, among others, a deregulation in its expression in mutants involved in heat stress (HSFs), as well as the onset of flowering, the coordination of plant development and the promotion or repression of the transcription of photosynthesis associated nuclear genes (*ein3/eil1*, *cop1* and *hy5*). A construct to obtain complementation and overexpression lines was developed and transformed into *egy3-1* and wild-type plants, respectively. Constructs to express recombinant *EGY3* in *Escherichia coli* have also been performed in order to produce, purify and analyse its catalytic activity. Additional functional studies are underway, including GFP-mediated subcellular localisation, which will be used to characterise the role of *EGY3* in development and heat stress responses.

GENES ENVOLVIDOS NO STRESSE EM *ARABIDOPSIS THALIANA*: INVESTIGANDO O PAPEL DAS ESQUALENO EPOXIDASES NA BIOSÍNTESE DOS ESTERÓIS, E DE *EGY3* COMO DETERMINANTE PLASTIDIAL DA RESPOSTA AO CALOR

Resumo

Nos últimos anos, a investigação em biologia vegetal tem-se focado em responder ao grave impacto que os stresses ambientais (e.g. disponibilidade de água nos solos, calor, frio, seca, salinidade) provocam na produção e rendimento agrícolas. Esta investigação é crucial para permitir a compreensão dos mecanismos que permitem o crescimento e desenvolvimento das plantas, assim como a resposta aos estímulos externos que determinam a sua sobrevivência/tolerância. Neste trabalho, recorrendo à espécie modelo *Arabidopsis thaliana*, foi estudada a função dos genes *SQE2*, *SQE3* e *EGY3*, dado o possível papel no desenvolvimento da planta e, particularmente, como determinantes na resposta a stresses abióticos.

O gene *SQE1* codifica a principal esqualeno epoxidase da via biossintética dos esteróis, tendo-se já demonstrado ser um gene determinante para o stresse abiótico, pois está envolvido na tolerância à seca e na regulação de espécies reativas de oxigénio. Os genes *SQE2* e *SQE3*, homólogos de *SQE1*, mostraram possuir atividade esqualeno epoxidásica em levedura, mas o seu papel no desenvolvimento vegetal, e como potenciais determinantes do stresse abiótico, ainda não foram elucidados em *Arabidopsis*. O estudo dos homólogos do *SQE1* revelou existir uma regulação complexa na biossintese dos esteróis. Os resultados demonstraram baixa expressão do *SQE2* em comparação com a do *SQE3*, em ensaio GUS (gene-repórter). O *SQE3* é muito expresso em plântulas e tecidos reprodutivos, o que é sustentado pelos dados de expressão dos *microarrays* de bases de dados públicas. Os mutantes de perda de função para ambos os genes foram tratados com terbinafina (inibidor específico da esqualeno epoxidase), tendo os mutantes *sqe3-1* apresentado elevada sensibilidade ao inibidor, o que sugere um papel *in planta* para o *SQE3*, enquanto que a falta de sensibilidade de *sqe2-1* poderá indicar um diferente papel para este gene. A recuperação do fenótipo de *dry2/sqe1-5* com o *SQE2* e o *SQE3* não foi obtida, pelo que se exclui uma total redundância dentro desta família. Ensaio para a localização subcelular permitiram evidenciar que *SQE1* e *SQE3* estão presentes no mesmo compartimento, o retículo endoplasmático. O perfil do esqualeno e de alguns esteróis revelou uma desregulação ao nível da parte aérea de *sqe3-1*, enquanto que em *dry2/sqe1-5* dados prévios evidenciaram uma desregulação importante ao nível das raízes, o que pode indicar algum grau de especificidade a nível do tecido. O duplo mutante *dry2/sqe3-1* é infértil, agravando o frágil fenótipo de *dry2/sqe1-5*. A investigação sobre o heterozigótico *dry2/dry2 SQE3/sqe3-1* revelou que este exibe uma taxa de aborto de 21%, contra a de 6% do *dry2/dry2 SQE3/SQE3*, o que sugere um papel parcialmente redundante do *SQE3* no desenvolvimento do embrião/semente, conjuntamente com *SQE1*. Uma série de ensaios fenotípicos foram efetuados envolvendo *SQE2* e *SQE3* no stresse abiótico. Alterações a nível do

choque térmico (*sqe2-1*) e tratamentos com sal, ABA e manitol (*sqe3-1*) foram averiguados com base nos padrões de indução face a estes stresses, a partir de dados de *microarrays*. No entanto, não se observaram diferenças significativas quando comparados com o ecótipo selvagem. Ensaio de infecção com a bactéria patogénica *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*) foram efetuados no *sqe3-1*, tendo por base a premissa de que o gene era sobre-expresso nestas condições e na hipótese de que SQE3 poderia fornecer 2,3-oxidoesqualeno à via do lanosterol, uma vez que LAS1, a seguinte enzima desta via, também demonstrou um padrão similar de indução. Contudo, os resultados de infecção foram inconclusivos.

Uma estratégia de prospeção de dados baseada em recursos bioinformáticos foi desenvolvida para identificar genes ainda não caracterizados que estivessem envolvidos na resposta ao stresse abiótico. Desta análise, foi selecionado *EGY3*, que codifica uma putativa metalloprotease do tipo EGY, prevista para o cloroplasto. O *EGY3* apresenta elevada expressão durante o stresse pelo calor e em resposta a outros stresses. Resultados do ensaio GUS demonstraram elevada expressão em plântulas, cilindro central e células do periciclo da raiz, primórdios das raízes laterais e em diversos órgãos florais. A análise da expressão do *EGY3* demonstrou a sua indução pelo calor e o seu mutante de perda-de-função apresentou um fenótipo associado à germinação pelo calor, com sementes mais tolerantes a altas temperaturas. Devido à expressão específica na raiz e a presença deste fenótipo pelo calor, uma análise da arquitetura da raiz foi conduzida no *egy3-1* em condições de crescimento normal e de choque térmico. Não houve diferenças entre genótipos idênticos, mas verificou-se indução pelo calor do comprimento das raízes laterais de plantas selvagens e do comprimento total da raiz do mutante. Por sua vez, a análise de fenótipos de desenvolvimento e de crescimento revelaram ocorrer floração e crescimento tardios, plantas mais robustas, e com ciclo de vida mais extenso. Estes resultados implicam o *EGY3* na resposta ao stresse pelo calor, nos mecanismos do ritmo circadiano e no início da floração. Isto é suportado pela análise de elementos *cis* no promotor do gene, onde os fatores de transcrição correspondentes estão envolvidos em processos como o desenvolvimento da semente, stresse abiótico, raízes laterais, desenvolvimento da roseta e indução da senescência. Análises *in silico* também sugerem que *EGY3* seja regulado por *HsfA2* e coexpresso com outro HSF e HSPs, o que reforça o seu envolvimento no stresse pelo calor. A expressão de *EGY3* em *microarrays* de mutantes conhecidos revelou uma desregulação da sua expressão em mutantes envolvidos no stresse pelo calor (HSFs), assim como no início da floração, na coordenação do desenvolvimento da planta e na promoção/repressão da transcrição de genes nucleares associados à fotossíntese (*ein3/eil1*, *cop1* e *hy5*). Para obter complementação e linhas de sobre-expressão foram obtidas construções para o efeito, transformadas no mutante *egy3-1* e em plantas selvagens, respetivamente. Para expressar *EGY3* num sistema heterólogo (*E. coli*) foram obtidas construções que no futuro serão utilizadas para a produção da proteína *EGY3*, sua purificação e subsequente análise da atividade catalítica. Estudos funcionais adicionais, incluindo a localização subcelular por GFP, estão a ser desenvolvidos no sentido de caracterizar o papel de *EGY3* no desenvolvimento e na resposta ao stresse pelo calor.

Table of contents

Acknowledgments	iii
Abstract	ix
Resumo	xi
Table of contents	xiii
Abbreviations	xvii
Chapter 1 - "General Introduction"	1
<hr/>	
1.1. The challenge of plant abiotic stress	3
1.2. <i>Arabidopsis thaliana</i> as a model organism	6
1.3. Aims and outline of the thesis	10
Chapter 2 - "Materials and Methods"	13
<hr/>	
2.1. Biological materials, culture media and reagents	15
2.1.1. Biological material	17
2.1.1.1. Plant material	17
2.1.1.2. Bacterial material	18
2.1.1.3. Vectors	18
2.1.1.4. Genes	20
2.1.2. Culture media	21
2.1.2.1. <i>In vitro</i> plant culture media	21
2.1.2.2. Microbiology culture media	22
2.1.3. Reagents	23
2.1.4. Bacteria handling	23
2.1.4.1. Growing of bacterial strains	23
2.1.4.2. Glycerol stock preparation	23
2.2. Plant methods	25
2.2.1. <i>Arabidopsis</i> sterilisation, germination and growth	27
2.2.2. <i>Nicotiana</i> germination and maintenance	27
2.2.3. Phenotype tests in <i>Arabidopsis</i>	27
2.2.4. <i>Arabidopsis</i> transformation by floral dipping	30
2.2.5. Selection of <i>Arabidopsis</i> transformants	30
2.2.6. Plant transient transformation	31
2.2.7. Infection of <i>Arabidopsis</i> by <i>P. syringae</i>	32
2.2.8. Stomata aperture response	32
2.2.9. Root hair measurement	32
2.2.10. Terbinafine inhibition of germination assay	33

2.2.11. Sterol and squalene analysis	33
2.2.12. Grafting	34
2.2.13. GUS histochemical staining	35
2.2.14. Seed analysis	36
2.2.15. Arabidopsis cross-fertilisation	36
2.3. Molecular biology methods	37
2.3.1. Primers	39
2.3.2. DNA methods	45
2.3.2.1. Plant genomic DNA isolation	45
2.3.2.2. Plasmid isolation	46
2.3.2.3. DNA fragment purification	47
2.3.2.4. DNA digestion with endonucleases	47
2.3.2.5. Amplification of DNA fragments by PCR	48
2.3.2.6. DNA sequencing	49
2.3.2.7. Gateway cloning	49
2.3.2.7.1. BP reaction	49
2.3.2.7.2. LR reaction	50
2.3.2.8. Cloning of PCR fragments into an expression vector	50
2.3.2.9. Transformation of bacteria	51
2.3.2.9.1. Transformation of <i>E. coli</i> cells	51
2.3.2.9.2. Transformation of <i>Agrobacterium</i> cells	52
2.3.3. RNA methods	53
2.3.3.1. RNA extraction	53
2.3.3.2. cDNA synthesis	53
2.3.3.3. Semi-quantitative RT-PCR	54
2.3.3.4. qRT-PCR	54
2.3.4. Quantification of nucleic acids	55
2.3.5. Nucleic acids electrophoretic separation	55
2.3.6. Protein methods	56
2.3.6.1. Protein expression in <i>E. coli</i>	56
2.3.6.2. Protein electrophoretic separation	56
2.4. Bioinformatic analysis	59
2.4.1. Informatic tools for data mining and selection of EGY3	61
2.4.1.1. NASCArrays and microarray data treatment	61
2.4.1.2. Multiple Array Viewer	61
2.4.1.3. Chloroplast 2010	61
2.4.2.4. BAR	61
2.4.1.5. TAIR	62

2.4.1.6. Genevestigator	62
2.4.2. Informatic tools for plant molecular biology	63
2.4.2.1. ATTED	63
2.4.2.2. GeneMANIA	63
2.4.2.3. AGRIS	63
2.4.2.4. Athena	64
2.4.2.5. SIGNAL	64
2.4.2.6. SUBA	64
2.4.2.7. EnsemblPlants	64
2.4.2.8. TMHMM	65
2.4.2.9. Signal P	65
2.4.2.10. ChloroP	65
2.4.2.11. UniProt	65
2.4.2.12. MitoProt	66
2.4.2.13. InterProScan	66
2.4.2.14. NCBI	66
2.4.2.15. DNASTAR	66
2.4.2.16. Software for primer design	67
2.4.2.17. Informatic tools for phylogenetic analysis	67
Chapter 3 - "Functional characterisation of SQEs"	69
<hr/>	
3.1. An introduction to SQEs and their role in plant development	71
3.1.1. Sterols	73
3.1.2. Biosynthetic pathway	74
3.1.2.1. Squalene biosynthesis	74
3.1.2.2. Oxidosqualene cyclases	75
3.1.2.3. Oxidosqualene to sterols	78
3.1.2.4. Brassinosteroids biosynthesis	80
3.1.2.5. Mutants of sterol and brassinosteroid pathway	82
3.1.3. Brassinosteroids signalling	86
3.1.4. Squalene epoxidase family	88
3.2. Functional characterisation of <i>SQE2</i> and <i>SQE3</i>	91
3.2.1. Characterisation of the SQE gene family	93
3.2.2. Gene expression analysis	97
3.2.3. Isolation of <i>sqe2</i> and <i>sqe3</i> loss-of-function mutants	101
3.2.4. Developmental and terbinafine sensitivity phenotypes of <i>sqe2-1</i> and <i>sqe3-1</i>	104
3.2.5. Defective sterol profile of the <i>sqe3-1</i> mutant	107
3.2.6. <i>SQE2</i> and <i>SQE3</i> do not complement <i>dry2/sqe1-5</i> mutants	111

3.2.7. Subcellular localisation of SQE1 and SQE3	113
3.2.8. <i>SQE1</i> and <i>SQE3</i> have a redundant role in embryo formation	118
3.2.9. Roots from <i>sqe3-1</i> restore defective aerial defects of <i>dry2/sqe1-5</i>	125
3.2.10. Abiotic stress-related phenotype analysis of <i>sqe2-1</i> and <i>sqe3-1</i>	127
3.2.11. ABA stomatal responses are not impaired in <i>sqe2-1</i> and <i>sqe3-1</i> mutants	130
3.2.12. Alternative lanosterol-pathway vs. biotic stress	131
3.2.13. Concluding remarks	136
Chapter 4 - Functional characterisation of <i>EGY3</i>	139
4.1. Metalloproteases and their role in plant abiotic stress	141
4.1.1. Proteases and protein degradation	143
4.1.2. Metalloproteases	146
4.1.3. S2P proteases and the EGY gene family	148
4.1.4. Protease involvement in chloroplast functioning and stress signalling	153
4.2. Functional characterisation of <i>EGY3</i>	157
4.2.1. Rationale behind a bioinformatics and web-based data-mining strategy for the identification of new abiotic stress determinants	159
4.2.2. Identification of <i>EGY3</i> as a putative heat-stress determinant	161
4.2.3. Characterisation of the EGY gene family	165
4.2.4. <i>EGY3</i> expression pattern analysis	170
4.2.5. Isolation of <i>egy3</i> loss-of-function mutants	173
4.2.6. In vivo developmental and heat stress phenotype characterisation of <i>egy3-1</i>	175
4.2.7. Heat stress-related phenotype analysis of <i>egy3-1</i>	178
4.2.8. Analysis on the regulation of <i>EGY3</i> expression	185
4.3.7. Overexpression, complementation and heterologous expression studies of <i>EGY3</i>	194
4.3.8. Concluding remarks	198
Chapter 5 - Final remarks and future prospects	203
Chapter 6 - Bibliography	209
Appendix I – Expression raw data	228
Appendix II – Squalene and sterol analysis raw data	244
Appendix III – Vector maps	246

Abbreviations

3D	three dimensional	HSF	heat shock factor
A _{###}	absorbance at ### nanometres	HSP	heat shock protein
a.a	amino acid	Hyg	hygromycin
ABA	abscisic acid	IAA	indole-3-acetic acid
Amp	ampicillin	IPTG	isopropyl β-D-1-thiogalactopyranoside
APS	ammonium persulfate	Kan	kanamycin
atm	atmosphere	L	litre
ATP	adenosine-5'-triphosphate	<i>lacZ</i>	part of <i>lac</i> operon that encodes the β-lactosidase
bp	base pair	LB	lysogeny broth
BR	brassinosteroid	Ler	Landesberg erecta
c.f.u.	colony forming unit	m	meter
CCD	charge-coupled device	M	molar
cDNA	complementary DNA	MASC	Multinational <i>Arabidopsis</i> Steering Committee
CDS	coding sequence	MES	2-(N-morpholino)ethanesulfonic acid sodium salt
Clx	cicloheximide	min	minute
Cm	chloroamphenicol	M-MLV RT	moloney murine leukemia virus reverse transcriptase
coA	coenzyme A	mol	mole
Col-0	Columbia	Mr	relative molecular mass
Ct	cycle threshold	mRNA	messenger RNA
CTAB	hexadecyltrimethylammonium bromide	MS	Murashige and Skoog culture medium
Da	Dalton	MTF	membrane-bound TF
dd water	double destilated water	NAD ⁺	nicotinamide adenine dinucleotide
DEPC	diethyl pyrocarbonate	NADPH	nicotinamide adenine dinucleotide phosphate
DMF	dimethylformamide	°C	degree Celsius
DMSO	dimethyl sulfoxide	Oligo (dT)	oligodeoxythymidylic acid
DNA	deoxyribonucleic acid	ON	overnight
DnaseI	deoxyribonuclease I	ORF	open reading frame
dNTPs	deoxyribonucleotides	OS	2,3-oxidosqualene
dpi	days post-inoculation	OSC	oxidosqualene cyclase
dsDNA	double stranded DNA	p.a.	pro analysis
DTT	dithiothreitol	PAR	photosynthetic active radiation
EDTA	ethylenediaminetetraacetic acid	PBS	phosphate buffered saline
eFP	electronic fluorescent pictographic	PCD	programmed cell death
EMS	ethyl methanesulfonate	PCR	polymerase chain reaction
ER	endoplasmatic reticulum	PEG	polyethylene glycol
F	Faraday	Pfam	protein family
FAD	flavin adenine dinucleotide	PIPES	1,4-piperazinediethanesulfonic acid
G	Golgi	Pq	paraquat
g	gram	PR	pathogen-related
<i>g</i>	relative centrifuge force	PSI	photosystem
GC-MS	Gas chromatography–mass spectrometry	PSII	photosystem II
gDNA	genomic DNA	PTGS	post-transcriptional gene silencing
GFP	green fluorescent protein	<i>Pto</i>	<i>Pseudomonas syringae</i> pv. <i>tomato</i> DC3000
GO	genome ontology	qRT-PCR	quantitative reverse transcriptase polymerase chain reaction
GOI	gene-of-interest	RFU	relative fluorescence unit
GST	glutathione S-transferase	Rif	rifampicin
GUS	beta-glucuronidase		
h	hour		
His	histidine		
HL	high light		
HMG-CoA	3-hydroxy-3-methyl-glutaryl-CoA		
HMGR	HMG-CoA reductase		
HS	heat shock		
HSE	heat shock element		

RIP	regulated intramembrane proteolysis	T-DNA	transfer DNA
RNA	ribonucleic acid	TE	tris-EDTA buffer
RNAi	interference RNA	TEMED	tetramethylethylenediamine
RNase A	ribonuclease A	TF	transcription factor
ROS	reactive oxygen species	Tm	melting temperature
rpm	rotations per minute	TM	transmembrane domain
S1P	site-1-protease	Tris	tris(hydroxymethyl) aminoethane
S2P	site-2-protease	Triton X-100	polyoxyethylene-p-isooctylphenol
SDS	sodium dodecyl sulfate	tRNA	transfer RNA
SE	squalene epoxidase domain	u.p.	ultra pure
sec	second	UV	ultraviolet
SOB	super optimal broth	V	Volt
SOC	super optimal broth with catabolite repression	v/v	volume per volume
SREBP	sterol regulatory element binding protein	Vis	visible
TAE	tris-acetate-EDTA-buffer	w/v	weight per volume
Taq	<i>Thermus aquaticus</i> polymerase	Ws	Wassilewskija
Tb	terbinafine	Wt	wild-type
TB	transformation buffer	X-Gal	bromo-chloro-indolyl-galactopyranoside
TBSV	tomato bushy stunt virus	X-Glu	5-Bromo-4-chloro-3-indolyl-beta-D-glucoside
		Ω	Ohm

Amino acids:

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

Nucleotides:

A - Adenine
C - Cytosine
G - Guanine
T - Thymine
U - Uracil
W - A or T
N-any base
dATP 2'-deoxyadenosine-5'-triphosphate
dCTP 2'-deoxycytidine-5'-triphosphate
dGTP 2'-deoxyguanosine-5'-triphosphate
dTTP 2'-deoxythymidine-5'-triphosphate
GDP Guanosine-5'-diphosphate
GTP Guanosine-5'-triphosphate

1.

"General Introduction"

1.1. *The challenge of plant abiotic stress*

Climate changes on Earth and the course of millions of years of evolution contributed to a high genetic diversity, demonstrating living creatures' capacity to adapt to the environment and its fluctuations (Zhu, 2002; Koiwa *et al.*, 2006). Environmental stresses can either be biotic when imposed by other organisms, or abiotic, when they are the result of a deficit or an excess in the physical or chemical environment. During their life span, plants are normally exposed to a variety of different conditions/stresses that affect their growth, development and productivity (Figure 1.1). As sessile organisms, plants are particularly vulnerable to abiotic stress challenges, and have developed an amazing array of responses to face stress imposition (Buchanan *et al.*, 2000). Independently of the stress factor, the impact of a given environmental stress on the plant's physiology is determined by a series of common characteristics, the most important being the intensity and duration of the stress. For instance, lethality of an imposed heat stress depends severely on the susceptibility, genotype, developmental stage, organ or tissue, and in this particular stress the duration and temperature at which it is imposed to the plant is of major importance (Buchanan *et al.*, 2000). Like in other stresses, survival to heat stress is determined by a complex gene network involving, in this particular case, heat shock proteins (HSP) and heat shock factors (HSF) among others (Kotak *et al.*, 2007); however, simple physiological responses also occur such as opening of the stomata to cool leaves by transpiration or even the reduction of the leaf angle to avoid exposure to high light and heat (Mittler, 2006). An important aspect to the stress response is the number or spacing of exposures, fundamental for allowing the possibility of a quicker response or a pre-adaptation of the plant, as is the case of plant pre-acclimation to heat stress prior to elevated temperatures (Larkindale *et al.*, 2005). This feature is, of course, highly correlated with the previous two aspects of stress imposition (Figure 1.1).

Finally, the combination of different stress factors simultaneously operating on a plant can substantially potentiate antagonistic or synergistic responses. For example, salt or heavy metal stress in combination with heat can become problematic since enhanced transpiration may result in enhanced uptake of salt or heavy metals. In another example, when plants already stressed by drought or cold, must face high light intensity, dark reactions are inhibited due to low temperature or insufficient availability of CO₂, and the high photosynthetic energy due to high light conditions triggers oxygen reduction and thus ROS production. On the other hand, some stress combinations could be advantageous to plants, when compared to individual stresses applied separately, as when a reduction in stomatal conductance in drought stress conditions enhances the tolerance of

plants to ozone stress (Mittler and Blumwald, 2010). When two stresses occur simultaneously, the corresponding response gene networks are not necessary identical. The regulatory mechanism, in a double response to stress, is tuned in such a way that the great majority of induced genes are related to the double response, and do not rely only in the genes induced of the single stress responses, being this last a minority in the pool of genes induced to overcome stress. This knowledge is of the utmost importance when the intent is to study stresses that affect worldwide crops, because in the field, crops will most likely face several combinations of stress. For instance, facing a prolonged exposure to abiotic stresses can turn plants more susceptible to pathogens (biotic stress) (Mittler and Blumwald, 2010).

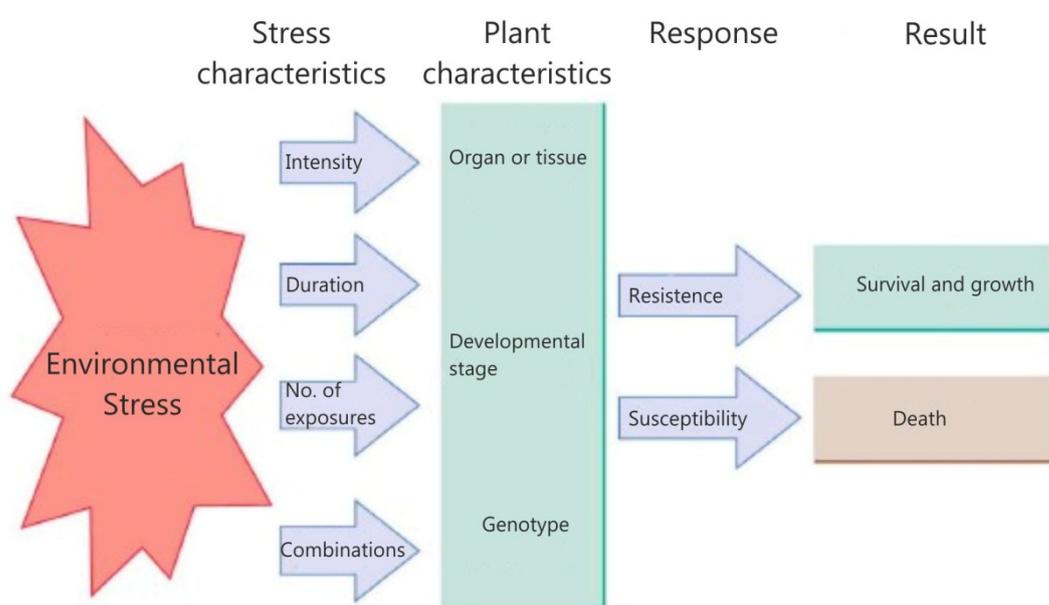


Figure 1.1. – Characteristics of environmental stress and the typified response of the plant. Adapted from Buchanan *et al.* (2000).

Plants stress responses can be grouped according to general- (basic protection) or stress-specific adaptive responses. The later occur as a response to unique factors in a certain circumstance, such as lowered oxygen tension, which is characteristic of a hypoxic stress in flooded roots (Baena-González, 2010). Basic responses are traditionally triggered by signals and signalling components shared by multiple pathways, which explains why adaptation to a particular stress/condition frequently ensures tolerance to other stresses (Baena-González, 2010). To achieve acclimation, an instantaneous response is needed to re-establish homeostasis, repair damaged cellular components and reprogram metabolism. Managing energy resources is of major importance, and takes place at the cellular level by arrest of biosynthetic processes, as well as at whole-plant level. Under stressful conditions resources are forced from reproductive activities into

metabolic reactions that increase stress tolerance, and this ability to re-direct nutrients to the necessary processes and to make the precise metabolic adjustments during stress is critical for plant survival. Initial responses are of uttermost importance since a failure to regulate the appropriate genes would result in increased sensitivity and cell-death, despite the important signals such as calcium fluxes, ROS and nutrient deprivation, which are involved in stress cross-talks. It is important to note that most environmental stresses have an impact on light absorption, carbon fixation or oxygen availability, reducing the efficiency of photosynthesis and/or respiration, and, as a consequence the overall energy status of the cell, and these are the core processes by which a plant preserves her survival, being the turning point for life or death (Baena-González, 2010).

Research on plant abiotic stress responses is certainly an increasingly demanding field. As a major limiting factor of plant growth, abiotic stress will likely become more severe as desertification advances and occupies more arable land (Vinocur and Altman, 2005; Century *et al.*, 2008). Another important aspect is the lack of water availability, which is increasing with the rise of salinity levels both in the water and in the soil. The greenhouse effect leading to global warming means that the average annual temperature may increase by 3°-5°C in the next 50-100 years, with concomitant drastic losses in worldwide crop production (Mittler and Blumwald, 2010). These stresses are ever more worrying and contemporary, since the rise in intensive farming and the use of land for biofuel places additional pressure on food supplies during a period of remarkable requirement for growth from developing countries (Johnston *et al.*, 2009).

Never before has it become so important to know more about the perception mechanisms and signalling responses that take place in a plant facing this range of abiotic stress conditions. It is therefore necessary to perform a comprehensive study of the plant's responses in order to better enhance plant tolerance to stress, since it is a complex and coordinated response that involves hundreds of genes (Borsani *et al.*, 2001; Mittler and Blumwald, 2010). Functional studies allow us to know the function of genes within the complex gene network that is the plant abiotic stress response. Modulating the response of these genes in crops and cultivars-of-interest is a most relevant strategy for plant improvement, and fundamental knowledge obtained in *Arabidopsis thaliana* has been systematically translated to plants of higher agronomic interest. Recent examples include: an easier and cheaper method to extract sugars from plant material developed in *Arabidopsis* to meet biofuel demands; a master regulator of plant root hair growth, as the nutrient mining machinery to enhance the plant root system; the extraction of petroleum precursors from plants to produce green plastic; an *Arabidopsis* gene that confers resistance in Brassica; insight into chromosome imbalances and predictable plant defects; and an *Arabidopsis* gene employed by Monsanto to improve soybean yields (MASC Report, 2011). These studies, in their gene-centric

approach, are best carried out in model organisms such as *Arabidopsis thaliana*. Therefore, they will continue to be pivotal tools in the extending of knowledge that will allow us to face the challenges ahead, increasing crop yield and tolerance, and ultimately diminishing hunger worldwide (MASC Report, 2011).

1.2. *Arabidopsis thaliana* as a model organism

Arabidopsis thaliana is a dicotyledon from the Brassicaceae family that is well distributed across temperate climates (Figure 1.2). *Arabidopsis* was adopted as a model organism because of its usefulness for genetic experiments. Characteristics such as a small and compacted genome, short generation time, small size that limited the requirement for growth facilities, and prolific seed production through self-pollination were fundamental for its selection as a model (Koornneef and Meinke, 2010). *Arabidopsis* is easily cultivated *ex vitro* and *in vitro* (Feng and Mundy, 2006), and is easily transformed by *Agrobacterium tumefaciens* (Zhang *et al.*, 2004; Leonelli, 2006) either by floral dipping (Clough and Bent, 1998) or more recently by transient transformation of seedlings (Marion *et al.*, 2008). This species has numerous ecotypes, but the most commonly used include Landsberg erecta (Ler), C24, Wassilewskija (WS) and Columbia (Col).

The Col ecotype was selected for the genome sequencing program which was concluded in 2000 by the *Arabidopsis* Genome Initiative (2000), becoming the first plant genome to be sequenced. Sequencing revealed a genome of about 125 Mbp, and allowed the elaboration of extensive physical and genetic chromosome mapping (The *Arabidopsis* Genome Initiative, 2000). This knowledge also led to a change of focus in genomic research from gene structure (structural genomics) to gene function (functional genomics). Functional genomics tries to understand the real function and role of a gene as well as the non-coding regions and repeats in the genome. When functional information is put together with structural genomics, statistical and bioinformatics tools, it becomes a valuable resource for unveiling the behaviour of biological systems (Feng and Mundy, 2006). Research in *Arabidopsis*, particularly after the sequencing of its genome, triggered a revolution in plant biology and unravelled important mechanisms in plant development, abiotic and biotic stress tolerance and adaptation (Feuillet *et al.*, 2011). As a hallmark in plant genomics, it paved the way not only for knowledge on this model plant itself, but also for the sequencing of other plant and particularly crop genomes.

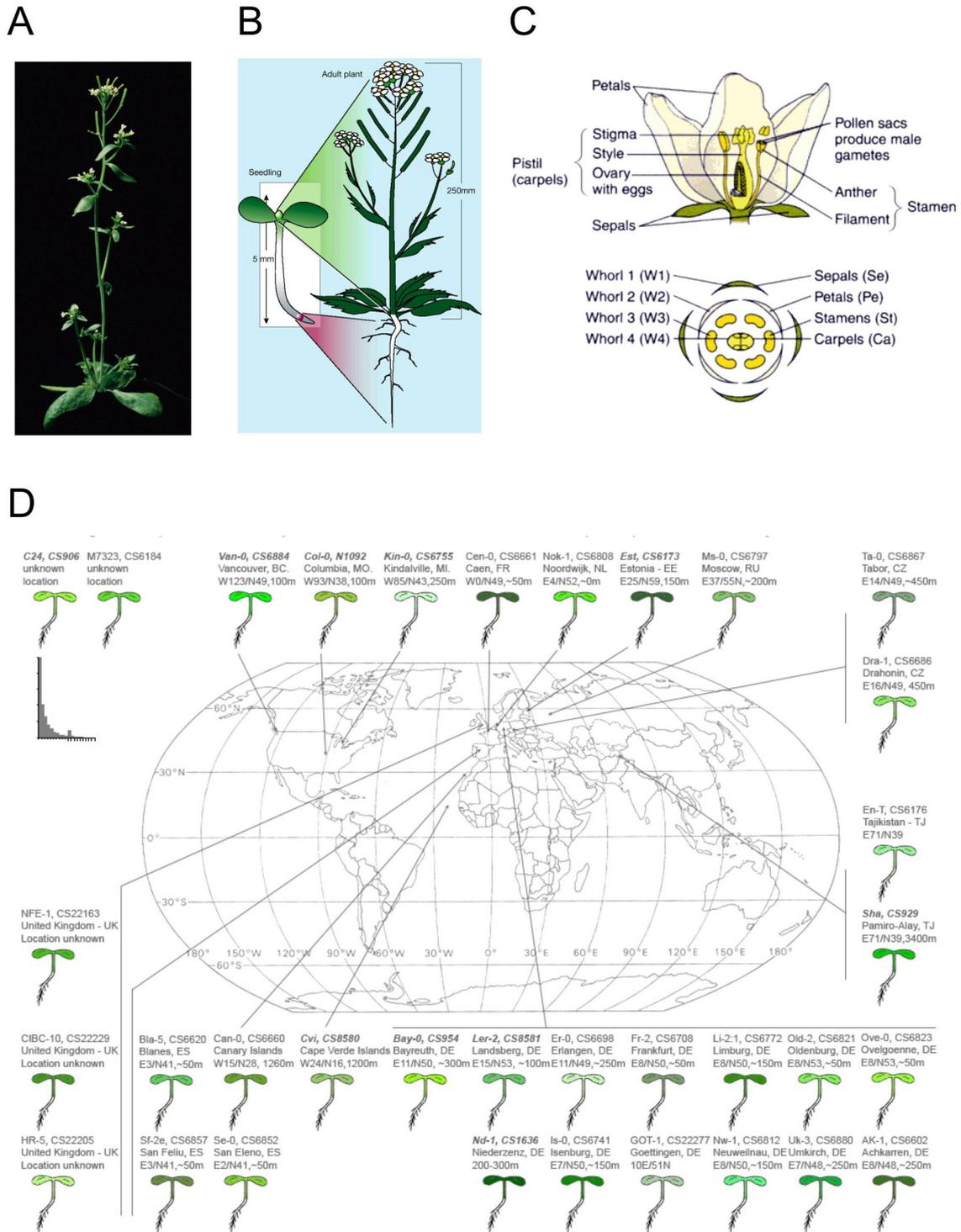


Figura 1.2. – Main developmental stages and geographical distribution features of the model plant *Arabidopsis thaliana*. A – *Arabidopsis thaliana* reproductive and vegetative tissue (Taiz and Zeiger, 2003). B – *Arabidopsis thaliana* schematics of growth from seedling to an adult plant (Taiz and Zeiger, 2003). C – Flower organs in detail (URL no.1). D – *Arabidopsis* ecotype distribution throughout the globe (URL no.2).

These included *Oryza sativa* ssp. *Japonica* (rice), *Sorghum bicolor* (sorghum), *Vitis vinifera* (grapevine), *Glycine max* (soybean), *Populus trichocarpa* (black cottonwood poplar), *Zea mays* ssp.

mays (maize), *Cucumis sativus* (cucumber), *Medicago truncatula* (barrel medic), *Carica papaya* (transgenic papaya) and *Solanum lycopersicum* (common tomato), among others (Feuillet *et al.*, 2011). In truth only Rice and Arabidopsis have finished genome sequences, as other sequenced genomes are still at a draft stage. They have been sequenced using different strategies, varying from standard, high-quality to even improved high-quality drafts of genome sequences.

There are now other features that could not have been foreseen 40 years ago and that have allowed Arabidopsis to remain the premiere model for plant biology. The total research effort that was conducted towards interdisciplinary, multi-investigator studies, requiring extensive community resources, allowed a major growth of Arabidopsis as a favoured organism (Koornneef and Meinke, 2010). The Arabidopsis model is beneficial because of the existence of extended genomic resources: mutant lines (transposons, T-DNA, RNAi), cDNA collections, more than 50 000 molecular markers. There are also a vast number of databases and molecular tools that provide a substantial amount of functional information or at least functional clues that help hypothesis generation. Features include subcellular localisation prediction, gene expression atlas (gene expression can be visualised in different tissues/organs and growth stages, in abiotic/biotic conditions, in response to hormones or chemical treatments, in specific known and important pathway mutants), protein interaction predictors, metabolome, methylome, cyclome (transcriptional landscape of plant circadian rhythms using genome tiling arrays) and proteome tools (phosphorylation databases, the lysine acetylome and many other tools of protein characterisation and post-translational modification), making it the most favourable model for the post-genomic era (Somerville and Koornneef, 2001; Zhang *et al.*, 2004; Wienkoop *et al.*, 2010; Feuillet *et al.*, 2011).

Since all the new molecular tools that sustain physiological studies have been a push forward into the advancement of knowledge on gene function, they permitted the use of genes either directly in heterologous systems or through their crop orthologs, so that many mechanisms were discovered and comprehended (Hilson *et al.*, 2003; Zhang *et al.*, 2004; Century *et al.*, 2008; Feuillet *et al.*, 2011). Also, the rapid development into the 'omics' era, with its different 'omics' tools, has helped clarify the genotype-phenotype relationships for fundamental and applied research. (Wienkoop *et al.*, 2010). In spite of all the *in silico* prediction tools, gene function is still best assessed by comparison of a visible phenotype between gain- or loss-of-function mutants and a wild-type plant (Feng and Mundy, 2006). Forward screenings act by choosing a biological process, followed by generating highly redundant mutant populations, screening the mutants for a desired phenotype and finally mapping and cloning the gene responsible for the phenotype. These screenings were traditionally the main functional genomics tool, based on the fact that genome mutations would produce the desired phenotype (Alonso and Ecker, 2006). Mutagenesis of a

certain population will produce different kinds of genome alterations depending on the methods used. Mutagenic agents include chemical mutagens like EMS that insert point mutations, physical mutagens that generate large rearrangements like insertions or deletions; and mutagens based on biological vectors, such as T-DNA, transposons and retrovirus (Feng and Mundy, 2006). Nowadays, because the many intensive forward screenings that have been carried out tend to saturate obvious visible phenotypes, reverse genetics has become increasingly popular. In reverse genetics, first a gene- or genes-of-interest are selected and then highly redundant mutant populations are generated. Nowadays, this is ensured by large collections of insertion mutant lines that cover ~96% of the Arabidopsis genes (MASC Report, 2010). The strategy then identifies mutants for the gene-of-interest and subsequently analyses the phenotype of the mutants (Alonso and Ecker, 2006). As stated, the strategy has been helped by the existence of extensive insertion mutant line collections. Using transposon or T-DNA lines, these mutants can be easily mapped to the genome by TAIL-PCR, with the advantage of having a mutation/genome ratio very close to 1 (Alonso *et al.*, 2003; Alonso and Ecker, 2006). Given that they are extensive, publicly available and gene-indexed collections of mutants, this means that nowadays an Arabidopsis gene-of-interest most likely possesses a knockout mutant allele.

Insertion mutants are not the only available strategy for loss- or gain-of-function studies. Gene silencing can be obtained through interference RNA strategies like the use of artificial microRNAs, to act as post-transcriptional regulators of target genes-of-choice (Bouché and Bouchez, 2001; Alonso and Ecker, 2006; Rubio-Somoza and Weigel, 2011). Disruption of gene function is normally associated to loss-of-function, but the same strategy can be adjusted to generate gain-of-function by activation tagging using transcription activators (Weigel *et al.*, 2000). Strong constitutive promoters can be used to regulate gene transcription to obtain an ectopic overexpression of the gene, also helping to elucidate its function. Also, promoter-reporter gene constructs can be used to obtain information of the gene's spatial and temporal expression pattern (e.g. luciferase and GUS). Information on subcellular protein localisation can be determined by chimeric fusions with reporter genes (e.g. GFP). All these strategies benefit immensely by profusion of vectors and cloning strategies available, including the systematic cloning of cDNA/ORF of Arabidopsis genes into cloning vectors, as well as the previously stated fact that Arabidopsis is easily transformed (Curtis and Grossniklaus, 2003; Alonso and Ecker, 2006).

1.3. Aims and outline of the thesis

Using separate strategies, functional characterisation of two sets of abiotic stress-related genes was initiated. *SQE2* and *SQE3* were singled out after *SQE1* mutant *dry2/sqe1-5* revealed a drought sensitive phenotype, therefore revealing a central role for sterols in drought tolerance and the regulation of ROS homeostasis. *EGY3* was singled out after a bioinformatics data-mining search of microarray data (*AtGenExpress-Heat series*), with the purpose of identifying a novel, functionally unresolved chloroplast-targeted gene that was induced, and specifically involved in the heat stress response. The main aim of the present work was to perform the functional characterisation of *SQE2*, *SQE3* and *EGY3*, in order to understand their role at both the molecular, cellular, and plant levels. More specifically, the objective was to determine their involvement in abiotic stress responses, also analysing their contribution to normal growth and development. Given that these genes are functionally unresolved, research aims included unravelling of subcellular localisation, tissue specificity and molecular function, which could enlighten its importance for the plant.

Concerning SQEs, specific aims included:

- characterisation of *SQE2* and *SQE3* at tissue level
- characterisation of *SQE3* at subcellular level
- profiling for changes in sterol contents
- molecular analysis of mutants and generated double mutants
- phenotype search of terbinafine sensitivity to understand or deepen a regulatory role of sterols in plants.
- phenotype search in abiotic-related stresses to understand *SQE2* and *SQE3* responses to those stresses.
- understand *SQE2* and *SQE3* relationship with *SQE1*

Concerning *EGY3*, specific aims included:

- characterisation of *EGY3* at tissue level
- molecular analysis of mutants
- molecular characterisation of *EGY3* inducible gene expression in the heat response
- identification of developmental and stress-related phenotypes in order to correlate its high expression in response to heat stress and a putative role in the heat response

- establishment of overexpression lines and heterologous expression constructs to respectively analyse its function *in planta* and determine its molecular function *in vitro*.

The present thesis is organised in an outline that reflects the parallel functional characterisation of both sets of genes. **Chapter 1** provides a general introduction to *Arabidopsis thaliana* as a model plant, to plant abiotic stress responses, and underlying research strategies that try to meet the worldwide demand of increased crop yield. **Chapter 2** provides the Material and Methods used to perform the experiments and analysis presented in the thesis. **Chapter 3** is devoted to the functional characterisation of SQEs. It includes an introduction to sterol and brassinosteroid pathways, giving an emphasis to their most relevant mutants and expanding on the present knowledge on SQEs (**subchapter 3.1**). It then presents the results and discussion of the work concerning the functional characterisation of *SQE2* and *SQE3* (**subchapter 3.2**). **Chapter 4** is devoted to the functional characterisation of *EGY3*, with an introduction to metalloproteases and their role in plant abiotic stress (**subchapter 4.1**), and presenting the bioinformatic analysis that led to the selection of *EGY3* along with the results and discussion of experimental work related to *EGY3* characterisation. **Chapter 5** includes the final remarks and future prospects concerning both research lines.

2.

"Materials and Methods"

2.1.

“Biological materials, culture media and reagents”

2.1.1. Biological material

2.1.1.1. Plant material

Arabidopsis thaliana (L.) Heynh. ecotypes and interruption mutant lines seeds were obtained from NASC (European Arabidopsis Stock Centre), a public seed stock centre (URL no.3). Transgenic lines were also generated in the course of the present work. All genotypes used in this work are described in table 2.1. Seeds were kept in the dark, under dry conditions, at room temperature, for a maximum of two years. *Nicotiana benthamiana* seeds were obtained from the University of Málaga seed stock (Rus, 2009). Seeds were kept in the dark, under dry conditions, at room temperature, for a maximum of two years.

Table 2.1. – *Arabidopsis thaliana* seeds used in the present work. Asterisk indicates plant material generated in the course of this work but not subjected to further analysis.

Genotype	Background ecotype	Mutant line	Mutation type	Origin
Col-0	-	-	-	NASC
Ler	-	-	-	NASC
<i>egy3-1</i>	Col-0	SALK_042231	T-DNA	NASC
<i>sqe2-1</i>	Col-0	SALK_064182	T-DNA Hm	NASC
<i>sqe3-1</i>	Col-0	SAIL_723_F01	T-DNA	NASC
<i>dry2/sqe1-5</i>	Ler	-	EMS	Posé et al. (2009)
<i>sqe2-1/sqe3-1</i>	Col-0	-	-	Present work
<i>dry2/sqe1-5</i> <i>SQE3/sqe3-1</i>	Col-0/Ler	-	-	Present work
<i>proEGY3::GUS</i>	Col-0	-	-	Present work
<i>proSQE2::GUS</i>	Col-0	-	-	Present work
<i>proSQE3::GUS</i>	Col-0	-	-	Present work
<i>proSQE1::SQE2</i>	<i>dry2/sqe1-5</i>	-	EMS	Present work
<i>proSQE1::SQE3</i>	<i>dry2/sqe1-5</i>	-	EMS	Present work
<i>35S::SQE3-GFP</i>	Col-0/ <i>sqe3-1</i>	-	-	Present work*
<i>35S::SQE1-GFP</i>	<i>dry2/sqe1-5</i> Ler	-	EMS -	Present work*
<i>las1-2</i>	Col-0	SAIL_676_A11	T-DNA	NASC
<i>las1-2/sqe3-1</i>	Col-0	-	T-DNA	Present work*
<i>las1-2/dry2</i>	Col-0/Ler	-	T-DNA/EMS	Present work*

NASC - European Arabidopsis Stock Centre (URL no.3)

2.1.1.2. Bacterial material

Escherichia coli and *Agrobacterium tumefaciens* strains used in molecular biology experiments are presented in table 2.2. *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*) was used in *Arabidopsis* infection studies.

Table 2.2. - Bacterial strains used in this work.

Species/Strain		Genotype	Reference
<i>Escherichia coli</i>	XL1 Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proABlacIqZΔM15 Tn10</i> (Tetr)]	(Bullock <i>et al.</i> , 1987)
	One Shot <i>ccdB</i> Survival T1	F- <i>mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araΔ139 Δ(ara-leu)7697 galU galK rpsL</i> (StrR) <i>endA1 nupG tonA::Ptrc-ccdA</i>	Gateway Technology with Clonase II (Invitrogen)
	DH5α	F, <i>endA1, hsdR17 (r_Km_K⁺), glnV44, thi-1, deoR, gyrA96, recA1, relA1, supE44, Δ(lacZYA-argF) U169, λ-, [φ80dlacΔ(lacZ) M15]</i>	(Griffith and Gietz, 2003)
	DB3.1	F- <i>gyrA462 endA1 glnV44 Δ(sr1-recA) mcrB mrr hsdS20(rB-, mB-) ara14 galK2 lacY1 proA2 rpsL20(Smr) xyl5 Δleu mtl1</i>	(Bernard and Couturier, 1992)
	TOP-10	F- <i>mcrA Δ(mrr-hsdRMS-mcrBC) phi 80lacZ ΔM15 ΔlacX74 deoR recA1 araD139 Δ(ara-leu) 7697 galU galK rpsL</i> (StrR) <i>endA1 nupG</i>	(Nelson <i>et al.</i> , 2007)
	BL21(DE3)p LysE	F- <i>ompT hsdSB(rB-,mB-) dcm galλ(DE3) pLysE, Cmr</i>	(Moffatt and Studier, 1987)
<i>Agrobacterium tumefaciens</i>	EHA105	EHA101 derivative (C58 pTiBo542;T-region::aph)	(Hood <i>et al.</i> , 1986)
	GV3101::pMP90	pMP90 (pTiC58DT-DNA)	(Koncz and Schell, 1986)
<i>Pseudomonas syringae</i>	DC3000	<i>Pseudomonas syringae</i> pv. <i>tomato</i> wild-type strain	(Cuppels, 1986)

2.1.1.3. Vectors

The vectors used in the present work are presented in table 2.3. For commercial vector maps look into Appendix III.

Table 2.3. – Vectors used in the present work.

Vector	Characteristics	Selection	Reference
pGEM-T-Easy	Subcloning vector for PCR fragments.	Ampicillin	(Marcus <i>et al.</i> , 1996)

2.1. - Biological materials, culture media and reagents

pDONR201	Gateway Entry vector for BP reaction.	Kanamycin	<i>Gateway Technology with Clonase II Manual</i> (Invitrogen)
pENTR/D-TOPO	Gateway Entry vector.	Kanamycin	<i>pENTR Directional TOPO Cloning Kits Manual</i> (Invitrogen)
pCAMBIA1303	35S:: <i>GUS</i> -GFP vector.	Kanamycin and Hygromycin	(URL no.4)
pMDC43/45	Gateway Destination vectors for GFP fusion in the N-terminus, under regulation of the 35S promoter.	Kanamycin and Hygromycin	(Curtis and Grossniklaus, 2003)
pMDC83	Gateway Destination vector for GFP fusion in the C-terminus, under regulation of the 35S promoter.	Kanamycin and Hygromycin	(Curtis and Grossniklaus, 2003)
Er-rb	Vector with fluorescent organelle markers. A sequence tagging to the ER fused with mcherry fluorescent CDS.	Kanamycin	(Nelson <i>et al.</i> , 2007)
G-rb	Vector with fluorescent organelle markers. A sequence tagging to the Golgi fused with mcherry fluorescent CDS.	Kanamycin	(Nelson <i>et al.</i> , 2007)
p19	Vector containing a viral-encoded suppressor of gene silencing, the p19 protein of tomato bushy stunt virus (TBSV), that prevents the onset of PTGS.	Kanamycin	(Voinnet <i>et al.</i> , 2003)
pET-25b(+)	Protein expression vector with an N-terminal <i>pelB</i> signal sequence for potential periplasmic localisation and a C-terminal Histidine tag.	Ampicillin	Novagen/Merck
pGEX-6P-1	Protein expression vector with C-terminal GST tag.	Ampicillin	GE Healthcare
p35S:: <i>Cerulean</i> -GWY	Gateway Destination vector for <i>Cerulean</i> fusion in the N-terminus, under regulation of the 35S promoter.	Ampicillin and Chloramphenicol	Paul Schulze-Lefert/Mark Kwaaitaal (personal donation)
pro <i>EGY3</i> :: <i>GUS</i>	<i>EGY3</i> promoter fused with <i>GUS</i> reporter gene in pCAMBIA1303.	Kanamycin and Hygromycin	Present work
pro <i>SQE2</i> :: <i>GUS</i>	<i>SQE2</i> promoter fused with <i>GUS</i> reporter gene in pCAMBIA1303.	Kanamycin and Hygromycin	Present work
pro <i>SQE3</i> :: <i>GUS</i>	<i>SQE3</i> promoter fused with <i>GUS</i> reporter gene in pCAMBIA1303.	Kanamycin and Hygromycin	Present work
pro <i>SQE1</i> :: <i>GUS</i>	<i>SQE1</i> promoter fused with <i>GUS</i> reporter gene in pCAMBIA1303.	Kanamycin and Hygromycin	Present work
pro <i>SQE1</i> :: <i>SQE2</i>	<i>SQE1</i> promoter fused with <i>SQE2</i> genomic DNA in pCAMBIA1303.	Kanamycin and Hygromycin	Present work
pro <i>SQE1</i> :: <i>SQE3</i>	<i>SQE1</i> promoter fused with <i>SQE3</i> genomic DNA in pCAMBIA1303.	Kanamycin and Hygromycin	Present work
<i>pSQE1CDS</i>	Entry clone in pDONR201 of <i>SQE1</i> CDS (with stop codon).	Kanamycin	Present work
<i>pSQE3CDS</i>	Entry clone in pENTR/D-TOPO of <i>SQE3</i> CDS (with stop codon).	Kanamycin	Present work
p35S:: <i>GFP-SQE1</i>	N-terminal fusion of GFP with <i>SQE1</i> CDS, under regulation of the 35S promoter, in pMDC43.	Kanamycin and Hygromycin	Present work
p35S:: <i>SQE1-GFP</i>	C-terminal fusion of GFP with <i>SQE1</i> CDS, under regulation of the 35S promoter, in pMDC83.	Kanamycin and Hygromycin	Present work

2.1. - Biological materials, culture media and reagents

p35S::GFP-SQE3	N-terminal fusion of GFP with <i>SQE3</i> CDS, under regulation of the 35S promoter, in pMDC45.	Kanamycin and Hygromycin	Present work
p35S:: <i>SQE3</i> -GFP	C-terminal fusion of GFP with <i>SQE3</i> CDS, under regulation of the 35S promoter, in pMDC83.	Kanamycin and Hygromycin	Present work
p35S::Cerulean- <i>SQE1</i>	N-terminal fusion of Cerulean fluorescent protein with <i>SQE3</i> CDS, under regulation of the 35S promoter in the p35S::Cerulean-GWY vector.	Ampicillin and Chloramphenicol	Present work
pEGY3-His	<i>EGY3</i> CDS fused with N-terminal <i>pelB</i> signal sequence for potential periplasmic localisation and a C-terminal HSV (herpes simplex virus) and histidine tag, in the protein expression vector pET25b(+).	Ampicillin	Present work
pGST-EGY3	<i>EGY3</i> CDS fused with C-terminal GST tag in the protein expression vector pGEX-6P-1.	Ampicillin	Present work
p35S:: <i>EGY3</i>	<i>EGY3</i> CDS controlled by 35S promoter in pCAMBIA1303. Vector for overexpression studies <i>in planta</i> .	Kanamycin and Hygromycin	Present work

2.1.1.4. Genes

The genes used in the experimental work are listed in table 2.4.

Table 2.4. – List of AGI codes and annotation for genes of *Arabidopsis thaliana* used in the present work. Annotation is according to TAIR 10 (URL no.5), and referenced publications.

Gene Name	ATG code	Annotation
<i>EGY3</i>	At1g17870	Ethylene dependent gravitropism-deficient and yellow-green-like 3. S2P-like putative metalloprotease (Chen <i>et al.</i> , 2005).
<i>SQE1</i>	At1g58440	Squalene epoxidase 1. Involved in sterol biosynthetic processes (Rasbery <i>et al.</i> , 2007; Posé <i>et al.</i> , 2009).
<i>SQE2</i>	At2g22830	Squalene epoxidase 2. Involved in sterol biosynthetic processes (Rasbery <i>et al.</i> , 2007).
<i>SQE3</i>	At4g37760	Squalene epoxidase 3. Involved in sterol biosynthetic processes (Rasbery <i>et al.</i> , 2007).
<i>ACT2</i>	At3g18780	Actin 2. Constitutively expressed in vegetative structures.
<i>LAS1</i>	At3g45130	Lanosterol synthase 1. Catalyzes the reaction from oxidosqualene to lanosterol. Oxidosqualene cyclase (OSC).
<i>CAS1</i>	At2g07050	Cycloartenol synthase 1. Catalyzes the reaction from oxidosqualene to cycloartenol. OSC.
<i>LUP1</i>	At1g78970	Lupeol synthase 1. Converts oxidosqualene to multiple triterpene alcohols and a triterpene diols. OSC.
<i>LUP2</i>	At1g78960	Lupeol synthase 2. Encodes a multifunctional 2-3-oxidosqualene-triterpene cyclase that can cyclize OS into lupeol, α - and β -amyrin.
<i>CAMS1/LUP3</i>	At1g78955	Camelliol C synthase 1. OSC.

<i>BAS/LUP4</i>	At1g78950	β -amyrin synthase activity. OSC.
<i>LUP5</i>	At1g66960	Lupeol synthase activity. OSC.
<i>BARS1/PEN2</i>	At4g15370	Baruol synthase 1. Pentacyclic triterpene synthase 2. OSC.
<i>PEN3</i>	At5g36150	Putative pentacyclic triterpene synthase 3. OSC.
<i>THAS1/PEN4</i>	At5g48010	Thalianol synthase 1.OSC. <i>THAS</i> is part of a small operon-like cluster of genes (with At5g48000 (<i>THAH</i>) and At5g47990 (<i>THAD</i>)) involved in thalianol metabolism (Field and Osbourn, 2008).
<i>MRN1/PEN5</i>	At5g42600	Marneral synthase 1. OSC.
<i>PEN6</i>	At1g78500	Putative pentacyclic triterpene synthase 6. OSC. In addition to the compounds lupeol, α -amyrin and bauerenol, this enzyme was also shown to produce two seco-triterpenes: α - and β -seco-amyrin.
<i>PEN7</i>	At3g29255	Putative pentacyclic triterpene synthase 7. OSC.

2.1.2. Culture media

2.1.2.1. In vitro plant culture media

Arabidopsis seeds were germinated in a Murashige and Skoog-based (MS) medium (Table 2.5) (Murashige and Skoog, 1962).

Table 2.5. - Composition of MS-based culture media used for growing Arabidopsis seedlings.

	Medium components	MS	MS 0.5x
Macronutrients	MgSO ₄ .7H ₂ O	1.65 g L ⁻¹	0.825 g L ⁻¹
	CaCl ₂ .2H ₂ O		
	KNO ₃		
	KH ₂ PO ₄		
	NH ₄ NO ₃		
Micronutrients	MnSO ₄ .H ₂ O	1 g L ⁻¹	1 g L ⁻¹
	KI		
	CoCl ₂ .6H ₂ O		
	ZnSO ₄ .7H ₂ O		
	CuSO ₄ .5H ₂ O		
	H ₃ BO ₃		
	Na ₂ MoO ₄ .2H ₂ O		
Fe/EDTA	FeSO ₄ .7H ₂ O	0.5 g L ⁻¹	0.5 g L ⁻¹
	Na ₂ EDTA.2H ₂ O		
Buffer	MES	0.5 g L ⁻¹	0.5 g L ⁻¹
Carbon source	Sucrose	1.5% (w/v)	1.5% (w/v)
	Agar	1.2% (w/v) or 0.8% (w/v)	1.2% (w/v)
pH 5.7			

For phenotype analyses, the medium was used at half macronutrient strength. Concentration of agar was chosen depending on the orientation of growth: horizontal 0.8% (w/v); vertical 1.2% (w/v). The culture media were sterilised by autoclaving recipients for 15-20 min at 121°C and 1 atm. For the selection of transformants, seeds were plated onto a medium containing MS supplemented with the appropriate antibiotics, namely hygromycin (40 µg mL⁻¹) for vector selection and ticarcillin (250 µg mL⁻¹) for reducing *Agrobacterium* contamination.

2.1.2.2. Microbiology culture media

Different culture media were used while growing and maintaining bacterial strains (Table 2.6). Agarised media was obtained by adding 1.5% (w/v) agar to the media's broth. Recombinant selection using the *lacZ* gene was performed by supplementing the appropriate medium with 50 µg mL⁻¹ of IPTG (from 50 mg mL⁻¹ in water) and 40 µg mL⁻¹ X-Gal (from 20 mg mL⁻¹ in DMF). The culture media were sterilised by autoclaving for 20 min at 121°C and 1 atm.

Table 2.6. - Composition of culture media used for growing bacteria strains.

Culture medium	Composition	Purpose	Species/Strain
LB	1% (w/v) NaCl 1% (w/v) bacto-tryptone 0.5% (w/v) yeast extract pH 7.0	Growth and maintenance	<i>E. coli</i> <i>A. tumefaciens</i>
LB*	1% (w/v) NaCl 1% (w/v) bacto-tryptone 0.5% (w/v) yeast extract pH 5.4	Growth for floral dipping transformation	<i>A. tumefaciens</i>
LB-Kan	LB 50 µg mL ⁻¹ Kanamycin	Selection/maintenance of transformants carrying Kan resistance mark	<i>E. coli</i>
LB-Amp	LB 100 µg mL ⁻¹ Ampicillin	Selection/maintenance of transformants carrying Amp resistance mark	<i>E. coli</i>
LB-Hyg	LB 50 µg mL ⁻¹ Hygromycin	Selection of transformants of the Gateway LR reaction when using pMDC vectors	<i>E. coli</i>
LB-Cm	LB 34 µg mL ⁻¹ Chloramphenicol	Growth and maintenance of <i>E. coli</i> strains for heterologous expression.	<i>E. coli</i>
LB-Amp-Cm	LB 100 µg mL ⁻¹ Ampicillin 34 µg mL ⁻¹ Chloramphenicol	Growth and maintenance of <i>E. coli</i> strains containing vectors for protein heterologous expression.	<i>E. coli</i>
LB-Rif-Kan	LB 50 µg mL ⁻¹ Rifampicin 50 µg mL ⁻¹ Kanamycin	Maintenance of transformants carrying Kan resistance mark	<i>A. tumefaciens</i>
LB-Rif-Clx	LB 15 µg mL ⁻¹ Rifampicin 2 µg mL ⁻¹ Cycloheximide	Growth /estimation of infection level	<i>Pseudomonas syringae</i> DC3000

SOB	0,5 % (p/v) yeast extract 2% (p/v) tryptone 10 mM NaCl 2,5 mM KCl 10 mM MgCl ₂ 10 mM MgSO ₄ pH 7.5	Production of competent cells	<i>E. coli</i>
SOC	SOB 20 mM glucose	Production of competent cells	<i>E. coli</i>

2.1.3. Reagents

All chemicals used for molecular biology methods and nucleic acid extractions were *Molecular Biology* grade. The remaining chemicals were *p.a.* grade.

2.1.4. Bacteria handling

2.1.4.1. Growing of bacterial strains

E. coli and *Agrobacterium* strains were grown in the appropriate medium as indicated in table 2.6. To isolate single colonies, the strains were stricken onto appropriate agarised medium and incubated overnight at 37°C (*E. coli*) or 28°C (*Agrobacterium*). Liquid cultures were obtained by inoculating the medium with a single colony and incubating at 37/28°C, with agitation (150-250 rpm).

2.1.4.2. Glycerol stock preparation

Long term viable stocks of *E. coli* and *A. tumefaciens* strains were prepared by inoculating single colonies in the appropriate liquid medium (table 2.6), followed by growth until the late exponential growth phase was achieved. Aliquots were then added to sterile glycerol-containing cryotubes to a final 20% (v/v) glycerol concentration. The vials were immediately stored at -80°C. To maintain cell viability, stocks were recovered by scraping off splinters of solid ice with a sterile wire loop.

2.2.

"Plant methods"

2.2.1. Arabidopsis sterilisation, germination and growth

Arabidopsis handling includes specific experimental steps and conditions necessary to promote seed germination and growing, which are described in this section.

Seed sterilisation. Arabidopsis seed dormancy was broken by stratification in water at 4°C, for 2-4 days. Superficial sterilisation was performed in a horizontal flow chamber (*BBH4 BRAUN Horizontal*). Seeds were consecutively transferred to 70% (v/v) ethanol, for 5 min; 20% (v/v) commercial bleach, for 10 min and a five time wash with sterile ultra-pure water. Between incubations, seeds were resuspended by vortexing, and collected by centrifugation for 1 min and 10000 g. Seeds were resuspended in sterile 0.20% (w/v) agarose.

Sowing and growth. Seeds were sowed onto plates containing MS agar medium, using a 1 mL micropipette tip for individual seed dispersion. Petri plates were sealed with *Parafilm*, and placed either vertically or horizontally, in a culture chamber at 23°C, under cool white light (70-100 $\mu\text{mol PAR m}^{-2} \text{sec}^{-1}$), and a long-day photoperiod (16 h light and 8 h dark regime). Plants were occasionally grown in a short-day photoperiod (8 h light and 16 h dark regime). Seedlings were transferred to soil after 7-10 days of *in vitro* growth. Soil was prepared by mixing turf (*Floragard*) and vermiculite in a 4:1 ratio. Plants were covered with plastic for 2-4 days, to acclimate to the new environment, preventing dehydration. Seeds were collected at the end of the life cycle, using an appropriate sieve with metallic mesh.

2.2.2. *Nicotiana* germination and maintenance

Nicotiana benthamiana seeds were placed directly on turf-rich soil (*Floragard*) and covered with plastic for one week to favour high humidity conditions. Plants were grown at 23°C, with 70% relative humidity, under cool white light (70-100 $\mu\text{mol PAR m}^{-2} \text{sec}^{-1}$), and a long-day photoperiod (16 h light and 8 h dark regime). Leaves from 3-week-old plants were used for transient transformation.

2.2.3. Phenotype tests in Arabidopsis

In order to access gene function, mutants were submitted to distinct imposed conditions, and their phenotype analysed and compared to the one exhibited by the wild-type. Germination analysis was performed after heat shock treatments. *In vitro* root growth was analysed in heat-, salt- and osmotic-stressed plants and also in plants grown in the presence of ABA. Seedling viability was also assessed after heat shock treatment, as described in the following segments.

Seed germination after heat shock. The germination rate was determined after heat shock stress. Following sterilisation, seeds were resuspended in 1 mL of sterile ultra pure water. Microtubes were sealed with *Parafilm*, and placed in a thermal bath at a pre-determined time, and temperature setting (Table 2.7). Subsequently, seeds were collected by centrifugation for 1 min and 10000 *g* and resuspended in sterile 0.20% (w/v) agarose. For imposing heat shock conditions described in table 2.7, 0.5x MS medium plates were (i) divided in eight equally sized segments, and seeds plated at a density of 25 seeds per segment, (ii) divided in twelve equally sized segments, and seeds plated at a density of 20 seeds per segments, half wild-type, half mutant . Four replicas of each condition were performed, and seed germination was estimated every day for 9-10 days, considering germination as the emergence of the hypocotyls. Plates were grown horizontally under cool white light (70-100 $\mu\text{mol PAR m}^{-2} \text{sec}^{-1}$), and a long-day photoperiod. Tests were made using mutant seeds and the background wild-type ecotype as control.

Table 2.7. – Heat shock conditions applied to Arabidopsis seeds used for assessing heat shock-induced phenotype alterations.

Experiment	Temperature	Time
Temperature range	i) 23, 38, 41, 44, 47, 50, 53 and 56°C ii) 23, 47, 49, 51, 53 and 55°C	60 min
Time course	i) 50°C	0, 15, 30, 60, 120, 180, 240 and 300 min

In vitro root growth. After seed plate germination (section 2.2.1), 7-day-old seedlings of both mutant and background ecotypes were transferred in a horizontal flow chamber to new plates containing 0.5x MS with or without supplementation. Media supplementation and growth conditions are depicted in table 2.8. Plates were sealed with *Parafilm* and grown vertically under cool white light (70-100 $\mu\text{mol PAR m}^{-2} \text{sec}^{-1}$), and a long-day photoperiod. Growth was estimated by measuring root growth, every two days, after stress imposition (day 0).

Seedling phenotype test – seedling in vitro stress test. Seeds of both mutant and background ecotypes were plated in 0.5x MS medium, sealed with *Parafilm* and grown horizontally under cool white light (70-100 $\mu\text{mol PAR m}^{-2} \text{sec}^{-1}$), and a long-day photoperiod. After seven days, plates were subjected to heat shock by submersion in a thermal bath for different times or at different temperatures (Table 2.8). Plates were then returned to the normal growth conditions, and three days after the heat shock seedling viability (presence of green cotyledons) or root growth was evaluated.

Table 2.8. - MS medium (0.5x) supplementation and temperatures used in seedling growth for phenotype analysis.

Test	Concentration	Temperature (°C)	Type of test
Mannitol	0, 1, 2, 5 % (w/v)	23°	Root measurement
NaCl	0, 50, 100, 150 mM	23°	Root measurement
ABA	0, 20, 40, 60, 80 µM	23°	Root measurement
Heat (1)	-	37° (2 h) / 37° (2 h) + 23°C (2 h) + 44° (2 h) / 44° (2 h)	Root measurement
Heat (2)	-	2 h at 23°, 30°, 34°, 37°, 40°, 42°, 44° and 45°	Root measurement
Heat (3)	-	45° for 0, 15, 20, 25 and 30 min	Viability/survival assessment
Heat (4)	-	45° for 0, 15, 20, 22.5 and 25 min	Viability/survival assessment

In soil plant stress test - HS+HL+Pq. Phenotype tests were also performed *in vivo*, using soil-grown plants to test for combined heat shock (HS), high light (HL) and oxidative stress (Paraquat). Plants were planted in soil, each pot containing two mutant plants and two plants from the background ecotype, in a total of nine pots per experiment. Plants were grown under long-day photoperiod at 23°C until they were 4-week-old (Nishizawa *et al.*, 2006). Plants were subjected to three different stress conditions: placed at 45°C with high light intensity (300 µmol PAR m⁻² sec⁻¹) and sprayed with 50 µM of Paraquat (Pq). They were photographed each hour for a period of eight hours.

Mini-Boyes developmental phenotypic test. Plants were evaluated throughout their life cycle, with specific timings of growth being analysed following a simplified version of Boyes *et al.* (2001). Plants were grown in a culture chamber at 23°C, under standard cool white light (70-100 µmol PAR m⁻² sec⁻¹), and a long-day photoperiod, with 14 plants attributed to each genotype (wild-type and *egy3-1* mutant). Developmental traits were evaluated within three and six weeks for the number of leaves per rosette and the radius of the plant's rosette. Plants were also monitored to identify the first day of flowering for each plant.

EZ-Rhizo phenotypic test. Arabidopsis seeds were grown vertically in a culture chamber at 23°C, under cool white light (70-100 µmol PAR m⁻² sec⁻¹), and a long-day photoperiod, in MS medium containing 0.5% phytoagar. After four days, the experiment was initiated by placing half of the plates to grow at 27°C. Plants were then photographed daily for a total of nine days. The EZ-Rhizo software was used to make a fast and accurate measurement of the root system architecture (Armengaud *et al.*, 2009). EZ-Rhizo is a reliable, semi-automated and easy-to-use software that

measures multiple parameters of young plants grown in vertical plates, namely the main root length, no. of root length per cm, average lateral root length and total root length. Due to the increasingly complex architecture of the root, photographs of plants within four days of the beginning of the experiment were selected for analysis.

2.2.4. Arabidopsis transformation by floral dipping

Arabidopsis floral dip transformation was performed according to a modification of the procedure described by Clough and Bent (1998). Arabidopsis plants were grown in four large pots, at a density of 4 plants per pot, under normal conditions, until the early bolting stage (bolt with 10-15 cm) was achieved. *Agrobacterium* single colonies were obtained by growing cells in appropriate selection medium, at 28°C, and used to inoculate 7 mL of LB liquid media, which were incubated ON at 28°C and 200 rpm. Aliquots of 500 µL of these cultures were used to inoculate 200 mL of LB* (LB pH 5.4) liquid medium (supplemented with the antibiotic of the corresponding construct and acetosyringone to a final concentration of 19.6 mg mL⁻¹), and incubated ON at 28°C and 200 rpm. Cells were collected by centrifugation for 12 min at room temperature and 5000 rpm, and the pellet was resuspended in a 500 mL cup containing 250 mL of 5% (w/v) sucrose. After addition of 125 µL of Silwett L-77 to the *Agrobacterium* suspension, the aerial part of plants (from which the siliques were previously removed) was dipped in the solution for 20 sec. Plants were laid down in a tray, covered with plastic and placed in the shade/dark for one day. The plastic was then removed, and plants were grown normally for the rest of their life cycle.

2.2.5. Selection of Arabidopsis transformants

After transformed plants completed the life cycle, T1 seeds were collected and grown in 0.8% (w/v) agar MS-medium containing hygromycin (40 µg mL⁻¹) and ticarcillin (250 µg mL⁻¹) for transformant selection. Positive control (resistant) seeds were also sowed onto the plate. Positive T1 transformants (similar to control plants) were transferred to soil after 10 days to complete their life cycle. A total of 20-40 T2 seeds per T1 plant were germinated in identical selective medium. Plants with only one T-DNA insertion, 3:1 (positive:negative) ratio were selected and grown on soil. Seeds were again collected (T3) and germinated in the same conditions to verify if the line was homozygous (1:0) or heterozygous (3:1) for the T-DNA insertion. Homozygous plants were selected for further analysis.

2.2.6. Plant transient transformation

In this section plant transformation experimental procedures are described for transient protein expression, using *Agrobacterium* infiltration in either tobacco or *Arabidopsis* plants.

***Nicotiana benthamiana* transient transformation.** Single colonies of *Agrobacterium* strains were obtained from -80°C stocks, by streaking onto plates and growing for 2 days at 28°C. A single colony was then transferred to 12 mL of LB liquid medium and grown ON at 28°C and 200 rpm. Cells were centrifuged for 15 min at 4000 *g*. The pellet was resuspended in 1 mL agro-infiltration buffer and grown until a final A_{600} of 0.5 was obtained. Agro-infiltration of constructs was co-infiltrated with a suppressor of gene silencing, the p19 protein of tomato bushy stunt virus (TBSV) to prevent the onset of PTGS, and so any silencing of the inserted transcript intended to be expressed (Voinnet *et al.*, 2003). When using *Agrobacterium* with the p19, ideal A_{600} was of 0.25. The resuspended pellet plus the agro-infiltration buffer were incubated for 2 hours and subsequently infiltrated with a 1 mL syringe in the abaxial side of the leaf. After 3 days, leaf tissue was seen under a *TCS SP5 II* confocal microscope (Leica).

Agro-infiltration buffer: 10 mM MgCl₂; 10 mM MES, pH 5.6; 150 μM acetosyringone.

***Arabidopsis thaliana* transient transformation.** Transient transformation of *Arabidopsis* was performed according to a modification of the method described by Marion *et al.*, (2008). Seeds were germinated and grown horizontally for 5 days before agro-infiltration. *Agrobacterium* strains were refreshed in a plate, containing appropriate selection medium, from -80°C stocks, and grown for 2 days at 28°C. For each plate of seedlings to transform, biomass of a single colony was then transferred to 5 mL LB liquid medium and grown ON at 28°C and 200 rpm. The following day, 30 mL of LB liquid medium were inoculated with the 5 mL of the pre-inoculum and grown ON at 28°C and 200 rpm. When the $A_{600} > 2$, cells were collected by centrifugation for 15 min and 4000 rpm, and resuspended in 15 mL of 5% (w/v) sucrose containing 200 μM acetosyringone. In a flow chamber, the solution was spread onto the plates covering the seedlings and agro-infiltrated in vacuum (400 mmHg) for two times 1 min per plate. The liquid was removed and plants were grown for another 3 days under standard conditions. Seedlings were observed under the *Epifluoresce microscope MZ FLIII* (Leica) to check for positive transient transformants, before observation in a *TCS SP5 II* confocal microscope (Leica).

2.2.7. Infection of Arabidopsis by *P. syringae*

Arabidopsis plants subjected to infection studies were grown at 23°C, under standard cool white light (70-100 $\mu\text{mol PAR m}^{-2} \text{sec}^{-1}$), short-day photoperiod and were used at ~4 weeks of age (rosette without bolt). *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*) strain, was refreshed 2 days earlier in agarised LB medium, supplemented with appropriate selection medium (15 $\mu\text{g mL}^{-1}$ rifampicin and 2 $\mu\text{g mL}^{-1}$ cycloheximide), and grown for 48 h at 28°C. Biomass was taken from the plate into a 2 mL solution of 10 mM MgCl_2 , vortexed and its A_{600} was measured. For expression analysis after 0, 8 and 24 h infection, plants were inoculated with $5 \cdot 10^7$ c.f.u. mL^{-1} ($A_{600} = 0.1$). For the elicitation experiments, three leaves per plant (mutants and wild-type) were inoculated with $5 \cdot 10^4$ c.f.u. mL^{-1} of *Pto* DC3000 in 10 mM MgCl_2 , in the abaxial side. Samples were taken at time zero and every two days, for a maximum of four days. Leaf discs (10 mm diameter) were cut from elicited leaves with a cork borer (Sigma) and homogenised with 1 mL of 10 mM MgCl_2 . Subsequently, aliquots of 100 μL of the homogenates were plated onto LB agarised medium with the appropriated antibiotics. Appropriate serial dilutions were made to correctly access c.f.u. number. Plates were grown for 48 h at 28°C and the number of c.f.u. subsequently determined (adapted from Macho *et al.*, 2010).

2.2.8. Stomata aperture response

Arabidopsis plants (2-3 week-old) were used to access the stomata aperture response under normal conditions and after ABA treatment. Leaves were cut and placed in a microtube containing 1 mL of stomata aperture buffer, and incubated in a short-day culture chamber at 23°C, under light of 70-100 $\mu\text{mol PAR m}^{-2} \text{sec}^{-1}$ for 3 hours. For ABA experiments, ABA was added into the samples to a final concentration of 5 μM , which were incubated for one additional hour in the same conditions as before. After incubation, leaf epidermis was removed using a two face adhesive tape, stained with 0.2% (w/v) toluidine blue and washed two times with water. Leaf epidermis was mounted in a microscope slide in water and covered with a cover slip. Stomata were observed in a microscope TCS NT (Leica) and photographed to allow measurements using the *Image J software*.

Stomata aperture buffer: 10 mM KCl; 50 μM CaCl_2 ; 5 mM MES; 0.01% (v/v) Tween-20; pH 5.6 (autoclave)

2.2.9. Root hair measurement

Root hair measurements were performed according to a modification of the procedure described by Posé *et al.* (2009). Seeds were surface sterilised and plated on solid MS medium

containing 0.5% (w/v) Phytigel (Sigma), at a pH of 5.8. Four days later, photos were taken with a photographic camera (*Nikon Coolpix 4500*) attached to a stereomicroscope *MZ FLIII* (Leica). Photographs of the intermediary region of the root were taken (about 20 plants per genotype). Measurement of root hair length was made using Image J software.

2.2.10. Terbinafine inhibition of germination assay

The experiment was performed in Dr. Albert Ferrer's group (Departament de Bioquímica i Biologia Molecular, Facultat de Farmàcia, Universitat de Barcelona). Terbinafine is a specific inhibitor of squalene epoxidases (SQE) enzymes. This inhibitor belongs to the class of allylamines and is a specific non-competitive inhibitor of fungal SQE, and a less potent and competitive inhibitor of the mammalian enzyme (Ryder, 1992; Nieto *et al.*, 2009). Wentzinger *et al.* (2002) demonstrated its effects on tobacco cells and later Nieto *et al.* (2009) in *Arabidopsis* plants. Seeds were sterilised and sown on solid MS supplemented with terbinafine (kindly supplied by Novartis Farmacéutica), to the final concentrations of 0, 0.1, 0.25, 0.5 and 1 μM from different terbinafine solutions (1 mM, 2.5 mM, 5 mM and 10 mM), prepared from a 100 mM terbinafine stock in DMSO. The final concentration of DMSO was kept constant in all plates regardless of the final concentration of the inhibitor. Plates were grown at 22°C under short-day photoperiod. Quantification of plant sensitivity to terbinafine was measured by the percentage of seedlings that developed true green leaves after 23 days.

2.2.11. Sterol and squalene analysis

Sterol analysis was conducted in seedlings (14-day-old) and in adult plants (1-month-old). For the seedling phase, plants were grown in agarised MS-medium for 14 days and frozen in liquid nitrogen after separating root from shoot tissue (an average of 250 mg of roots and 500 mg of shoot per sample). At the adult stage, leaves were placed in microtubes and frozen in liquid nitrogen (~500 mg per sample). Sterol levels were determined by Dr. Albert Ferrer's group (Departament de Bioquímica i Biologia Molecular, Facultat de Farmàcia, Universitat de Barcelona). Briefly, frozen plant material (seedling/leaves) was ground to a fine powder and lyophilized overnight (14-16 hours). A sample of 10 μL of 1.2 $\mu\text{g } \mu\text{L}^{-1}$ 5 α -cholestane in hexane was added to each sample (15-30 mg) as an internal standard. For sterol extraction, samples were suspended in 5 mL of (2:1, v/v) dichloromethane:methanol, sonicated for 10 min in a water-bath sonicator, and centrifuged for 5 min and 5000 *g*. The supernatant fraction was recovered to a new tube and the pellet was re-extracted twice. The combined extracts were then dried in a vacuum rotary evaporator

(*Mivac Duo Concentrator*, GENEVAC). The sterol ester fraction in the residue was saponified in 1.5 mL of 7.5% (w/v) KOH in 95% (v/v) methanol ON at room temperature. The saponification reaction was stopped by adding 1.5 mL of water. The resulting free sterols were extracted three times in 5 mL hexane and the combined extracts dried in a vacuum rotary evaporator (*Mivac Duo Concentrator*, GENEVAC). To derivatize sterols prior to GC-MS analysis, the extract was resuspended in 150 μ L bis(trimethylsilyl)-trifluoroacetamide (Merck) and incubated for 1 h at 150°C. Derivatized sterol samples were dried under nitrogen stream and resuspended in 200 μ L iso-octane. Samples (1 μ L) were submitted to gas chromatography-mass spectrometry (GC-MS) analysis using a *Trace GC Ultra gas chromatograph* coupled to an *ITQ900 mass spectrometer* (Thermo Scientific) with a capillary column DB-5MS (5% phenylpolysiloxane, 95% methylpolysiloxane) of 30 m length, 0.25 mm diameter, and 0.25 μ m film thickness (AWScientific). Helium was utilized as a carrier gas at a constant flow rate of 2 mL min⁻¹. A Triplus injector was used with a split flow of 12 mL min⁻¹. Injector temperature was kept at 280°C. The initial oven temperature was 245°C. It was programmed at a rate of 2°C min⁻¹ until 265°C and then at a rate of 3.5°C min⁻¹ until 290°C. The mass selective detector was run under standard electron impact conditions. Integration of chromatograms resulting from GC-MS analysis was performed using *Xcalibur software* (Thermo Scientific) and quantification of sterols was based on the relative peak area of 5 α -cholestane.

2.2.12. Grafting

Arabidopsis seeds were germinated on solid 0.5x MS medium with 0.6% (w/v) phytigel (Sigma), grown vertically for 4 days and then transferred to a 0.22 μ m sterile filter (Millipore) previously placed in contact with the medium. Seedlings were grown under long-day photoperiod and standard conditions. Three days later, seedlings were grafted in a flow chamber with the aid of a VWR stereomicroscope, a sterile razor blade (to cut), and a sterile tweezer (to move plants). The type of grafting was a wedge graft (Y shape) as described in Turnbull *et al.* (2002; 2010). Afterwards, plates were wet with sterile u.p. water, sealed with *Parafilm* and grown vertically for another seven days without moving the plates. Selected grafts were put in soil with a high content of water and covered with plastic to maintain a high humidity content. Three days later, the plastic cover was punctured to allow plants a slow, but efficient adaptation to a normal humidity condition. After ~3 days, the plastic was completely removed, and 2-3 days later watering was halted for four consecutive days to enhance phenotypic characteristics, since *dry2/sqe1-5* plants have a peculiar drought phenotype. Aerial and root tissue were removed and DNA extraction was performed to

confirm successful grafting. Since the *SQE1* promoter in Ler (background of *dry2/sqe1-5*) has a In/Del when compared to Col-0 (background of *sqe3-1*), a PCR analysis was carried out to find the root genotype. The difference between Ler and Col-0 ecotypes is listed in the Primers section 2.3.1.

2.2.13. GUS histochemical staining

Promoter regions were determined based on the analysis of the upstream intergenic region and the AGRIS database (Davuluri *et al.*, 2003). Promoter regions were subsequently cloned into the pCAMBIA1303 vector, transformed in *Agrobacterium* and used to transform *Arabidopsis* by the floral dipping method. Positive homozygous transformants were used for GUS histochemical staining, as previously described (Jefferson, 1987). Seedlings were subjected to vacuum infiltration of the GUS staining solution for four times 10 min, and incubated ON at 37°C. Chlorophyll discoloration was made with sequential extractions in 50, 75 and 90% (v/v) ethanol. Plants in different developmental stages were observed using bright field analysis in a *Leica DM5000 B* microscope coupled with a CCD colour camera (*Leica DFC 320*) or a *Wild M8* (Wild Heerbrugg) stereomicroscope coupled to the same CCD camera. GUS histochemical staining was also performed in adult tissues in response to 0, 1 and 3 h of heat stress (37°C) prior to the staining, to see GUS induction through time of exposure.

Roots were cross-sectioned to determine histochemical GUS staining localisation. After GUS staining, roots were fixed by incubation for 15 min with 4% (w/v) paraformaldehyde, transferred to another solution in 4% (w/v) paraformaldehyde and vacuum infiltrated for 2 min. Roots were again transferred to a new 4% (w/v) paraformaldehyde and incubated ON at 4°C. Three or four roots were aligned in a plate with a 1 mm layer of 1.5% (w/v) agarose and then covered with melted 1.5% (w/v) agarose. After solidification a cube of 0.5-1 cm was cut and washed two times with deionised water for 15 min. Embedding in methacrylate and subsequent sectioning was performed at ICVS (University of Minho). Embedding was carried out using the *Technovit 7100 kit* (Heraeus Kulzer), according to the manufacturer's instructions. Microtome thin sectioning was performed with 10-20 µm thickness. Sections were visualised using a *Leica DM5000 B* microscope coupled with a CCD colour camera (*Leica DFC 320*).

GUS staining solution: 190 µM X-Glu; 20% (v/v) methanol; 0.5 mM potassium ferricyanide; 0.5 mM potassium ferrocyanide; 0.3% (v/v) Triton X-100; 0.1 M sodium-phosphate buffer pH 7.

4% (w/v) paraformaldehyde: from a 10% (w/v) paraformaldehyde in PBS, pH 7.4.

PBS: 0.8% (w/v) NaCl; 0.02% (w/v) KCl; 0.144% (w/v) Na₂HPO₄; 0.024% (w/v) KH₂PO₄.

2.2.14. Seed analysis

Seeds/embryos were analysed prior to full silique development. Siliques of *dry2/dry2* *SQE3/sqe3-1* plants were collected, opened and the total number of aborted or regular seeds/embryos was estimated. The analysis was performed in a *Leica Zoom 2000* bench stereomicroscope with the help of tweezers. Analysis was made in 5-11 siliques per plant in five independent plants.

2.2.15. Arabidopsis cross-fertilisation

Plants were subjected to artificial fertilisation in order to obtain double mutants (e.g. *sqe2-1/sqe3-1*). Crosses were made with the assistance of a *Leica Zoom 2000* bench stereomicroscope and special crossing tweezers. Siliques, flowers, and opened buds were excised from female donors. The closed buds were opened and all organs were removed with the exception of the carpel. In the male donor, a flower was removed with a closing tweezer, and the pollen from the anthers was placed at the surface of the carpel stigma of the female donor to promote fertilisation. Crosses were carried out using both genotypes as female and male, since infertility of one mutant might be observed. After crossing, the stem of the plant was signalled and formation of a viable silique was observed within two days. Seeds were later recovered. Double mutants were identified by diagnostic PCR genotyping of the F2 generation.

2.3.

"Molecular biology methods"

2.3.1. Primers

Primer design followed Griffin and Griffin (1994) parameters using the 6th version of *OLIGO Primer Analysis Software* (URL no.6) or *Primer3* (URL no.7). Primers were synthesised by Metabion or Sigma. Primers used in this work are listed in table 2.9.

Table 2.9. - List of oligonucleotides used in this work.

Name	Sequence (5'-3')	Primer size (bp)	Primer T _m (°C)	Amplification product size (bp)	Use
Egy3proFx1	TTGAATTCATCTAATGACTTAAGGTTT	28	57.8	285	EGY3 promoter amplification for <i>GUS</i> fusion in pCAMBIA1303.
Egy3ProRx1	AGCCATGGTTTACGAAAACCTC	21	55.9		
Egy3dpcrLP	GTGGGTTTTCTGAAACCC	20	64.2	1012	Genotyping SALK_042231 mutant line interrupting <i>EGY3</i> .
Egy3dpcrRP	GGAAGCAAATGAGTCCAAGG	20	63.5		
LBb1SALKdpcr	GCGTGGACCGCTTGCTGCAACT	22	63.6	-	Left border primer for genotyping SALK T-DNA lines.
LBb1.3SALKdpcr	ATTTTGCCGATTTTCGGAAC	19	63.6	-	Left border primer for genotyping SALK T-DNA lines.
LB3SAILdpcr	TAGCATCTGAATTCATAACCAATCTCGATACAC	34	69	-	Left border primer for genotyping SAIL T-DNA lines.
Egy3sobFw	CCATGGCTTCTCTTTGTTTCTA	24	62	1730	EGY3 CDS amplification for overexpression in pCAMBIA1303.
Egy3sobRv	CACGTGTTAGAAATCATCACACGGA	26	66		
Egy3rtpcrFw	GCAGCAAGAGATGGATTGGAA	21	59	gDNA- 388; cDNA-295	EGY3 amplification for semi-quantitative RT-PCR
Egy3rtpcrRv	GCGAACATCCGTAGCGAAAA	20	58		

2.3. - Molecular biology methods

Egy3GC1	AAAAAGCAGGCTTAACCATGGCTTCTCTCTTTGT TTCT	38	76.4	gDNA- 2180; cDNA-1752	EGY3 CDS amplification for first step PCR for Gateway BP recombination. Forward primer.
Egy3GE1	AGAAAGCTGGGTAGAAATCATCACCACGGAAGA AAG	36	76.9		EGY3 CDS amplification without stop codon for GFP fusion in pMDC83
pET25begyFw	AACCATGGCTTCTCTCTTTGTTTC	24	65.4	1733	EGY3 CDS amplification for protein expression with His tag in pET-25b(+).
pET25begyRv	ACGCGGCCGCGAAATCATCACCACGGAAGAA	31	86.2		
pGEX6P1egyFw	CGGGATCCATGGCTTCTCTCTTTGTTTCTAC	31	74.6	1740	EGY3 CDS amplification for protein expression with GST tag in pGEX-6P-1.
pGEX6P1egyRv	TAGCGGCCGCTTAGAAATCATCACCACGGA	30	80		
pC1303L35S #2	GTTGGCCGATTCATTAATG	19	60.3	pCAMBIA1303 - 1071	Left primer of 35S in pCAMBIA1303 for colony PCR and insert sequencing.
pC1303R35S	AGTTTTTTGATTCACGG	18	55		Right primer of 35S in pCAMBIA1303 for colony PCR and insert sequencing.
Primer NOS pC1303Rv	GATAATCATCGCAAGACCG	19	55		NOS terminator reverse primer in pCAMBIA1303 for colony PCR and insert sequencing.
SQE1ProFw	TAAAGCTTTGCTGCTCGCTCG	21	59.8	Ler - 620	SQE1 promoter amplification replacing 35S in pCAMBIA1303.
SQE1ProRv	AACTAGTTATCGTTCTGCTTTTACAGAGAT	30	61.3		
SQE2ProFw	GACTGCAGAATGGTGTCTC	20	58	1854	SQE2 promoter amplification for GUS fusion in pCAMBIA1303
SQE2ProRv	TTCCTGATTACGAATGGTCCATGGTTT	28	66		
SQE3ProFw	CGGAATTCGCTGCTTG GGCTGATTC	25	69.5	1930	SQE3 promoter amplification for GUS fusion in pCAMBIA1303
SQE3ProRv	AGCCATGGGCGTACACAGAGATTTTCAGG	30	70.6		
SQE2gFw	AGGATCCACTAGTATGAAACCATTGTAATCAGG	34	67.1	2350	SQE2 genomic amplification for fusion in pCAMBIA1303 containing SQE1 promoter
SQE2gRv	ACTCGAGCACGTGTTAAGGAGGAGCACGGTATA TG	35	71.8		

SQE3gFw	AGGATCCACTAGTATGGCTCCGACGATATTCGTT	34	69.5	2368	SQE3 genomic amplification for fusion in pCAMBIA1303 containing SQE1 promoter
SQE3gRv	ACTCGAGCACGTGTCATTGAGGAGAAGAAGAAG AAG	36	70.6		
SQE2dpcrLP	TCATCGTCCGAGCTGGTGTC	20	63	995	Genotyping SALK_064182 Hm mutant line interrupting SQE2.
SQE2dpcrRP	GTGCCCCGTGATTTGCAAATG	20	58		
SQE3dpcrLP	TGAATTACTTCAGCCTGGTGG	21	59	1064	Genotyping SAIL_723_F01 mutant line interrupting SQE3.
SQE3dpcrRP	AGACAATGCAACGGTCATACC	21	59		
SQE3 dpcr L2	TGGCCAAGAGCTTAAGTCAT	20	56	703	Genotyping SAIL_723_F01 mutant line interrupting SQE3. Better pair of primers.
SQE3 dpcr R2	CGCCTGCGAGTGATTGATA	20	58		
SQE1GC	AAAAAGCAGGCTCGATAATGGAGTCACAATTATG GAAT	38	72	1626	SQE1 CDS amplification for first step PCR for Gateway BP recombination. Forward primer.
SQE1GD	AGAAAGCTGGGTACTIONATGAACATTTGGTTTCTCC AAC	37	74.6		SQE1 CDS amplification for first step PCR for Gateway BP recombination. Reverse primer with stop codon.
attB(Fw)	GGGGACAAGTTTGTACAAAAAAGCAGGCT	29	72.8	-	Gateway second step PCR with universal adaptor for BP recombination.
attB(Rv)	GGGGACCACTTTGTACAAGAAAGCTGGGT	29	74.5		
pdon201SeqFw	TCGCGTTAACGCTAGCATGGATCTC	25	72.3	2493	Amplification of products of Gateway BP reaction in pDONR201. Colony PCR and sequencing.
pdon201SeqRv	GTAACATCAGAGATTTTGAGACAC	24	58.3		
SQE1GfpNTFw	AGTTAATTAATAATGGAGTCACAATTATGGAA	32	64	1616	SQE1 CDS amplification without stop codon for GFP fusion in pMDC83
SQE1GfpNTRv	ATGGCGCGCCCTGAACATTTGGTTTCTCCAAC	32	75		
SQE3stopENTRFw	AAGCGCCGCCATGGCTCCGACGATATTC	29	83.8	1600	SQE3 CDS amplification for pENTR/D-TOPO cloning. Reverse primer with stop codon.
SQE3stopENTRRv	AAGCGCGCCCTCATTGAGGAGAAGAAGAAG	31	79.9		

2.3. - Molecular biology methods

SQE3GfpNTFw	TCTTAATTAATAATGGCTCCGACGATATTC	31	67	1599	SQE3 CDS amplification without stop codon for GFP fusion in pMDC83
SQE3GfpNTRv	CTGGCGCGCCCTTGAGGAGAAGAAGAAGGA	33	78		
pMDC-35S	TTCATTTCAATTTGGAGAGGACC	22	63.9	-	Flanking primer of <i>attR</i> regions inside pMDC's 35S promoter. Colony PCR and sequencing.
pMDC32R2flank	CGGCCGCTCTAGAACTAGTTAA	22	63.4	-	Flanking primer of <i>attR</i> regions in pMDC's. Colony PCR and sequencing.
pMDCgfpriht	GGATTACACATGGCATGGATG	21	64.8	-	Flanking primer of <i>attR</i> region of GFP in pMDC43/45.
pMDCgpleft	TTGGGACAACCTCCAGTGAAAAG	22	65	-	Flanking primer of <i>attR</i> region of GFP in pMDC83.
InDelPromSqe1Fw	TGCTCGCTCGTACTTTTGAG	20	58	Col-0 – 628; Ler - 365	SSLP SQE1 promoter amplification of an InDel.
InDelPromSqe1Rv	GAATCAAATAACGCGAGGTGA	21	57		
SQE1RTFw	ATCTTTGCTTTTCGGGTTTGA	20	54	gDNA-411; cDNA-329	SQE1 amplification for RT-PCR.
SQE1RTRv	CACCAAATGCAATCCACAAC	20	56		
SQE2RTFw	GAGCTTTTGGCTTGGAGCTA	20	58	gDNA-397; cDNA-320	SQE2 amplification for RT-PCR.
SQE2RTRv	TTCCATTGATGCACTTTCA	20	54		
SQE3 RTFw	CCACGACCTATGAGCCTTGT	20	60	gDNA-409; cDNA-320	SQE3 amplification for RT-PCR.
SQE3 RTRv	CCCGTTTCATTCGTACCATA	20	58		
Act2Fw	AAGATCTGGCATCACACTTTCT	22	58	gDNA-354 ; cDNA-276	ACT2 amplification for RT-PCR
Act2Rv	GATGGCATGAGGAAGAGAGA	20	58		
ActinqF	CTAAGCTCTCAAGATCAAAGGCTTA	25	62.8	cDNA - 211	ACT2 amplification for qRT-PCR
ActinqR	ACTAAAACGCAAACGAAAGCGGTT	25	69.6		

LAS1RTFw	ATCTCACCTCTCGCGGATTA	20	63.8	gDNA-421; cDNA-259	LAS1 amplification for RT-PCR.
LAS1RTRv	AGTGGAGAAACCCCATCCTC	20	64.2		
LAS1LPSAIL	TCACCATATCGATTGGGACAC	21	59	1029	Genotyping SAIL_676_A11 mutant line interrupting LAS1.
LAS1RPSAIL	ATTTGTGACAATAGCGCC	21	57		
ATLUP1RTFw	GGAGTTTGTGGAATGCACCT	20	63.9	gDNA-434; cDNA-248	LUP1 amplification for RT-PCR and SALK_088764 mutant line genotyping.
ATLUP1RTRv	TGCGCATAGCTAAACAATCG	20	63.7		
ATPEN3RTFw	AGCGTTGCAATCTGATGATG	20	63.9	gDNA-1606; cDNA-393	PEN3 amplification for RT-PCR and SALK_029567 mutant line genotyping.
ATPEN3RTRv	GGTTTGCAACCAGACCATTC	20	64.3		
BARS1 RTFw	CGCGAGTGATAGTGGAGGAT	20	64.2	gDNA-837; cDNA-396	BARS1 amplification for RT-PCR and SALK_071455 mutant line genotyping.
BARS1 RTRv	GGTCGAGGATCCATTTACGA	20	63.7		
THAS1RTFw	CGATTGCAGCATTGACTCAG	20	64.6	cDNA-297	THAS1 amplification for RT-PCR and SALK_138058 mutant line genotyping.
THAS1RTRv	TGCTTGACAAGAGAGAAAGC	21	63.5		
MRN1RTFw	ACATTACCGGACACCTGGAG	20	63.8	gDNA-620; cDNA-273	MRN1 amplification for RT-PCR and SALK_152492 mutant line genotyping.
MRN1RTRv	GCCAAGCTTTAGCCACCATA	20	63.9		
ATLUP2 RTFw	TCGCCATATTTCCAAAGGAG	20	63.7	gDNA-577; cDNA-293	LUP2 amplification for RT-PCR and SALK_024920 mutant line genotyping.
ATLUP2 RTRv	CATACTCGCGTTCAGCCATA	20	63.7		
ATLUP5RTFw	GATGCTCTCCACACCTTCGT	20	64.3	gDNA-849; cDNA-409	LUP5 amplification for RT-PCR and SALK_022044 mutant line genotyping.
ATLUP5RTRv	TGAAGTCACCTGATGGGTTG	20	63.7		

2.3. - Molecular biology methods

ATPEN6RTFw	AGCCAAGTGTGGGATACAGC	20	64.1	gDNA-368; cDNA-267	PEN6 amplification for RT-PCR and SALK_060682 mutant line genotyping.
ATPEN6RTRv	GCTTCCAAAGACAAGGCAAC	20	63.7		
ATLUP4RTFw	GGGTGGGGATTACACATTGA	20	64.4	gDNA-530; cDNA-293	LUP4/BAS amplification for RT-PCR and SALK_073455 mutant line genotyping.
ATLUP4RTRv	CGGCAGTAGCTCCACATTTT	20	64		
ATLUP3RTFw	GTGTTACTGCCGGTTGGTTT	20	63.7	gDNA-332; cDNA-251	LUP3 amplification for RT-PCR and SALK_137624 mutant line genotyping.
ATLUP3RTRv	TAAATGGCCAACATGCAAGA	20	63.9		
ATPEN7RTFw	CCCCTCATTCTACAGCTTCG	20	63.6	gDNA-723; cDNA-356	PEN7 amplification for RT-PCR and SALK_142842 mutant line genotyping.
ATPEN7RTRv	CCAGAGACACGAGCAAGATG	20	63.7		
CAS1RTFw	ACGTGGTTTGGAGTGAAAGG	20	63.9	gDNA-455; cDNA-247	CAS1 amplification for RT-PCR.
CAS1RTRv	GTGGTTTCCGGTCTACCTCA	20	63.9		

2.3.2. DNA methods

2.3.2.1. Plant genomic DNA isolation

Distinct methods were used to obtain genomic DNA from *A. thaliana* tissues, according to the nature of tissue sample and the degree of purity needed: CTAB-based method and Fast DNA extraction method.

CTAB method. Leaf tissue was harvested from each plant to obtain a high integrity and high purity genomic DNA with CTAB method extraction (Doyle J.J. and Doyle J.L., 1987). Tissue was grinded with liquid nitrogen and subsequently added 700 μL of CTAB buffer, vortexed and incubated for 25 min at 65°C. Samples were centrifuged for 5 min at room temperature and 12000 *g*. The aqueous phase was recovered, precipitated with 1:1 vol of cold (-20°C) isopropanol and centrifuged for 20 min at room temperature and 12000 *g*. The pellet was washed with 300 μL of 70% (v/v) ethanol and centrifuged for 5 min at room temperature and 12000 *g*. The pellet was then dried for 10 min at 37°C, solubilised in 30 μL of 0.1x TE with RNase A (100 $\mu\text{g mL}^{-1}$), and incubated for 20 min at 37°C. Genomic DNA was kept for 24 h at 4°C to allow complete dissolution of the pellet and stored at -20°C.

CTAB buffer: 2% (w/v) CTAB; 1.4 M NaCl; 0.1 M Tris-HCl (pH 8.0); 0.02 M EDTA (pH 8.0); add 0.1% (v/v) β -mercaptoethanol before using.

TE: 10 mM Tris-HCl (pH 8.0); 1 mM EDTA.

Fast DNA extraction method. Leaf tissue was harvested from each plant to perform a rapid DNA extraction (Edwards *et al.*, 1991). Plant tissue was transferred to a microtube and 400 μL of extraction buffer was added prior to grinding the tissue with polypropylene pestles. Microtubes were centrifuged for 5 min at room temperature and 14000 rpm, and the supernatant was transferred to a new microtube containing 300 μL of isopropanol. After another centrifugation of 5 min and 14000 rpm in identical conditions, DNA pellet was rinsed with 500 μL of 70% (v/v) ethanol, and spined down for 2 min. The pellet was air-dried and resuspended in 50-100 μL of ultra pure water.

Extraction buffer: 200 mM Tris-HCl pH 7.5; 250 mM NaCl; 25 mM EDTA; 0.5% SDS.

Root DNA extraction: Root tissue DNA was isolated using the *ZR Plant/ Seed DNA Kit* (Zymo Research). This procedure was performed in the scope of the grafting experiments in order to identify the root genotype.

2.3.2.2. Plasmid isolation

The isolation of plasmid DNA from *E. coli* was performed by the following methods, according to the purpose of the experiments and the degree of DNA purity needed.

Boiling method. The boiling method was used for a quick recovery of plasmids (Holmes and Quigley, 1981). An aliquot of 1.5 mL of *E. coli* liquid culture was centrifuged for 5 min and 8000 *g*. The pellet was resuspended in 400 μ L of STET supplemented with 25 μ L of fresh lysozyme solution. Cells were lysed by incubation at room temperature for 10 min, followed by boiling for 1 min at 95°C. After centrifugation for 15 min and 14000 *g*, the supernatant was recovered, and plasmid DNA precipitated by the addition of 300 μ L of isopropanol. Following centrifugation at 14000 *g*, the pellet was resuspended in 20-100 μ L of TE and stored at -20°C.

STET: 10 mM Tris-HCl, pH 8.0; 100 mM NaCl; 1 mM EDTA; 5% (v/v) Triton X-100.

Lysozyme solution: 10 mM Tris-HCl, pH 8.0; 10 mg mL⁻¹ lysozyme.

TE: 10 mM Tris-HCl, pH 8.0; 1 mM EDTA.

Four-step method. This method was used and developed to obtain a high purity plasmid (M. Botella, personal communication). An aliquot of 1.5 mL of *E. coli* liquid culture was centrifuged for 1 min and 14000 rpm and the supernatant was discarded. This step was repeated once, for a total of 3 mL of liquid culture. The pellet was resuspended in 100 μ L of GTE solution I and vortexed. Lysis was promoted by adding 200 μ L of GTE solution II (freshly prepared) and microtubes were gently inverted. Subsequently, 150 μ L of GTE solution III were added, mixed gently by inversion, incubated on ice for 15 min, and centrifuged for 15 min and 14000 rpm. The supernatant was carefully recovered, transferred to a new microtube and centrifuged in the same conditions. Precipitation of DNA was promoted by adding to the supernatant 1 mL of 100% (v/v) ethanol, followed by centrifuged for 15 min and 14000 rpm. The pellet was then dissolved in 50 μ L of TE with RNase A, incubated for 5 min at 37°C, vortexed, and incubated again for 15 min at 37°C. Subsequently, 30 μ L of GTE solution IV was added, the microtube was vortexed, and incubated for 2 h in ice. Microtubes were centrifuged for 15 min and 14000 rpm, the pellet was washed with 500 μ L of 70% (v/v) ethanol, and centrifuged as before. Finally, the pellet was air dried, resuspended in 30 μ L of TE and stored at -20°C.

GTE Solution I: 50 mM Glucose, 10 mM EDTA, 25 mM Tris-HCl pH 8.0.

GTE Solution II: 0.1 M NaOH, 1% (w/v) SDS.

GTE Solution III: 3 M potassium acetate, 5 M glacial acetic acid. Autoclaved.

GTE Solution IV: 20% (w/v) PEG, 2.5 M NaCl.

TE: 10 mM Tris-HCl (pH 8.0); 1 mM EDTA.

RNase A was added to a final concentration of 10 μ g mL⁻¹.

Plasmid isolation kit method. Plasmid isolation was also performed using *Wizard Plus SV Minipreps DNA Purification System* (Promega) and *Wizard Plus Midipreps DNA Purification System* (Promega) commercial kits, according to the manufacturer's instructions.

2.3.2.3. DNA fragment purification

DNA purification from agarose gels, PCR amplifications or endonucleases digestions were performed using the *Wizard SV Gel and PCR Clean-Up System Kit* (Promega), according to the manufacturer's instructions.

2.3.2.4. DNA digestion with endonucleases

DNA digestion with restriction endonucleases (table 2.10) was performed according to the procedures described by Sambrook and Russel (2001), and following the manufacturer's instructions.

Table 2.10. – List of endonucleases used to obtain constructs in this work outlining restriction enzymes and restrictions sites in each primer (already described in section) for each construct.

Restriction enzymes	Primer (5'-3')	Purpose
<i>EcoRI</i> <i>NcoI</i>	TTG [*] AATTCATCTAA TGA CTT AAG GTT T AGC [*] CATGGTTTACGA AAA CTC	EGY3 promoter cloning
<i>NcoI</i> <i>PmlI</i>	C [*] CATGGCTTCTCTCTTTGTTTCTA CAC [*] GTGTTAGAAATCATCACCACGGA	EGY3 cDNA cloning for gene overexpression
<i>NcoI</i> <i>NotI</i>	AAC [*] CATGGCTTCTCTCTTTGTTTC ACGC [*] GGCCGCCGAAATCATCACCACGGAAGAA	EGY3 cDNA cloning for protein expression with His tag
<i>BamHI</i> <i>NotI</i>	CGG [*] GATCCATGGCTTCTCTCTTTGTTTCTAC TAGC [*] GGCCGCTTAGAAATCATCACCACGGA	EGY3 cDNA cloning for protein expression with GST tag
<i>PstI</i> <i>NcoI</i>	GAC [*] TGCAGAAATGGTGTTC TTCCTGATTACGAATGGTTC [*] CATGGTTT	SQE2 promoter cloning
<i>EcoRI</i> <i>NcoI</i>	CGG [*] AATTCGCTGCTTG GGCTGATTC AGC [*] CATGGGCGTACACAGAGAGATTCAGG	SQE3 promoter cloning
<i>HindIII</i> <i>SpeI</i>	TAA [*] AGCTTTGCTGCTCGCTCG AA [*] CTAGTTATCGTTCTGCTTTTACAGAGAT	SQE1 promoter cloning
<i>SpeI</i> <i>PmlI</i>	AGGATCCA [*] CTAGTATGAAACCATTTCGTAATCAGG ACTCGAGCAC [*] GTGTTAAGGAGGAGCACGGTATATG	SQE2 genomic cloning
<i>SpeI</i> <i>PmlI</i>	AGGATCCAC [*] TAGTATGGCTCCGACGATATTCGTT ACTCGAGCAC [*] GTGCATTGAGGAGAAGAAGAAGAAG	SQE3 genomic cloning
<i>PacI</i> <i>AscI</i>	TCTTA [*] ATTAATAATGGCTCCGACGATATTC CTGG [*] CGGCCCTTGAGGAGAAGAAGAAGAAGGA	Cloning of SQE3 cDNA to a GFP C-terminal fusion
<i>NotI</i> <i>AscI</i>	AAGC [*] GGCCGCCATGGCTCCGACGATATTC AAGG [*] CGGCCCTCATTGAGGAGAAGAAGAAG	Cloning of SQE3 cDNA
<i>PacI</i> <i>AscI</i>	AGTTA [*] ATTAATAATGGAGTCACAATTATGGAA ATGG [*] CGGCCCTGAACATTTGTTTCTCCAAC	Cloning of SQE1 cDNA to a GFP C-terminal fusion

Reactions were performed in a total volume of 20-50 μL for 1.5 h to ON periods at 37°C. In plasmid linearisation reactions *Shrimp Alkaline Phosphatase* (SAP) (Fermentas) was added to promote dephosphorylation of the 5'-phosphorylated ends of DNA, and also to prevent re-ligation of the linearised plasmid DNA in cloning experiments. Endonucleases were heat inactivated according to their specification, or the reaction was purified as described in 2.3.2.3.

2.3.2.5. Amplification of DNA fragments by PCR

DNA amplifications by PCR were performed using *Mastercycler Gradient* (Eppendorf) or *MJ Mini Gradient Thermal Cycler* (BIO-RAD) devices.

Diagnostic PCR. Diagnostic PCR was performed mainly in uniplex but also as multiplex (more than 2 primers per reaction). Reactions (50 μL ; 0.2 mL microtube) were as follows: ~1 μg DNA, 0.6 μM of each primer, 0.2 mM of *dNTP mix* (Promega), 2 mM of MgCl_2 , 10 μL 5x *Green GoTaqReaction Buffer* (Promega) and 0.25 μL of *GoTaq DNA polymerase* (Promega). PCR steps: (1) denaturation for 5 min at 94°C; (2) 30 cycles of denaturation for 45 sec at 94°C, annealing for 45 sec at 55°C, extension for 45 sec at 72°C; (3) final extension of 10 min at 72°C. Extension time and annealing temperatures were adapted according to the size of the expectable fragments (1 min per 1 kb) and primer pairs T_m , respectively.

Promoter and CDS amplifications. A proofreading enzyme was used when correct amplified sequences were needed for cloning. Reactions (50 μL ; 0.2 mL microtube) were as follows: ~600 ng DNA, 1 μM of each primer, 25 μL of 2x *ACCUZYME MIX* (Bioline). PCR steps: (1) denaturation for 5 min at 94°C; (2) 35 cycles of denaturation for 45 sec at 94°C, annealing for 45 sec at 55°C, extension for 1 min at 68°C; (3) final extension of 10 min at 68°C. Extension time and annealing temperatures were adapted according to the size of the expectable fragments (1 min and 30 sec per 1 kb) and primer pairs T_m , respectively.

***E. coli* colony PCR.** Reactions (25 μL ; 0.2 mL microtube) were as follows: one *E. coli* colony, 0.6 μM of each primer, 0.2 mM of *dNTP mix* (Promega), 2 mM of MgCl_2 , 5 μL 5x *Colorless GoTaqReaction Buffer* (Promega) and 0.5 μL of *Taq DNA Polymerase*. PCR steps: (1) denaturation for 10 min at 94°C; (2) 30 cycles of denaturation for 45 sec at 94°C, annealing for 45 sec at 55°C, extension for 1 min at 72°C; (3) final extension of 10 min at 72°C. Extension time and annealing temperatures were adapted according to the size of the expectable fragments (1 min per 1 kb) and primer pairs T_m , respectively.

***Agrobacterium* colony PCR.** Reactions (25 μL ; 0.2 mL microtube) were as follows: one *Agrobacterium* colony, 0.6 μM of each primer, 0.2 mM of *dNTP mix* (Promega), 2 mM of MgCl_2 ,

5 μL 5x Colorless GoTaqReaction Buffer (Promega) and 0.5 μL of Taq DNA Polymerase. PCR steps: (1) denaturation for 15 min at 94°C; (2) 30 cycles of denaturation for 45 sec at 94°C, annealing for 45 sec at 55°C, extension for 1 min at 72°C; (3) final extension of 10 min at 72°C. Extension time and annealing temperatures were adapted according to the size of the expectable fragments (1 min per 1 kb) and primer pairs T_m , respectively.

Gateway PCR. Amplification of DNA fragments for insertion into Gateway recombination (*attB*) sites was performed through a two-step PCR strategy and using a proofreading enzyme. First step was made using specific primers for the sequence, altogether with some bases of the recombinant Gateway system at the 5' end of the primer, so that in the second step, using universal *attB* Gateway adaptor primers, they could anneal to the recombinant primer and hence the sequence. Reactions (50 μL ; 0.2 mL microtube) were as follows: ~600 ng DNA, 0.4 μM of each primer, 25 μL of 2x ACCUZYME MIX (Bioline). PCR steps: (1) denaturation for 5 min at 94°C; (2) 35 cycles of denaturation for 30 sec at 94°C, annealing for 30 sec at 50°C, extension for 4 sec at 68°C; (3) final extension of 10 min at 68°C. Second step was made using universal Gateway *attB* primers (table 2.9). Reactions (50 μL ; 0.2 mL microtube) were as follows: 10 μL of first-step reaction, 0.8 μM of each primer, 25 μL of 2x ACCUZYME MIX (Bioline). PCR steps: (1) denaturation for 1 min at 94°C; (2) 5 cycles of denaturation for 15 sec at 94°C, annealing for 30 sec at 45°C, extension for 4 min at 68°C; (3) 25 cycles of denaturation for 15 sec at 94°C, annealing for 30 sec at 55°C, extension for 4 min at 68°C; (4) final extension of 10 min at 68°C. Extension time and annealing temperatures were adapted according to the size of the expectable fragments (1 min and 30 sec per 1 kb) and primer pairs T_m , respectively.

2.3.2.6. DNA sequencing

Plasmid inserts were sequenced by STAB VIDA services, using universal or purposefully designed primers.

2.3.2.7. Gateway cloning

2.3.2.7.1. **BP reaction**

BP reaction (recombination between *attB* and *attP* Gateway recombination sites) was carried out to obtain the *attL*-flanked entry clone for *SQE1* CDS in the pDONR201 vector. The reaction was performed in a total volume of 10 μL , 0.2 mL microtubes containing: 50 fmol

pDONR201, 50 fmol *attB*-PCR product, 2 μ L of *BP Clonase II* (Invitrogen) and TE buffer (pH 8). Equal molarity was obtained using the following formula:

$$\text{ng}(\textit{attB}) = \text{fmol} \times N \text{ (nucleotide number)} \times (660 \text{ fg/1 fmol}) \times (1 \text{ ng}/10^6)$$

The reaction was incubated for 18 h at 25°C. Subsequently, 1 μ L of Proteinase K (*BP clonase II kit*, Invitrogen) was added and the reaction incubated for 10 min at 37°C. An aliquot of 5 μ L was used to transform *E. coli* XL1-Blue competent cells. Cells were plated in LB agarised medium containing kanamycin (50 μ g mL⁻¹). Colonies were used to perform an *E. coli* colony PCR with specific primers for pDONR201. Positive colonies were selected for plasmid isolation and insert sequencing.

2.2.2.7.2. LR reaction

LR reaction of recombination between *attL* (entry clone) and *attR* (destination vector) recombination sites was carried out to obtain the *attB* expression clones in pMDC's vectors. The reaction was performed in a total volume of 10 μ L in 0.2 mL microtubes containing: 50 fmol of a Gateway destination vector (pMDC43, pMDC45 or p35S::Cerluean-GWY), 50 fmol of an *attL*-entry clone (SQE1CDS or SQE3CDS), 2 μ L of *LR Clonase II* (Invitrogen) and TE buffer (pH 8). Equal molarity was obtained as in the BP reaction. The mix was incubated for 18 h at 25°C. Subsequently, 1 μ L of Proteinase K (*LR clonase II kit*, Invitrogen) was added and the reaction incubated for 10 min at 37°C. An aliquot of 5 μ L was used to transform *E. coli* XL1-Blue/DH5 α competent cells. Cells were transformed and plated in LB agarised medium containing hygromycin (50 μ g mL⁻¹). Colonies were used in a colony PCR with specific primers for the destination vectors, and positives ones were selected in kanamycin (50 μ g mL⁻¹) for plasmid isolation and sequencing.

2.2.2.8. Cloning of PCR fragments into an expression vector

PCR fragments and digested plasmid vectors were separated by agarose gel electrophoresis and the fragments of interest recovered using the *Wizard SV Gel and PCR Clean-Up System Kit* (Promega). Subsequently, DNA fragments were cloned into a subcloning vector (*pGEM-T Easy vector*, Promega), or into the final destination vector. Protocol was performed according to the manufacturer's instructions. However, to subclone into pGEM-T vector (Marcus *et al.*, 1996) with PCR fragments non-obtained with a proofreading *Taq*, a prior step to cloning, called adenylation, was necessary. The DNA suffered adenylation of the 3' end by incubation with *Taq*

polymerase, allowing for the generation of the A-tail necessary to anneal the T-overhangs of the pGEM-T vector in question. An aliquot of the ligation reaction (5 μ L) was used to transform *E. coli* XL1-Blue/DH5 α cells. Cells were transformed as described in later section 2.3.2.9.1, and plated onto selective LB agarised medium with appropriate antibiotics. pGEM clones were selected in plates with the antibiotic ampicillin (100 μ g mL⁻¹) and also with X-Gal (40 μ g mL⁻¹) and IPTG (50 μ g mL⁻¹) for white/blue screening (cloning products interrupt proLacZ regulation of *lacZ*).

2.3.2.9. Transformation of bacteria

2.3.2.9.1. Transformation of *E. coli* cells

The protocol followed to perform *E. coli* competent cells and its transformation was based on Inoue *et al.* (1990).

Competent cells. *E. coli* competent cells were obtained by inoculating 250 mL of SOB medium with a single colony of *E. coli* strains. Cells were grown at 18°C with vigorous shaking (200-250 rpm) until A_{600} was of 0.6 (2-3 days). The medium was cooled on ice for 10 min and cells were collected by centrifugation for 10 min at 4°C and 2500 *g*. The pellet was resuspended in 80 mL of ice-cold TB buffer, and incubated on ice for 10 min. Cells were centrifuged for 10 min at 4°C and 2500 *g*, gently resuspended in 20 mL of ice-cold TB buffer with 7% (v/v) DMSO. It was subsequently incubated for 10 min on ice and 100 μ L were aliquoted into microtubes. Competent cells were immediately placed in liquid nitrogen and stored at -80°C.

Transformation. *E. coli* cell transformation was initiated by thawing competent cells on ice. DNA sample (1-20 μ L) was added to 100 μ L of cells by gentle mixing, and the mixture was incubated for 30 minutes at 4°C. Cells were then heat-shocked for 90 sec at 42°C, followed by incubation on ice for 1 min. After addition of 1 mL of SOC (or LB) liquid medium, cells were incubated for 1 hour at 37°C with vigorous shaking (200-250 rpm), spined down for one second at 10000 *g* and the pellet resuspended in 100 μ L of the supernatant. Finally, cells were transferred to agarised LB medium plates containing appropriate antibiotics, and grown overnight at 37°C.

TB: 10 mM Pipes; 15 mM CaCl₂; 250 mM KCl; 55 mM MnCl₂. Mix all components except MnCl₂ and adjust pH to 6.7 with KOH. Dissolve MnCl₂ and sterilize solution through a 0.45 μ m filter.

2.3.2.9.2. Transformation of *Agrobacterium* cells

Two procedures were used to obtain *Agrobacterium* (EHA105 and GV3101::pMP90 strains) competent cells, able to be transformed.

Electrocompetent cells. *Agrobacterium* strains were inoculated in agarised LB medium, from a -80°C glycerol stock, and grown for 2 days at 28°C. A colony was then resuspended in 5 mL LB liquid medium and grown ON at 28°C and 200 rpm. Cells were harvested by centrifugation for 1 min at 13000 rpm and resuspended in 1 mL of 300 mM sucrose per microtube. The pellet was resuspended in 100 µL of 300 mM sucrose. Aliquots of 100 µL of competent cells were used for vectors electroporation.

Chemically competent cells. Single colonies of *Agrobacterium* strains were obtained in agarised LB medium from a -80°C glycerol stock and grown for 2 days at 28°C. A colony was then transferred to 5 mL LB liquid medium and grown ON at 28°C and 200 rpm. An aliquot of 50 µL from the 5 mL culture was transferred to a freshly prepared 50 mL LB liquid culture and grown ON at 28°C and 200 rpm. When A_{600} was between 0.5-1, the culture was cooled on ice for 10 min and centrifuged for 6 min at 4°C and 3000 g. The pellet was rinsed with 1 mL of 20 mM cold CaCl₂ and spined briefly. The pellet was resuspended in 1 mL of 20 mM cold CaCl₂, aliquoted (100-150 µL) into microtubes and kept on ice for subsequent transformation, or alternatively frozen with liquid nitrogen for long-term storage.

Two transformation cell procedures were performed to transform the *Agrobacterium* competent cells.

Electroporation method. Transformation of *Agrobacterium* by electroporation was performed by mixing 100 µL of electrocompetent cells with 100 ng of the vector. After careful mix, an electric pulse was given in an electroporator *Gene Pulser II* (BIO-RAD), which was set to 2.5 kV and 400 Ω, with a capacitance of 25 µF. Subsequently, 1 mL of LB liquid medium was added to cell suspension, shaken, incubated in a microtube for 1 h at 28°C and 200 rpm, and plated onto appropriate selection medium and grown for 48 h at 28°C.

Chemical method. Transformation of *Agrobacterium* was performed by adding 1 µg of the vector into a microtube containing 100-150 µL of competent cells, mixed by inversion, and frozen in liquid nitrogen for 5 min. The microtube was incubated for 10 min at room temperature, followed by the addition of 1 mL of LB liquid medium was incubated for 3 h at 28°C and 200 rpm. Cells were plated onto appropriate selection medium and grown for 48 h at 28°C.

2.3.3. RNA methods

RNA manipulation was carried out under specific conditions to prevent RNase contamination. Ultra pure water was used in all solutions, previously treated overnight with 0.1% (v/v) DEPC, and autoclaved to destroy DEPC. Ultra pure water and disposable material was autoclaved for 20 min at 121°C and 1 atm.

2.3.3.1. RNA extraction

RNA from plant tissues was isolated using the commercial reagent *TRIZOL* (Invitrogen) and following the manufacturer's instructions. Tissue was grinded to a fine powder in liquid nitrogen and 1 mL *TRIZOL* (Invitrogen) was added. Samples were incubated for 5 min at room temperature, and 200 μ L of chloroform were added, followed by a 3 min incubation at room temperature. The top aqueous phase was recovered after a centrifugation of 15 min at 4°C and 12000 *g*. RNA was precipitated after adding 500 μ L of isopropanol, and incubating for 10 min at room temperature. A centrifugation for 10 min at 4°C and 12000 *g* was performed, and the supernatant discarded. The pellet was washed with 1 mL of 75 % (v/v) ethanol, vortexed and centrifuged for 5 min at 4°C and 7500 *g*. The pellet was subsequently dried in a flow chamber and then dissolved in 30-50 μ L of DEPC-treated water. The RNA's concentration and purity was determined spectrophotometrically (Nanodrop *ND-1000 Spectrophotometer*, Alfacene). An aliquot of 1 μ g RNA was run on a 1% agarose gel to confirm RNA integrity. RNA samples were immediately frozen in liquid nitrogen and stored at -80°C.

RNA of leaves elicited by *Pto* was obtained using the *RNeasy Plant Mini Kit* (Qiagen) and following the manufacturer's instructions.

2.3.3.2. cDNA synthesis

A 2 μ g RNA sample was treated with DNase I (Sigma) prior to cDNA synthesis. This treatment involved a 15 min period incubation with DNase I at 37°C, followed by an inactivation of the enzyme, and a heat denaturation for 10 min at 70°C of both DNase I and the RNA. To synthesize the first-strand cDNA, 1 μ g of RNA was primed with 1 μ L of Oligo(dT)₁₅ (0.5 μ g μ L⁻¹) (Promega) and DEPC-treated water was added to a final volume of 17.75 μ L. This mixture was heated for 5 min at 70°C and cooled quickly on ice for 5 min. Reverse transcription was promoted using the enzyme *M-MLV RT* (Moloney murine leukemia virus reverse transcriptase). A mixture of 1 μ L of *M-MLV RT (H-)* (Promega) and its 5x buffer were added to 1.25 μ L of *dNTP mix* (10mM)

(Promega), and transferred to the first mixture containing the RNA. Reverse transcriptase reaction was carried for 60 min at 50°C, followed by 15 min at 70°C, and the cDNA was stored at -20°C.

2.3.3.3. Semi-quantitative RT-PCR

Semi-quantitative RT-PCR was first optimised in terms of amplification cycle number, to identify the mid-exponential amplification cycle phase (data not shown). *ACT2* was optimised for 27 cycles, *EGY3* for 30 cycles, *SQE1* for 30 cycles, *SQE2* for 34 cycles and *SQE3* for 30 cycles. For RT-PCR, the cDNA was diluted to 1/10 (used for constitutive genes) and 1/50 (used for the remaining genes). An aliquote of 1 µL of 1/50 cDNA was used in a PCR reaction mixture for constitutive gene expression. PCR steps were as follows: (1) denaturation for 5 min at 94°C; (2) n cycles of denaturation for 45 sec at 94°C, annealing for 45 sec at 55°C and extension for 30 sec to 1 min and 30 sec at 72°C; (3) final step of 5 min extension at 72°C. Samples were run in a 1% (w/v) agarose gel.

PCR reaction mixture (final volume 50µL):

5 units of GoTaq DNA polymerase (Promega), 10 µL 5x Green GoTaq Reaction Buffer, 0.2 mM dNTP mix (Promega), 2 mM MgCl₂ (Promega), 0.4 µM of each primer, u.p. water and cDNA.

2.3.3.4. qRT-PCR

Quantitative gene expression analysis were performed by quantitative real time PCR (qRT-PCR) with a *Rotor gene Q series* system (Qiagen) using *SYBR Green Master Mix* (Quanta Biosciences) with a specific dsDNA *SYBR Green I* dye to detect fluorescence. Three biological replicas and three technical replicas per sample were considered. Melting curves were used to discard the presence of primer dimmers or more than one amplification product. The melting curves were calculated in the end of the qRT-PCR program by increasing one degree of temperature from 55° to 95°C. This was plotted with the rate of change of the relative fluorescence units (RFU) with time (T) on the Y-axis vs. the temperature on the X-axis, and this will peak at the melting temperature (T_m). A primer dimer artefact would give a peak with a lower melting temperature (because it is such a short DNA) and so two peaks would appear, in opposite of the only one desired, corresponding to the product of the amplification expected. Data analysis allowed the determination of the Ct's (cycle threshold) for each gene using calibration curves with an efficiency of 0.95-0.99, estimated after analysing samples of known concentrations. Relative expression was calculated with the *Rotor gene Q Series Software* of the device relative to constitutive expression using *ACT2* as the constitutive expression gene and the samples of the mock treatment were used

as calibrator (value one of relative expression) to observe up- or down-regulation through time of infection with *Pto*.

PCR reaction mixture (final volume 25 μ L):

12.5 μ L (SYBR Green Master Mix), 0.25 μ M of each primer, u.p. water and cDNA.

Program: 40 times: 30 sec at 95°C, 30 sec at 60°C. 1 min at 95°C. 1 min at 65°C. Melting curve – one degree increase from 55-95°C.

2.3.4. Quantification of nucleic acids

Nucleic acid quantification was performed spectrophotometrically in a Nanodrop *Spectrophotometer ND-1000* (Alfagene), a micro-volume UV-Vis spectrophotometer for nucleic acid and protein quantitation. A minimum volume of 1.5 μ L per sample was used. Nucleic concentration was determined considering that 1 A_{260} = 50 ng DNA μ L⁻¹ and 1 A_{260} = 40 ng RNA μ L⁻¹. To determine the purity of the nucleic acid samples, A_{260}/A_{230} , and A_{260}/A_{280} ratios were also estimated (Sambrook and Russel, 2001).

2.3.5. Nucleic acids electrophoretic separation

DNA fragments were resolved by electrophoretic separation using a horizontal gel apparatus. Gels were made by melting 0.8-1.2 % (w/v) agarose in 0.5x TAE. TAE (0.5x) was also used as running buffer. DNA was stained by adding 1 μ L of ethidium bromide (1 mg mL⁻¹) to the melted agarose gel. DNA samples, except those from PCR with *GoTaq Green buffer* were mixed with 0.20 vol. of loading buffer (6x *MassRuler DNA Loading Dye*). *MassRuler DNA Ladder Mix* (an 80-10,000 bp molecular weight standard; Fermentas) and λ DNA digested with *Pst*I were used as molecular weight markers. Alternatively, DNA staining was carried out with the fluorescent intercalating agent *GelRed* (Biotium). *GelRed* was used after the gel run, so the gel was incubated for 25 min in 0.5x TAE solution to a final 0.2x *GelRed*. Gel was visualised under UV light.

50x TAE buffer: 2 M Tris; 0.95 M acetic acid; 50 mM EDTA_{Na2} pH 8.0.

6X MassRuler DNA Loading Dye: 10 mM Tris-HCL (pH 7.6); 0.03% bromophenol blue; 60% (v/v) glycerol; 60 mM EDTA.

Loading buffer: 30% (w/v) glycerol; 0.1 M EDTA; 0.25% (w/v) bromophenol blue.

2.3.6. Protein methods

2.3.6.1. Protein expression in *E. coli*

The vectors pEGY3-His and pGST-EGY3 were transformed into *E. coli* BL21(DE3)pLysE expression strain, as described in section 2.3.2.9.1. Colonies were obtained in LB medium supplemented with ampicillin and chloramphenicol. An experiment was designed to obtain the optimal inductor (IPTG) concentration, plus the better time of induction, in order to obtain samples with the maximum protein expression. One colony was inoculated into 10 mL of LB liquid medium plus proper antibiotics and incubated ON at 37°C and 200 rpm. An aliquot of the overnight culture (500 µL) was inoculated into 25 mL of LB liquid medium with the antibiotics, and incubated at 37°C and 200 rpm, until A_{600} reached 0.4-0.6. At this time, it was considered time zero of induction, and 1 mL of the culture was harvested and centrifuged for 2 min and 14000 *g*. The supernatant was discarded and the pellet frozen at -20°C. IPTG was then added to the liquid medium to a final concentration of 0.4 mM and 1 mM in the different flasks. One mL of culture was harvested each hour for a maximum of 3 hours for pGST-EGY3 and 4 hours for pEGY3-His, centrifuged and frozen as for the time zero of induction.

2.3.6.2. Protein electrophoretic separation

SDS-polyacrylamide electrophoresis gels were used to separate proteins in a discontinuous buffer system (Laemmli, 1970), through a 12% (w/v) acrylamide resolving gel, in a Mini-PROTEAN Cell (BIO-RAD) apparatus. Frozen protein expression of *E. coli* BL21(DE3)pLysE induced cells (previously transformed with the protein expression vectors) were added to 100 µL lysis buffer and the microtubes were vortexed until completely resuspended. Samples were then boiled for 5 min at 95°C after adding 2x SDS sample buffer (1x in final concentration), and centrifuged at 14000 *g* for 15 min to pellet cellular debris. Supernatant was loaded into the stacking gel, and electrophoresis was carried out at 80 V. Gels were then incubated with Coomassie gel staining solution for 3 h and 100 rpm (Meyer and Lamberts, 1965), and transferred to a destaining solution, under agitation, until clear bands were seen.

Lysis buffer: 0.02 M HEPES pH 7.8; 0.8 M KCl; 0.1% (v/v) Triton X-100.

SDS buffer (1x): 0.125 M Tris-HCl pH 6.8; 0.5% (v/v) SDS; 20% (v/v) glycerol; 5% (v/v) DTT; 0.025% (w/v) bromophenol blue).

Stacking gel: 50 µL of 10% (w/v) SDS; 630 µL of 1 M Tris pH 6.8; 830 µL of 30% Acrylamide and bis-acrylamide solution (Bio-Rad); 50 µL of 10% (w/v) APS; 5 µL of TEMED and 3.435 mL of dd water.

Running gel 12%: 90 μ L of 10% (w/v) SDS; 2.3 mL of 1.5M Tris pH 8.8; 3.6 mL of 30% Acrylamide and bis-acrylamide solution (Bio-Rad); 90 μ L of 10% (w/v) APS; 3.6 μ L of TEMED and 2.916 mL of dd water.

Running buffer (1 L): 48 mM Tris; 39 mM glycine; 20% (v/v) methanol pH 9.2; 3.75 mL of 10% (w/v) SDS solution and dd water.

Coomassie gel staining solution: 30% (v/v) methanol, 8% (v/v) acetic acid; 0.1% (w/v) Coomassie Brilliant Blue R-250 (Sigma).

Destaining solution: 30% (v/v) methanol; 8% (v/v) acetic acid.

2.4.

“Bioinformatic analysis”

2.4.1. Informatic tools for data mining and selection of *EGY3*

2.4.1.1. NASCArrays and microarray data treatment

NASCArrays (URL no.8) is an online repository of information of Arabidopsis experimental microarrays run by the NASC Affymetrix Facility (Craigon *et al.*, 2004). The present thesis utilised the *AtGenExpress-Heat Series* of microarray data, which was downloaded as bulk files from NASCArrays. This uncompressed data (signal values, annotation, probe set) was treated in an Excel spreadsheet to group in a single file data concerning 0, 0.25, 0.5, 1 and 3 h of heat shock stress, with corresponding controls. The relative expression ratio was calculated by dividing pixel count from stress data with the corresponding pixel count from control data, thus determining up- and down-regulation. “Data sorting” and “duplicate removal” capabilities of Excel software were also used during data mining.

2.4.1.2. Multiple Array Viewer

Multiple Array Viewer software was used to perform a hierarchical clustering analysis of the previously established 147 up-regulated genes, in order to visualise expression patterns.

2.4.1.3. Chloroplast 2010

Chloroplast 2010 project (URL no.9) is a project of large-scale reverse genetics recurring to screenings of T-DNA insertion mutants, a high-throughput genotyping of putative chloroplast-targeted genes in order to encounter phenotypes to those genes (Ajjawi *et al.*, 2010). This tool was used to retrieve functional information concerning *EGY3* as well as the list of putative plastid-targeted genes in Arabidopsis (5208 genes), which was used in the data-mining strategy to identify of *EGY3*.

2.4.2.4. BAR

BAR (URL no.10) is the Bio-Array Resource. It is a collection of user-friendly web-based tools to work with functional genomics and other data. Most features were designed for plants, though an eFP browser exists for the mouse model organism. Primarily, tools were created for Arabidopsis, but now some tools are also available for poplar, rice, *Medicago truncatula*, maize, soybean and barley. Despite the complexity of the data, these web-based tools are user-friendly, taking into consideration the basic researcher. BAR database comprises genomic tools and

proteomic tools, some of which are widely used for monitoring expression and localisation prediction. More specifically, the tool Venn selector was used to cross-reference the 5208 plastid-targeted genes with the 147 selected genes. Arabidopsis eFP browser and the e-Northerns w. Expression Browser, was used to visualise expression patterns for both developmental and stress-imposing conditions (Toufighi *et al.*, 2005; Winter *et al.*, 2007). Cell eFP browser was used in the *in silico* prediction of subcellular localisation, based mainly on the SUBA database (Heazlewood JL *et al.*, 2007), but with a more appealing graphic layout.

2.4.1.5. TAIR

The Arabidopsis Information Resource (TAIR) (URL no.5) sustains a database of genetic and molecular biology data for *Arabidopsis thaliana*, with constant updating. Data available from TAIR includes the complete genome sequence along with gene structure, gene product information, metabolism, gene expression, DNA and seed stocks, genome maps, genetic and physical markers, publications, and information about the Arabidopsis research community. TAIR also provides extensive links to other Arabidopsis resources. The variety of tools is impressive, from Sequence viewer, Genome Browser, Arabidopsis Tiling Array Transcriptome Express Tool, the CSB.DB (Comprehensive Systems-Biology Database), Expression Angler (BAR), Expression Browser (BAR), MapMan, NASCArrays Gene Swinger, NASCArrays Two Gene Scatter Plot, NASCArrays Spot History Pathway Tools Omics Viewer, Correlated gene (ATTED-II) and Cluster Cutting, among others. TAIR was specifically used to perform the GO categorisation (Biological Processes and Molecular Function) and the literature/state-of-the-art analyses of the final 36 genes-of-interest that led to *EGY3* identification. TAIR was also a constantly used resource throughout the present thesis, namely as a research tool for genes, sequence annotation (introns, exons, T-DNA lines) and GO annotation.

2.4.1.6. Genevestigator

Genevestigator (URL no.11) expression database includes a group of user-friendly tools that convert high-quality microarray data into easily interpretable results. Analysis of gene expression and regulation can be observed in a variety of conditions with the access to systematic transcriptomics data such as *AtGenExpress* (Hruz *et al.*, 2008). Genevestigator was used throughout the work to see expression profiles of genes-of-interest, both in mutant and wild-type backgrounds, as well as stress and development conditions. This user-friendly platform allowed for a quick insight into the particularities of expression of the analysed genes.

2.4.2. Informatic tools for plant molecular biology

2.4.2.1. ATTED

ATTED (URL no.12) is a publicly available database of co-expressed genes and *cis*-elements for identifying co-regulated gene groups in Arabidopsis (Obayashi *et al.*, 2007), which also provides co-expressed gene networks (Obayashi *et al.*, 2009). ATTED has been updated with two new features: condition-specific co-expression and homologous co-expression with rice. The development of an interactive visualization system, using the Cytoscape web system, improved the network representation (Obayashi *et al.*, 2011). The ATTED tool was used to identify co-expressed genes, and analyse networks to spot putative interactor partners or stress related genes.

2.4.2.2. GeneMANIA

GeneMANIA (URL no.13) is a tool with a Cytoscape plug-in that allows a fast gene function prediction by displaying available networks of co-expression, co-localisation, genetic interactions, pathways, physical interactions and prediction relationships between genes-of-interest based on published studies. GeneMANIA recognises the most related genes to the query gene set using a guilt-by-association approach. It has over 800 networks from six organisms and each related gene is traceable to the source network used to make the prediction (Montejo *et al.*, 2010). This tool was used to complement information about genes-of-interest, and spot any related network of value to the work.

2.4.2.3. AGRIS

AGRIS (URL no.14) is the Arabidopsis Gene Regulatory Information Server, an information resource of Arabidopsis *cis*-regulatory elements and transcription factors. Three interlinked databases are available: AtTFDB, AtcisDB and ReIN (Regulatory Networks Interactions Module), they provide clear and updated information on transcription factors, predicted and experimentally verified *cis*-regulatory elements and their interactions, respectively (Davuluri *et al.*, 2003). Promoter gene sequences were retrieved using AGRIS databases' predicted sequences for the genes-of-interest.

2.4.2.4. Athena

Athena (URL no.15) is a web-based application that stores data related to the control of gene expression. It contains a large set of data visualisation, mining, and analysis tools. Novel tools were added to facilitate the analysis of promoter sequences: a promoter visualisation tool to enable a rapid search of regulatory sequences; a TF binding site enrichment tool to identify statistically over-represented TF sites in queried promoters and a data-mining tool to select promoter sequences (O'Connor *et al.*, 2005). Athena database was used to complement the information obtained by the AGRIS database, analysing *cis*-elements and putative binding TFs to the promoter sequences.

2.4.2.5. SIGNAL

SIGNAL T-DNA Express (URL no.16) is the centralised database to search and localise Arabidopsis mutants from the large collections of insertion mutants that have been generated throughout the years by a considerable number of consortia. It also allows identification of available cDNA sequences for Arabidopsis, using a simple interface (Alonso *et al.*, 2003). The SIGNAL T-DNA Express was used to search for mutant lines for genes-of-interest. SIGNAL T-DNA Primer Design (URL no.17) was used to calculate, for each mutant line, primer sequences for diagnostic PCR purposes, with the outcome including insertion site location, primer sequence and estimated product size. The SIGNAL T-DNA Primer Design was used in some cases in the primer design for diagnostic PCR.

2.4.2.6. SUBA

SUBA (URL no.18), a SUB-cellular location database for Arabidopsis proteins, gathers and summarises information from various web-based subcellular localisation predictors. It is also useful for obtaining other protein features such as molecular weight, isoelectric point, grand average of hydropathy value, length, sequence and hydropathy plot. It can also redirect to various other databases of choice (Heazlewood JL *et al.*, 2007). This database was used in the present work to predict *in silico* localisation of proteins-of-interest.

2.4.2.7. EnsemblPlants

EnsemblPlants (URL no.19) is a portal that provides information concerning annotation of genes as well as regulatory regions, conserved base pairs across species (variation), microarray probeset mapping, comparative genomics (e.g. phylogenetic trees), and detailed information about

transcripts and proteins (domains and features) (Kersey *et al.*, 2010). EnsemblPlants database was used to gather functional information for genes-of-interest, particularly in the protein domain and features option for Pfam domain search.

2.4.2.8. TMHMM

TMHMM v 2.0 (URL no.20) is a web-based tool for the prediction of transmembrane helices in protein topology, based on the Hidden Markov Model (Krogh *et al.*, 2001). This predictor was used for the analysis made for the proteins EGY1, EGY2, EGY3, SQE1, SQE2 and SQE3. Putative transmembrane domains are defined by thresholds, and information also includes the presence of a signal anchor or a cleavable signal peptide.

2.4.2.9. Signal P

SignalP 3.0 server (URL no.21) is a web-based tool to predict the presence and location in proteins of a signal peptide/anchor and corresponding cleavage sites in the a.a. sequence (Emanuelsson *et al.*, 2007). This tool was used to identify the presence of transit peptides, their localisation and cleavage sites in both SQEs and EGYs.

2.4.2.10. ChloroP

ChloroP 1.1. server (URL no.22) is a web-based tool aiming at the identification of chloroplast transit peptides and their cleavage sites, with a high performance level (Emanuelsson *et al.*, 1999). This tool was used to identify the presence of transit peptides, their localisation and cleavage sites, for EGY1, EGY2 and EGY3.

2.4.2.11. UniProt

UniProt (URL no.23) is the Universal Protein Resource, a comprehensive resource for protein sequence and annotation data. It gives a brief but complete summary of the names and origin, protein attributes, ontology, sequences, references in the literature and portals to other databases related to 3D structure, protein-protein interaction, proteomic information, genome annotation, phylogenomic, gene expression and family and domain databases (The UniProt Consortium, 2011). This tool was used to gather important protein information for various proteins-of-interest.

2.4.2.12. MitoProt

MitoProt (URL no.24) is a web-based tool that calculates the N-terminal protein region harbouring a mitochondrial targeting sequence and its cleavage site (Claros and Vincens, 1996). MitoProt was used to identify the cleavage site and the length of the mitochondrial targeting sequence of SQE2 protein.

2.4.2.13. InterProScan

InterProScan (URL no.25) is a web-based tool that integrates documentation resource for protein families, domains, regions and sites. It combines a number of databases (referred to as member databases), each with their own methodologies and using the biological information of well-characterised proteins to derive protein signatures to queried proteins. By centralising information, this tool capitalises the individual potential of each database, producing a powerful and integrated database and diagnostic tool (Hunter *et al.*, 2009). This tool was used to determine protein topology, namely putative Pfam domains that could enlighten protein function.

2.4.2.14. NCBI

NCBI (URL no.26) is the National Center for Biotechnology Information that provides access to extensive genomic information, among other features. NCBI has various tools useful for genomic and proteomic research. In the present work, NCBI was used mainly to search for literature through PubMed (URL no.27), and to perform protein or nucleotide BLAST searches (URL no.28). Basic Local Alignment Search Tool (BLAST) encounters regions of local similarity between sequences. The program can compare nucleotide or protein sequences to sequences present in the databases and calculate a statistical significance of matches defined by query coverage, score and the E-value. BLAST can be used to infer functional and evolutionary relationships between sequences as well as help identify members of gene families.

2.4.2.15. DNASTAR

Nucleotide sequence analysis was performed using *DNASTAR software* (Lasergene). Within this pack of programmes, EditSeq was used for editing, translating and inverting sequences as well as finding ORFs. Contig analysis of sequenced products was made using the SeqMan feature. MegAlign was used to perform ClustalW alignments on protein data.

2.4.2.16. Software for primer design

Primer design, except for the diagnostic PCR primers, was carried out using *OLIGO Primer Analysis Software v6.0* (URL no.6) or *Primer3* online application (URL no.7). Primer design generally took in consideration the following principles for each pair: matching and optimal Tms, correct primer length, avoidance of primer dimer and hairpin structures, optimisation of GC content (40-60%), presence of GC clamp, estimation of optimal annealing temperature (Griffin and Griffin, 1994).

2.4.2.17. Informatic tools for phylogenetic analysis

In order to establish phylogenetic relationships between SQE1 or EGY1 with their corresponding homologues, sequences were retrieved following blastp homology search (NCBI) using the SQE1 or EGY1 a.a. sequences. The MEGA 5 software (Tamura *et al.*, 2011) was used to perform a ClustalW (BLOSUM) alignment of all sequences, allowing also for the exclusion of previously retrieved partial proteins. Phylogenetic trees were then performed by Phylogeny à la Carte (Dereeper *et al.*, 2008) (URL no.29), using the Maximum-likelihood algorithm, with subsequent Bootstrap analysis (500 trees). The rate of aminoacid substitution was empirically calculated using the WAG model. Rates among sites were Gamma distributed (G) with 4 categories. The ML Heuristic Method used was Nearest-Neighbor-Interchange (NNI).

3.

*“Functional characterisation of
SQEs”*

3.1.

*“An introduction to SQEs and their role in
plant development”*

3.1.1. Sterols

Sterols are isoprenoid-derived compounds essential for growth and development in eukaryotes. Sterols are major constituents of the biological membranes and also provide precursors for the biosynthesis of steroid hormones in plants, the brassinosteroids (Benveniste, 2004; Boutté and Grebe, 2009). In vertebrates, cholesterol is the major sterol present however, this varies in other organisms. In fungi and some unicellular algae the major sterol is ergosterol, and in higher plants, there is a complex mixture of cholesterol (minor component), 24-ethyl sterols (> 60% of sitosterol and stigmasterol) and 24-methyl sterols (< 40%) (Benveniste, 2002). An example, of a plant sterol profile for the wild-type *Arabidopsis thaliana* ecotype Columbia could be of about 64% sitosterol, 11% 24-methyl cholesterol (an epimeric mixture of campesterol + 22-dihydrobrassicasterol), 6% stigmasterol, 3% isofucosterol, and 2% of brassicasterol, a unique sterol present in Brassicaceae (Benveniste, 2004).

Sterols can be present in three forms: free sterols, steryl esters, and steryl glucosides. These free sterols are constituents of the membrane lipid bilayer and its interaction with phospholipids are functionally important to regulate membrane fluidity and permeability. (Benveniste, 2004). Phytosterols, for instance, have been shown to increase cohesion in the membrane, to maintain a state of plant membranes dynamics less susceptible to temperature shocks (Dufourc, 2008). Additionally, other sterol regulatory functions have been reported: (a) HEDGEHOG covalent binding of cholesterol correlated to embryonic development of vertebrates, (b) cholesterol interaction with caveolin inducing membrane microdomains formation (caveoleae/rafts), which may constitute signaling centers for multiple pathways, (c) transport of plant sterols by elicitors from *Phytophthora spp.* leading to a hypersensitive like response in tobacco. These processes can activate signalling pathways implicated in important mechanisms, such as cell division, development or resistance to pathogens (Benveniste, 2002).

Studies over the years showed the importance of a correct sterol composition in plants because of its role in embryonic pattern formation, cell division, cell elongation, cell polarity, cellulose accumulation, as well as genetic interactions between the sterol biosynthesis and the ethylene signalling pathways. However, little is known about the mechanisms and the downstream targets, by which sterols influence these processes (Boutté and Grebe, 2009).

3.1.2. Biosynthetic pathway

3.1.2.1. Squalene biosynthesis

Sterol biosynthesis is complex and involves at least 25 steps from IPP (isopentenyl diphosphate) to the end of the pathway (Benveniste, 2002). IPP is formed through the mevalonate pathway (also known as HMG-CoA reductase pathway or isoprenoid pathway) (Figure 3.1), in which acetyl-CoA is the precursor of this cytosolic pathway, and IPP is the intermediate of triterpenoids and terpenes biosynthesis (Phillips *et al.*, 2006; Boutté and Grebe, 2009).

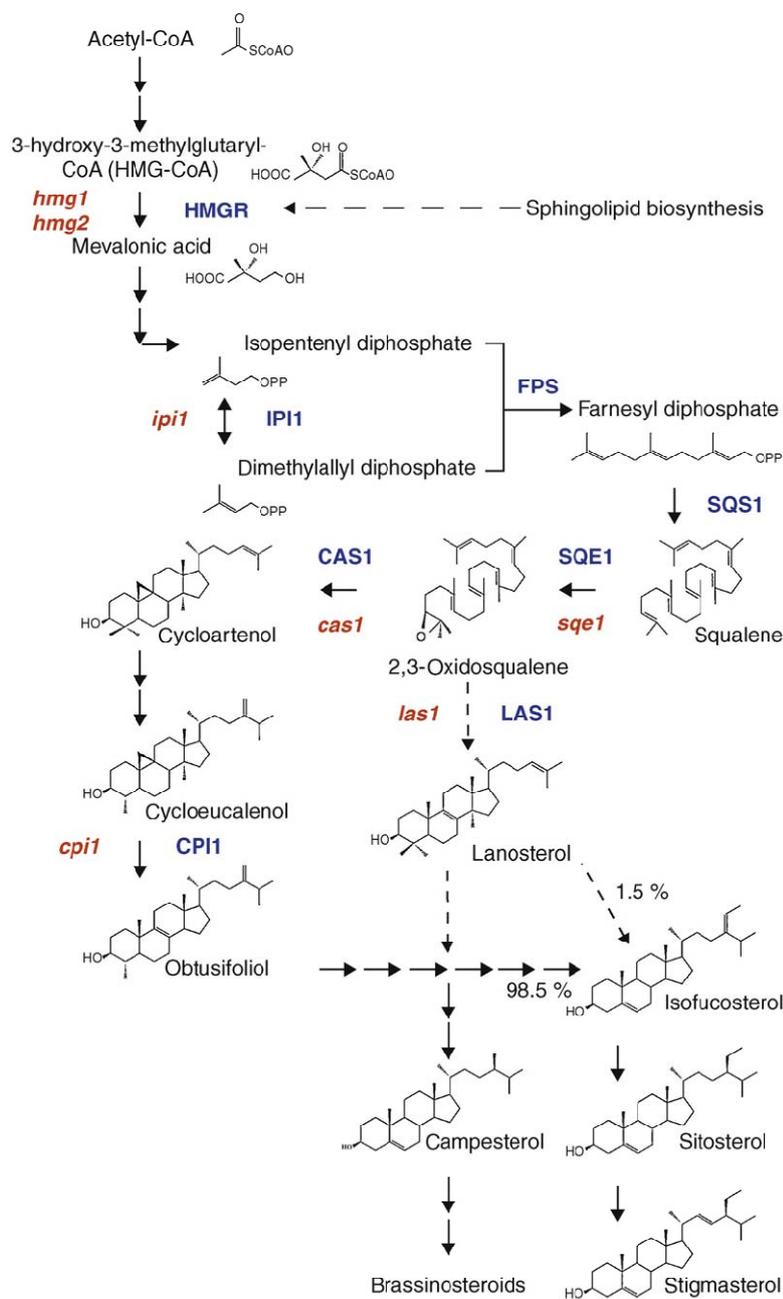


Figure 3.1. – Diagram of the sterol biosynthetic pathway. Adapted from Boutté and Grebe (2009).

The HMG-CoA is catalysed by HMGRs, encoded in Arabidopsis by the two isoforms coded by *HMG1* and *HMG2* genes, and is a very important step in the mevalonate pathway (Benveniste, 2002). Isopentenyl isomerase (IPI1) catalyses the isomerisation between isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) in the mevalonate pathway. One molecule of dimethylallyl pyrophosphate and two molecules of isopentenyl pyrophosphate condense to form farnesyl pyrophosphate (FPP). The tail-to-tail coupling of two molecules of farnesyl pyrophosphate yields squalene by action of squalene synthase 1 (SQS1), and its epoxidation is catalyzed by squalene epoxidase 1 (SQE1), converting squalene to 2,3-oxidosqualene, which is the first oxygenation step in the sterol biosynthetic pathway (Benveniste, 2002; Berg *et al.*, 2002; Boutté and Grebe, 2009).

3.1.2.2. Oxidosqualene cyclases

The 2,3-oxidosqualene cyclases (OSC), form a family of biocatalysts that convert 2,3-oxidosqualene (OS) to polycyclic triterpenes. Triterpenoids, synthesised from IPP via squalene, include sterols, steroids and triterpenoid saponins (Figure 3.2). Apart from sterols that derive from cycloartenol or lanosterol, other triterpenes may have different roles, including in plant defence. The high diversity of triterpenes is mainly due to the fact that the OSC family has greatly extended in plants (Phillips *et al.*, 2006). A diversity of triterpenoids is biosynthesised in plants and multiple OSC enzymes are encoded. The major OSC is CAS (cycloartenol synthase 1) which is involved in the cyclization of 2,3-oxidosqualene (OS) into cycloartenol through the protosteryl cation intermediate, and it is assumed that CAS is the point of derivation for the others OSC (Phillips *et al.*, 2006; Boutté and Grebe, 2009). Another product of cyclization, via the protosteryl cation intermediate, is catalysed by lanosterol synthase (LSS/LAS1) to produce lanosterol, known for being the initial carbocyclic sterol precursor in animals, fungi, and trypanosomatids. It was proven that cycloartenol, rather than lanosterol, is the major plant sterol precursor. However, lanosterol biosynthesis has been confirmed in some plants, as in the latex of *Euphorbia sp.* In Arabidopsis, LAS1 protein is the most similar to CAS1, demonstrating its possible importance despite CAS1 predominance and major role in plants (Phillips *et al.*, 2006).

3.1. – An introduction to SQEs and their role in plant development

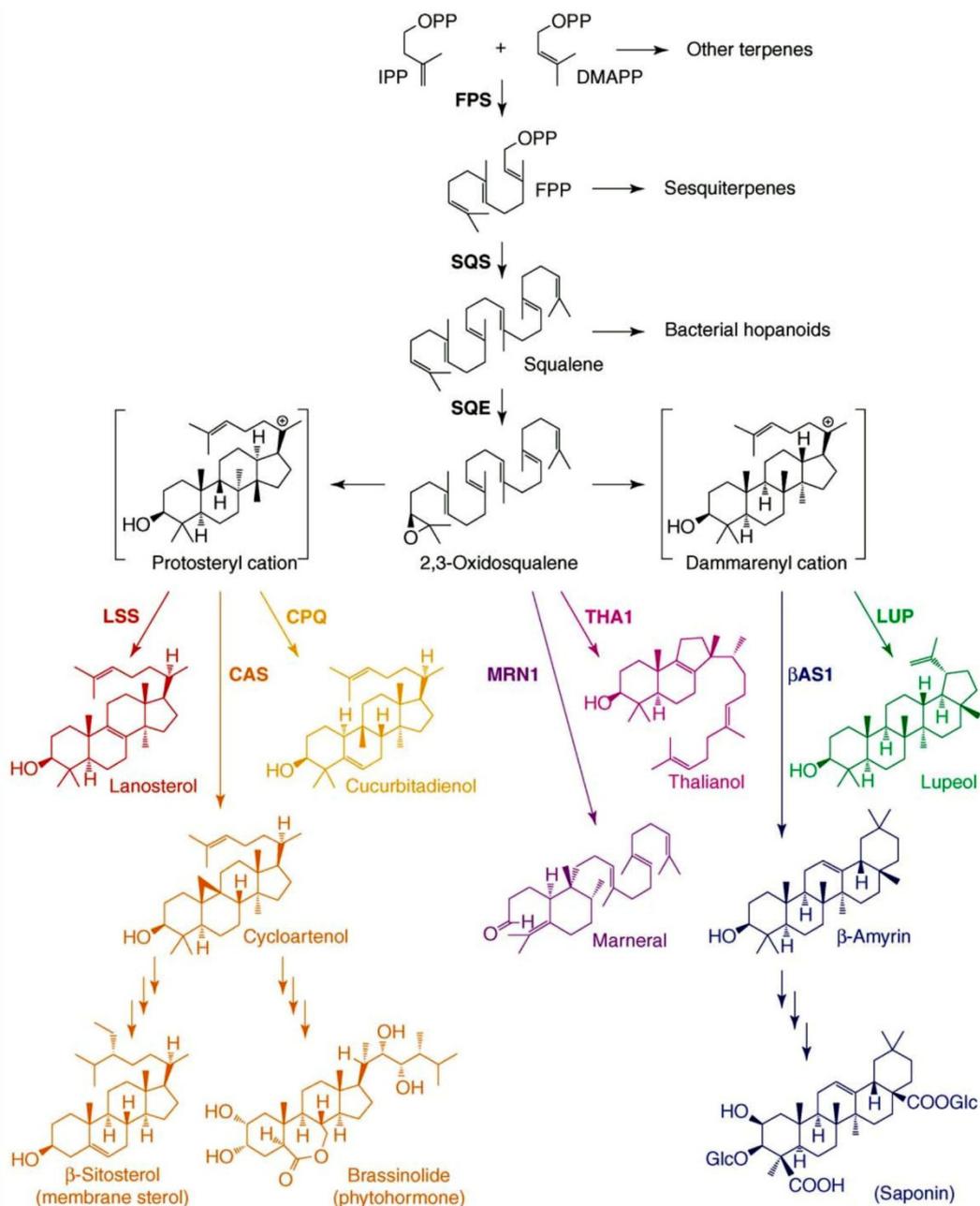


Figure 3.2. - Schematic of the cyclization of 2,3-oxidosqualene through different OSC (oxidosqualene cyclases) with different end products, in pathways parallel to the main sterol biosynthetic pathway. Adapted from Phillips *et al.* (2006)

Recent studies using loss- and gain-of-function *Arabidopsis* mutants of the *LAS1* gene, substrate feeding and subsequent metabolite detection revealed that the lanosterol pathway contributes to 1.5% of the phytosterol production in *Arabidopsis* (Ohyama *et al.*, 2009). Although, this may not seem essential in standard conditions, it is necessary for the phytosterol production, and it is also suggested that the lanosterol pathway may be important for secondary metabolite synthesis, related to plant defence compounds. Future studies would be relevant to access

conditions and biological processes that require the lanosterol pathway for the total phytosterol production (Boutté and Grebe, 2009; Ohyama *et al.*, 2009).

The cyclization of 2,3-oxidosqualene (OS) can also be achieved through the dammarenylation (Figure 3.2), performed by lupeol synthase (LUP) and β -amyrin synthase (BAS). Some of their end-products, such as lupeol and β -amyrin, are reported in various plant processes. The β -amyrin is a precursor of saponins, which are triterpene glycosides, such as the antifungal saponin avenacin found in *Avena* roots. The β -amyrin synthase genes are thought to have a tissue-specific expression that influences the specificity of β -amyrin and its metabolites. On the other hand, LUP is connected to root nodulation in several plants such as *Glycyrrhiza* and *Vicia faba*, but the mechanism is not known (Phillips *et al.*, 2006). Arabidopsis encodes several multifunctional enzymes paralogs that biosynthesize β -amyrin, lupeol, in fact, Arabidopsis has 13 putative OSC (Husselstein-Muller *et al.*, 2001)¹ in which there are some various putative LUPs previously reported to be multifunctional, meaning that they can produce minor products others than lupeol: At1g78970/LUP1, At1g78960/LUP2, At1g66960/LUP5, and At1g78500/PEN6 (Phillips *et al.*, 2006).

Other OSC have also been previously catalogued in the LUP gene family, thought presently their information has been updated by their end-products in Arabidopsis: the At1g78955/CAMS1 producing mainly camelliol (Kolesnikova *et al.*, 2007) and the At1g78950/BAS producing β -amyrin (Shibuya *et al.*, 2009). The At5g42600/MRN1 catalyzes an unusual cyclization reaction: oxidosqualene is converted to a bicyclic cation that undergoes rearrangement, and an A-ring cleavage to generate a monocyclic aldehyde, marneral (mainly) (Phillips *et al.*, 2006; Xiong *et al.*, 2006), and At5g48010/THAS1 that converts OS into thalianol (Phillips *et al.*, 2006). Most of them have been characterised by cloning and heterologous expression in yeast. The PEN (Pentacyclic triterpene synthase) gene family At4g15340/PEN1, At4g15370/PEN2, At5g36150/PEN3 and At3g29255/PEN7 are thought to be OSC, with PEN2 having other name - BARS1, for its conversion of OS into baruol (Lodeiro *et al.*, 2007). Nevertheless, more *in vivo* studies in Arabidopsis must be conducted to determine all products of the putative OSC. Putative OSC are in figure 3.3, where they are grouped according to their phylogeny, with CAS1 and LAS1 extremely close and grouping PEN's and LUP's together.

¹ In this study they exclude LAS1 from the 13 putative OSC: CAS1, AtLUP1-AtLUP5 and AtPEN1-AtPEN7.

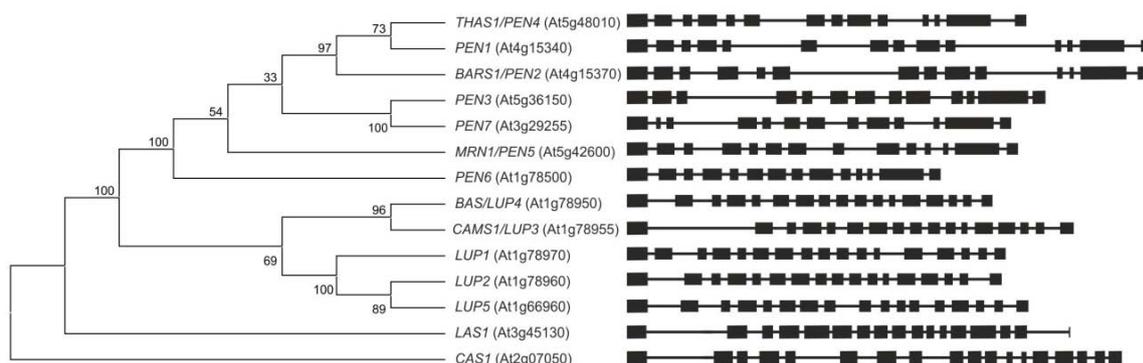


Figure 3.3. – Cladogram of the *Arabidopsis thaliana* putative oxidosqualene cyclases. The cladogram of the amino acid sequences was rooted with CAS1. At the right the exons (dark boxes) and introns (lines) are depicted. Data was analysed using MEGA 5 (Tamura *et al.*, 2011) first with a Clustal W alignment and afterwards with a phylogeny analysis using a maximum likelihood tree with a bootstrap of 500.

Interestingly, it was reported an operon-like gene cluster in *Arabidopsis thaliana* required for the triterpenoid synthesis, which they named the thalianol pathway. They studied the region of four contiguous genes predicted to encode an OSC (At5g48010/THAS1), two CYP450s and a BADH family acyltransferase. All were correlated in terms of expression, specifically in the root epidermis. They proved THAS1 function, and also that both CYP450 are involved in the conversion of thalianol to downstream products, being these contiguous genes involved in three consecutive steps, in the synthesis, and modification of thalianol. The other gene is hypothesised to be required for modification of the desaturated thalian-diol, though not proven. It is also reported the possibility that other triterpenoid biosynthetic genes are also in clusters (Field and Osbourn, 2008).

3.1.2.3. Oxidosqualene to sterols

As previously referred, 2,3-oxidosqualene can be converted to cycloartenol by CAS, being used in the production of plant sterols. One characteristic of higher plants is that they contain both 24-methyl and 24-ethyl sterols, and this is accomplished by SMTs (S-Adenosylmethionine-sterol-C-methyltransferases). In *Arabidopsis*, the SMT1, catalyses the conversion of cycloartenol into 24-methylene cycloartanol (Figure 3.4). The 24-methylene cycloartanol is then converted to cycloeucalenol which involves a C-4 α -methyl oxidase, a C4-decarboxylase/C3-dehydrogenase, and a 3-keto reductase (Benveniste, 2002).

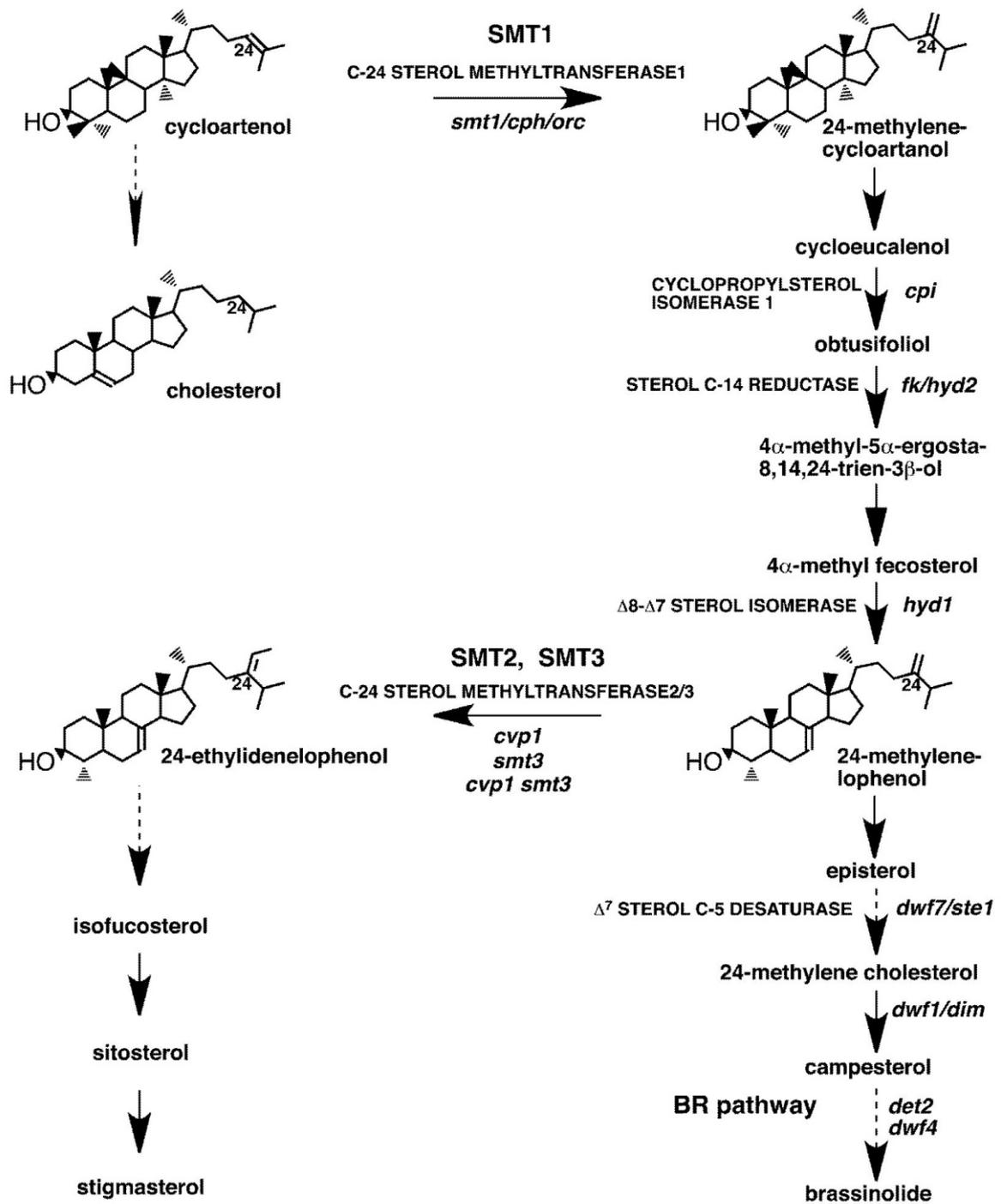


Figure 3.4. – Phytosterol biosynthetic pathway. Relevant mutants are represented parallel to the enzyme affected in the biosynthetic steps. Multiple steps are indicated in dashed lines. Adapted from Carland *et al.* (2010).

Obtusifoliol is then formed by the opening on the cyclopropane ring, described has a restricted step to the plant kingdom and catalyzed by CPI (cyclopropyl sterol isomerase 1). The product is then converted to 4 α -methyl-5 α -ergosta-8,14,24(24¹)-trien-3 β -ol by CYP51, an obtusifoliol 14 α -demethylase, proceed by a $\Delta^{8,14}$ -sterol- Δ^{14} -reductase, and a $\Delta^8\Delta^7$ sterol isomerase to the 24-methylene-lophenol end-product (Benveniste, 2002). Thus, this product can undergo 24-methyl end product sterols or be transformed by another SMT (SMT2 and SMT3) to undergo 24-ethyl end products (Figure 3.4) (Carland *et al.*, 2010). Though the end-products are different, mutants affecting those enzymes inferred that they work in parallel for both 24-methyl and ethyl pathways, involving desaturases and reductases to have campesterol and sitosterol. Sitosterol can be converted to stigmasterol, both the most predominant sterols in plants. On the other hand, campesterol enters the Brassinosteroid pathway to the production of brassinolide (Benveniste, 2002; Carland *et al.*, 2010).

3.1.2.4. Brassinosteroids biosynthesis

Brassinosteroids (BRs) include over 40 polyhydroxylated sterol derivatives, ubiquitously spread through the plant kingdom. BRs are the phytohormones that resemble most the animal steroid hormones, and are functionally involved in the regulation of embryonic and post-embryonic development, and in the adult homeostasis. BRs regulate the expression of numerous genes, are involved in many complex metabolic pathways, and contribute to the regulation of morphogenesis, cell division and differentiation (Clouse, 2002).

Brassinolide is formed by the conversion of the membrane sterol campesterol, in a series of reductions, hydroxylations, epimerizations, and oxidations represented in figure 3.5 (Clouse, 2002).

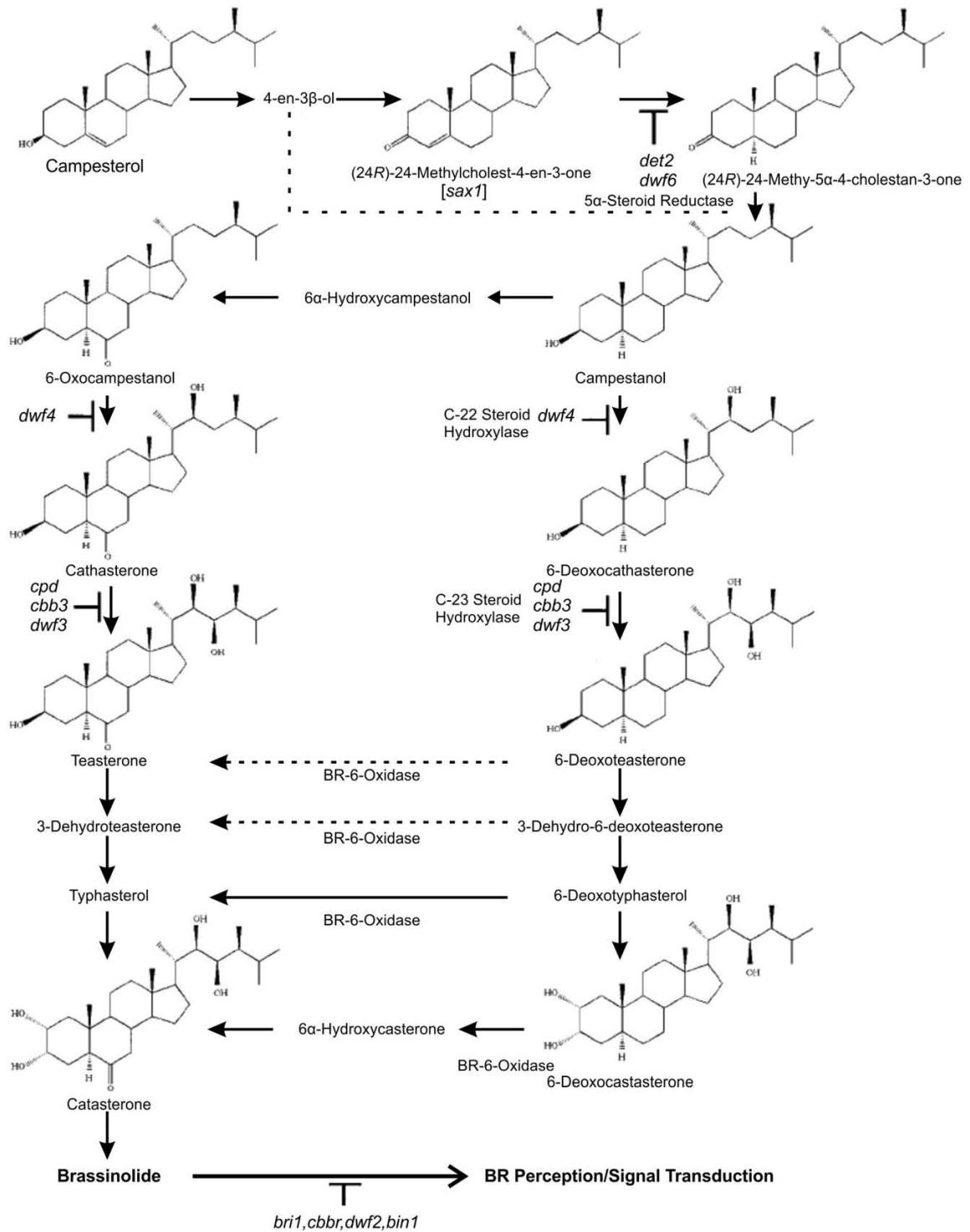


Figure 3.5. – Brassinosteroid biosynthetic pathway. The pathway is shown from campesterol to brassinolide with relevant known Arabidopsis mutants that affect the pathway. Mutants relative to perception or signalling are also represented. Adapted from Clouse (2002).

3.1.2.5. Mutants of sterol and brassinosteroid pathway

Mutants affected in sterol and brassinosteroid (BR) pathways have already been described in the literature. Here, a brief summary of the most relevant sterol and BR mutants, the corresponding affected proteins, and the functional consequences of the mutation will be described.

***hmg1* and *hmg2*.** HMG-CoA reductase (HMGR) is a key regulatory enzyme converting HMG-CoA into mevalonic acid (MVA) and has been described as the main rate-limiting step in isoprenoid biosynthesis (Boutté and Grebe, 2009). In mammals, the level of this enzyme is regulated at transcriptional and post-translational levels. An inhibition in this enzyme will reduce biosynthesis of cholesterol (Suzuki *et al.*, 2004), but it has also been reported a feedback regulation of this enzyme in plants (Nieto *et al.*, 2009). In Arabidopsis, there are two encoded genes, *HMG1* and *HMG2*. These two genes display a different profile of expression, *HMG1* being expressed in all tissues and *HMG2* in meristematic and floral tissues (Enjuto *et al.*, 1994; Enjuto *et al.*, 1995). The *hmg1* mutant presents early senescence, sterility and a dwarf phenotype (cell elongation deficit), resulted in part by a decrease in metabolites downstream of squalene, and by sterol levels that are 50% lower than wild-type. This phenotype is not related to cytokinins and BRs, excluding the hypothesis of being a BR-deficient mutant (Suzuki *et al.*, 2004). Although under normal growth conditions *hmg2* mutant do not show any phenotype, it responds more severely than wild-type to the inhibitor lovastatin (Suzuki *et al.*, 2004). The *hmg2* mutant presented a sterol content 15% lower than that of the wild-type, and triterpenoid quantification of β -amyrin, α -amyrin and lupeol, evidenced a difference in the levels of triterpenoids in both mutants, being 65% and 25% lower than in the wild-type, for *hmg1* and *hmg2*, respectively. These results indicate that both HMGRs are responsible for the biosynthesis of triterpenes despite lack of visible phenotype in the *hmg2* mutant (Ohyama *et al.*, 2007). It has been recently reported that *HMG2* controlled by *HMG1* promoter, complemented HMGR activity, sterol accumulation, gene expression, and morphology in the *hmg1*. The importance of this step is reinforced by the male gametophytic lethality of *hmg1/hmg2* double mutant, suggesting a complete loss-of function of the MVA pathway (Suzuki *et al.*, 2009).

***dry2/sqe1-5*.** The *dry2/sqe1-5* mutant, named for being drought hypersensitive, is affected in the enzyme squalene epoxidase 1. This mutant cannot be rescued by BR and presents pleiotropic developmental defects, such as altered root architecture and root hairs, diminished shoot size and chlorophyll content. The mutants are dwarfed with long lifespan, have pale-green leaves and short number of seeds *per* silique (Posé *et al.*, 2009). Other alleles have been reported to have

phenotypes more severe, such as the production of shrivelled and unviable seeds (Rasbery *et al.*, 2007).

smt1. S-Adenosylmethionine-sterol-C-methyltransferase catalyses a step in sterol biosynthesis unique to higher plants and fungi. In fact, they are present in two different steps, as previously stated. The *smt1-3* mutant was isolated and present shorter petioles and stems, smaller and rounder leaves, and dwarfed siliques. They also presented growth and fertility deficiency, exhibit root sensitivity to calcium, defective embryo morphogenesis, and the phenotype is not rescued by brassinosteroids. However, it seems to be some promiscuity between the three Arabidopsis SMT enzymes, though SMT1 preferentially converts cycloartenol in the first step, and SMT2 and 3 in the second step. These last two can perform the first step in the absence of SMT1, however, at a slower metabolic rate (Diener *et al.*, 2000; Benveniste, 2004).

cvp1. Cotyledon vascular patterning 1 is the name of the mutant for the *SMT2* resulting in some developmental defects, which do not affect viability, such as irregular cotyledon vein patterning, serrated floral organs, and reduced plant size. It is suggested that *SMT3* can substitute *SMT2* activity. The *smt3* mutant itself is apparently wild-type, but the double mutant *cvp1/smt3* presents an enhanced phenotype relative to the *cvp1*, such as irregular cotyledon vein pattern, and even defective root growth, no apical dominance, sterility, and homeotic floral transformations. Sterol profiles in these mutants are affected, though not in what concerns brassinosteroid profiles, and this different profile appears to affect auxin response of *cvp1* plants (Carland *et al.*, 2010).

fackel and hydra. The *fackel* (*fk*) mutant was isolated among other dwarf mutants and corresponded to a C-14 reductase in sterol biosynthesis (Figure 3.4). This mutant shows body disorganisation of the seedling, sustaining the hypothesis of the gene's involvement in cell division and expansion, as in the embryo (Schrack *et al.*, 2000). In fact, it was reported another mutant for the same gene, which was called *hydra2*, identified together with *hydra1* for its seedling lethality. Both *hyd1* and *hyd2/fk* are dwarfed, have multiple cotyledons and have abnormal vascular patterning; however, *hyd2/fk* mutants display a more severe root-defective phenotype. The *hyd1* is a mutant for the subsequent enzyme in the pathway after the *fackel* gene, encoding a $\Delta^8\Delta^7$ sterol isomerase. Both phenotypes, for the hydra mutants, can be rescued partially by auxin inhibition and ethylene signalling, but not with applied sterols or brassinosteroids (Souter *et al.*, 2002; Pullen *et al.*, 2010).

dwf7/ste1. The *dwarf7/ste1* mutant corresponds to a defect in the Δ^7 sterol C-5 desaturase, has a dwarf phenotype, prolonged lifespan, short inflorescences, dark-green plants with round leaves, reduced fertility, and short pedicels and siliques. Though not mechanically sterile the mutant has lesser number of seeds per silique. In resume, cell elongation is the reason for drastic reduction of

so many organs. Sterol biosynthesis deficit results in poor brassinosteroid outcome and *dwf7* phenotypes are rescued by BRs (Choe *et al.*, 1999; Clouse, 2002).

dwf5. The *dwf5* mutant is disrupted in a sterol Δ^7 reduction step. Phenotype is similar to other dwarf BR mutants including small, round and dark-green leaves, and short stems, pedicels, and petioles. It can also be rescued by BRs (Choe *et al.*, 2000).

dwf1/dim. The *dwf1* mutant was isolated prior to its allele *dim*. However, only when another allele *cabbage1 (cbb1)* was reported, it was shown to be rescued by BR treatment (Clouse, 2002). The *dim* mutant has very short hypocotyls, petioles, stems, and roots as other dwarf mutants due to reduced size of cells, in the longitudinal plan of these organs (Takahashi *et al.*, 1995). DWF1 protein is reported to be involved in both steps, in the conversion of 24-methylenecholesterol to campesterol, and from isofucosterol to sitosterol (Klahre *et al.*, 1998; Clouse, 2002).

det2/dwf6. The *de-etiolated-2* and *dwarf6* mutants are affected in a steroid-5- α -reductase in the brassinosteroid biosynthesis. The *det2*, in Arabidopsis, presents the same aspect in the light or dark as for the hypocotyl growth inhibition, cotyledon expansion, primary leaf initiation, anthocyanin accumulation, and derepression of light-regulated gene expression. When in the light, however, the mutant is dwarf with dark-green leaves, delayed senescence and flowering, and also have lesser male fertility and apical dominance (Chory *et al.*, 1991). Brassinolide can rescue the *det2* phenotypes (Fujioka *et al.*, 1997). While *det2* is a typical BR-deficient mutant, its phenotype is not as severe as *cpd* and *bri1* mutants (Clouse, 2002).

dwf4. The *dwf4* mutant is characterised by a dwarf phenotype as previously described for mutants of BR biosynthesis and can be rescued by supplemented BRs. The effects of the mutations are reflected in cell elongation defects, limited growth in almost all organs, which results in sterility, and also present delayed flowering and a prolonged lifespan (Azpiroz *et al.*, 1998). This mutant is affected in an encoded cytochrome P450 that mediates multiple 22 α -hydroxylation Steps in the BR biosynthesis (Choe *et al.*, 1998). This enzyme seems to play an important role in BR biosynthesis because it is product of feedback regulation (along with the *CPD* gene) by transcription factors (BZR1) of the Brassinosteroid signalling (Kim and Wang, 2010).

cpd. *CPD* gene encodes a cytochrome P450 steroid side-chain hydroxylase (CYP90), and is an essential enzyme in the brassinolide hormone biosynthesis, and acts as a C-23 steroid hydroxylase (Mathur *et al.*, 1998). This dwarf mutant presents de-etiolation and derepression of light-induced genes in the dark, male sterility, and activation of stress-regulated genes in the light. This mutants are rescued by BRs and are feedback regulated by brassinolide (Szekeres *et al.*, 1996; Kim and Wang, 2010).

bri1. This mutant was identified as a brassinosteroid insensitive, although retains sensitivity to auxins, cytokinins, ethylene, abscisic acid, and gibberellins. The mutant presents multiple deficiencies in development, including dwarfism, dark-green and thickened leaves, male sterility, reduced apical dominance, and de-etiolated and dark grown seedlings (Clouse *et al.*, 1996). Several other alleles were described for this locus, and it was reported to be an ubiquitously expressed putative receptor kinase (Li and Chory, 1997). In fact, it encodes a putative Leucine-rich repeat receptor Serine/Threonine kinase involved in brassinosteroid signalling (Friedrichsen *et al.*, 2000; Oh *et al.*, 2000). Some of the previous described mutants are depicted in figure 3.6.

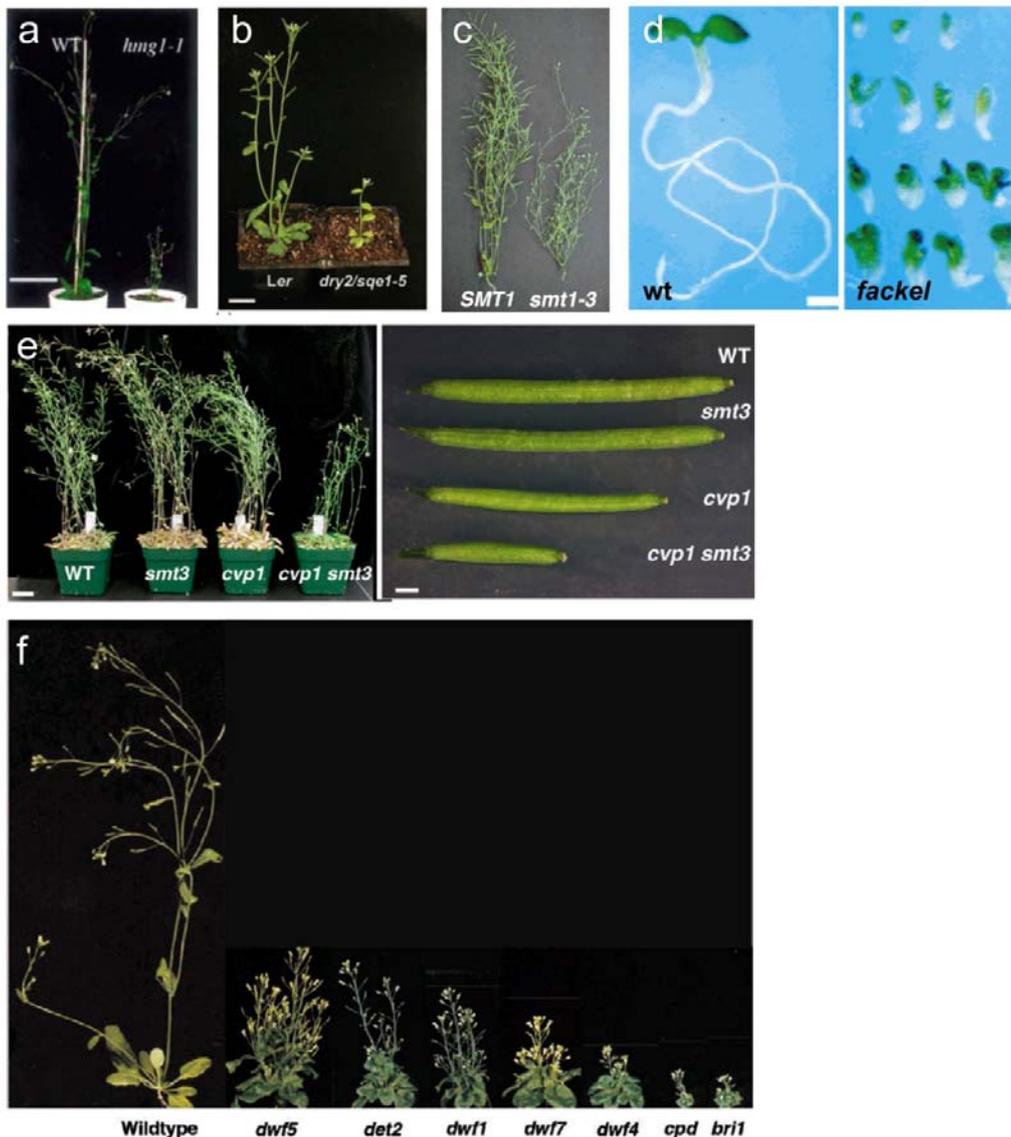


Figure 3.6. – Sterol and brassinosteroid mutant phenotypes. a) The *hmg1-1* mutant, scale bar indicates 5 cm. Suzuki *et al.* (2004). b) The *dry2/sqe1-5* mutant, scale bar indicates 2 cm. Posé *et al.* (2009). c) The *smt1-1* mutant. Diener *et al.* (2000). d) The *fackel* mutant, scale bar indicates 0.5 mm. Schrick *et al.* (2000). e) The *smt3*, *cvp1* and the double *cvp1/smt3* mutant in adult stages in soil (scale bar indicates 1 cm) and representation of their siliques (scale bar indicates 1 μ m). Carland *et al.* (2010). f) The *dwf5*, *det2*, *dwf1*, *dwf7*, *dwf4*, *cpd* and *bri1* mutants in comparison to wild-type plants. Clouse (2002).

3.1.3. Brassinosteroids signalling

Brassinosteroids are essential hormones with importance in a variety of processes, from development to plant physiology, such as cell elongation, vascular differentiation, root growth, light responses, stress resistance and senescence. As described before, most affected mutants of BR pathway include growth defects (dwarfism), dark-green leaves, delayed flowering, male sterility, and photomorphogenesis when in the dark (Kim and Wang, 2010).

Brassinolide is perceived by the BRI1 receptor kinase at the cell surface, activating the complex BRI1/BAK1 by transphosphorylation, and subsequent phosphorylation of the other BSKs kinases (Figure 3.7). Then, BSU1 phosphatase is activated by the BSKs, resulting in the desphosphorylation and inactivation of the BIN2 kinase. BIN2, being inactive, is incapable of constitutively phosphorylating BZR1 and BZR2/BES1, which would have led to cytoplasm retention by the 14-3-3 proteins. So, with BIN2 inactive, BZR transcription factors accumulate in an unphosphorylated form in the nucleus. Moreover, it is suggested that this BZR transcription factors are responsible for the regulation of BR genes, modulating plant growth and development (Kim and Wang, 2010).

Microarray analyses led to the discovery of hundreds of BR-responsive genes, involved in processes such as cell wall synthesis and modification, cytoskeleton formation, and the biosynthesis, signalling and transport of plant hormones, particularly auxin. However, there are some putative transcription factors in the list suggesting an even more complex regulatory network (Li and Jin, 2006). BZR1 binds to BRRE elements (BR-Response Element) present in the BR biosynthetic *CPD* and *DWF4* promoters, suppressing their expression in what is known as a regulatory feedback inhibition (Kim *et al.*, 2006; Li and Jin, 2006). BES1, however, was found to bind the *SAUR-AC1* (auxin inducible gene) promoter to activate gene expression (Yin *et al.*, 2005; Li and Jin, 2006).

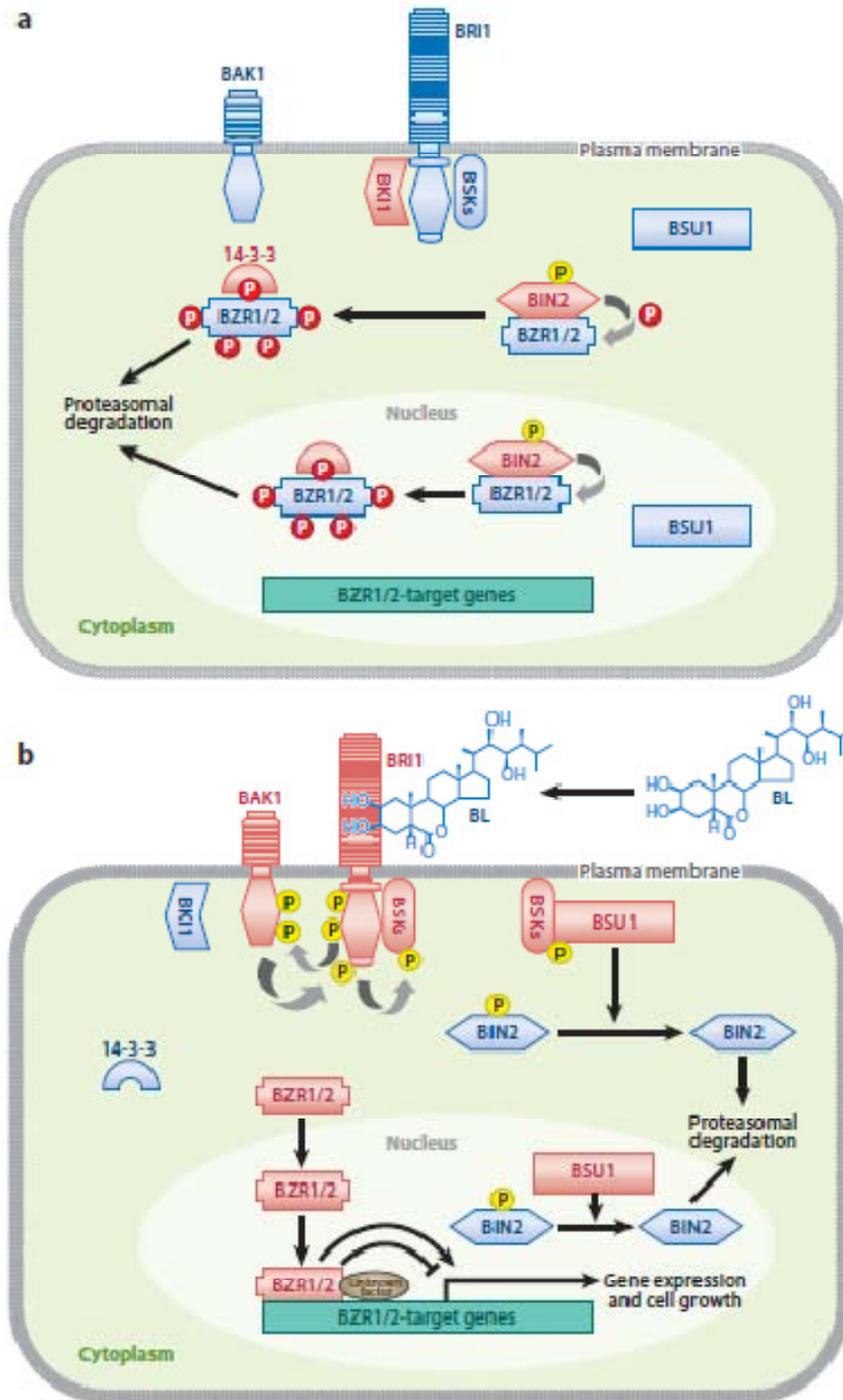


Figure 3.7. – Brassinolide signalling pathway. a) In the absence of brassinolide, BAK1 cannot bind to BRI1 because of BK1 interaction. BRI1 is also bound to inactive forms of BSKs. BSU1 is inactive and the active BIN2 constitutively phosphorylates BZR1 and BZR2/BES1, leading to an export from the nucleus to the cytoplasm, where is retained by the 14-3-3 proteins. There is also loss of DNA-binding activity, and proteasomal degradation of BZR1 and BZR2/BES1. b) In the presence of brassinolide, the hormone binds to the receptor BRI1, which in turn provokes association with BAK1 and disassociation with BK11. Sequential transphosphorylations between BRI1 and BAK1, activation of BRI1 is achieved and it phosphorylates BSKs. BSKs are then released to bind BSU1, activating it to inhibit BIN2 through dephosphorylation. This leads to an accumulation of unphosphorylated BZR1 and BZR2/BES1 in the nucleus. These transcription factors can bind the genomic DNA to regulate BR-target gene expression, thereby modulating growth and development of plants. Adapted from Kim and Wang (2010).

3.1.4. Squalene epoxidase family

Squalene epoxidase enzymes convert squalene into 2,3-oxidosqualene, in a mediated oxidation process important in sterol biosynthesis. In mammals and yeast, only one squalene epoxidase has been found (Landl *et al.*, 1996; Nagai *et al.*, 1997). However, several genes, putatively encoding SQEs, have been described in plants, suggesting a possible step of unique regulation in plants. In *Arabidopsis thaliana*, by BLAST search with known sequences of *Medicago*, six putative SQE genes were identified (Rasbery *et al.*, 2007). *SQE1*, *SQE2* and *SQE3* cDNAs were able to complement the yeast *erg1*, a mutant that lacks the SQE, functionally demonstrating enzymatic SQE activity for the proteins encoded by the three genes (Rasbery *et al.*, 2007). The other *Arabidopsis* genes that show homology to SQEs (*SQE4*, *SQE5* and *SQE6*) may be catalytic distinct, which is evidenced by its distinct phylogeny with the others and the inability to complement the yeast *erg1* mutant. Expression analysis, based on microarray databases, suggest that *SQE1* and *SQE3* are expressed in most plant tissues, whereas *SQE2* and *SQE4* present a lower expression. The *SQE5* expression appears higher in adult plant tissues, while *SQE6* is more expressed in seeds, hypocotyls, and rosettes (Rasbery *et al.*, 2007). Phenotypic analyses of *sqe1-3* and *sqe1-4*, mutants affected in *SQE1*, indicated defects in development, including reduced and highly branched roots, hypocotyl elongation and production of unviable seeds leading to sterility. These defects were not rescued by brassinolide, suggesting that their phenotypes are not associated with BR (Rasbery *et al.*, 2007). Later, a hypomorphic allele for *SQE1*, *dry2/sqe1-5*, was identified and analysed in with detail, although sharing many of the phenotypes described for the null *sqe1-3* and *sqe1-4* (Posé *et al.*, 2009). The *dry2/sqe1-5* mutant shows reduced size under standard growing conditions, is hypersensitive to dehydration, has pale green leaves (reduced chlorophyll content), smaller and branched root, and prolonged lifespan, but in contrast to the null alleles, has viable seeds (Posé *et al.*, 2009). Experiments involving stomata closure showed that *dry2* is ABA insensitive and is required for ABA-induced stomatal closure. Due to this, ROS production (H_2O_2) was analysed, showing in *dry2/sqe1-5* very low ROS levels compared to wild-type (Posé *et al.*, 2009). ROS is also a well known second messenger involved in for example polarized root hair growth, which in turn needs a precise localised production at the root tip of ROS by the RHD2/AtrbohC NADPH oxidase (Gapper and Dolan, 2006; Takeda *et al.*, 2008). Based on that, root hair was analysed in *dry2/sqe1-5* and they were smaller and branched than wild-type. Subsequent experiments showed that these defects in root hair morphology were due to a de-localization of the RHD2 NADPH oxidase protein. Therefore, analysis of *dry2/sqe1-5* showed a previously unrecognised role for sterols in the regulation of NADPH oxidases and ROS production

(Posé *et al.*, 2009). GUS histochemical studies revealed that *SQE1* is expressed in the shoot and in the elongation zone of the root. Expression studies of the *SQE* genes, encoding functional enzymes, revealed that *SQE3* expression was higher in the mutant, when comparing it to the wild-type, compensating in part the *SQE1* deficiency and the similar sterol profile in shoots (*SQE3* is expressed higher in shoots than in roots), while in roots the unbalanced sterol profile is very clear (Posé *et al.*, 2009).

Although mutant characterisation has been instrumental in the elucidation of the sterol and brassinosteroid biosynthetic pathways, it is still far from totally understand all the aspects and the processes in which they participate. Triterpenoid mutants in the biosynthesis show in most cases severe pleiotropic defects, demonstrating the importance of triterpenoid biosynthesis for plant development (Rasbery *et al.*, 2007). There are, however, important questions to address such as: (1) why some mutants of this pathway have phenotypes more severe than others? (2) How some unbalanced sterol profiles do not result in visible changes? (3) Do specific proteins need particular interactions to be functional? (Lindsey *et al.*, 2003).

This thesis chapter is focused in the *Arabidopsis* *SQE* family of genes in order to investigate the role of other *SQE* genes other than the previously characterised *SQE1*. Question such as the organ or tissue-specificity, redundancy, regulation, localisation, and their role *in vivo* arise. These questions reveal the necessity of performing further genetic and biochemical analysis in order to understand the role of these genes.

3.2.

*“Functional characterisation of SQE2 and
SQE3”*

3.2.1. Characterisation of the SQE gene family

Squalene epoxidase (monooxygenase) enzymes (SQEs) catalyse the conversion of squalene to 2,3-oxidosqualene, a step of the sterol pathway that is ubiquitous to eukaryotes (Pearson *et al.*, 2003). Squalene epoxidase importance in plants has been demonstrated in recent reports that characterise several *SQE1* mutant alleles (Rasbery *et al.*, 2007; Posé *et al.*, 2009). The *sqe1* mutants develop several defects similar to other mutants of this pathway, including dwarfism, long lifespan, and pale-green leaves. These mutants also present altered root architecture and root hairs, diminished shoot size, and chlorophyll content (Posé *et al.*, 2009). The *sqe1-1* and *sqe1-2* T-DNA insertion lines located upstream the ATG codon did not present any morphological differences. Meanwhile, the null *sqe1-3* and *sqe1-4* alleles with the T-DNA inserted in the 6th and 7th exon, respectively, had severe developmental phenotypes. Although *sqe1-3/sqe1-3* embryos developed normally on a heterozygous parent plant, *sqe1-3/sqe1-3* embryos developing on *sqe1-3/sqe1-3* plants were completely unviable, suggesting that maternal tissue contributes with SQE1 product(s) to developing embryos (Rasbery *et al.* 2007). With studies by Posé *et al.* (2009), a novel role for sterols was proposed, in which sterols would have an essential role in the localisation of NADPH oxidases, which are themselves required for the regulation of reactive oxygen species (ROS), stomatal responses and drought tolerance. This focus and importance of *SQE1* reinforced the need to understand the roles of the other SQEs in Arabidopsis. Rasbery *et al.* (2007) suggested the existence in Arabidopsis of six putative genes encoding SQE proteins based on homology (At1g58440-SQE1, At2g22830-SQE2, At4g37760-SQE3, At5g24140-SQE4, At5g24150-SQE5, At5g24160-SQE6). In the sterol biosynthetic pathway, yeast has been well studied and shares some steps with the plants sterol biosynthetic pathway, although yeast is a diverged phylogenetical organism (Lovato *et al.*, 2000). Yeast mutants affected in each step of the sterol pathway have been engineered and the yeast mutant *erg1* that is deleted for the endogenous SQE gene has no squalene epoxidase gene, and thus cannot grow in a medium lacking exogenous sterols. *erg1* was transformed with all six putative Arabidopsis SQEs. Only SQE1, SQE2 and SQE3 from Arabidopsis were able to complement *erg1*, suggesting that SQE4, SQE5 and SQE6 are not *bona fide* squalene epoxidases (Rasbery *et al.*, 2007). Based on these indications, and the fact that functional characterisation of *SQE1* and a suppressor screening has been taking place within the research group (Posé and Botella, 2009; Posé *et al.*, 2009), a research effort was initiated to functionally characterise remaining SQEs, namely SQE2 and SQE3. To begin with, an updated phylogenetic analysis using SQEs from different organisms was performed. Figure 3.8 shows that plant SQEs

3.2. – Functional characterisation of SQE2 and SQE3

diverge from animal SQEs, and that dicotyledons and monocotyledons essentially form distinct phylogenetic groups.

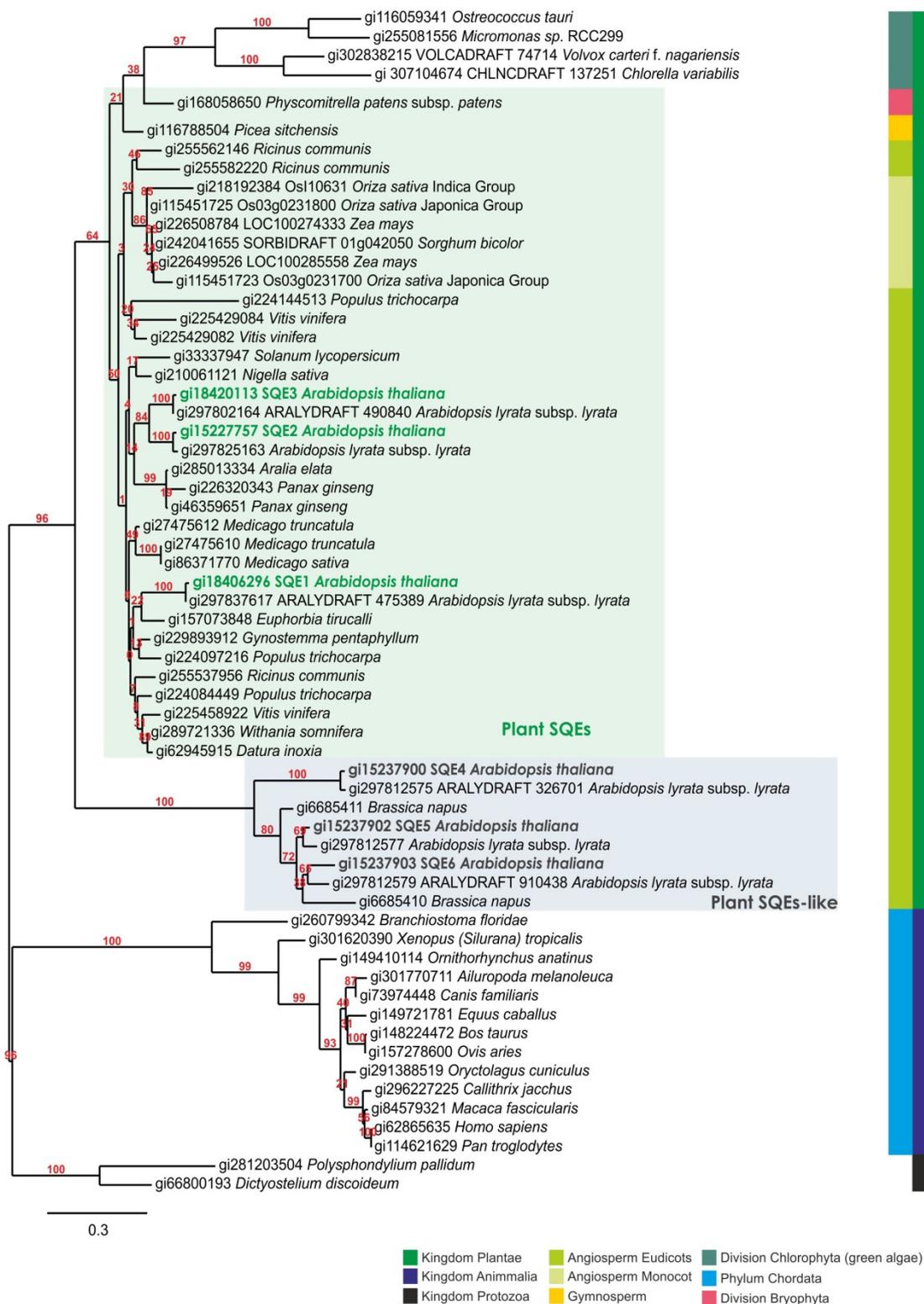


Figure 3.8. – Phylogenetic tree of sequences homologous to squalene monoxygenases/epoxidases. Sequences were retrieved following blastp homology search (NCBI) using the SQE1 a.a. sequence. The MEGA 5 software was used to perform a ClustalW (BLOSUM) alignment of all sequences after exclusion of previously retrieved partial proteins. Phylogenetic reconstruction was performed using Maximum-likelihood with subsequent Bootstrap analysis (500 trees). The rate of amino acid substitution was empirically calculated using the WAG model (URL no.29).

Moreover, a distinct clade is formed by the non functional SQEs, previously designated SQE-like (Rasbery *et al.*, 2007), that includes *Arabidopsis thaliana* SQE4, SQE5 and SQE6, along with *Arabidopsis lyrata* orthologs, as well as members of the Brassicaceae family. SQE1, SQE2 and SQE3 are part of the main clade of plant SQEs identified in this homology search, which are likely to portray the group of functional SQEs. As expected due to their phylogenetic proximity, *Arabidopsis thaliana* SQE1, SQE2 and SQE3 display each an *Arabidopsis lyrata* ortholog.

Arabidopsis SQE-like members (SQE4, SQE5 and SQE6) that did not prove to be squalene epoxidases are in tandem in the genome, which suggests recent duplication events. Gene clusters that have been assembled from plant genes are likely products of gene duplication, neofunctionalisation, and genome reorganisation (Field and Osbourn, 2008). This new function may reside simply in different acceptance of substrates, since the yeast heterologous system indicated that they could not convert squalene to 2,3-oxidosqualene, although this does not discard the possibility that squalene is converted into other substrates, so biochemical analysis of putative substrates or products would be needed. SQE-like enzymes have the conserved Pfam domains of the SQE family, however, they seem to be more specific to certain tissues: SQE4 is very low expressed, SQE5 is more expressed in adult tissues and SQE6 in seeds, hypocotyls and rosettes (Rasbery *et al.*, 2007), suggesting a degree of differential regulation at the tissue level. Their unknown subcellular localisation is predicted *in silico* to be very diverse but still associated to membranous organelles.

Phylogenetic analysis indicates that SQE2 and SQE3 are more related to SQE1 than to SQE4, SQE5 and SQE6. In addition to the phylogenetic studies a topological characterisation of SQE1, SQE2 and SQE3 was performed (Figure 3.9). Interestingly, and based on topological analysis of their sequences, SQE3 and SQE1 might contain putative transmembrane domains that are not identified in SQE2, assuming the *in silico* prediction defined by the TMHMM server 2.0 (Figure 3.9). This may suggest the existence of subfunctionalisation or even nonfunctionalisation (Briggs *et al.*, 2006) of SQE2 after a duplication event that originated SQE2 and SQE3. As shown in figure 3.9, all three encoded proteins possess Pfam domains for squalene epoxidase (SE), FAD dependent oxidoreductase and FAD binding, which are part of the FAD/NAD(P)-binding Rossmann fold Superfamily. This superfamily represents redox enzymes which have a catalytic domain that confers substrate specificity (in these cases the SE domain); and the other is a Rossmann domain, that normally binds to NAD⁺, so that this co-factor reversibly accepts the hydride ion, which is lost or gained by the substrate in the redox reaction.

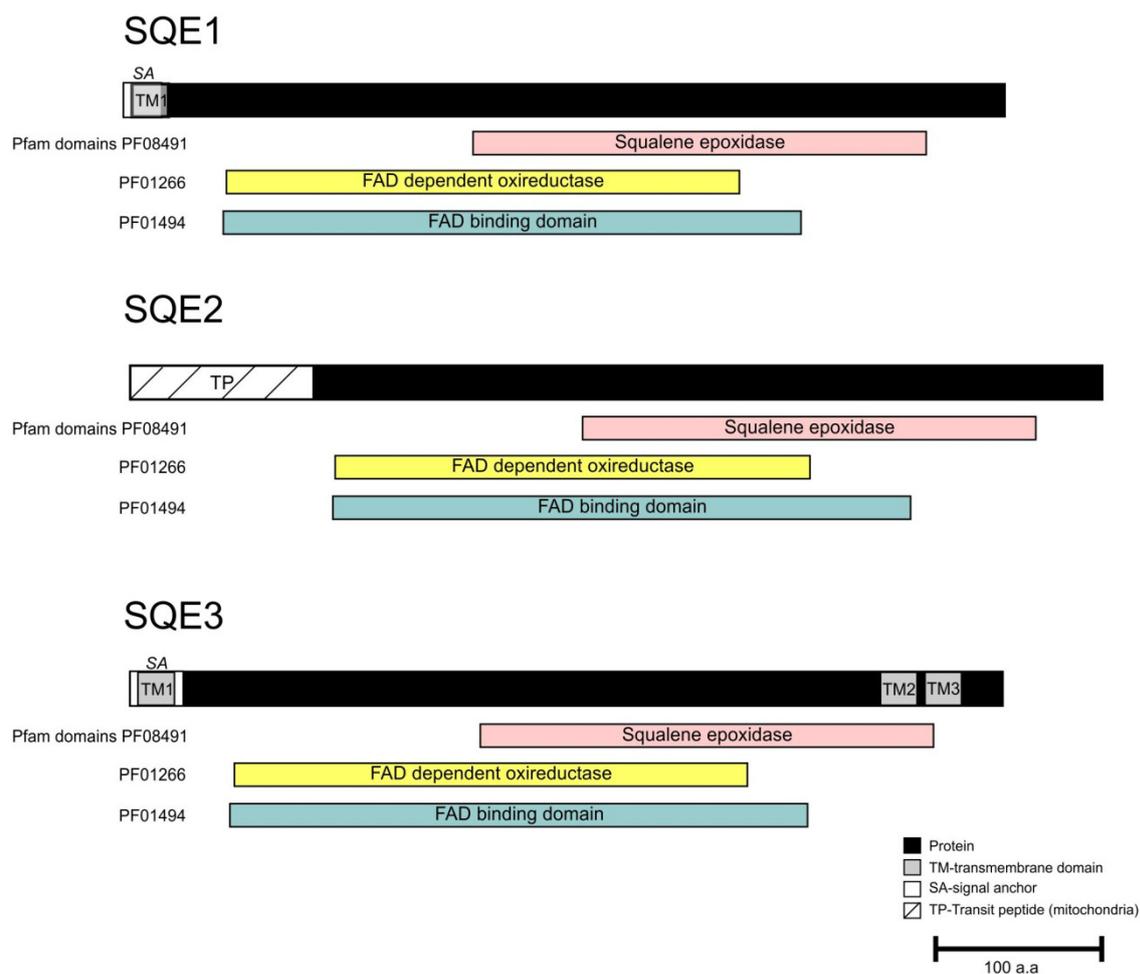
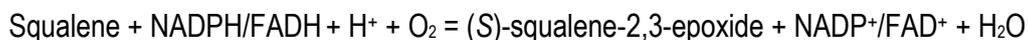


Figure 3.9. – Topological analysis of *Arabidopsis thaliana* SQE family members evidencing squalene epoxidase activity in yeast. Topological analysis depicts putative transmembrane domains, signal anchors, transit peptides and Pfam domains of the SQE1, SQE2 and SQE3 proteins (URL no.20,21,24,30,31).

This Rossmann domain can also bind co-factor FAD, and the other FAD related domains described in the Pfam database (URL no.31). So for squalene epoxidases the stoichiometry reaction would be (URL no.32):



In figure 3.9 it is also depicted the most likely hydrophobic transmembrane domains as well as signal/transit peptides for subcellular targeting according to the Ensembl Plants database (URL no.30) and MitProt database (URL no.24).

Using SUBA database (Heazlewood JL *et al.*, 2007), the number of predictors which putatively identify the subcellular localisation of the proteins was analysed. For SQE1 protein, localizations varied from plastid (MultiLoc), ER (Predotar), mitochondria (SubLoc), plasma membrane (WoLFPSORT) and extracellular (TargetP). SQE2 was predicted to be localised in the

mitochondria (iPSORT, Mitoprot2, MultiLoc, Predotar, SubLoc and TargetP), plastid (LocTree) and plasma membrane (WoLFPSORT). Moreover, SQE3 was predicted to be localised in the plastid (MultiLoc and Predotar), extracellular (LocTree and TargetP), mitochondria (SubLoc) and plasma membrane (WoLFPSORT). All these predictors are used when forming the Cell eFP browser of the BAR database, and so the prediction scores in this database are not very high given the variability in predictions, which may indicate some movement along the secretory pathway.

Models for targeting prediction normally indicate SQE2 as directed to the mitochondria, and the MitProt (Prediction of mitochondrial targeting sequences server) recognise a presence of a transit peptide targeted to the mitochondria with a probability over 0.9, and a cleavage site at the 111 a.a. protein position (URL no.24). For both SQE1 and SQE3 there is more than a 0.9 probability of having a signal anchor rather than a signal peptide, as predicted by SignalP 3.0 server (URL no.21). Since both signals are in transmembrane domains, this likely indicates that they are not cleaved, and remain membrane anchor proteins. The presence of N-terminal signal peptides or analogous transmembrane domains in the protein may represent that they engage the ER-translocation machinery (Bassham *et al.*, 2008). The presence of a target peptide most likely to the mitochondria, and the lack of a transmembrane domain in SQE2 *in silico* predictions, may suggest separate intracellular localisation, a different function or even a nonfunctionalisation of this ortholog (Briggs *et al.*, 2006).

3.2.2. Gene expression analysis

When comparing whole tissue expression of all three SQEs, SQE1 is more expressed in the roots (Posé *et al.*, 2009), SQE3 is expressed highly in almost every tissue except for roots and SQE2 show low expression, as deduced by the gene expression microarray data available for *Arabidopsis* (AtGenExpress/BAR database) (Figure 3.10).

Expression analysis using the available microarray data of SQE2 and SQE3 indicate that SQE3 has a higher expression than SQE2 in most tissues, and its expression is particularly high in mature pollen, the guard cells and the stigma (Figure 3.11). A detailed expression analysis of SQE1 using SQE1 promoter-GUS has already been reported (Posé *et al.*, 2009).

Arabidopsis eFP Browser-Developmental map

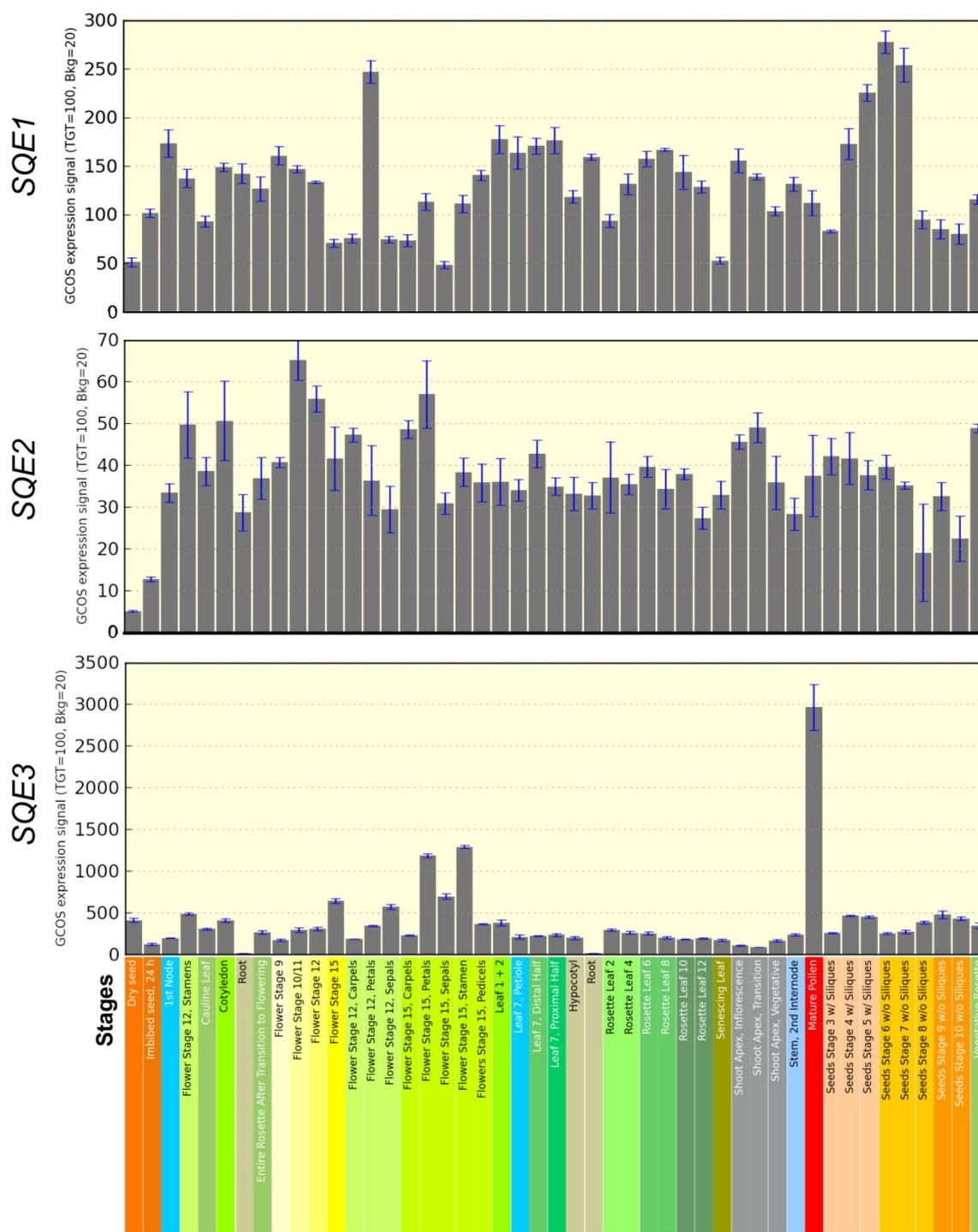


Figure 3.10. – Arabidopsis eFP Browser from BAR database (URL no.10). Developmental expression map of SQE1, SQE2 and SQE3 are depicted. Stages are described at the bottom of the chart (Winter *et al.*, 2007). Note the differences in the scale for the 3 genes. Total expression values and standard deviations are listed in table A.1 (SQE1), A.2 (SQE2), A.3 (SQE3) of the Appendix I.

Therefore in order to analyse spatial expression patterns of *SQE2* and *SQE3* genes at the tissue level, promoter::GUS constructs were generated in a plant expression vector. Stable transformants were obtained for each construct, homozygous lines were isolated after three generations, and GUS histochemical analysis in different tissues was subsequently performed (Figure 3.12). Results confirmed that *SQE3* expression is higher than *SQE2* and is mostly consistent with the microarray data. *SQE2* promoter-driven GUS staining appears in the shoot, and root meristematic areas (in detail in figure 3.12-C), whereas *SQE3* is expressed in the entire seedling, with predominance in the cotyledon (namely the stomata, the vasculature and meristematic tissues). In leaves of 1-month-old adult plants there is no GUS staining in transgenics for both constructs. In pro*SQE3*::GUS plants, staining is observed in silique edges and flowers, particularly sepals, tips of the stigma, anthers and filaments of the stamen.

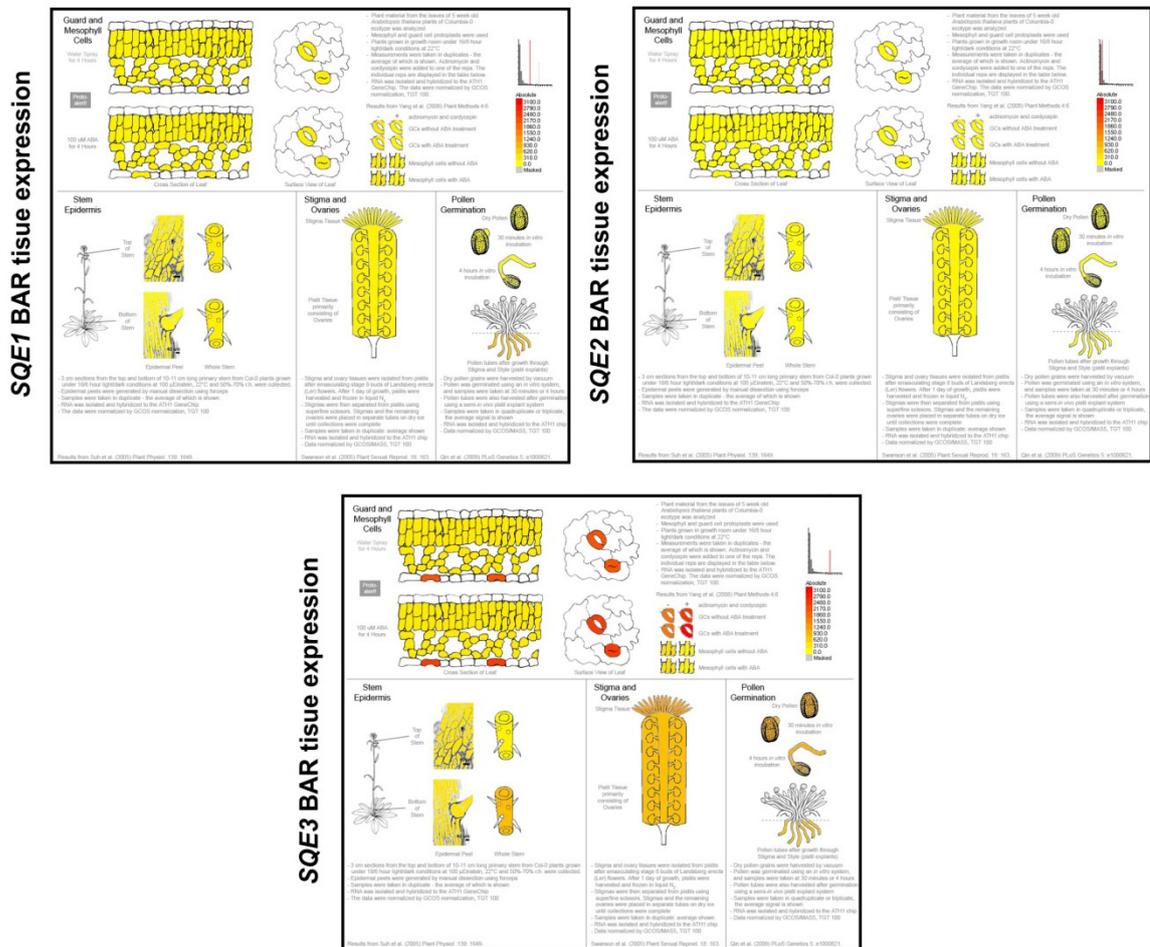


Figure 3.11. – Arabidopsis eFP Browser from BAR database (URL no.10). Tissue expression map of *SQE1*, *SQE2* and *SQE3* are depicted. Colorimetric scale was adjusted with a threshold so to be compared (Winter *et al.*, 2007). Total expression values and standard deviations are listed in table A.1 (*SQE1*), A.2 (*SQE2*), A.3 (*SQE3*) of the Appendix I.

3.2. – Functional characterisation of *SQE2* and *SQE3*

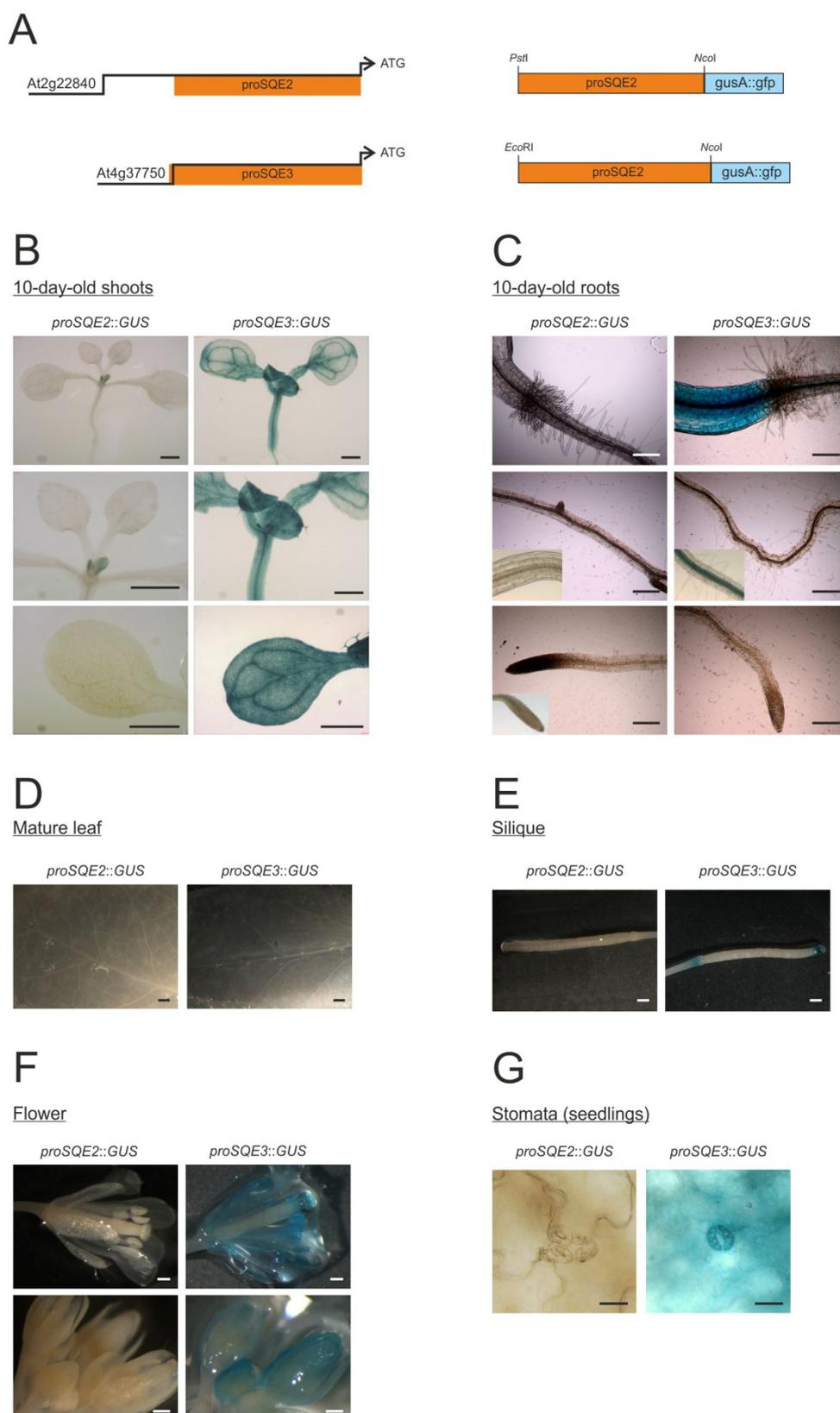


Figure 3.12. – Histochemical analyses of GUS activity in *proSQE2::GUS* and *proSQE3::GUS* plants. A – Schematic representation of the constructs *proSQE2::GUS* and *proSQE3::GUS* in pCambia1303. B – Shoots of 10-day-old seedlings. Scale bar represents 1 mm. C - Roots of 10-days-old seedlings. Scale bar represents 0.5 mm. D – Mature leaves of 1-month-old plants. Scale bar represents 1 mm E – Siliques of 1-month-old plants. Scale bar represents 1 mm F- Flowers of 1-month-old plants. Scale bar represents 1 mm G- Stomata detail of 10-day-old seedlings. Scale bar represents 20 μ m.

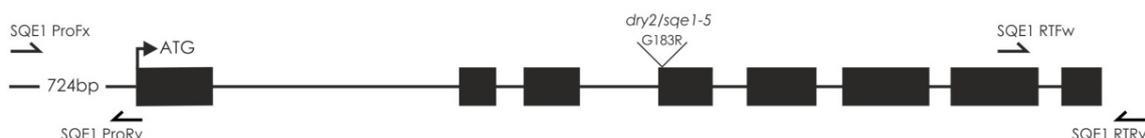
3.2.3. Isolation of *sqe2* and *sqe3* loss-of-function mutants

To functionally characterise *SQE2* and *SQE3*, a reverse genetic approach was conducted in order to gain insight into their function. This genotype to phenotype strategy goes in opposite ways of forward genetics, example of the EMS point mutant *dry2/sqe1-5* (Posé *et al.*, 2009), in which a mutant with a specific phenotype is isolated, and subsequently map-based-cloning is conducted in order to identify the gene responsible (Alonso and Ecker, 2006). Reverse genetics is becoming a common strategy since there is a large availability of T-DNA and transposons insertion mutants. In fact, insertions have been identified for 27543 of the 28691 genes in Arabidopsis (MASC Report, 2010). The localisation of the T-DNA insertion or of the transposon may induce a knockout, a knockdown, or in some cases a knockon, that triggers an ectopic or increased expression of the gene, if for example it is located in the promoter (Feng and Mundy, 2006).

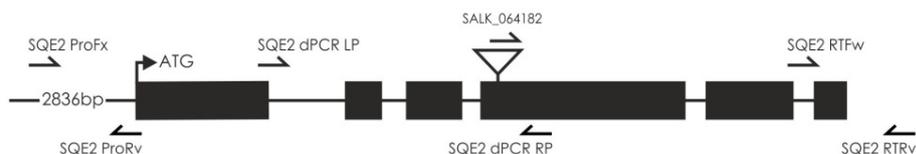
Lines were searched in T-DNA Express: Arabidopsis Gene Mapping Tool (URL no.16) and subsequently ordered from TAIR for insertions in both *SQE2* and *SQE3*. For selecting the mutant line, it was taken into consideration the ecotype (preferably Col-0) and the location of the T-DNA, aiming to avoid promoter and intron regions in order to increase the probability of a knockout. For the case of *SQE3*, with no alternatives, a T-DNA line in a 4th intron was selected. However, intron insertions in most cases do not produce a viable protein as the T-DNA insert is too big to be correctly spliced.

Schematic representations of *SQE1*, *SQE2*, and *SQE3* are shown in figure 3.13, in which ordered T-DNA lines and other aspects that are considered relevant to the experimental work and the discussion are also depicted. *SQE1* contains eight exons, while *SQE2* and *SQE3* have six. The *SQE1* gene was already functionally characterised using the *dry2/sqe1-5* mutant (Posé *et al.*, 2009), and this mutant was the one used in the present work, as a mutant for *SQE1*. This EMS mutant has a point mutation in the 4th exon that alters the amino acid in the 183 position, from a glycine (G), a nonpolar neutral amino acid, to an arginine (R), a nonpolar, positively charged and hydrophilic amino acid (Figure 3.13). This position is consistent with the FAD dependent oxidoreductase and FAD binding domains location within the protein, being so fundamental for the function of the enzyme. This aminoacidic change generates pleiotropic phenotypes previously described (Posé *et al.*, 2009). The main advantage of this mutant is that it is fertile, therefore allowing easier manipulation.

SQE1



SQE2



SQE3

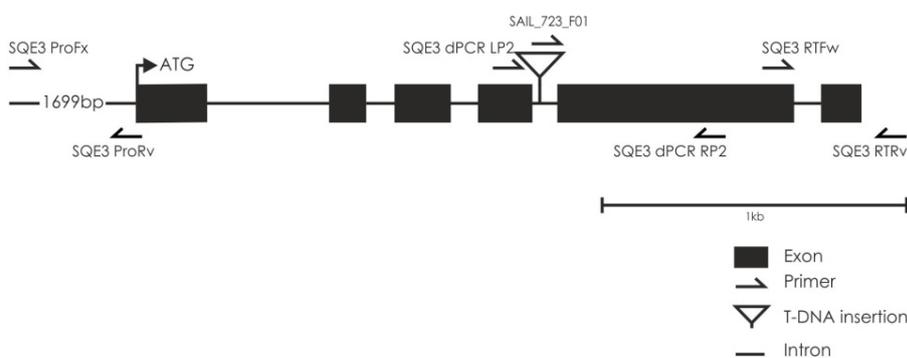


Figure 3.13. – Schematic representation of squalene epoxidase genes *SQE1* (At1g58440), *SQE2* (At2g22830) and *SQE3* (At4g37760). Disruption by the T-DNA occurs downstream of the ATG codon at 1179 and 1343 bp for *SQE2* and *SQE3*, respectively. The positions of gene-specific PCR primers used for diagnostic-PCR genotyping of the T-DNA insertion (SQEx dPCR LP and RP) and RT-PCR analysis (SQEx RTFw and Rv) are represented by arrows. Exons are represented by dark boxes, introns by dark lines and the T-DNA insertion by a triangle. Upstream of the ATG is the indication of the estimated promoter size, as predicted by The Arabidopsis Gene Regulatory Information Server (AGRIS) database (URL no.14) (Davuluri *et al.*, 2003).

For *SQE2* it was selected a T-DNA line from the SALK collection, located in the 4th exon, likely to generate a gene knockout. A diagnostic PCR was performed in order to find homozygous plants for the insertion allele (Figure 3.14-A). In light of the results, it was demonstrated that all plants genotyped (designated by coordinates C4, C5 and so forth) contained the insertion (LB-RP reaction), and could not amplify the genomic segment of 995 bp (LP-RP), because the T-DNA insertion was present in homozygosis. Columbia served as the control, and as expected no T-DNA were amplified, and only amplifications with the LP-RP primers were obtained. A SAIL line was selected for *SQE3* in which the T-DNA is located in the 4th intron. Plants were genotyped by diagnostic PCR. Since the seeds are of a mixed T2 population, the presence of different genotypes

for the T-DNA insertion was expected. LB-RP primers produced amplification in all plants with the exception of Col-0 as shown in figure 3.14-B, indicating the presence of the T-DNA in all these lines. However, when primers for amplifying the gene (LP-RP) were used, the pattern of the bands was difficult to interpret. For that reason, PCR products were analysed with better resolution in a new gel (highlighted in the figure 3.14-B). In the control (Columbia) there are two visible amplification products. Since the one of higher molecular weight is the expected (1064 bp), the lower one was considered to be of low primer specificity, so the plant G1 was considered to be an homozygous plant for the T-DNA. The higher weight band is more intense in the control likely because of primer competition between two possible amplification products, while in a sample where there is no possible LP-RP amplification, competition does not exist and the lower band (unspecific) is more intense as is the case of G1.

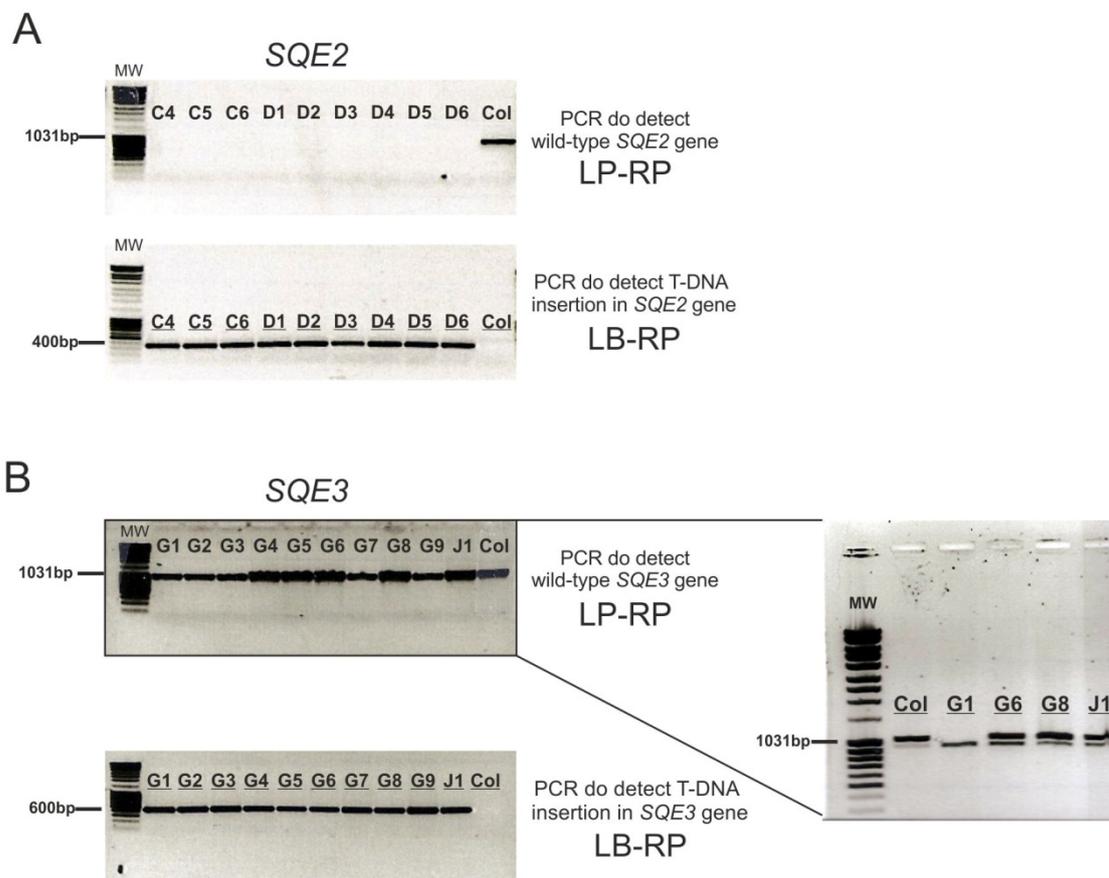


Figure 3.14. – Diagnostic PCR genotyping of *sqe2-1* (SALK_064182Hm) and *sqe3-1* (SAIL_723_F01) insertion mutants from a heterogeneous mutant population. Electrophoretic analysis of *SQE2* (A) and *SQE3* (B). Selected samples from *SQE3* analysis LP-RP were better resolved to highlight differences between closer sized fragments. LP and RP - left and right primers, respectively (as indicated in figure 3.13); LB - left border specific for the T-DNA used (SALK – LBb1; SAIL – LB3). MW - Molecular Marker MassRuler DNA Ladder Mix.

After a blast search of the LP and RP primer, it was determined that the primer pair could overlap a very high conserved region within *SQE2* gene giving therefore the possibility of an unspecific

amplification of *SQE2* on putative *sqe3* mutants. Hereupon new primers LP2 and RP2, depicted in figure 3.13, were designed and used for further analysis (e.g. *SQE3* genotypic analysis while generating double mutants involving *sqe3-1* mutant) and presented specificity for the genotypic analysis.

RT-PCR was performed to ensure the knockout of both lines and a newly generated *sqe2-1/sqe3-1* double mutant (Figure 3.15). Primers used for RT-PCR analysis are depicted in figure 3.13 (*SQExRTFw* and *Rv*). The three pair of primers were designed so that an intron was within the primer amplicon to determine possible gDNA contamination in the samples. They were also designed at the non-conserved 3'UTR of *SQE1*, *SQE2* and *SQE3* (Posé et al., 2009). Both *SQE2* and *SQE3* have amplicons of 320bp, while *SQE1* is of 329bp (Table 2.9, section 2.3.1). Amplification cycles applied to the RT-PCR reaction varied from 30 for *SQE1* and *SQE3*, and 34 for *SQE2* due to its low expression. Being a semi-quantitative analysis, expression must be taken with caution. Nevertheless, looking at these results we can infer that in 1-month-old leaves *SQE3* is the most expressed, followed by *SQE1*, and being *SQE2* the less expressed. Most importantly, all mutants analysed are most likely knockouts, since there is no detectable expression of the genes in their corresponding mutants (Figure 3.15).

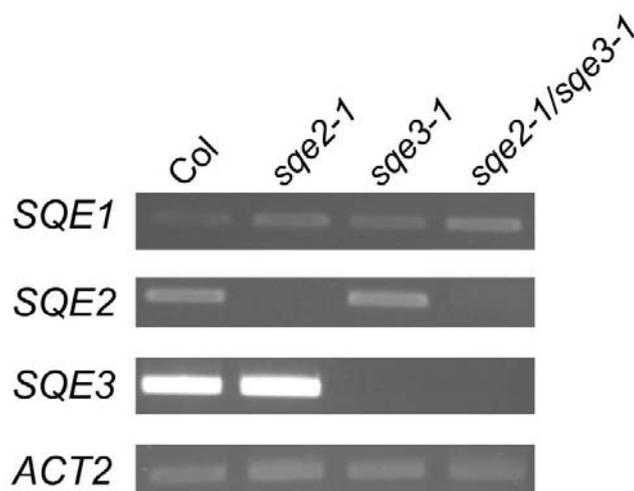


Figure 3.15. – Semi-quantitative RT-PCR analysis of *sqe2-1* and *sqe3-1* single and *sqe2-1/sqe3-1* double mutants, with background Col-0 as a control. *SQE1*, 2 and 3 expression was analysed using specific primers (Figure 3.13) and *ACT2* as a control gene displaying constitutive expression.

3.2.4. Developmental and terbinafine sensitivity phenotypes of *sqe2-1* and *sqe3-1*

In order to functionally characterise *SQE2* and *SQE3*, the knockout mutant lines for both genes were subjected to a morphological analysis and phenotypic characterisation. Analysis of

sqe2-1 and *sqe3-1* indicates the absence of visible developmental differences to wild-type in standard growing conditions (exemplified in figure 3.16), as well as for the double *sqe2-1/sqe3-1* mutant. As shown, there are no differences in shape and size in 7-day-old seedlings (Figure 3.16-A), or one month-old plants (Figure 3.16-B), when comparing to wild-type plants.

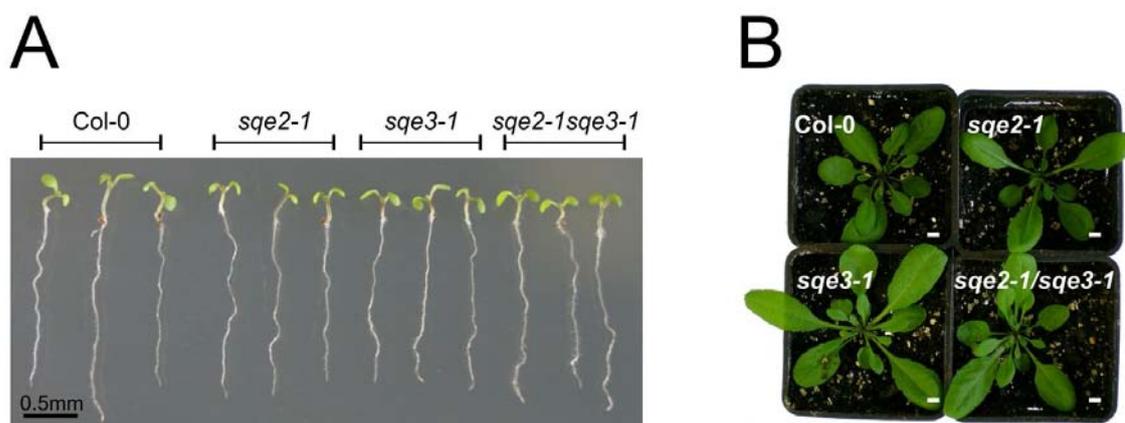


Figure 3.16. – Morphological characterisation of *sqe2-1*, *sqe3-1* and *sqe2-1/sqe3-1* mutants and their wild-type ecotype Col-0. A - 7-day-old seedlings. B - 1-month-old plants. (Bars indicate 0.5 mm).

SQE1, *SQE2* and *SQE3* genes were shown to encode squalene epoxidases based on complementation in yeast (Rasbery *et al.*, 2007). *dry2/sqe1-5* was found to be hypersensitive to terbinafine, an inhibitor of squalene epoxidase activity, which is consistent with a decreased squalene epoxidase activity in this mutant (Posé *et al.*, 2009). This inhibitor belongs to the class of allylamines and is a specific non-competitive inhibitor of fungal and plant SQEs (Ryder, 1992; Nieto *et al.*, 2009). Wentzinger *et al.* (2002) demonstrated its effects on tobacco cells, and later Nieto *et al.* (2009) in Arabidopsis. Therefore we analysed the effect of terbinafine on *sqe2-1* and *sqe3-1* seedlings by germinating them in solid MS-medium supplemented with increasing concentrations of terbinafine and calculating the percentage of seedlings that developed true leaves after three weeks. As shown in figure 3.17, *sqe3-1* mutants exhibited an increased sensitivity to terbinafine relative to the wild-type Col-0, while *sqe2-1* seedlings did not show significant differences. These results support a role for *SQE3* *in planta*, however the absence of visible morphological differences during normal development and growth for both *sqe2-1* and *sqe3-1*. This suggests that *SQE2* and *SQE3* do not have an essential role for normal plant growth but may have a role under particular conditions. An example of this is the HMGR (3-hydroxy-3-methyl-glutaryl-CoA reductase) activity in Arabidopsis. HMGR is encoded by two genes *HMG1* and *HMG2* in the Arabidopsis genome (Ohyama *et al.*, 2007). These enzymes are responsible for the conversion of HMG-CoA into mevalonate, in a rate-limiting step of the mevalonate (MVA) pathway. Experiments with mutants for

the two genes (*HMG1* and *HMG2*) elucidated a greater role than the one expected for HMGR2, despite the absence of phenotype in normal conditions. *hmg2-1* mutants showed high sensitivity to lovastatin, an inhibitor for HMGR activity (Suzuki *et al.*, 2004). Further analysis showed that the *hmg2-1* mutant contain 15% less sterol content than its wild-type, highlighting an important role for HMGR2, despite the lack of visible phenotype in the mutant (Ohyama *et al.*, 2007). The terbinafine sensitivity analysis of *sqe1-2* and *sqe1-3* suggest an important role for SQE3 and a marginal specific role for SQE2 in squalene epoxidase activity.

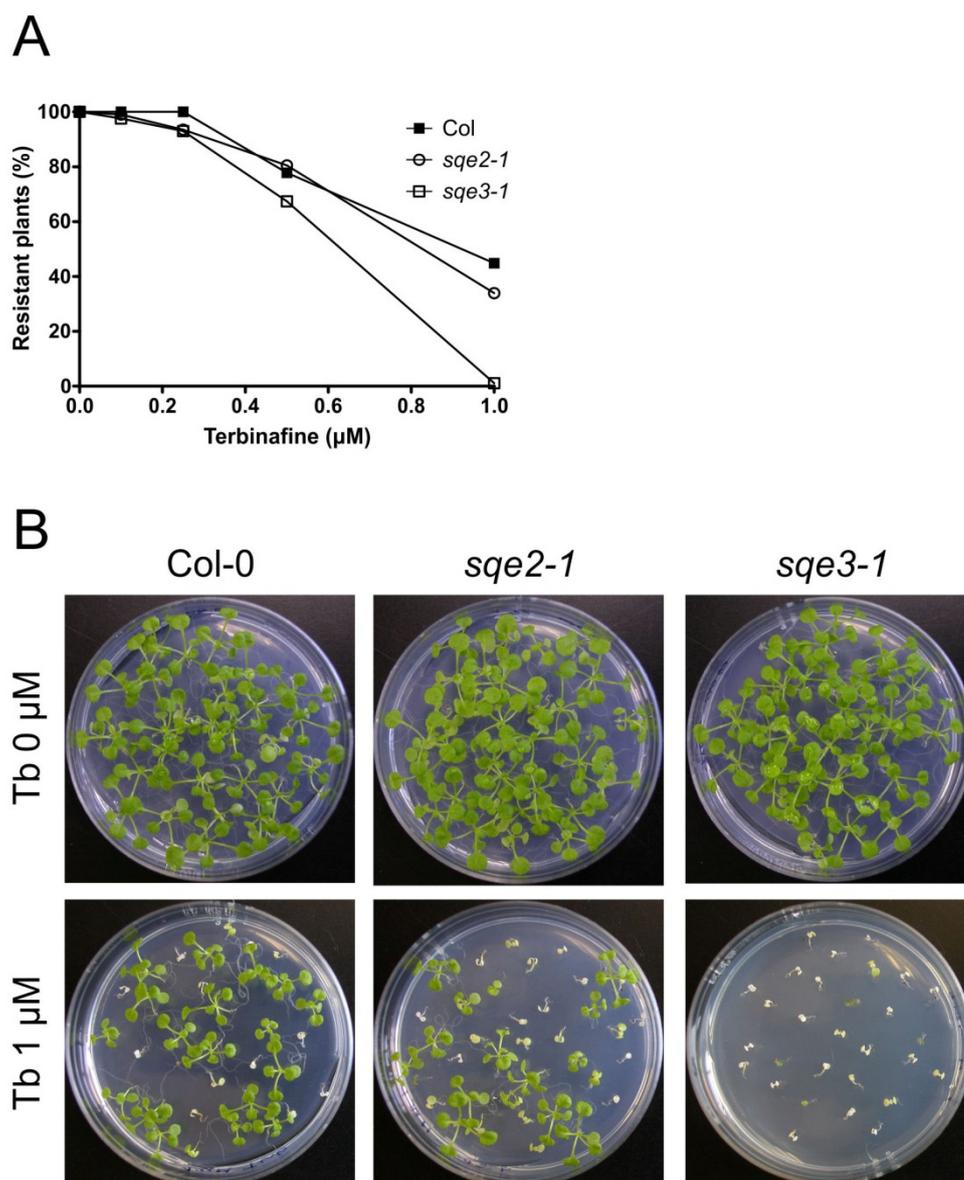


Figure 3.17. – The *sqe3-1* mutant displays a terbinafine sensitivity phenotype. Seedling viability of Col-0, *sqe2-1* and *sqe3-1* was evaluated *in vitro* in the presence of different concentrations of terbinafine (Tb), a known inhibitor of squalene epoxidase activity. Resistant plants represent the ones that developed green leaves.

3.2.5. Defective sterol profile of the *sqe3-1* mutant

In order to investigate a role for *SQE3* in sterol biosynthesis as suggested by the increased sensitivity to terbinafine, sterol content was analysed in *sqe3-1*. Squalene and free sterols were quantified in leaves of 1-month-old *SQE* mutants (*dry2/sqe1-5*, *sqe2-1* and *sqe3-1*), as well as their corresponding wild-types (Ler for *dry2/sqe1-5*, and Col-0 for *sqe2-1* and *sqe3-1*). *dry2/sqe1-5* has reduced quantities of sterols, with the exception of sitosterol, when compared with its wild-type Ler (Figure 3.18-A). Moreover, differences on *sqe2-1* and *sqe3-1* mutants are only marginally evident for *sqe2-1* in which stigmasterol, an interconvertible sitosterol, is marginally higher. Overall, these results indicate no significant changes in the sterol profile for *sqe2-1* and *sqe3-1*, at this developmental stage.

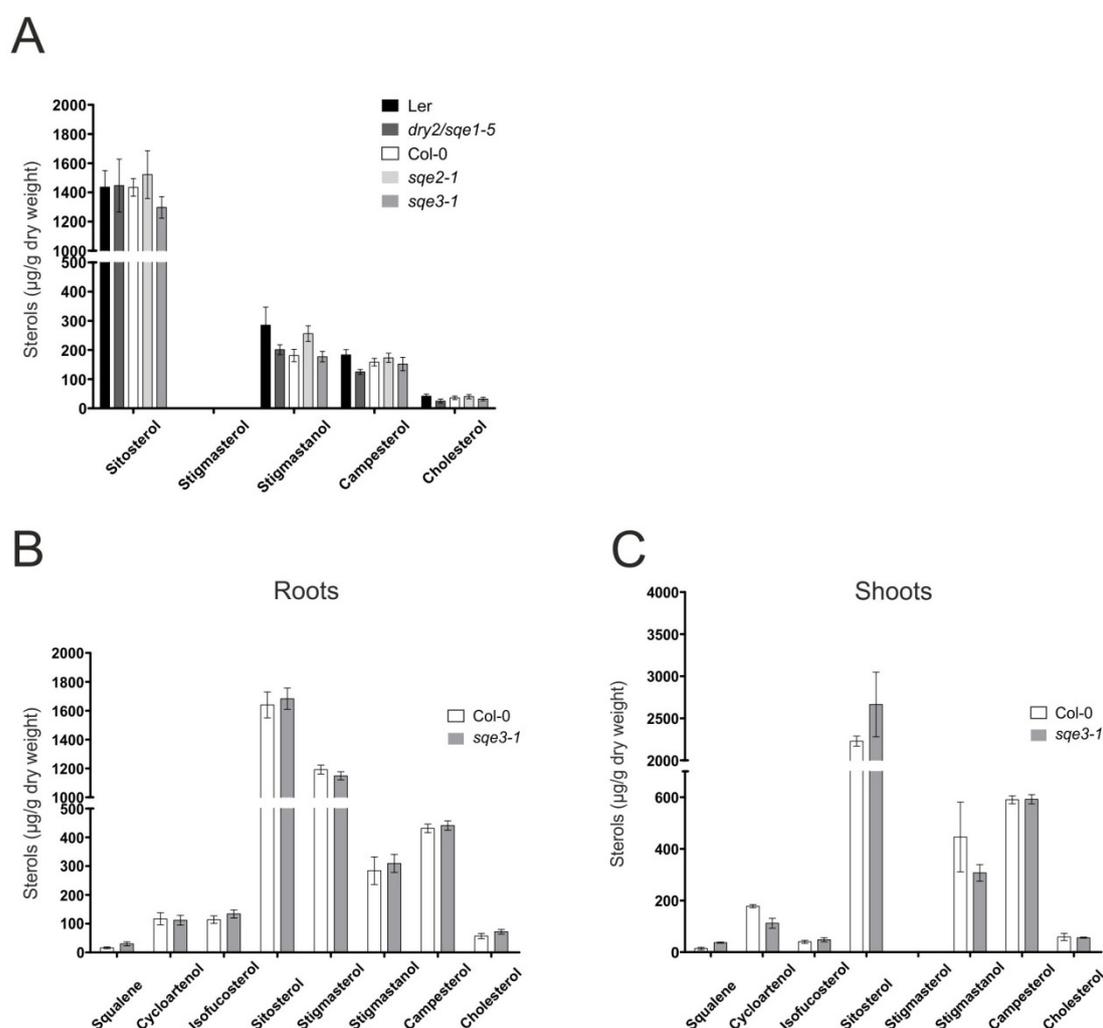


Figure 3.18. – Analysis of the profile of squalene and a few sterols. Content of individual sterols in 1-month leaves in Ler, *dry2/sqe1-5*, Col-0, *sqe2-1* and *sqe3-1* genotypes (A), 14-day-old roots (B) and 14-day-old shoots (C) for Col-0 and *sqe3-1* genotypes. Error bars depicts means ± SEM (N=3 for all except Col-0 in figure B which N=2). Total expression values and standard deviations are listed in table A.8 (A), A.9 (B), A.10 (C) of the Appendix II.

SQE2 and SQE3 expression was not detected in 1-month-old plants based on promoter::GUS staining (Figure 3.12). Based on SQE3 gene expression, that evidences a higher expression in cotyledons and young leaves (see Figure 3.12-B) the profile of squalene and sterols was evaluated in both shoots and roots of 14-day-old *sqe3-1* mutant seedlings as well as squalene, the predicted substrate for SQE3 (Figure 3.18-B,C). The number of sterols that can be detected in seedlings is significantly higher than in 1-month-old leaves. *sqe3-1* roots accumulate ~2 times more squalene, consistent with a reduction of SQE activity, while the other sterols analysed in this tissue remained mainly unaltered, except for a little increase in cholesterol and isofucosterol (Figure 3.19). Previous studies supported that SQE1 has a main role in channelling squalene towards sterol biosynthesis in the root, where the sterol profile in this organ in both *dry2/sqe1-5* mutant (Posé *et al.*, 2009) and *sqe1-3* (Rasbery *et al.*, 2007) is significantly altered. *dry2/sqe1-5* accumulates some intermediate compounds such cycloartenol and campesterol, but contain less end-pathway sterols such as sitosterol and stigmasterol. However, no differences in sterol composition or squalene were found in the shoot, (Posé *et al.*, 2009) supporting a main role in roots as previously indicated. Interestingly, in addition to the squalene accumulation in the *sqe3-1* shoots tissue, there are some changes in sterols such as a reduction of cycloartenol (~1-2), the product of the action of the cycloartenol cyclase on 2,3 oxidosqualene, a slight increase in the main sterol sitosterol, and a decrease in stigmastanol (Figure 3.20). Taking together the analysis of sterol content of *dry2/sqe1-5* and *sqe3-1* mutants, it is clear that the regulation is very complex and probably compensation in the activity among SQE genes can be envisaged. It also opens the possibility that in addition to 2,3-oxidosqualene, SQE3 is converting squalene to a different product.

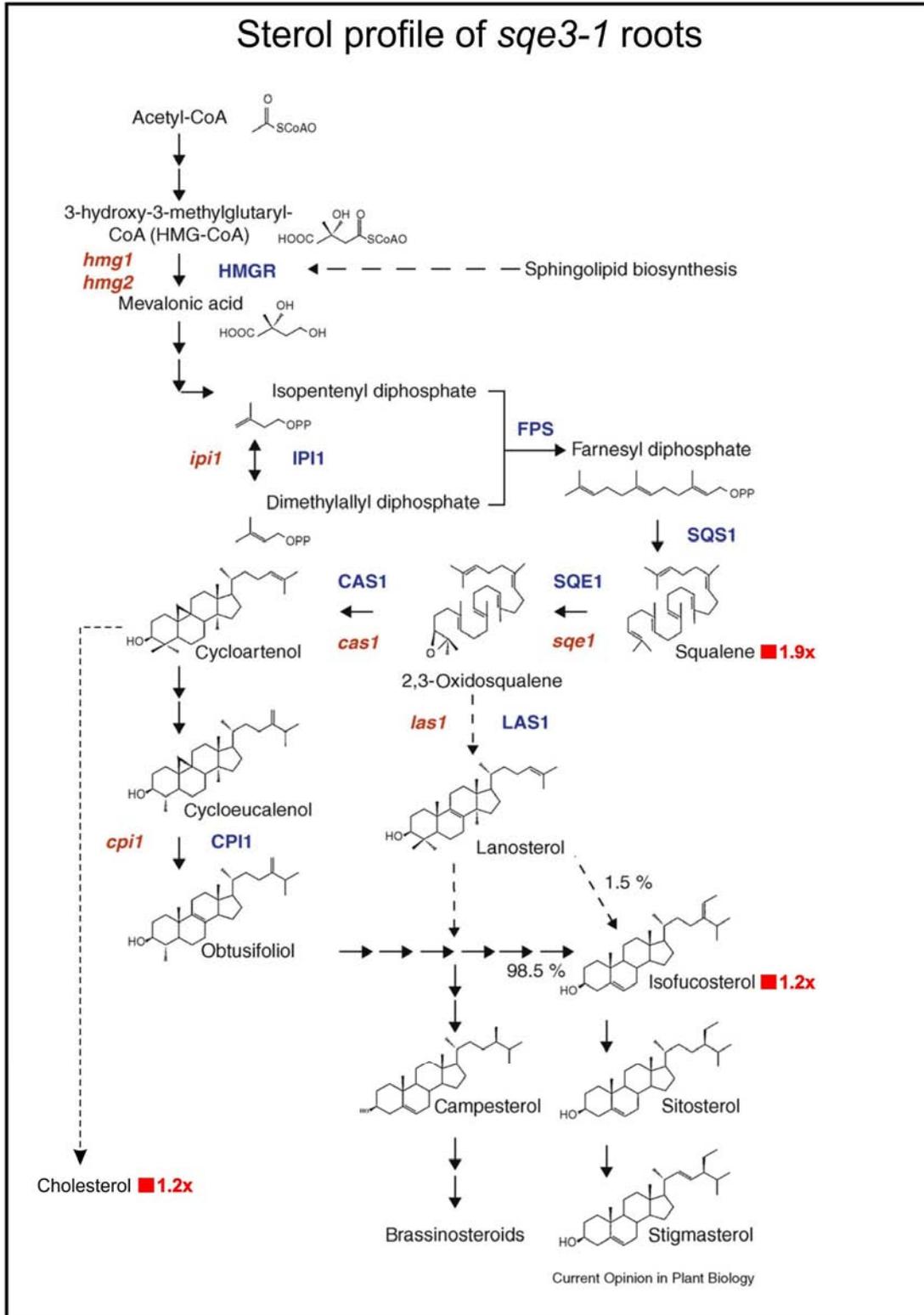


Figure 3.19. – Sterol biosynthetic pathway depicting changes in squalene and sterols in *sqe3-1* 14-day-old roots relative to Col-0 data (based on data from Figure 3.18-B). Up-regulations are in red while down-regulations are by definition in green. Adapted from Boutté and Grebe (2009).

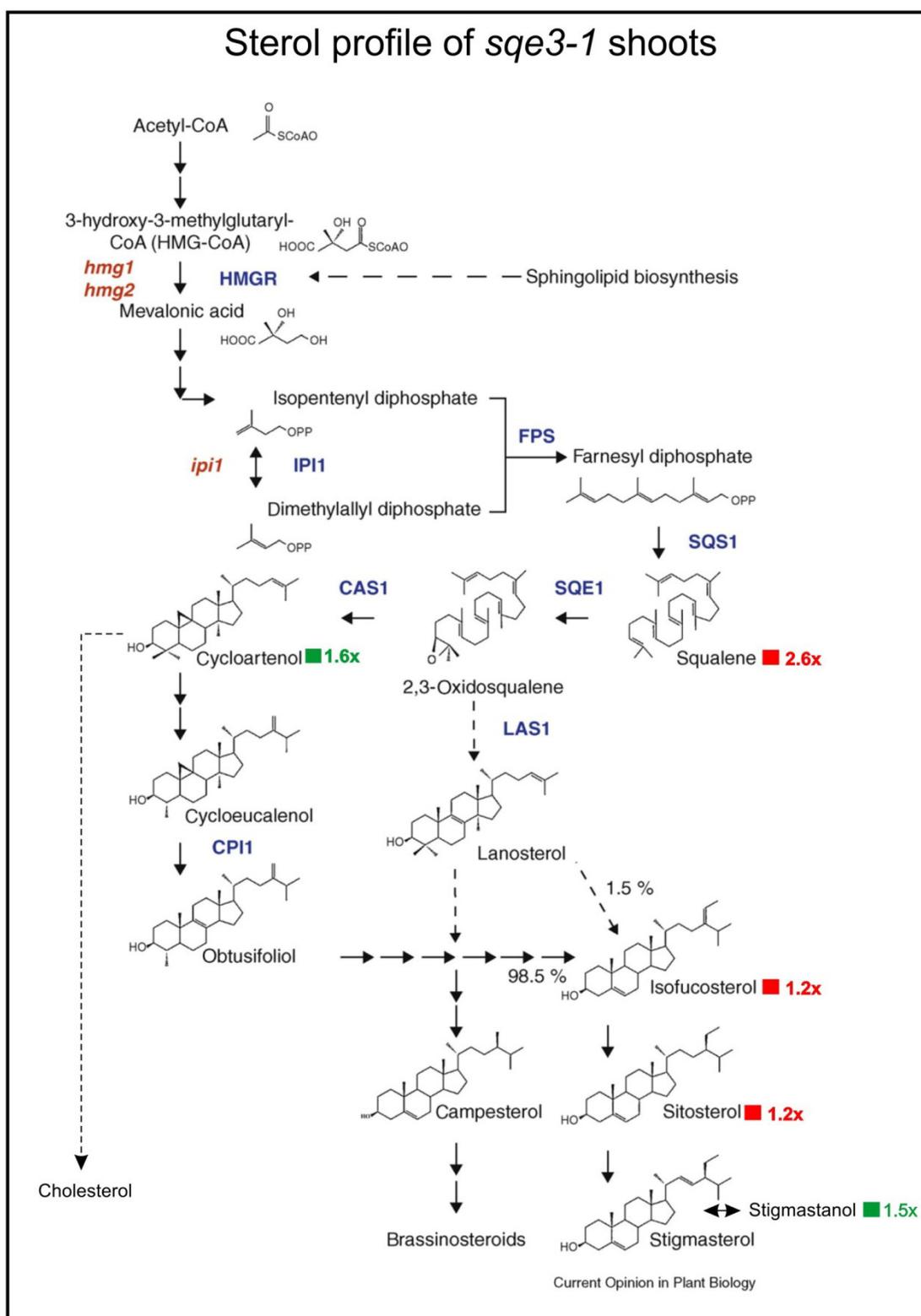


Figure 3.20. – Sterol biosynthetic pathway depicting changes in squalene and sterols in *sqe3-1* 14-day-old shoots relative to Col-0 data (based on data from Figure 3.18-C). Up-regulations are in red while down-regulations are by definition in green. Adapted from Boutté and Grebe (2009).

3.2.6. *SQE2* and *SQE3* do not complement *dry2/sqe1-5* mutants

SQE1, *SQE2* and *SQE3* are capable to convert squalene to 2,3-oxidosqualene, based on yeast complementation. However, this is the result of overexpression, which makes difficult to predict that the real function *in planta* is to epoxidase squalene to 2,3-oxidosqualene. One possibility is that *SQE3* provides the main SQE activity in shoots and *SQE1* provides the main SQE activity in roots, and that this is solely due to their different expression pattern. In order to address this possibility, a promoter-swap strategy was developed to determine whether *SQE2* or *SQE3* could functionally complement *dry2/sqe1-5* under the control of the *SQE1* promoter (Figure 3.21). Two constructs, *proSQE1::SQE2* and *proSQE1::SQE3* were generated and subsequently transformed in the *dry2/sqe1-5* background, with at least three independent transformants being obtained. The promoter size was identical to that previously used to complement *dry2/sqe1-5* with the *SQE1* gene (Posé *et al.*, 2009).

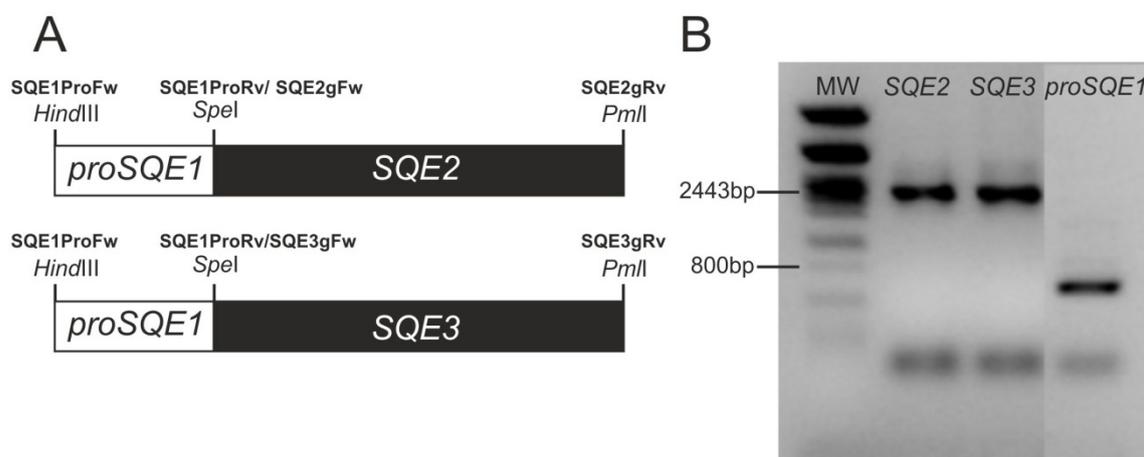
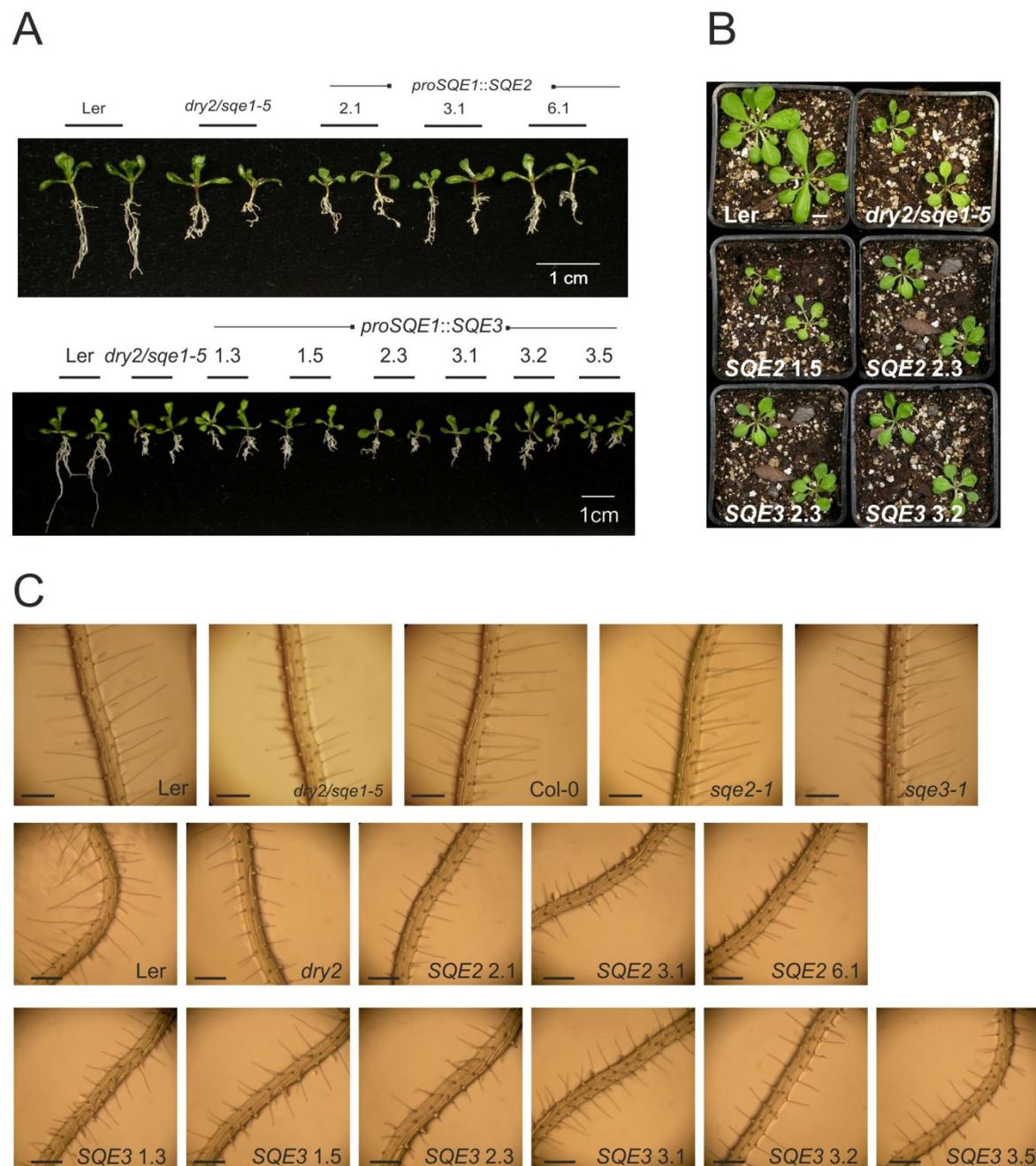


Figure 3.21. – Promoter swap constructs. A - *proSQE1* was fused to *SQE2* or *SQE3* genomic sequence in the pCAMBIA1303 vector using the primers and restriction enzymes indicated (see Table 2.9, section 2.3.1, for more information on the primers). B – Amplification of both genes and the promoter for construct purposes. *SQE1* promoter was amplified with *SQE1ProFw* and *SQE1ProRv*, while *SQE2* and *SQE3* genomic sequences were amplified with *SQE2gFw* and *SQE2gRv* as indicated in A. MW – Molecular Weight Marker (λ DNA digested with *Pst*I).

Plants were then analysed for those phenotypes that were shown to be defective in *dry2/sqe1-5* (Posé *et al.*, 2009). The *dry2/sqe1-5* mutant is characteristic for having smaller pale green leaves with reduced chlorophyll content, branched and small roots, and prolonged life span. Independent *dry2/sqe1-5* transgenic lines with both *SQE2* and *SQE3* were morphologically undistinguishable of *dry2/sqe1-5* (Figure 3.22-A,B) indicating a lack of complementation at first glance. A specific phenotype of *dry2/sqe1-5* is the smaller size of root hairs, which is due to a

de-localization of RHD2 NADPH oxidase (Posé *et al.*, 2009). Therefore, root hairs of the complemented lines were investigated using a stereomicroscope, and subsequently analysed using ImageJ software. Results depicted in figure 3.22-C,D show that all complemented lines have a distribution of root hair size almost identical to that of *dry2/sqe1-5*, suggesting that SQE2 and SQE3 cannot functionally complement SQE1. Moreover, *sqe2-1* and *sqe3-1* do not show abnormal root hairs when compared to the wild-type Col-0 (Figure 3.22-C). Taking together, these results indicate that SQE2 or SQE3 cannot restore *dry2/sqe1-5* defects independent of their expression patterns.



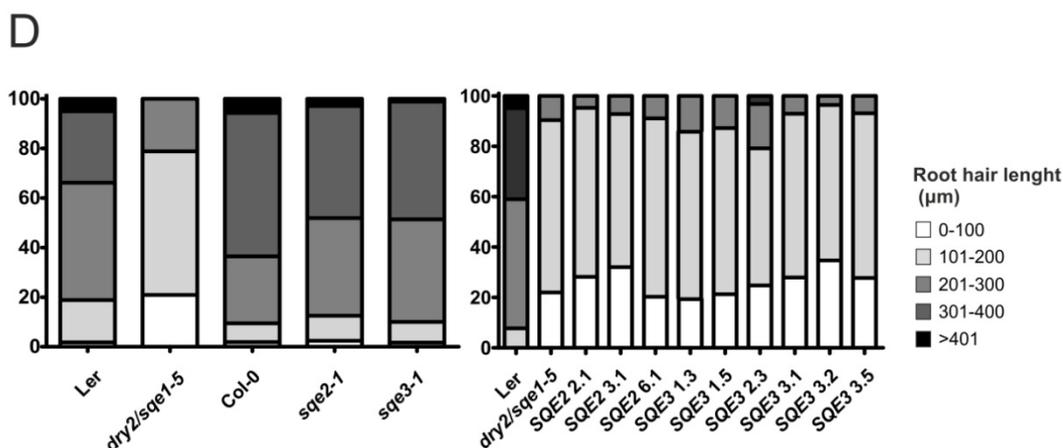


Figure 3.22. – Phenotype characterisation of promoter swap constructs following permanent plant transformation. Constructs were used to transform the *dry2/sqe1-5* background and homozygous T2 plants were subsequently obtained. Morphology of 13-day-old seedlings (A) and 1-month-old plants (B) (Scale bar represents 1 cm). C - Phenotypic analysis of root hairs, including *sqe2-1* and *sqe3-1* mutants (Scale bar represents 200 µm). D - Quantification of root hair length.

3.2.7. Subcellular localisation of SQE1 and SQE3

The finding that SQE2 and SQE3 cannot complement SQE1 function despite being driven by the same promoter suggests that either the protein is located in different cellular locations or that it cannot interact *in planta* with SQE1 partners. *In silico* prediction indicate that SQE proteins might have important differences in localisation, i.e. SQE1 has one putative transmembrane domain, SQE3 has three and SQE2 none (see figure 3.9). Since SQE2 is predicted to be a soluble protein and possibly likely targeted to the mitochondria, is possible that has a different role altogether in this other organelle, as sterol biosynthesis have been shown to occur mainly at the endoplasmic reticulum (Benveniste, 2004). We then performed subcellular localisation for SQE1 and SQE3 by generating an in frame fusion of the protein with GFP using Gateway technology. The constructs 35S::GFP-SQE1 and 35S::SQE3-GFP were used to transiently transform Arabidopsis cotyledons followed by confocal microscopy analysis (Marion *et al.*, 2008). As shown in figure 3.23, strong GFP fluorescence was observed in vesicles, evenly spread throughout the cell. In the SQE3-GFP fusion, the GFP is concentrated in spots, particularly larger, and in more quantity than in the GFP-SQE1 fusion, where the fluorescence seems to accumulate near the periplasmic space. Since both constructs tested have GFP at different terminus comparison is not feasible, however, stable transformations in Arabidopsis are underway. Overall results seem to suggest a targeting of the protein to the ER, Golgi or some component of the secretory pathway, however, the different pattern suggest different subcellular localisation.

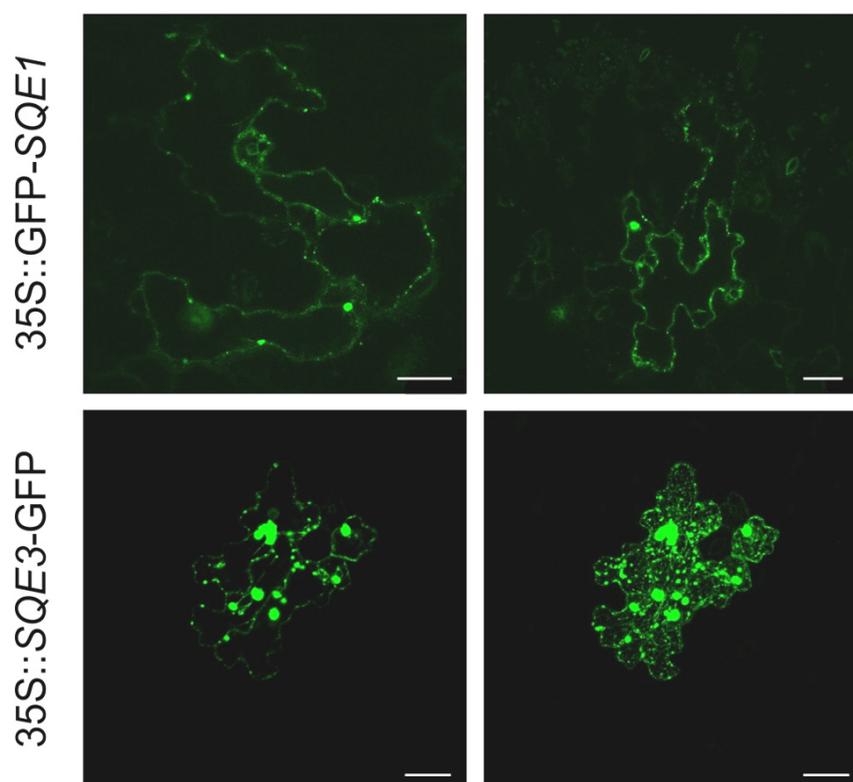


Figure 3.23. – Subcellular localisation of SQE1 and SQE3 proteins. Transient expression of 35S::GFP-SQE1, 35S::SQE3-GFP in *Arabidopsis thaliana* cotyledons. *Arabidopsis* seedlings were transformed with the GFP construct using pMDC's Gateway destination vectors (Curtis and Grossniklaus, 2003). Scale bar represents 25 μ m. Confocal analysis was performed with GFP excitation at 488 nm, and emission at 492-587 nm.

In order to better compare SQE1 and SQE3 localisation and to resolve subcellular targeting we analysed the constructs 35S::GFP-SQE1, 35S::SQE3-GFP, and a newly generated 35S::GFP-SQE3 for a transiently co-transformation in tobacco leaves with an mCherry ER marker (Figure 3.24) or a mCherry Golgi marker (Figure 3.25) (Nelson *et al.*, 2007). As shown in figure 3.24, GFP-SQE1, GFP-SQE3 and SQE3-GFP co-localise with the Er-mCherry (orange overlay signal). In contrast, GFP-SQE1, GFP-SQE3 and SQE3-GFP did not co-localise with the G-mCherry fusion (Figure 3.25). This analysis suggest that SQE1 and SQE3 are both likely to be targeted to the ER, but not to the Golgi apparatus, consistent with previous data indicating that plant sterol biosynthesis takes place at the ER (Boutté and Grebe, 2009).

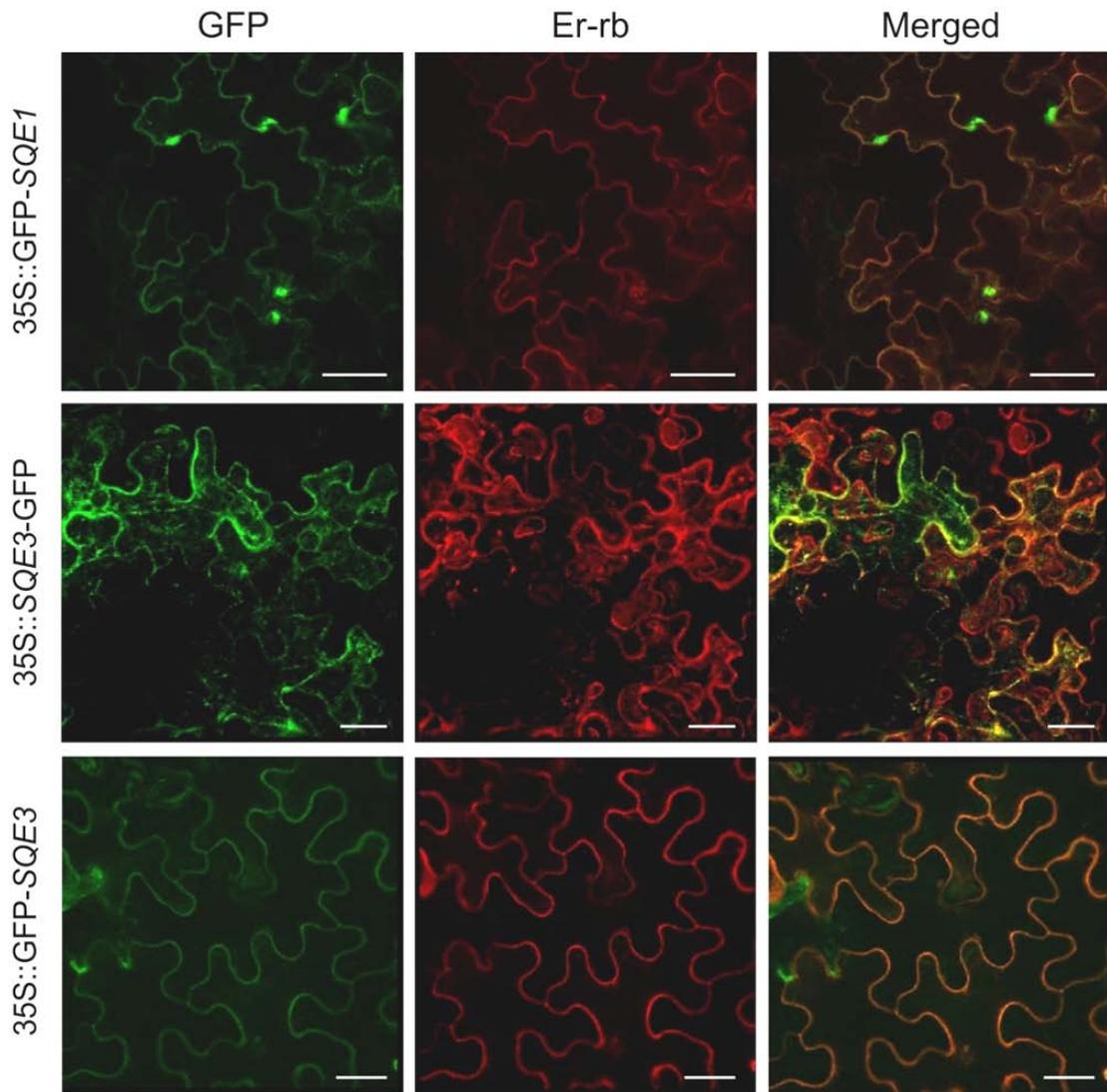


Figure 3.24. – Subcellular localisation of SQE1 and SQE3. Transient expression of 35S::GFP-SQE1, 35S::SQE3-GFP and 35S::GFP-SQE3 with an Endoplasmatic Reticulum subcellular marker in *Nicotiana benthamiana* leaves. Tobacco was co-transformed with the GFP construct using pMDC's Gateway destination vectors (Curtis and Grossniklaus, 2003), a binary vector pFGC expressing Er-rb marker (Er - endoplasmatic reticulum; rb - mCherry fluorescent protein) (Nelson *et al.*, 2007) and a 35S::p19 vector expressing p19 protein (Voinnet *et al.*, 2003), a viral-encoded suppressor of gene silencing. Scale bar represents 25 μ m. Confocal analysis was performed with: GFP excitation at 488 nm, and emission at 492-587 nm; mCherry excitation at 594 nm, and emission at 600-685 nm.

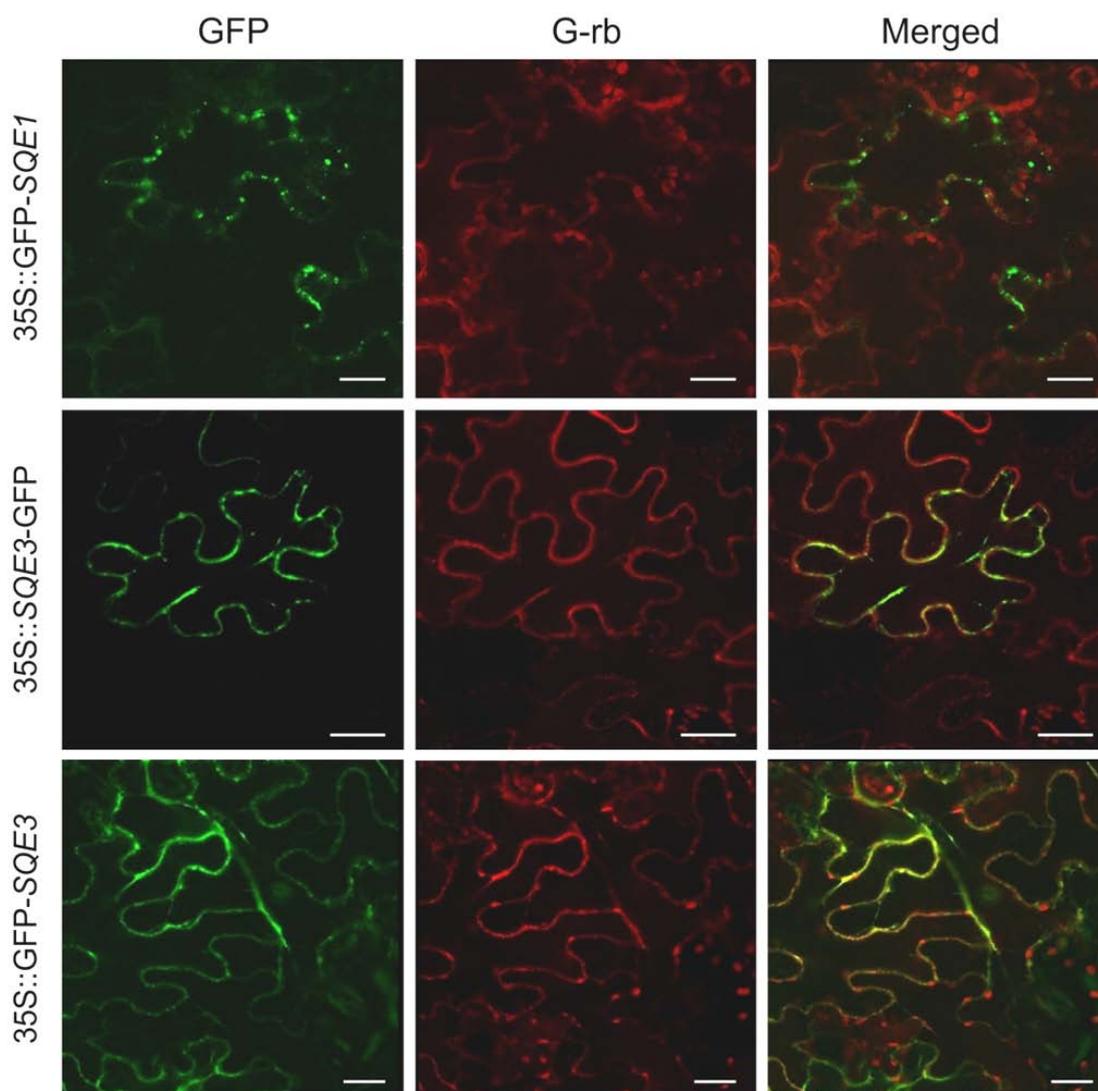


Figure 3.25. – Subcellular localisation of SQE1 and SQE3 proteins. Transient expression of 35S::GFP-SQE1, 35S::SQE3-GFP and 35S::GFP-SQE3 with a Golgi apparatus subcellular marker in *Nicotiana benthamiana* leaves. Tobacco was co-transformed with the GFP construct using pMDC's Gateway destination vectors (Curtis and Grossniklaus, 2003), a binary vector pFGC expressing G-rb marker (G - Golgi apparatus; rb - mCherry fluorescent protein) (Nelson *et al.*, 2007) and a 35S::p19 vector expressing p19 protein (Voinnet *et al.*, 2003), a viral-encoded suppressor of gene silencing. Scale bar represents 25 μ m. Confocal analysis was performed with: GFP excitation at 488 nm, and emission at 492-587 nm; mCherry excitation at 594 nm, and emission at 600-685 nm.

The process of introducing these fusions into *Arabidopsis* mutants to determine whether these constructs complement the endogenous genes was performed and the selection of transformants is underway. In future experiments these lines will be crossed with existing lines that constitutively express fluorescent proteins targeted to the organelles of interest, thus allowing for confirmation of previous co-localisation studies. Notwithstanding, transient expression suggest that SQE1 and SQE3 are likely to be present in the same compartment. Analysis of *in silico* membrane topology (Figure 3.26) predict that SQE1 have its catalytic domains located in the outside of the membrane while the catalytic domains of SQE3 are predicted in the inside, which may explain the lack of

complementation, and therefore they may have separate functions or suffer different regulation. However, the existence of the transmembrane domains defined by this prediction is only putative, since looking at the figure 3.26, there is also a high probability of more TMs for SQE1 and SQE2. Predictions are based on thresholds and algorithms, and so should be looked only as putative since *in planta* data is not available yet.

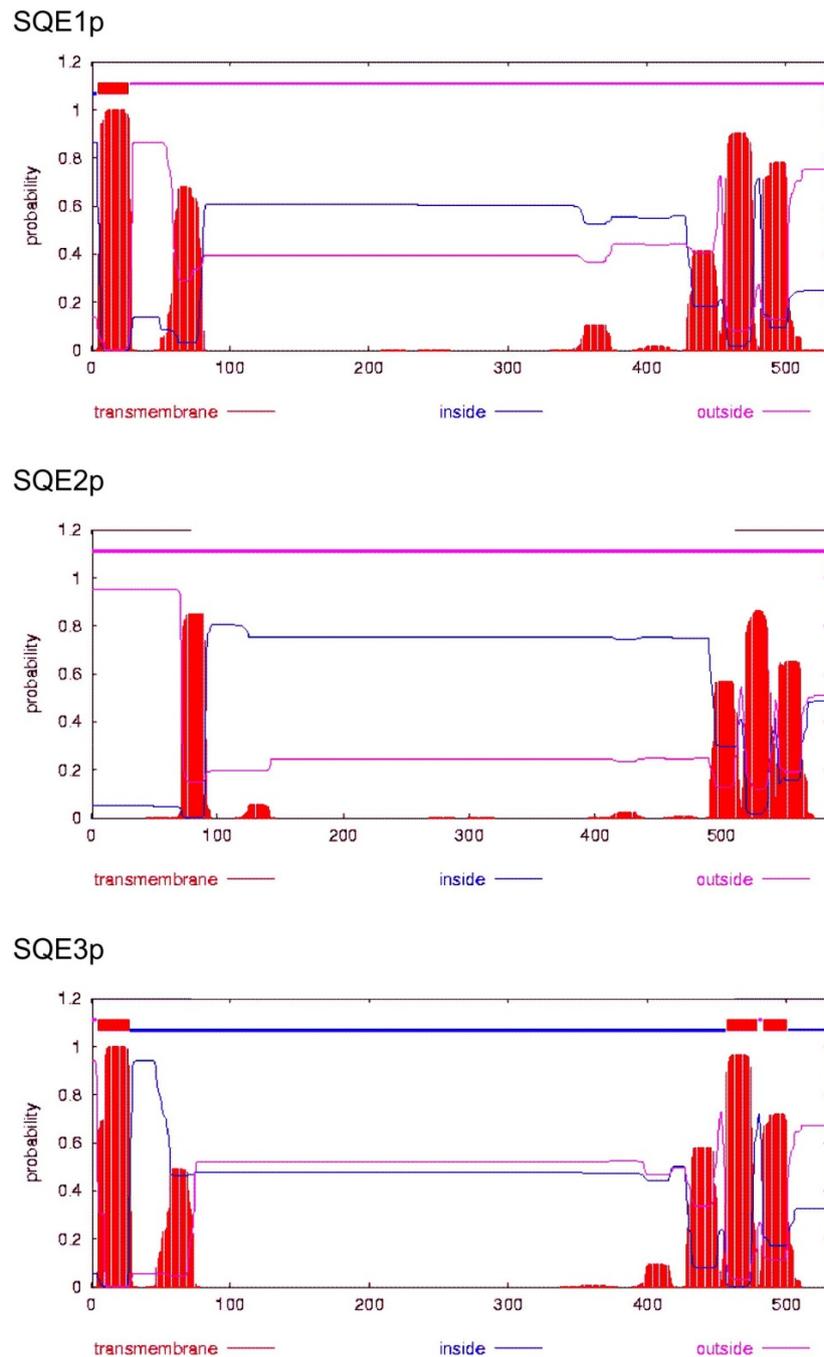
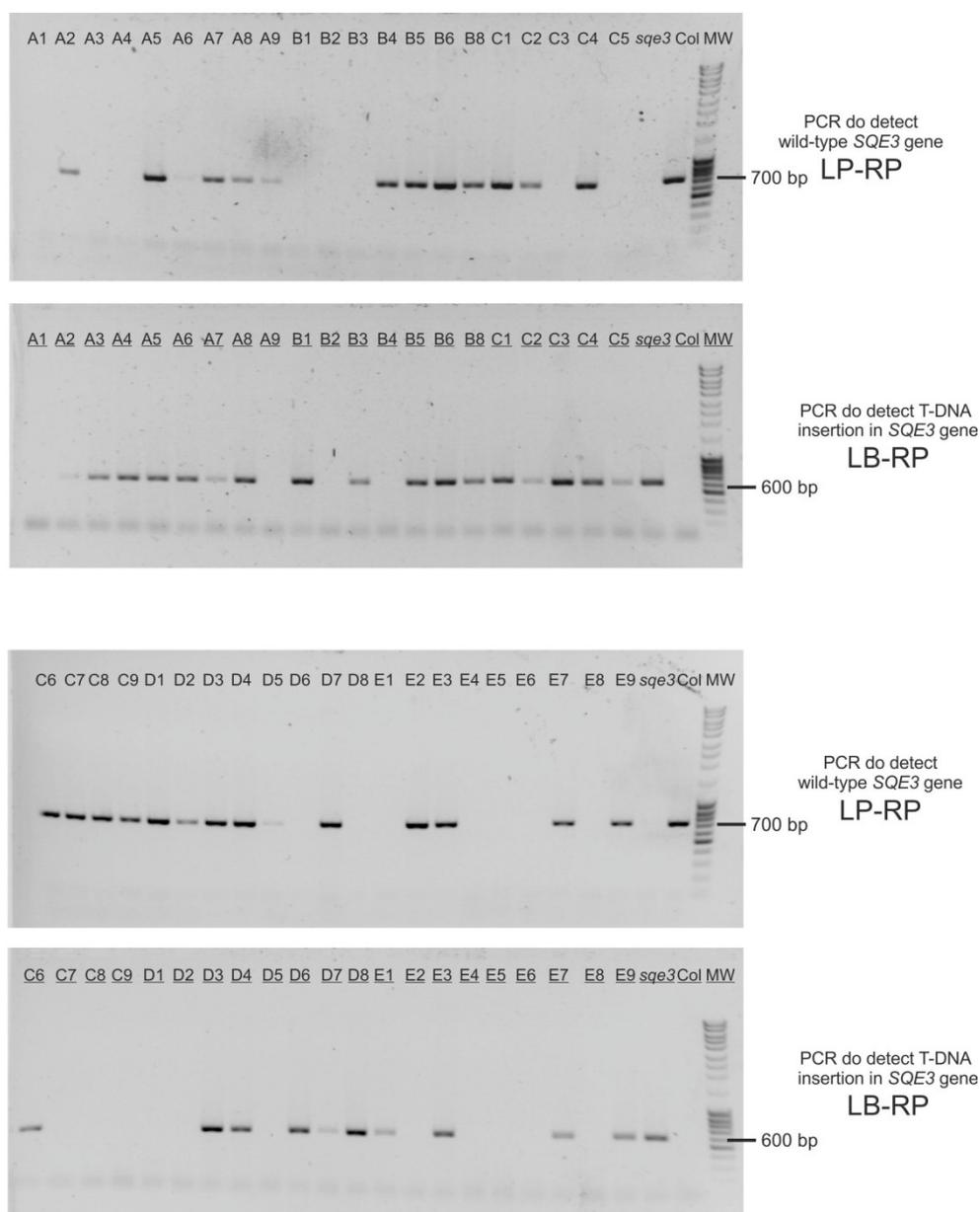


Figure 3.26. – Protein topology of SQE1, SQE2 and SQE3 proteins from TMHMM server (URL no.20).

3.2.8. *SQE1* and *SQE3* have a redundant role in embryo formation

Due to the lack of a visible phenotype for *sqe3-1* and *sqe2-1* mutants, and the results for the subcellular localisation, which indicate the possibility of both *SQE1* and *SQE3* being present in the same compartment (the ER), double mutants were generated. Progeny of a *SQE2/sqe2-1* *SQE3/sqe3-1* was analysed using diagnostic PCR and double *sqe2-1/sqe3-1* mutant plants were identified (Figure 3.27).

A



B

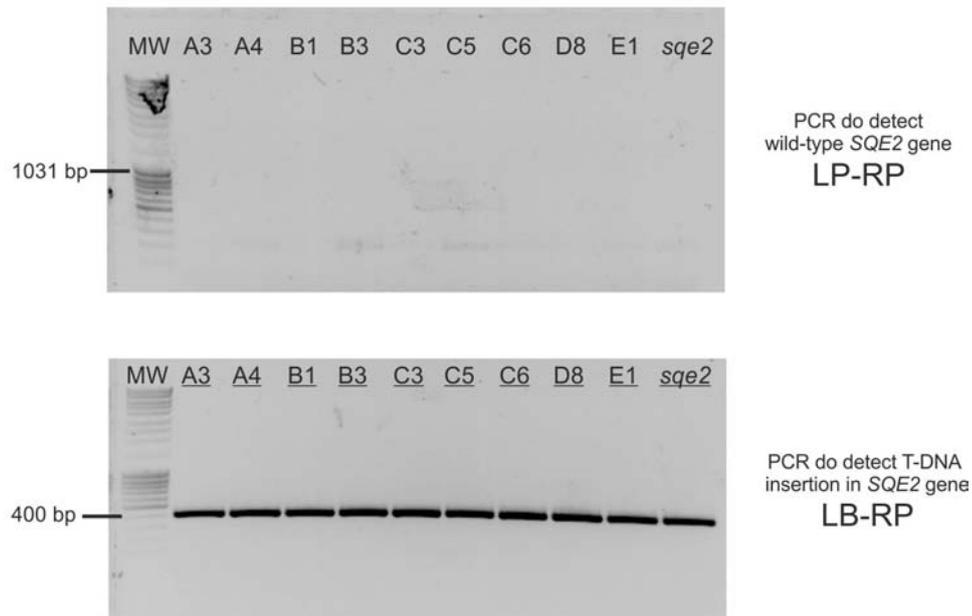


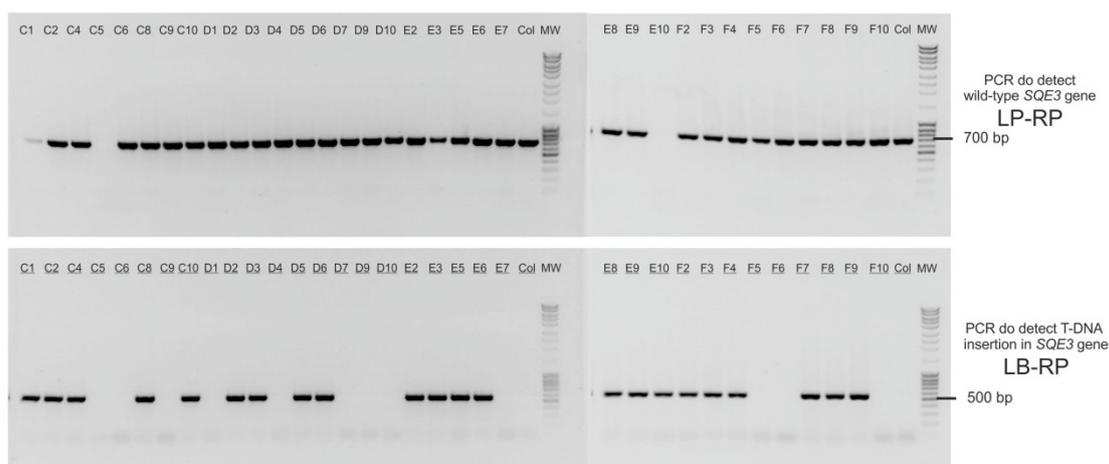
Figure 3.27. – Diagnostic PCR for the isolation of *sqe2-1/sqe3-1* double mutants. A - SQE3 amplification. B- SQE2 amplification of positive *sqe3-1/sqe3-1* plants. MW - Molecular Marker MassRuler DNA Ladder Mix.

No phenotypical differences of *sqe2-1/sqe3-1* compared to the wild-type or single mutants were observed (Figure 3.16). The absence of alterations in the phenotype of the double mutant reinforces the idea that SQE2 may play a different role, maybe in specific conditions and in a different compartment and/or tissue. In the future is planned the analysis of terbinafine sensitivity in the double *sqe2-1/sqe3-1*, so that we could corroborate the hypothesis that SQE2 may have a different function altogether.

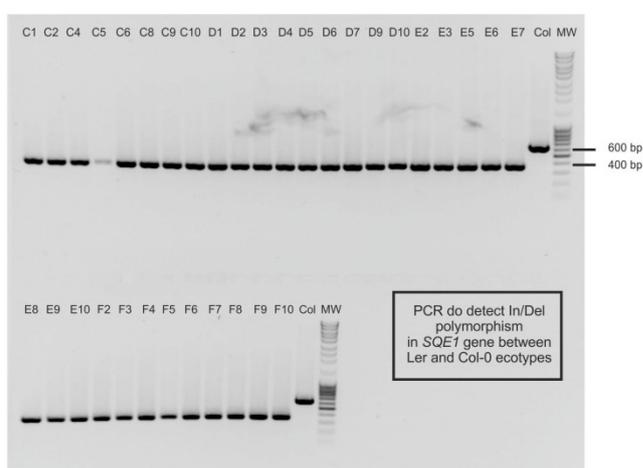
Seeds obtained from a double heterozygous *SQE1/dry2/sqe1-5* x *SQE3/sqe3-1* plant were grown on plate in order to readily visualise the distinguishable *dry2/sqe1-5* root phenotype (smaller and highly branched). Thirty plants that were *dry2/sqe1-5* based on their phenotype were transferred to soil, and therefore it was expected a quarter of these plants (~7 or 8 plants) to be double mutants. Diagnostic PCR confirmed that all plants were *dry2/sqe1-5* homozygous, but only *SQE3/sqe3-1* or *SQE3/SQE3* genotypes were found, strongly suggesting that the double mutant is not viable. To further confirm this, a *dry2/dry2* *SQE3/sqe3-1* plant was selected and its resulting progeny analysed in more detail. Only one double mutant was identified out of the 60 plants analysed, instead of the ~15 plants (1/4) expected (Figure 3.28). These plants were initially genotyped for SQE3 (Figure 3.28-A), and then confirmed for the *dry2/sqe1-5* mutation using a polymorphism of SQE1 promoter between Ler and Col-0 (Figure 3.28-B). In figure 3.28-C, the

different genotypes are depicted in plants that are 1-month-old, including the unique double mutant identified. As previously reported, *dry2/sqe1-5* plants show reduced growth and size, and the double mutant is clearly smaller than the rest and was infertile, indicating an important role of *SQE3*, particularly in a *dry2/sqe1-5* background. Additional search of 60 plants, progeny from a *dry2/dry2* *SQE3/sqe3-1* plant of did not yield any other double mutant.

A



B



C

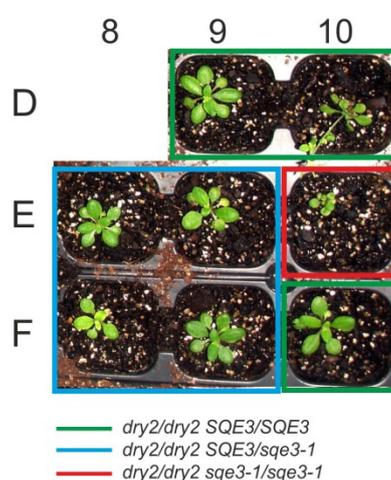


Figure 3.28. – Diagnostic PCR for the isolation of *dry2/sqe3-1* double mutants. A - *SQE3* amplification. B - *SQE1* promoter amplification. *SQE1* promoter in Ler (*dry2/sqe1-5* background) is 263 bp shorter than in Col-0 (*sqe3-1* background). C- Plants genotyped for the isolation of *dry2/sqe3-1* double mutants at ~1-month-old stage (a sample of the 60 analysed). Plants are labelled according to their genotypes. MW - Molecular Marker MassRule DNA Ladder Mix.

The double mutant identified showed the typical characteristics of *dry2/sqe1-5* during development: dwarfed plants with pale-green leaves, long lifespan, short siliques, except that it had a more slighted retarded growth (Figure 3.28-C and 3.29-A). Moreover, siliques from the double

mutant were similar to those of *dry2/sqe1-5*, but did not contain viable seeds (Figure 3.29-B). These non viable seeds looked shrivelled and they resembled, when comparing to the literature (Meinke and Sussex, 1979) the result of embryos that failed to produce seeds rather than unfertilized ovules, which would be much smaller and colourless. This led us to assume the double mutant is embryo-lethal, since the fertilised ovule did not complete its development into a normal seed.

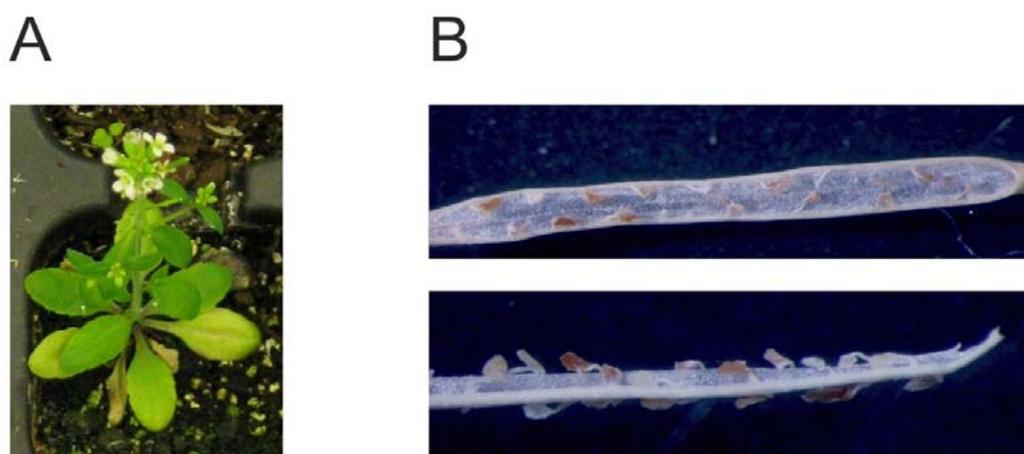


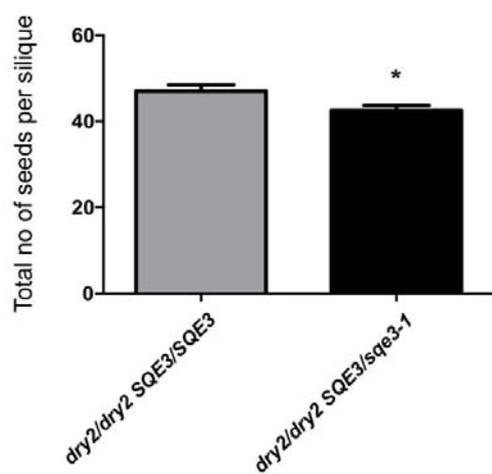
Figure 3.29. – The infertile *dry2/sqe3-1* double mutant. A - Plant morphology of ~2-month-old stage. B - Siliques with aborted seeds at the end-life cycle.

Siliques of *dry2/sqe1-5* *SQE3/sqe3-1* plants were further analysed to investigate seed set and were compared to *dry2/sqe1-5* siliques. Interestingly, we observed a number of aborted seeds in the *dry2/sqe1-5* mutant. This result explains the deviation of the segregation of the *dry2/sqe1-5* mutation that was obtained during the map-based-cloning of *dry2/sqe1*, with around 1/5 of the plants being *dry2/sqe1-5* mutants and not the expected 1/4 (D. Posé, personal communication). Importantly, *dry2/sqe1-5* *SQE3/sqe3-1* siliques present a significantly higher number of aborted seeds, which are represented by white triangles in figure 3.30-A. Following this result, the number of seeds, number of viable seeds and percentage of embryo abortion was estimated in both genotypes (Figure 3.30-B,C,D). *dry2/sqe1-5* heterozygous *SQE3/sqe3-1* have reduced number of seeds per silique (Figure 3.30-B), and importantly the number of viable seeds and the embryo abortion rate is very different (Figure 3.30-C,D), highlighting the importance of *SQE3* in a *dry2/sqe1-5* background, and suggesting an important sterol contribution of *SQE3* for seed or embryo formation. In summary, *dry2/sqe1-5* display a higher seed abortion rate (6%) than wild-type Col-0 which is around 1% (Meinke and Sussex, 1979), while seed abortion raised to ~21% in *dry2/dry2* *SQE3/sqe3-1* mutant. This number is close to the 25% seed abortion rate that is typical of heterozygous plants for embryo-defective mutants affected in seed development (Meinke *et al.*, 2008).

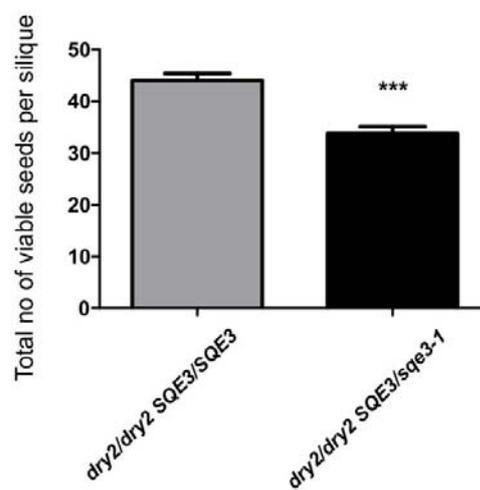
A



B



C



D

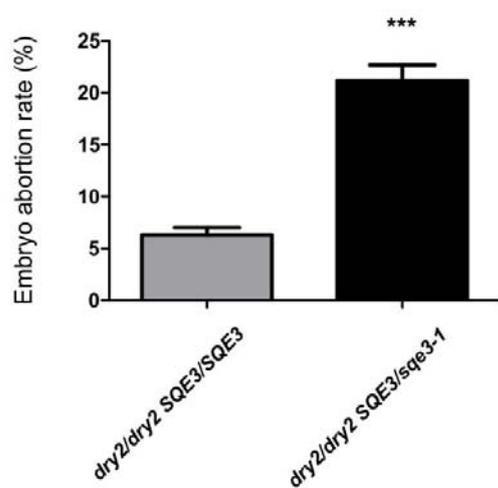


Figure 3.30. – Seed analysis on the *dry2/dry2 SQE3/sqe3-1* and *dry2/dry2 SQE3/SQE3* genotypes. Silique morphology (white triangles represent aborted seeds) (A). Total number of seeds per silique (B). Total number of viable seeds per silique (C). Seed abortion rate (D). Error bars represent SEM (N ≥ 30). Asterisks represent significantly different levels between genotypes (student t-test; *, P < 0.05; **, P < 0.01; ***, P < 0.001).

The finding that seed number is lower in the heterozygous *dry2/dry2* *SQE3/sqe3-1* likely explain why a less than 25% rate is obtained, suggesting that impairment occurs very early during seed formation or even during embryo development. In further studies, it would be important to determine in which embryonic developmental stage growth arrest occur using Hoyer's solution to clear seeds/embryos and visualising them with Differential Interference Contrast Light Microscopy (DIC) (Meinke *et al.*, 2008).

The observed existence of nearly 25% seed abortion rate, rather than unfertilised ovules, is an indication of that embryo development is compromised in seed formation (Meinke *et al.*, 2008). Using available microarray data the expression of *SQE1* and *SQE3* was analysed during seed development using BAR (Arabidopsis eFP browser - URL no.10). As shown in figure 3.31, both genes are expressed at various stages during seed development. *SQE1* is more expressed in the interior (chalazal and embryo structures) and *SQE3* in the exterior (outer coat and endosperm) indicating a complementary gene expression during development (Figure 3.31).

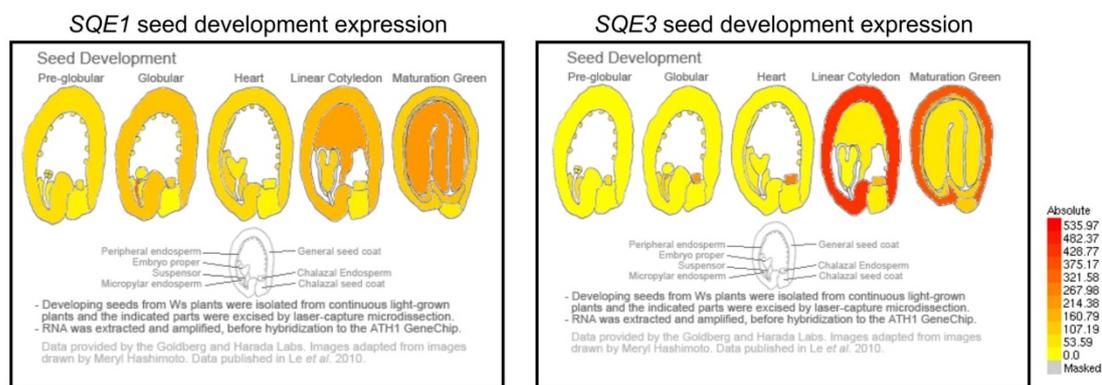


Figure 3.31. – Arabidopsis eFP Browser from BAR database (URL no.10). Seed development expression of *SQE1* and *SQE3* are depicted. Colorimetric scale was adjusted with a similar threshold to allow comparison (Winter *et al.*, 2007; Bassel *et al.*, 2008). Total expression values and standard deviations are listed in table A.1 (*SQE1*) and A.3 (*SQE3*) of the Appendix I.

Statistical occurrence of phenotypes can be associated with penetrance or expressivity, being penetrance the proportion of genotypes that show expected phenotypes, and expressivity the extent to which trait expression is different between individuals (Miko, 2008). The expression of *SQE1* and *SQE3* may explain the very low rate of double mutants, indicating a high penetrance of this phenotype. Our results indicate that double mutants can, very occasionally, develop a fully viable seed; however, since seed viability is also compromised, it can be speculated that the ~4% (from 21% to 25%) of seeds that should (in case of full embryo-lethality) but do not abort, are most likely nonviable. Plant fitness should be taken into account, as it might influence seed development. The unique double mutant identified showed enhanced *dry2/sqe1-5* phenotypes, with the smaller

and branched roots, dwarf phenotype, pale-green leaves and long lifespan; however, their siliques are full of aborted seeds, and with some different coloration. Altogether the knockout of *SQE3* and the hypomorphic *SQE1* allele in the double mutant likely reduce the sterol levels that make unviable embryos. Sterols have been reported to be required for embryonic pattern formation, cell division, cell elongation, cell polarity and cellulose accumulation (Boutté and Grebe, 2009). It was already reported that sterol profiles of *dry2/sqe1-5* and *sqe3-1* are different, with more importance in the root and shoot respectively.

The almost complete absence of a double mutant could be fairly explained by the dramatic loss of squalene conversion into 2,3-oxidosqualene that occurs in this mutant. In the case of *sqe1-3* and *sqe1-4* T-DNA mutants that contain nonviable seeds the authors suggest that specific cells or tissues rely more on *SQE1* function, which together with gene expression data indicates a major role of this protein in embryo/seed development (Rasbery *et al.*, 2007). On the other hand, *dry2/sqe1-5* produces viable seeds, and it could be hypothesised that the presence of a defected protein is perceived by the cell with subsequent recruitment of *SQE3*, since it was shown by qRT-PCR that this gene increases in *dry2/sqe1-5* plants (Posé *et al.*, 2009). As previously shown (Figure 3.31), levels of expression for both genes do not coincide in later developmental stages, whereas *SQE3* is more expressed in the seed coat, *SQE1* is in the embryo in the linear cotyledon and mature green stage. Therefore, if *SQE3* is not present maybe cannot compensate in tissues where *SQE1* should be expressed and it may interfere with normal seed development in the *dry2/sqe1-5* double mutant.

Other relevant aspect may be the influence of the tissue genotype within the seed: while the embryo is diploid (♀♂), the endosperm is triploid (♀♀♂) and the seed coat is entirely maternal (Haughn and Chaudhury, 2005). It would be important to determine whether the embryos are arrested at specific stages in *dry2/dry2 SQE3/sqe3-1* plants. This may help to understand the roles of *SQE1* and *SQE3* in seed or embryo development. The seed coat, where *SQE3* is highly expressed, plays also a very important role to seed development, with a series of cell divisions and expansions (Haughn and Chaudhury, 2005). In the analysed *dry2/dry2 SQE3/sqe3-1* mutant, 75% of embryos will possess at least a wild-type *SQE3* allele in the seed coat. Phenotype analysis of a segregating *DRY2/dry2 sqe3-1/sqe3-1* mutant should be interesting to analyse with the results obtained so far, as this will generate a seed coat without a functional *SQE3* allele. In a similar fashion, comparison of segregation and seed phenotype in both segregating genotypes should be interesting to study *SQE1* role, particularly, when it comes to early developmental stages, where *SQE1* seems to have a more preponderant expression. An interesting aspect to take into account in the phenotypic analysis is seed colour as this is highly dependent on seed coat development.

Regarding this tissue, it has been reported that proanthocyanidin is necessary for seed colour coat (Haughn and Chaudhury, 2005), and sterols have been implicated in the embryonic pattern formation in *hydra1* and *fackel/hydra2* mutants. These mutants are defective in a sterol C-14 reductase gene (*fackel* and *hydra2*) and in the sterol C-14 isomerase (*hydra1*), six and seven steps below SQE in the sterol biosynthetic pathway, and possess a pleiotropic phenotype with defective embryonic and seedling cell patterning, morphogenesis, and root growth (Berleth and Chatfield, 2002; Souter *et al.*, 2002). It is possible that defects in SQE1/SQE3 could alter embryonic development more seriously, or earlier in the embryo formation stages, generating a phenotype which is even more severe than *hydra1* and *fackel/hydra2* mutants. Notwithstanding, overall results suggest an important role for SQE3 in sterols formation during seed development, sterols that are essential in the absence or reduced function of SQE1.

3.2.9. Roots from *sqe3-1* restore defective aerial defects of *dry2/sqe1-5*

Grafting experiments in the *dry2/sqe1-5* mutant are presently being carried out to explore the importance of tissue specificity to the *dry2/sqe1-5* phenotype (Amorim-Silva *et al.*, unpublished data). A similar approach was conducted using the *sqe3-1* mutant, with the following combinations tested: *dry2/sqe1-5* scion and *dry2/sqe1-5* rootstock; *dry2/sqe1-5* scion and *sqe3-1* rootstock; *sqe3-1* scion and *dry2/sqe1-5* rootstock; *sqe3-1* scion and *sqe3-1* rootstock. Shoots of *sqe3-1* with *dry2/sqe1-5* roots failed to promote a good grafting and hence plants did not survive in soil, likely due to the roots of *dry2/sqe1-5* that are very small and dwarfed, which make it developmentally compromised. As shown in figure 3.32, *sqe3-1* roots complemented the phenotype of *dry2/sqe1-5* shoot. Grafting was confirmed by amplification of a polymorphism in the *SQE1* promoter, which resulted in a 365 bp amplification product in Ler (*dry2/sqe1-5* background), and 628 bp in Col-0 (*sqe3-1* background). Results suggest (1) the presence of a mobile signal in roots that can move to the aerial part and restore normal development of the *dry2/sqe1-5* mutant or (2) a toxic compound that accumulates in *dry2/sqe1-5* roots and that is absent in *sqe3-1*. Similar results were previously obtained by the complementation of *dry2/sqe1-5* shoots with wild-type roots (Amorim-Silva *et al.*, unpublished data). The fact that *sqe3-1* did not compromise this complementation suggests that no functional redundancy is likely to take place in root between these two mutants.

Given present results, a *SQE1* wild-type allele in roots is more important for the development of a normal plant, since a substitution for a normal root will produce a healthier plant.

Roots, especially root hairs (defective in *dry2/sqe1-5*), have an important role in sustentation and in nutrient uptake. (Hopkins and Huner, 2008), so it is not surprising that they influence so greatly the aerial part. It is crucial to unveil the role of *SQE1* in this process, and whether the phenotype manifests by a reduction in sterol content or an up- or down-stream deregulation of gene signalling. This aspect is presently being pursued in the lab by other members (Amorim-Silva *et al.*, Gonzalez *et al.*, unpublished data). Additional information should be obtained by extending grafting experiments to other mutant lines of enzymes of the pathway and in particular to this work a grafting experiment with the *sqe2-1*, and with double mutant *sqe2-1/sqe3-1* could add more information to which factor contributes to this complementation.

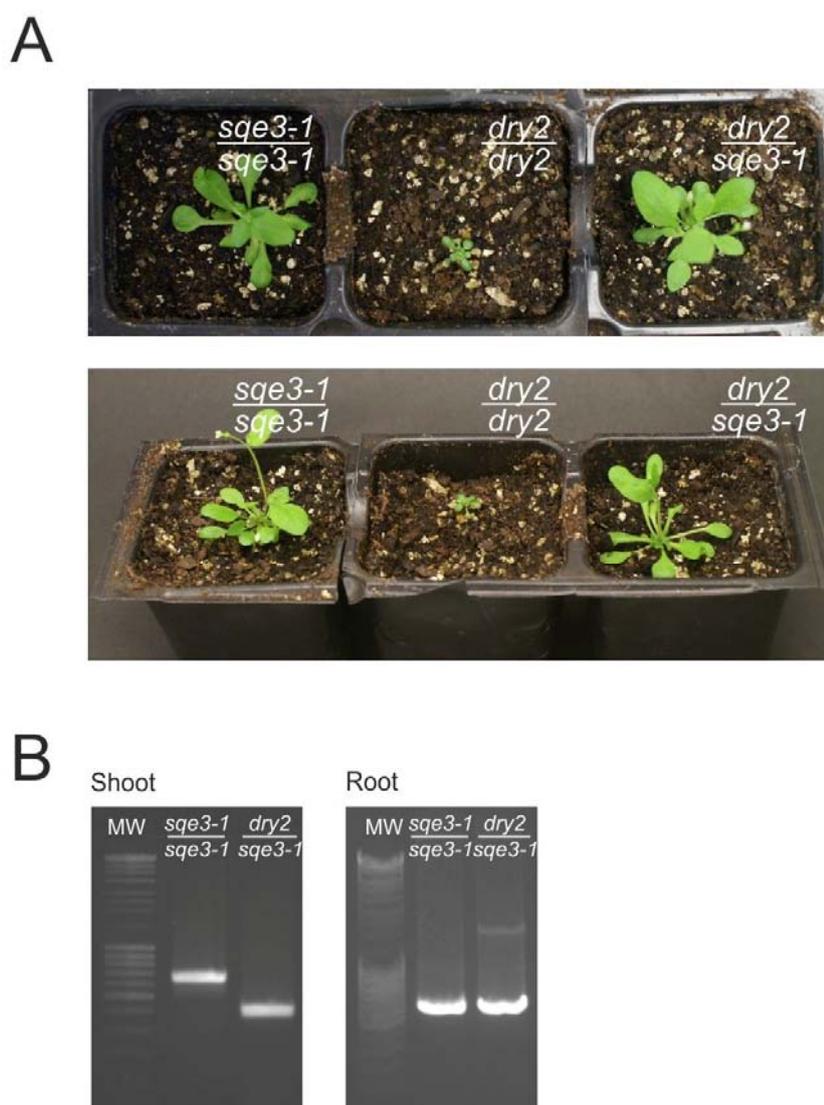


Figure 3.32. – Grafting experiment between *dry2/sqe1-5* and *sqe3-1* genotypes in 1-month-old plants. A – Plants morphology. The numerator indicates the scion and the denominator the stock. B - Root tissue PCR analysis of *SQE1* promoter in graftings.

3.2.10. Abiotic stress-related phenotype analysis of *sqe2-1* and *sqe3-1*

A plethora of resources has been available for Arabidopsis researchers since its outbreak with the sequencing of the genome in 2000 (The Arabidopsis Genome Initiative, 2000). Development over the years produced collections of homozygous T-DNA insertion mutant lines, RNAi and microRNAs resources, cDNA and ORF clones, large-scale microarray data, proteome, metabolome and methylome. All this was also facilitated by a number of web-based databases and browsers that permit a user-friendly research (MASC Report, 2010). Nowadays, web-based databases are fundamental to research in the post-genomic era, where gene function prediction is much facilitated by the surmountable information available, including putative predictors of subcellular localisations, the creation of putative interactor profiles for each protein/gene, co-regulated gene networks, promoter *cis*- and *trans*-analysis, topology analysis of proteins, amongst others.

Databases have compiled extensive microarray data and placed them in user-friendly platforms such as BAR (URL no.10) and Genevestigator (URL no.11) that allow meta-data analysis. *SQE2* and *SQE3* expression analysis was focused first on abiotic stress, based on the fact that *dry2/sqe1-5* was drought sensitive and ROS defective (Posé *et al.*, 2009). Abiotic stresses such as water deficit (leading to drought), high salinity, high and low temperatures are widely studied because of their impact on crop production, and, as a consequence, in the economy (Vinocur and Altman, 2005).

Expression profiles of *SQE2* and *SQE3* were analysed at the time in the Genevestigator Response Viewer tool (Hruz *et al.*, 2008) to uncover differential expression in abiotic stress-imposing conditions. However, we present here a more systematic image from BAR database (Figure 3.33 and 3.34).

3.2. – Functional characterisation of *SQE2* and *SQE3*

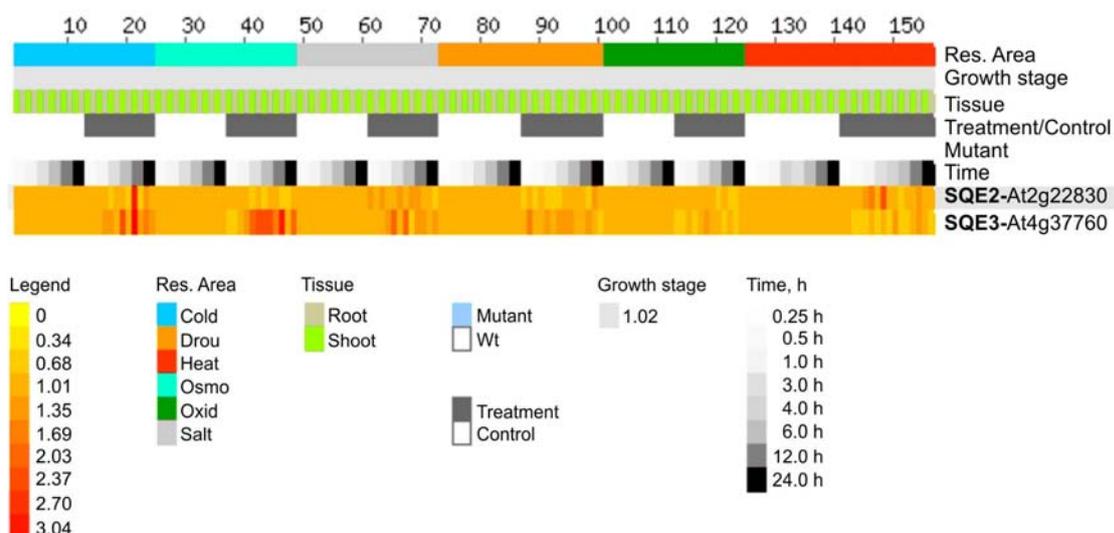


Figure 3.33. – Expression profile of abiotic stress responses of *SQE2* and *SQE3* in the Expression browser of the BAR database (URL no.10) (Toufighi *et al.*, 2005).

As shown in figure 3.33 and 3.34, *SQE2* expression is induced by heat and cold stress, while *SQE3* is induced by abscisic acid (ABA), osmotic, salt, and cold stresses. Therefore, phenotypic analysis was performed for both mutants under those stress conditions that most affected their gene expression.

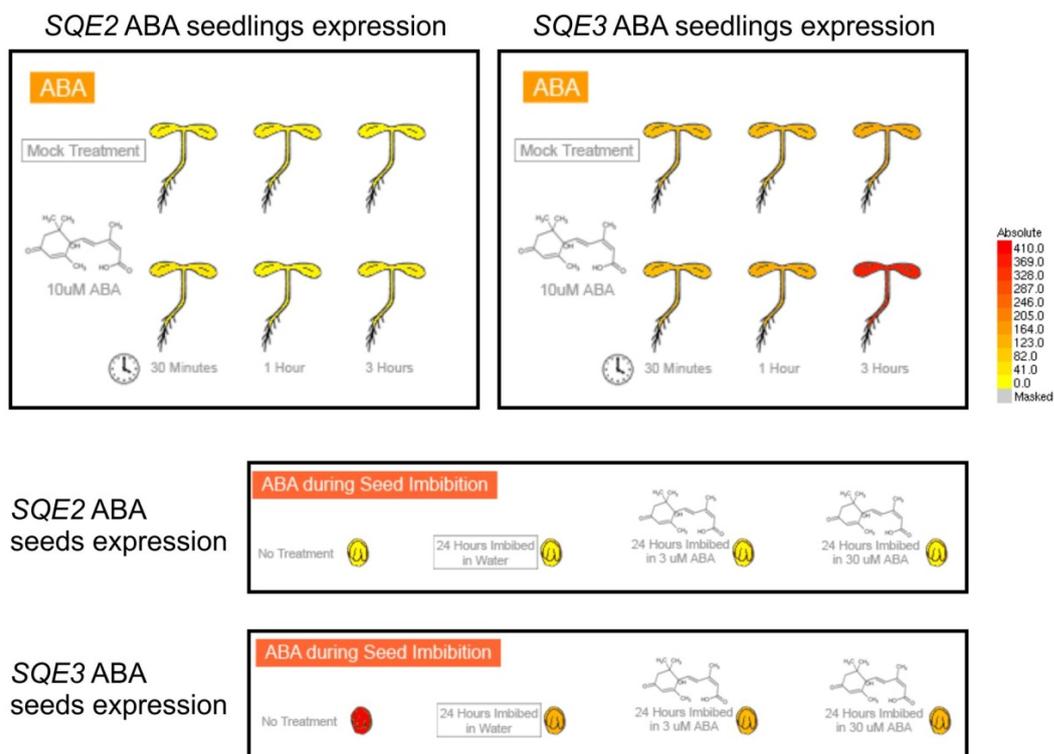


Figure 3.34. – Expression profile of seedlings and seeds response to ABA of *SQE2* and *SQE3* in the Arabidopsis eFP Browser from BAR database (URL no.10). Colorimetric scale was adjusted with a similar threshold to allow comparison (Winter *et al.*, 2007; Goda *et al.*, 2008). Total expression values and standard deviations are listed in table A.2 (*SQE2*) and A.3 (*SQE3*) of the Appendix I.

The *sqe2-1* mutant was subjected to a heat shock experiment, in which *in vitro*-grown 7-day-old seedlings were exposed to 45°C, during a maximum period of 30 min (Figure 3.35). Initial results suggested increased sensitivity of *sqe2-1* when compared to the wild-type after 20 min of heat exposure time (Figure 3.35-A). However, when the experiment was repeated with a larger number of seedlings and a wider range of time (0, 15, 20, 22.5, 25 min), no differences were found (Figure 3.35-B).

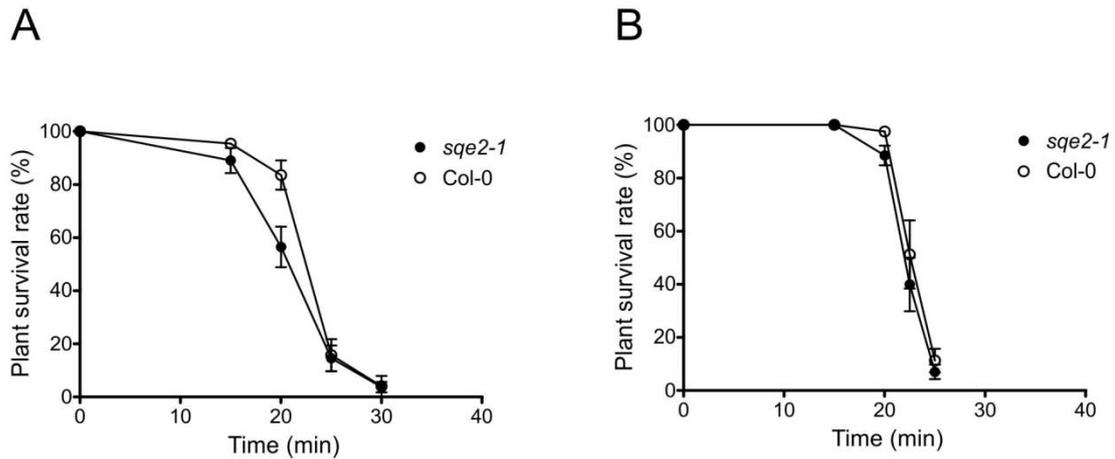


Figure 3.35. – *In vitro* phenotype analyses of *sqe2-1* by estimation of plant survival after heat shock at 45°C of 7-day-old plants. A - Heat shock at 0, 15, 20, 25 and 30 min. B – Heat shock at 0, 15, 20, 22.5 and 25 min. Error bars represent SEM (N=3 for A, N=5 for B).

Exogenous ABA treatment, salt and osmotic stresses were selected for analysis in *sqe3-1*. Seedlings were grown on agar plates during 7 days and then transferred to different concentrations of NaCl (0–200 mM). Root growth was measured and is represented in figure 3.36-A, with no differences between wild-type Col-0 and *sqe3-1*. Similar experiments were performed using mannitol and exogenous ABA, and no significant differences were found either (Figure 3.36-B, 3.36-C). Therefore, none of the mutants referred presented a phenotype in the evaluated conditions.

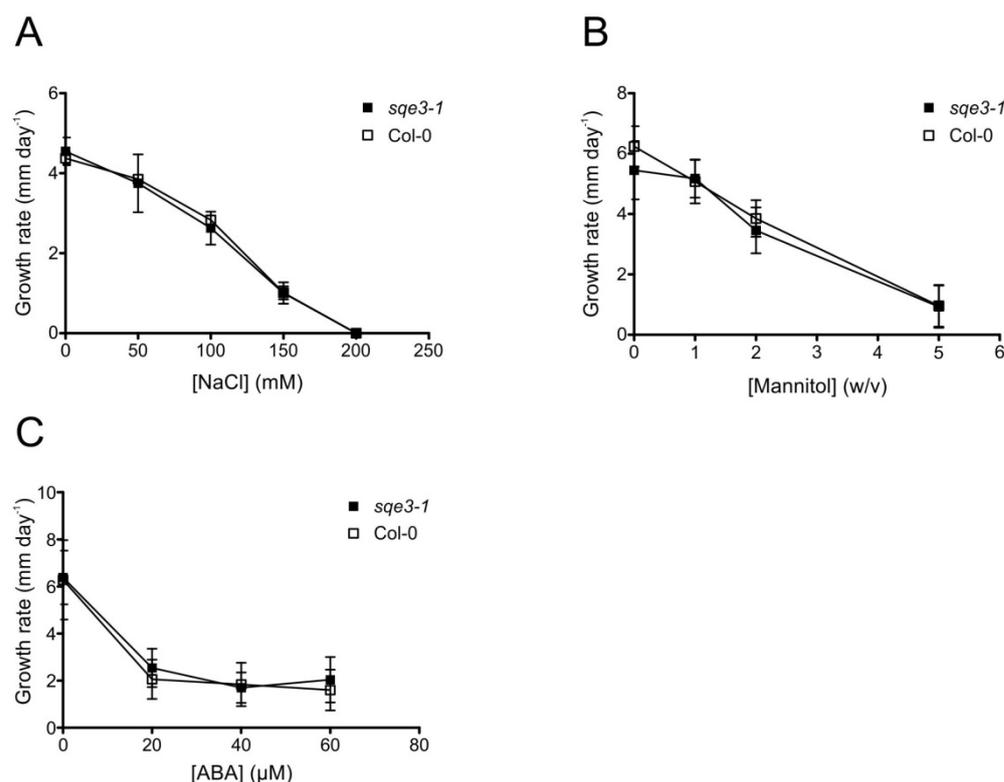


Figure 3.36. – *In vitro* phenotype analyses of *sqe3-1* by estimation of the root growth rate of 7-day-old plants. A - Salt stress induced by NaCl supplementation to the medium. Supplementation was of 0, 50, 100, 150 and 200 mM. Root measurement for 13 days, every 2-3 days. B - Osmotic stress induced by mannitol supplementation to the medium. Supplementation was of 0, 1, 2, and 5%. Root measurement for 9 days, every 2 days. C - Growth inhibition by ABA supplementation to the medium. Supplementation was of 0, 20, 40, 60 μ M. Root measurement for 9 days, every 2 days. Error bars represent SEM (N>15).

3.2.11. ABA stomatal responses are not impaired in *sqe2-1* and *sqe3-1* mutants

SQE1 has been associated with ABA-mediated regulated mediation of stomatal aperture, with stomatal aperture defects being present in *dry2/sqe1-5* (Posé *et al.*, 2009). As previously shown, *SQE3* is highly expressed in the stomata (Figure 3.11, Figure 3.12-G). Therefore stomatal responses were analysed in both *sqe3-1* and *sqe2-1* mutants. Epidermis of three-week-old plants was submerged in stomatal aperture buffer and stomata were stained for microscope analysis, followed by measurement of their aperture. To induce an ABA-dependent stomatal closure this hormone was added to a final concentration of 5 μ M and the solution was analysed one hour later. As shown in figure 3.37, both *sqe2-1* and *sqe3-1* have standard responses to ABA, leading to stomatal closure. The mutants for *SQE2* and *SQE3* presented no visible defects in this response, reinforcing the idea of separate functions within this gene family.

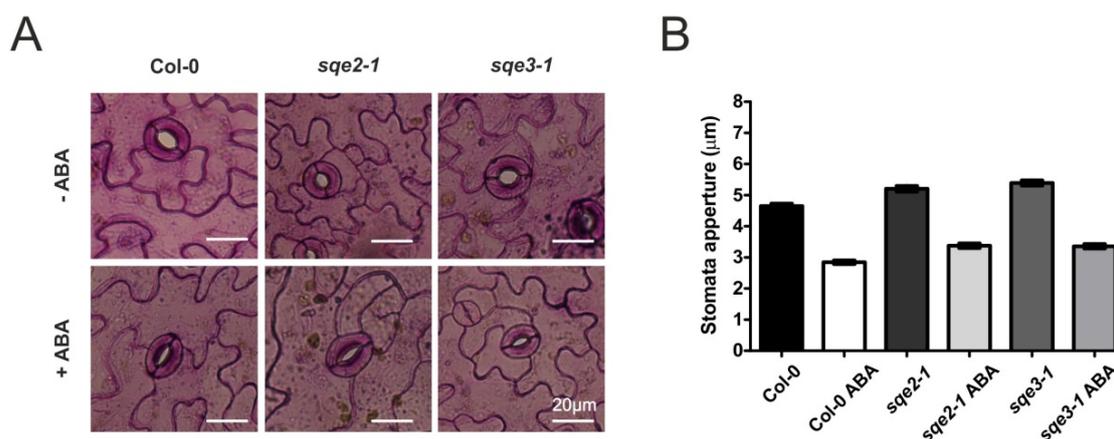


Figure 3.37. – Stomata aperture response to ABA in Col-0, *sqe2* and *sqe3*. A - Stomata images of Col-0, *sqe2-1* and *sqe3-1* mutants in response to ABA. B - This response was evaluated in two consecutive days using standard bright-field microscopy. One hour after ABA (to a final 5 µM concentration) was added, stomata were photographed and aperture was subsequently measured using the ImageJ software. Error bars depict means of two-day analysis ± SEM (N>130) of 3 independent samples for each day. No significant differences (student t-test, 95% confidence interval).

3.2.12. Alternative lanosterol-pathway vs. biotic stress

Recent studies reported an alternative biosynthetic pathway to the formation of phytosterols, different from the cycloartenol-based pathway. The alternative pathway is through lanosterol and is estimated to contribute to 1.5% of total sterols. Conversion of 2,3-oxidosqualene to cycloartenol or lanosterol is catalysed by the specific oxidosqualene cyclases (OSC) CAS1 and LAS1, respectively (Ohyama *et al.*, 2009). There are various OSC described in plants that can have diverse functions, including plant defence (Phillips *et al.*, 2006). The *sqe3-1* mutant does not have visible phenotype alterations, in standard conditions, which could suggest a minor role for SQE3 in comparison to SQE1. Given the recent studies that support an alternative pathway, we hypothesised that SQE3 might be involved in the formation of 2,3-oxidosqualene, to be used by an OSC other than CAS1, being LAS1 a possible candidate.

A microarray analysis of some OSCs suggest that their expression can be modulated in biotic stress conditions, namely in response to *Pseudomonas syringae*. This pathogenic bacteria is referred in the literature as biotrophic, but sometimes also as hemi-biotrophic, and some specific strains such as *Pseudomonas syringae* pv *tomato* DC3000 are particular virulent in the Arabidopsis host (Glazebrook, 2005).

In order investigate a possible association of SQE3 to LAS1 or other OSCs, microarray database analysis was performed for the available 12 (of the 14 predicted) OSCs, together with the three SQEs, in response to biotic stress so to determine commonalities in their pattern of expression (Figure 3.38). Within OSCs, there was no available information on the expression of

CAMS1 and *PEN7* (two OSCs that convert 2,3-oxidosqualene: Camelliol C synthase 1 and Putative pentacyclic triterpene synthase 7, respectively), due to lack of spotted probes in the ATH1 Affymetrix genechip, in which most microarray analyses are performed. The analysis showed a higher induction to *Pseudomonas sp.* of *SQE3*, when compared to *SQE1* and *SQE2*. However, this expression seemed to be depending on the *P. syringae* strains employed. *SQE3* seems to be up-regulated in studies involving the virulent strains *P. syringae* DC3000 and DC3000 *hrpA* mutant. This last strain is deficient in the major structural protein of the Hrp pilus, a component of the type III secretion system (TTSS), required for effector protein delivery, parasitism, and pathogenicity (Roine *et al.*, 1997; Collmera *et al.*, 2002). *SQE3* is also up-regulated in avirulent *P. syringae* *avrRpm1* and down-regulated in studies involving plants infected with *P. syringae* DC3000 in Ler, when grown in short days, and with a non-host pathogen *P. syringae* pv. *syringae* (challenging of 6 and 12h). It appears to be also down-regulated when plants are infected with *P. syringae* DC3000 COR-*hrpS* (mutant without the phytotoxin coronatine and also without the a protein of the hrp pillus complex), and with *P. syringae* DC3000 in the *penta* mutant (Zhang *et al.*, 2008), a mutant defective in a gibberellin biosynthesis gene (*GA1*) and in gibberellin signalling genes (*GAI*, *RGA*, *RGL1* and *RGL2*). Similar to *SQE3*, expression of *LAS1* is induced in response to the virulent *P. syringae* DC3000 infection and since we hypothesise that *SQE3* (and/or *SQE2*) could be involved in the lanosterol pathway through *LAS1*, the finding that both genes are co-regulated by *P. syringae* DC3000 is of particular interest. Therefore, it was determined whether *sqe3-1*, *sqe2-1* or the double mutant showed differences in their sensitivity to the *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*) virulent strain.

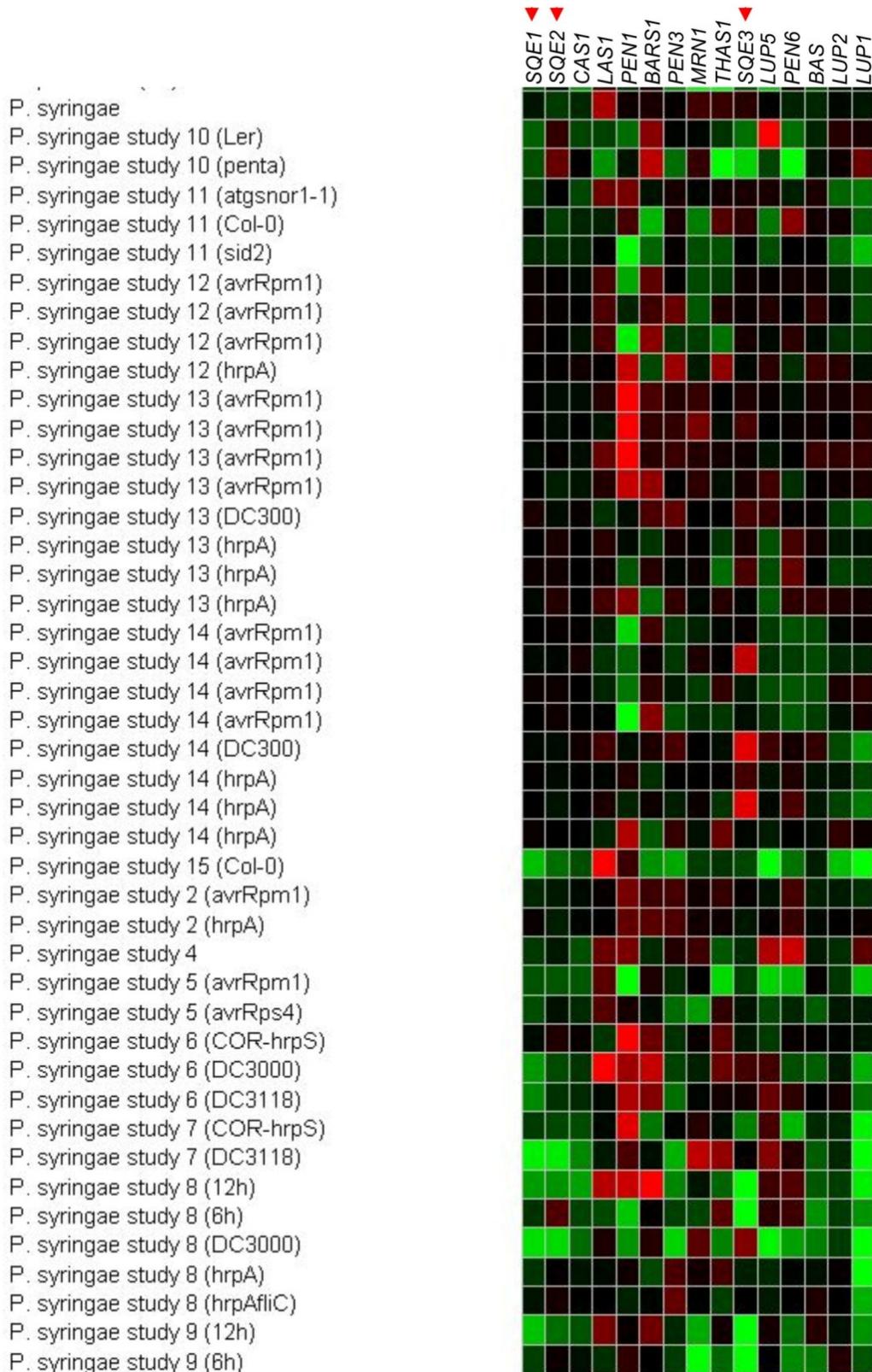


Figure 3.38. – Web-based gene expression data of OSCs in response to *Pseudomonas syringae* challenging. Genevestigator (URL no.11) (Hruz *et al.*, 2008). Red triangles highlight SQEs.

In order to confirm the *P. syringae* gene induction observed in the microarray, gene expression was analysed on a semi-quantitative RT-PCR with all the putative OSC and the SQEs genes, thus pre-eliminating some OSC and confirming SQE3 expression facing this stress (data not shown). Subsequently, gene expression was analysed after 8 and 24 hours of challenging with *Pto* by quantitative RT-PCR (qRT-PCR), with those genes likely to have an important role in sterol biosynthesis, i.e. SQE1, SQE3, CAS1 and LAS1 (Figure 3.39). Only SQE3 showed a clear induction after 24 h of treatment, consistent with the microarray data.

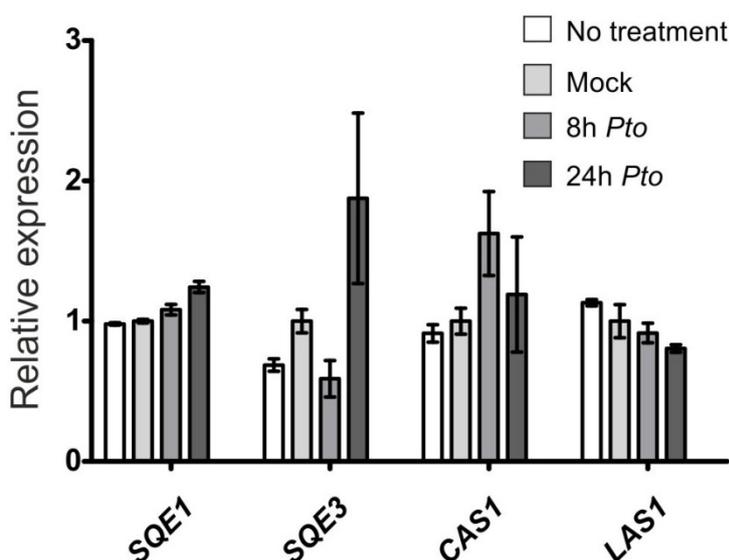
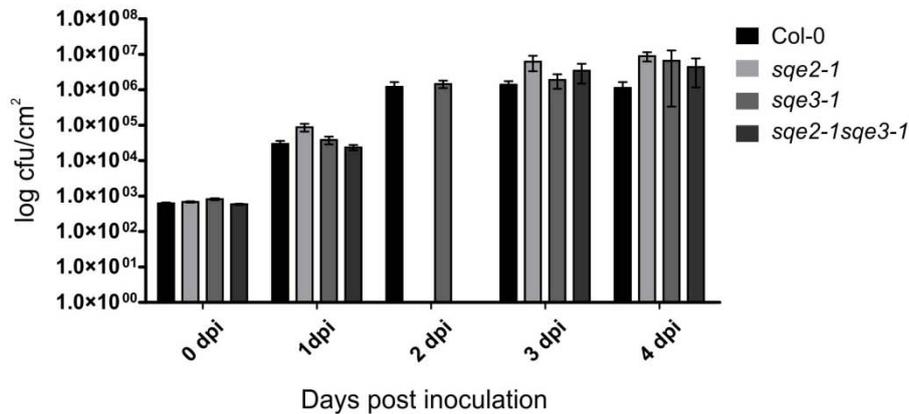


Figure 3.39. – Expression analyses of Arabidopsis Col-0 plants inoculated with *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*) wild type with respective controls. qRT-PCR of SQE1, SQE3, CAS1 and LAS1 (these last two are genes involved in cyclization of squalene). Error bars represent SEM; N=3. Relative expression was calculated using ACT2 gene as reference and normalizing it with the mock treatment results.

Afterwards, a time course experiment was conducted with Col-0, *sqe2-1*, *sqe3-1*, and *sqe2-1/sqe3-1* double mutant after challenge with *Pto* (Figure 3.40-A). First results from this experiment showed no differences in *P. syringae* growth 4 days-post-inoculation (dpi) in the three genotypes analysed. However, when the number of replica was increased (from three to five) in a second experiment, and the samples were collected at time points zero and four days after inoculation (Figure 3.40-B), these new results suggested a slight, though statistically significant sensibility to infection in the *sqe3-1* mutant, (student t-test). Therefore, was felt the need to perform a subsequent experiment with even higher replicas (10), in order to resolve with certainty the possible presence of a biotic stress-related phenotype associated with *sqe3-1*, which was discarded because no significant differences were found (data not shown).

A



B

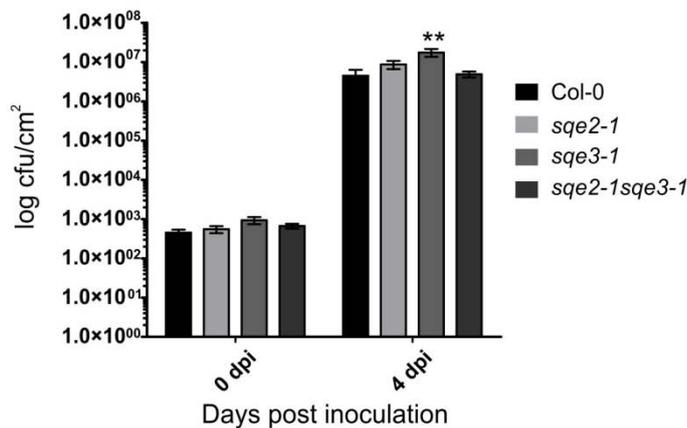


Figure 3.40. – *Arabidopsis thaliana* elicitation assays of *Pseudomonas syringae* pv. *tomato* DC3000 in Col-0, and mutants *sqe2-1*, *sqe3-1* and *sqe2-1/sqe3-1*. Bacterial numbers (means \pm SEM) are reported from an initial infiltration dose of 5.10^4 c.f.u. mL⁻¹. Error bars represent SEM; N=3 for (A); N=5 for (B). Asterisks represent significantly different levels between genotypes (student t-test; *, P < 0.05; **, P < 0.01; ***, P < 0.001).

In this work, it was already shown that the double mutant *dry2/sqe1-5 sqe3-1* is not viable. It is possible that if SQE3 is involved in sterol biosynthesis through the lanosterol pathway, it is expected that the double mutant *dry2/las1* would not be viable since its addition to the contribution of the canonical pathway would block the 1.5% of final contribution of the lanosterol pathway to sterols. The *las1* mutant does not present any morphological phenotype, and double mutants were generated to be analysed in the future.

3.2.13. Concluding remarks

In the *Arabidopsis* genome, there are six genes whose products show high homology to squalene epoxidases, but only three (*SQE1*, *SQE2* and *SQE3*) have been shown to be able to convert squalene into 2,3-oxidosqualene in yeast (Rasbery *et al.*, 2007). Present work focused on the functional characterisation in *Arabidopsis* of the *SQE* gene family, particularly the homologues of *SQE1*, an enzyme described as playing an essential role in sterol composition and plant development, and reporting a previously unrecognised role for sterols in the regulation of ROS, through mislocalisation of RHD2 NADPH oxidase (Posé *et al.*, 2009). With two more functionally unresolved homologues in *Arabidopsis* (*SQE2* and *SQE3*) it was important to address their roles and their possible functional redundancy. The finding that a null allele of *SQE1* was able to produce plants (although infertile) favours this hypothesis because singly copy genes of the sterol pathway have shown embryonic defects (Schrick *et al.*, 2000; Souter *et al.*, 2002). Knockout mutants for *SQE2* (*sqe2-1*) and *SQE3* (*sqe3-1*) were obtained and used to analyse their sensitivity to terbinafine, the specific inhibitor of the enzyme squalene epoxidase. *sqe2-1* and *sqe3-1* do not show visible phenotypic defects, although *sqe3-1* has an increased sensitivity to terbinafine, as was previously reported for the *dry2/sqe1-5* mutant (Posé *et al.*, 2009), suggesting that the *in planta* squalene epoxidase activity provided by *SQE3* has an important role in particular circumstances. The same terbinafine experiment was carried out for *sqe2-1* mutants with no significant differences, reducing *SQE2* to a minor role for bulk sterol biosynthesis in plants. The *sqe1-3* mutants show an accumulation of squalene, reduction of cycloartenol and an increase in sitosterol in shoots but not in roots, implicating a more prominent role for *SQE3* in shoots. The *SQE1* was already associated with sterol composition in the roots and with a high accumulation of HMGR activity (a rate-limiting step enzyme) in roots (Posé *et al.*, 2009). The higher relevance of *SQE3* in shoots is consistent with the expression data gathered, either through the microarray information or the promoter-*GUS* studies.

GUS staining suggest that *SQE3* may have a role in flowering and a role in the male reproductive development and hence perhaps fertilisation. This is supported by crosses of *sqe3-1* with *dry2/sqe1-5* that were performed to obtain double mutants. However, a double mutant was found only once and was infertile. Heterozygous *dry2/dry2 SQE3/sqe3-1* siliques presented about 21% embryo abortion rates, which indicate an aggravated phenotype than the one already presented by *dry2/sqe1-5*, which indicates an essential role for *SQE3* in seed development. Since both *SQE1* and *SQE3* have different expression in various tissues of the seed embryo, they can both be vital for the proper development of a mature seed. A double mutant *sqe2-1/sqe3-1* was

also generated but showed no morphological differences, however, a more detailed phenotypic analysis is needed in order to analyse the role of *SQE2*.

Considering the different expression pattern of *SQE1*, *SQE2*, and *SQE3*, one important issue to address was whether *SQE2* and *SQE3* could substitute *SQE1* *in planta*, if they were expressed under the same regulatory regions. However, promoter swaps and detailed phenotypic analysis indicated that neither *SQE2* nor *SQE3* could complement *SQE1* function.

This result reinforced the necessity to determine the subcellular locations of the *SQE* proteins. Because *SQE2* seems to have a minor role in sterol biosynthesis, analysis was focused in *SQE1* and *SQE3* locations. GFP fusions and transient expression suggest an endoplasmic reticulum localisation for both proteins, discarding rough differences in their intracellular localisations. These results suggest that functional redundancy does not seem to be occurring, not only because expression is observed in different tissues but also because at protein level, *SQE3* does not complement *SQE1*, despite their apparently identical subcellular localisation. It might be possible that in plants, both proteins are regulated in different manners. Post-translational modifications may be regulating these proteins in response to the plants needs, so that they can be present in the same compartment, but yet fulfilling different roles. From the topological analysis *SQE3* and *SQE1* seem to be different, while *SQE1* has only one transmembrane domain and has its N-terminal facing the inside and the C-terminal the outside of the membrane of the organelle, *SQE3* have three putative transmembrane domains with the N-terminal facing the outside and the C-terminal inside the membrane. If this is the case *in vivo*, it indicates that the activity domains of the enzymes are facing in opposite directions, while *SQE1* acts on the outside, *SQE3* acts inside the ER. This may explain how they can be in the same compartment, yet not being functionally redundant. Besides, tissue specificity can be an issue, and thus they could likely be more specific in certain types of tissues as is the case of roots for *SQE1*.

Since *dry2/sqe1-5* was very sensitive to drought stress, abiotic stress responses were analysed in mutants of *SQE2* and *SQE3*. Drought sensitivity of *dry2/sqe1-5* and high levels of expression of *SQE3* in stomata led to the analysis *in planta* of the mutants by calculating the stomatal aperture in response to ABA. However, no differences were obtained relative to wild-type plants. Other phenotypes were investigated for *sqe2-1* and *sqe3-1* mutants based on expression analysis (exogenous ABA, heat, osmotic and salt stress), however, no differences were found either.

A recent report indicated that ~1.5% of total phytosterols is provided by an alternative lanosterol pathway (Ohyama *et al.*, 2009). Therefore we hypothesised that *SQE3* could provide 2,3-oxidosqualene mainly to this pathway by interacting with *LAS1*, while *SQE1* would account for

the other 98.5% through the well described cycloartenol pathway, through CAS1. There are 14 predicted OSC in Arabidopsis, and they can convert 2,3-oxidosqualene into many different and other unknown/putative *in planta* compounds. These compounds were identified in other plant species besides Arabidopsis and with several methods of identification. Cycloartenol, lanosterol, β -amyirin, lupeol, thalianol and marneral were confirmed *in planta* (Phillips *et al.*, 2006). When expression profiles were analysed some OSC were overexpressed in response to biotic stress, in particular *LAS1* by the virulent *P. syringae* DC3000. Since the aim was to investigate a possible connection between *LAS1* and SQE3, expression profile on SQE3 was investigated and found that it was also up-regulated in *P. syringae* DC3000. Therefore changes in sensibility after *P. syringae* DC3000 challenge in *sqe3-1*, *sqe2-1* and the double mutant *sqe2-1/sqe3-1* were analysed. However, no clear differences were found indicating that if they have a role is very small.

Several aspects must be further explored. These include observing expression or activity of key enzymes of the pathway, namely HMGR (a rate-limiting step enzyme), and other enzymes near the SQEs that have important phenotypes associated. Also, a thorough analysis of the seed/embryo of the *dry2/dry2 SQE3/sqe3-1* and *DRY2/dry2 sqe3-1/sqe3-1* should be performed, including an analysis of the embryo stages to see the time point in which development of the embryo is arrested. Important are also expression studies in seeds to confirm microarray analysis, at least in the whole seed, together with tissue specific expression in the seeds using promoter-GUS lines. Meanwhile, additional subcellular localisation studies are underway, with stable transformants being generated. These steady transformants can also be used to cross with plants expressing fluorescent constructs that are specific to certain organelles, and can be used to complement mutant phenotypes to determine if the GFP fusions are functional.

In summary, it was shown that SQE3 is important for plant development particularly in conditions of low SQE1 activity, likely the main enzyme involved in phytosterol biosynthesis in Arabidopsis. The chemical analysis and terbinafine sensitivity phenotype are not sufficient to assign a more specific role for SQE3 and many open questions remain. For example, whether SQE1 and SQE3 interact with different partners, or whether this interaction is less efficient in SQE3. It is also possible that in addition to the residual 2,3-oxidosqualene that became important in particular cases for the phytosterol biosynthesis (as demonstrated in the double mutant), the SQE3 might produce a different compound that is important for other unknown processes.

4.

*“Functional characterisation
of EGY3”*

4.1.

*“Metalloproteases and their role in plant
abiotic stress”*

4.1.1. Proteases and protein degradation

Proteins are essential for an organism's homeostasis as they determine enzymatic, structural and regulatory processes. Homeostasis is achieved by a balance in protein levels, allowing for growth and development of organisms. Due to this balance, proteins can have a short or extended life, which will depend on the needs of a cell or organism. The excess of protein is harmonised with its absence, in a very coordinate manner. This happens in response to a number of significant changes and external stimuli, necessary during the development of an organism, particularly at a cellular level (Kato and Sakamoto, 2010). Proteases, by degrading polypeptides, are an essential part of protein homeostasis. The present review will focus on the categorisation of plant proteases, and will subsequently overview functional knowledge on chloroplastic proteases, with emphasis on the EGY family of putative SP2 metalloproteases.

Proteins first need to be synthesised correctly, through a very coordinated process that ranges from the nucleus to the cytoplasm. DNA is transcribed into mRNA, which in turn is translated into polypeptides, with the help of ribosomes and tRNAs (Nelson and Cox, 2000). Proteins have yet to suffer maturation, so, within cells, most polypeptides go through subsequent folding, processing, modification and proper assembly as a complex in order to become functional proteins (Kato and Sakamoto, 2010). Eventually, proteins need to be degraded, and protein turnover has been proven to be fundamental for development, either for re-location of amino acids for *de novo* synthesis, or for the repairing of errors (Schaller, 2004). Proteins involved in turn-over cycles can have a housekeeping function, meaning that they act in the repair of biosynthesis or translation errors that originate improper folding, or in the repair of thermal, oxidative and chemical damage (Clarke, 2005). They can also act as key regulators of different processes in response to developmental and environmental signals (van der Hoorn, 2008). One of the appointed examples of proteolysis is the extremely important ubiquitin/proteasome pathway, in which proteins linked to polyubiquitin chains are "marked to die", and go through the 26S proteasome for degradation (Figure 4.1) (Hellmann and Estelle, 2002).

In plants, protein degradation plays an important role in germination, differentiation, morphogenesis, senescence and programmed cell death, being carried out by proteases that degrade nonfunctional proteins into amino acids (Palma *et al.*, 2002). For instance, during seed germination, protein degradation plays a vital role in the breakdown of storage proteins, while in the beginning of senescence it helps remobilise proteins and reallocate nitrogen resources to reproductive plant organs (Schaller, 2004).

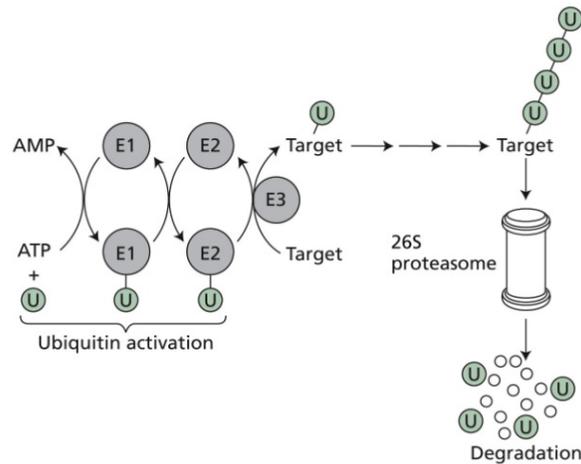


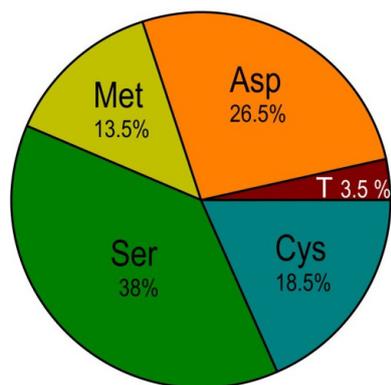
Figure 4.1. – The Ubiquitin/proteasome pathway. ATP is necessary in the activation of the enzyme E1 (Ubiquitin-activating enzyme), which transfers a ubiquitin (U) molecule to the E2 enzyme (Ubiquitin-conjugating enzyme). The E3 enzyme (Ubiquitin-ligase enzyme) mediates the final transfer of U to the target protein, which might already be polyubiquitinated. Finally the target protein is degraded by the 26S proteasome. Adapted from Taiz and Zeiger (2003).

Protease function is present in almost all activities of living cells (Kato and Sakamoto, 2010). Proteases are thought to be substrate specific, and have their activity regulated in time and space (van der Hoorn, 2008). Proteases are designated as endopeptidases or exopeptidases if they cleave peptide bonds internally or externally, respectively. Exopeptidases can be classified according to the cleavage extremity as aminopeptidases (N-terminal) or carboxypeptidases (C-terminal). Aminopeptidases are more ubiquitous in plant tissues, and are referenced in many processes such as development, growth, senescence and defence against pathogens. Carboxypeptidases are the major proteases found in seeds, with serine as their most common active site (Palma *et al.*, 2002). All proteases polarise the carbonyl group, of the substrate peptide bond, by stabilising the oxygen in an oxyanion hole. This process makes the carbon atom more vulnerable for attack by an activated nucleophile. Functionally, proteases are classified according to their catalytic mechanism as cysteine proteases, serine proteases, metalloproteases or aspartic proteases (Figure 4.2; Table 4.1) (van der Hoorn, 2008). Whereas in cysteine proteases and serine proteases the amino acid residues act as nucleophiles activated by histidine, aspartic proteases use water as the nucleophile activated by electrostatic interactions with the metal ions (metalloproteases) or with the aspartate molecule (aspartic proteases). Threonine proteases also exist as a category of proteases, but they are less referenced in the literature (van der Hoorn, 2008).

Table 4.1. - Classes of plant proteases, listing of related biological processes and estimation of the number of families and clans (adapted from Rawlings, 1999; Palma *et al.*, 2002; Schaller, 2004; van der Hoorn, 2008).

Protease category	Known involvement in	#Clans	#Families
Cysteine proteases	<ul style="list-style-type: none"> ▪ PCD ▪ Epidermal cell fate ▪ Flowering time ▪ Inflorescence architecture ▪ Pollen or embryo development ▪ Germination ▪ Xylogenesis 	5	15
Serine proteases	<ul style="list-style-type: none"> ▪ Senescence ▪ PCD ▪ Xylogenesis ▪ Tissue differentiation ▪ Infection of plants cells ▪ Pathogenesis in virus-infected plants ▪ Germination ▪ Brassinosteroid signalling 	9	14
Metalloproteases	<ul style="list-style-type: none"> ▪ Nodulation ▪ Plastid differentiation ▪ Thermotolerance ▪ Regulation of root and shoot meristem size ▪ Sensitivity to auxin conjugates ▪ Meiosis 	11	19
Aspartic proteases	<ul style="list-style-type: none"> ▪ Storage proteins degradation ▪ Extracellular PR proteins degradation ▪ Plant defence 	2	3

The plant *Arabidopsis thaliana* has over 800 proteases annotated in its genome, distributed in ~60 families and 30 different clans, as listed in the MEROPS protease database (Figure 4.2) (Rawlings, 1999; van der Hoorn, 2008). Cysteine and serine proteases take up the majority of proteases, whereas only 13.5% are predicted to be metalloproteases.



Arabidopsis: 880 proteases

Figure 4.2. – Protease distribution in *Arabidopsis thaliana* according to the catalytic class listed in the MEROPS protease database. Asp - aspartic, T - threonine, Cys – cysteine, Ser – serine and Met – metallo proteases. Adapted from Rawlings (1999).

4.1.2. Metalloproteases

Metalloproteases comprise 51 families, and are characterised by a zinc metal-binding site, though some metalloproteases can rely on other divalent cations (e.g. cobalt, manganese) for their activity (Hooper, 1994; Lewis and Thomas, 1999; Schaller, 2004). These enzymes present a common motif, HEXXH, for metal ion-binding, though some variations of the motif exist (HXXE, HXXE, HXH, and HEXXHXXGXXH). A common and important feature for the protease is the presence of histidine(s) (Hooper, 1994). We can find diversification in metalloprotease function in plants as observable in table 4.1, and though it is known that they participate in certain processes, their direct substrates remain mostly unknown. Significant metalloproteases are briefly summarised, with emphasis given to chloroplastidial metalloproteases.

MPA1 (meiotic prophase aminopeptidase 1) is a member of clan MA, family M1, and is essential for chromosome pairing and recombination during meiosis (van der Hoorn, 2008). Null T-DNA mutants for this gene have reduced fertility, and the pollen is mostly nonviable, deformed, and smaller than wild-type (Sánchez-Morán *et al.*, 2004). Subcellular localisation was detected in the cytoplasm and the nucleus, and its activity occurs at an early stage in the recombination pathway. MPA1 may be required for the assembly, or disassembly, of RAD51 (RecA homolog), or MSH4 (mismatch repair protein) protein complexes, since its activity is limited to a narrow time frame, between RAD51 and MSH4 loading into the chromatin (Sánchez-Morán *et al.*, 2004; van der Hoorn, 2008).

FtsH is a member of clan MA, family M41, and a membrane-bound ATP-dependent metalloprotease, belonging to the AAA (ATPase associated with diverse cellular activities) protease subfamily (Kato and Sakamoto, 2010), first identified in *E. coli* (Tomoyasu *et al.*, 1993). The Arabidopsis genome has 12 functional FtsH genes, plus four FtsH gene homologs that lack the

zinc-binding motif, but still maintain ATPase function. Nine of the twelve functional FtsHs are found in the chloroplast, particularly in the thylakoid membranes. AtFtsH1, AtFtsH2, AtFtsH5 and AtFtsH8 are the most abundant, and form heterocomplexes. AtFtsH2 and AtFtsH8 are interchangeable, and capable of stabilising AtFtsH1 and AtFtsH5 in the thylakoid membrane, to form heterocomplex type A (AtFtsH1 and AtFtsH5), and type B (AtFtsH2 and AtFtsH8) (Sakamoto *et al.*, 2003; Kato and Sakamoto, 2010). Mutants for AtFtsH5 and AtFtsH2, *var1* and *var2*, respectively, exhibit a leaf-variegated phenotype, in which white tissue contains undifferentiated plastids, that lack typical thylakoids (Martínez-Zapater, 1993; Takechi *et al.*, 2000; Sakamoto *et al.*, 2002). They are important for chloroplast biogenesis, regulation of thylakoid formation and repair of PSII by removal of the photodamaged D1 protein (Sakamoto *et al.*, 2003; van der Hoorn, 2008; Kato and Sakamoto, 2010).

AMP1 (altered meristem program 1) is a member of clan MH, family M28, and promotes differentiation, by restricting the size of the shoot apical meristem (van der Hoorn, 2008). Several mutants were found for this gene, and phenotypes vary from normal growth in the dark, growth with more than two cotyledons, leaves in whorls, bushier plants, early flowering, and male and female semi-sterility, resulting in smaller siliques and lesser seeds (Chaudhury *et al.*, 1993; Conway and Poethig, 1997; Mordhorst *et al.*, 1998). Mutants are impaired in the timing events of embryogenesis, cell division is perturbed at the globular stage, the size of shoot apex dramatically increases, producing enlarged leaf primordia during seed development (Chaudhury *et al.*, 1993; Conway and Poethig, 1997; Mordhorst *et al.*, 1998). The subcellular location is unknown (van der Hoorn, 2008).

SPP (stromal processing peptidase) is a member of clan ME, family M16, which has an inverted zinc-binding motif HXXEH. It is present in a soluble form in the stroma, and participates in the cleavage of most precursors targeted to the chloroplast. Transit peptides are bound to SPP, and the mature protein is released upon cleavage, though the transit peptide remains for one further cleavage (Richter *et al.*, 2005; Kato and Sakamoto, 2010). An antisense construct of pea SPP expressed ectopically in tobacco showed chlorosis and retarded growth (Wan *et al.*, 1998), whereas in Arabidopsis it showed lethality in seedlings, reinforcing its fundamental role in chloroplast biogenesis and plant viability (Zhong *et al.*, 2003).

PreP (presequence protease) is a member of clan ME, family M16, which has an inverted zinc-binding motif HXXEH, and is ATP-independent (Kato and Sakamoto, 2010). In higher plants, it was first identified as being involved in the degradation of the mitochondrial presequence (Ståhl *et al.*, 2002), but in Arabidopsis two homologues (AtPreP1 and AtPreP2), have been targeted to both the mitochondria and the chloroplast (Bhushan *et al.*, 2003). Single T-DNA mutants for these genes

showed no phenotype, but double-knockout mutants showed chlorosis in leaves, diminished chlorophyll a and b content, and malformed chloroplast and mitochondria (Cederholm *et al.*, 2009). It is suggested that PreP1 and PreP2 play a role in the degradation of transit peptides released by SPP, and an important role in *in vivo* proteolytic events because of the double-knockout mutant's phenotype (Kato and Sakamoto, 2010). Some of the chloroplastidial metalloproteases are typified in figure 4.3.

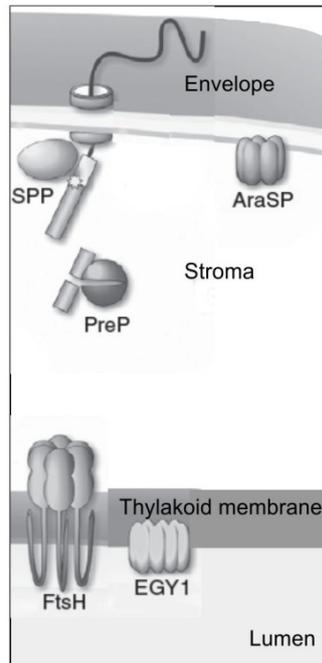


Figure 4.3. – Localisation and structural representation of various chloroplastidial metalloproteases. Adapted from Kato and Sakamoto (2010).

4.1.3. S2P proteases and the *EGY* gene family

Regulated intramembrane proteolysis (RIP) regulates transmembrane signalling. Different types of intramembranously cleaving proteases (I-CLiPs) can carry out this conserved mechanism, and are commonly grouped according to their substrates, catalytic mechanisms, and properties. There are serine, aspartic and metalloproteases involved in these processes. RIP allows the inactivation or activation of transcription factors or signal peptides that will coordinate the response between cellular compartments. These proteases have multiple transmembrane domains, within which the conserved residues often reside. These residues are specialised in peptide bond cleavage normally embedded in hydrophobic membranes (Chen and Zhang, 2010).

The first protease shown to perform the RIP function was the human S2P (site-2 protease) (Rawson *et al.*, 1997), which is involved in feedback regulation of sterol and fatty acid synthesis and uptake, by controlling the activity of SREBP (Sterol regulatory element binding protein)

transcription factors (Brown and Goldstein, 1997; Chen and Zhang, 2010). S2P proteases are included in the larger family of metalloproteases, and orthologs have been identified and studied mostly in bacteria (Rudner *et al.*, 1999; Alba *et al.*, 2002). The model for the human S2P pathway (Figure 4.4) involves two membrane proteins, site-1 protease (S1P), and site-2 protease (S2P); their substrates, and other possible regulatory proteins. While SREBPs reside in the ER membrane, S1P and S2P are located in the Golgi apparatus. In a sterol depleted cell, conformational changes occur that sort the SREBP complexes to the Golgi apparatus, in COPII-coated transport vesicles. In the Golgi, a sequential double cleavage occurs, first by S1P and then by the zinc metalloprotease S2P, leading to the release of the N-terminus of the SREBP. This N-terminal bHLH-Zip domain goes to the nucleus and promotes transcription of sterol and fatty acid synthesis, by binding to promoter SRE (sterol regulatory sequences) elements (Chen and Zhang, 2010).

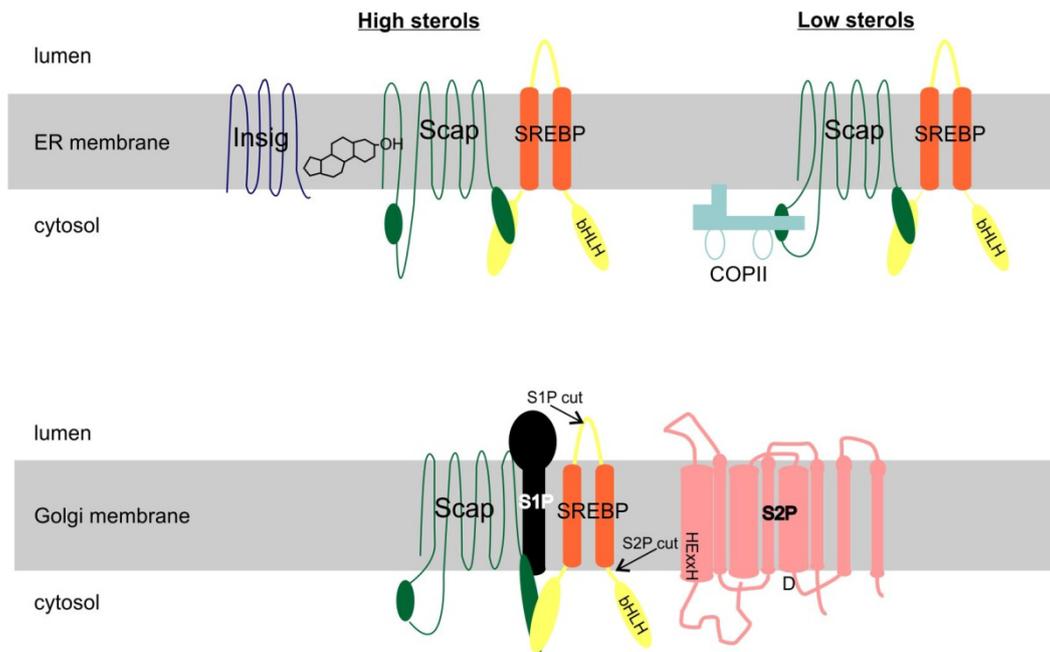


Figure 4.4. – Schematics of the S2P cascade in humans. Substrates are shown in orange while Site-2 proteases are shown in pink. Site-1 and -2 proteases cleavage are represented by arrows in the lower image. The metal chelating motif that includes HExxH is shown in black in the S2P. Under low levels of sterols, human SREBP is transported from the ER membrane to the Golgi to be sequentially cleaved by S1P and S2P. Adapted from Chen and Zhang (2010).

In *Arabidopsis* only one S1P (Liu *et al.*, 2007) was characterised, though there is no known S2P pair for this enzyme. In this species, members of the M50 family of metalloproteases are considered putative S2P peptidases. According to the MEROPS protein database, the following AGI codes are considered M50 members: At1g05140, At2g32480 (*ARASP*), At5g05740 (*EGY2*), At5g35220 (*EGY1*), and At4g20310. All the corresponding proteins are putatively located in the

chloroplast, except the last one which might be located in the endomembranous system. Of these S2P homologues in Arabidopsis, EGY1 and AraSP are the best characterised.

AraSP is a member of clan MM, family M50, which has the zinc-binding motif HEXXH and an additional highly conserved NPDG motif (Bölter *et al.*, 2006). This NPDG motive was already described in the SREBP S2P protease, involved in the mammalian sterol biosynthesis, as well as in other homologues of different organisms (Lewis and Thomas, 1999; Rudner *et al.*, 1999). AraSP is located in the chloroplast's inner envelope, presenting 4-5 transmembrane helices. Knockout mutants plants and N-terminus 525 bp antisense constructs of the gene transformed into plants showed impairment in the development of chloroplasts, suggesting an important though not yet clear role in this organelle (Bölter *et al.*, 2006; van der Hoorn, 2008).

EGY1, also a member of clan MM, family M50, is an ATP-independent protease with the zinc-binding motif HEXXH, the highly conserved motive NPDG (both canonical for homology to the S2P proteases), and eight putative transmembrane helices in its C-terminus. Its catalytic centre is embedded inside the membrane. The mutant *egy1* (ethylene-dependent gravitropism-deficient and yellow-green 1) mutant was isolated by Chen and co-workers (2005), from a screening of mutagenised plants with both deficient pigmentation and defective ethylene-dependant gravitropic responses. The same study showed the presence of abnormal plastids with an undeveloped lamella structure and reduced thylakoid membrane stacking. Results suggest that a deficient EGY1 blocks normal chloroplast development and interferes with the differentiation of plastids into amyloplasts in the endodermis, thus desensitizing gravity perception, essential for a normal gravitropic response (Chen *et al.*, 2005). Subsequent studies showed that EGY1 was necessary for the biogenesis and replication of endodermal plastids which are important for the ethylene-dependent gravicurvature response to light (Guo *et al.*, 2008; Kato and Sakamoto, 2010). It was experimentally shown that EGY1 has proteolytic activity degrading β -casein *in vitro* in a ATP-independent manner (Chen *et al.*, 2005).

While unravelling EGY1 function, two other homologues in Arabidopsis were found and named EGY2 (50%) and EGY3 (42%) for their aminoacidic similarity. Interestingly, Chen *et al.* (2005) found a unique signature motif to these EGY-like proteins, the GNRL motif, also found in cyanobacteria orthologs. This motif distinguished EGY-like proteins from other S2P homologues in Arabidopsis (Chen *et al.*, 2005). However, EGY3 does not exhibit the characteristic HEXXH motif, which is replaced by a SEIAT sequence, and even the NPDG motif is changed to NPEG (Figure 4.5), only maintaining intact the GNLR motif unique to EGY-like proteins. The lack of the characteristic HEXXH motif in EGY3 may indicate a loss of metalloprotease activity so it may not be catalytically active. Studies on SpoIVFB, a S2P family member from *Bacillus subtilis*,

demonstrated that a substitution of the aspartic acid (D) on the NPDG motif for an alanine (A), asparagines (N), or a glutamic acid (E) impaired the processing of pro- σ^k (the inactive form of the prokaryotic transcription factor) to the mature form σ^k (Rudner *et al.*, 1999). This TF is important for the regulation of gene expression during the sporulation process (Rudner *et al.*, 1999). In human S2P, it was also demonstrated that a substitution on the NPDG motif of an aspartic acid (D) for a asparagine (N) abolished the activation of the SREBP, and subsequently its cleavage by S2P (Zelenski *et al.*, 1999). These mutagenesis studies infer that the importance of the conserved aspartic acid is consistent with the possibility of it being the third metal ligand, present in the NPDG motif, whereas the other two metal ligands (the histidines; H) are present in the HEXXH motif (Rudner *et al.*, 1999; Zelenski *et al.*, 1999).

As previously stated, the EGY3 has neither the HEXXH motif, the hallmark of metalloproteases, nor the putative third metal ligand. Apart from not having these characteristics, EGY3 importance has yet to be revealed. It has been suggested that some metalloproteases might act as chaperones or fulfil other important tasks in related proteolytic processes, as mentioned in the case of some FtsHs that also lack the zinc-binding-motif (Kato and Sakamoto, 2010; Olinares *et al.*, 2010). These FtsHs, predicted to the chloroplast and only harbouring the ATP-binding domains, have been viewed as putative catalytically inactive chaperones. Mutants for the FtsHi2 and FtsHi3 proteins were shown to be embryo lethal (URL no.33) (Olinares *et al.*, 2010). EGY3 may also be catalytically inactive, but the role of the specific EGY-like protein motif GNLR has yet to be discovered. The protein is putatively chloroplastidial but according to Chen and co-workers (2005) it is phylogenetically distant, along with EGY2, from the EGY1 protein. EGY3 has less transmembrane domains suggesting a different topology, which might also support its involvement in different processes. Like EGY3, EGY2 functional characterisation is still lacking, but it is likely a true metalloprotease, since it has 50% a.a. similarity to EGY1, but the more strikingly it has intact HEXXH and NPDG motifs. Differences include the presence of only three transmembrane domains, which could implicate some functional difference from the other family members. In summary, more functional information is necessary regarding the EGY family of putative metalloproteases.

4.1.4. Protease involvement in chloroplast functioning and stress signalling

Plants are subjected to several adverse conditions, and have developed various mechanisms to avoid or recover from their effects. Proteases are involved in responses to many external stimuli. Several changes, mostly related to stress, will need proteases to adjust the cellular proteome by promoting degrading/recycling of proteins and using cleavage to process new proteins or even activate TF that would lead to expression of other genes/proteins. Several mechanisms are known to have a particular protease involved, although the regulation and the precise substrates or mechanisms involved are not always completely understood (Kato and Sakamoto, 2010). Plants need sunlight, essential to perform photosynthesis. However, the absorption of light can be harmful, damaging the photosystem II (PSII), and provoking photoinhibition. This can lead to loss of photosynthetic activity, growth and productivity. Plants have therefore developed photoprotection mechanisms, such as light avoidance, photoradiation screening and ROS scavenging systems, among others. Interestingly, the chloroplast contains at least 11 different protease families (García-Lorenzo *et al.*, 2006). Plastidial proteases are of extreme importance throughout the organelle's life (plastid differentiation, homeostasis, senescence, conversion to plastid type), particularly because plastid function is highly affected by external conditions, and control of protein quality is extremely important during photosynthetic activity (Kato and Sakamoto, 2010).

The best known example of protease involvement in protein function is the repair of damaged photosystem II (PSII) protein D1, that is removed by the FtsH complex, with the help of Deg proteases (Kato and Sakamoto, 2010) (Figure 4.6). When photodamaged, the PSII has to be repaired. PSII damaged proteins lead to partial disassembly of the PSII complex (PSII repair cycle), and new proteins are synthesised and substitute the damaged ones, with proteases being important in the process (e.g. Deg, CtpA and FtsHs) (Takahashi and Badger, 2011). In particular, the D1 protein needs to be *de novo* synthesised, and cannot be simply repaired. During the PSII repair cycle, the PSII dimer is monomerised and subsequently partially disassembled. Degradation of the D1 protein is primarily promoted by action of a FtsH protease. Therefore, *de novo* synthesis of precursor D1 (pD1) protein is required, by transcriptional activation of the chloroplast-encoded *psbA* gene. Maturation of pD1 protein (cleavage of C-terminus amino acids), also requires protease activity (by a carboxyl-terminal peptidase; CtpA), and the PSII complex can subsequently be reassembled and photoactivated (Takahashi and Badger, 2011). Under high light conditions, the repair cycle is less efficient, because D1 protein synthesis is inhibited at the translation level, due to limitations on the Calvin Cycle caused by the environmental stress. ROS

levels also interfere in the PSII repair cycle. It is known that translation of D1 protein is regulated by both the ATP:ADP ratio and the stromal redox potential (Takahashi and Badger, 2011).

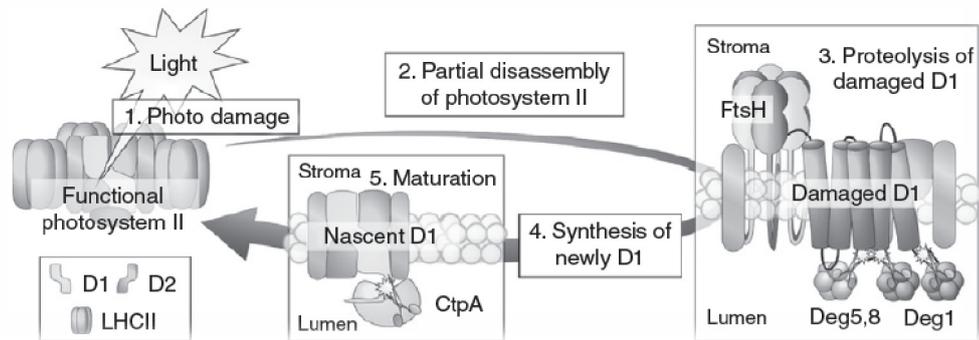


Figure 4.6. – PSII repair cycle. Photodamage of PSII leads to subsequent disassembly and proteolysis of damaged D1 protein by Deg and FtsH proteases and new synthesis of D1 to restore a functional PSII. Adapted from Kato and Sakamoto (2010).

Heat stress, one of the most relevant stresses to plants, can lead to senescence and death, and is well associated with stimulation of protein degradation for removal of abnormal proteins (Huang and Xu, 2008). Transcription factors (TFs) are very important in heat stress, since HSFs (heat stress TFs) are the terminal components of a signal transduction chain that coordinates activation of genes responsive to heat and chemical stressors (Nover *et al.*, 2001). In general, TFs are subjected to several regulatory steps, including transcriptional regulation, post-translational modification, and nuclear transport. After synthesis, they are mostly kept “dormant” in the cytoplasm, waiting for activation through non-covalent interactions with other factors, or even post-translational modifications. In some cases, TFs are activated by interaction with intracellular membranes, followed by proteolytic cleavage. This has been called controlled proteolytic cleavage of membrane-bound TFs (MTFs), allowing for quick responses to sudden environmental changes. MTFs can be activated by two different mechanisms: regulated ubiquitin/proteasome dependent processing (RUP) or by regulated intramembrane proteolysis (RIP) (Seo *et al.*, 2008; Chen and Zhang, 2010). Heat stress is a good example of an abrupt stress that can induce the proteolytic release of transcription factor domains of MTFs. For instance, Gao *et al.* (2008) has reported that bZIP28, a gene encoding a putative membrane-tethered transcription factor, is up-regulated in response to heat, and that a bZIP28 null mutant has a striking heat-sensitive phenotype. Chloroplastic ClpB/HSP100 regulates the activity of protein complexes, unfolds proteins for presentation to proteases, facilitates the refolding of denatured protein aggregates while using ATP energy (Lee *et al.*, 2006), and is involved in heat thermotolerance in chloroplast development (Kotak *et al.*, 2007). Other HSP are also targeted to the chloroplast where they protect the

compartment, with evidence being reported in tobacco of sHSP protecting photosystem II under some abiotic stress conditions (Neta-Sharir *et al.*, 2005; Kotak *et al.*, 2007). The FtsH11 chloroplastidic protease, when mutated, shows reduced photosynthetic capability after heat stress and contributes to plant tolerance to high temperatures (Chen *et al.*, 2006). Altogether, heat stress seems to be correlated to protease function in the chloroplast. More generally, plant MTFs include 3 bZIP members and 13 NAC members, some of which have been functionally characterised. It is thought that they intervene in various aspects of the stress response, and the activation of these MTFs is viewed as crucial for stress adaptation. NAC genes are transcriptionally regulated by various abiotic stresses and stress-related growth hormones, and are associated with several important processes, including floral development, apical meristem formation, stress response and growth hormone response. Several of the membrane-associated bZIP and NAC families in Arabidopsis are activated during a stress response, by proteases embedded in the membrane, in the endoplasmic reticulum. Given that the estimation of the number of transcription factors that are bound to membranes is of about 10%, their importance during stressfully impairing conditions cannot be disregarded (Seo *et al.*, 2008).

The bZIP are TFs associated with a wide variety of processes that include pathogen defence, light and stress signalling, seed maturation and flower development, and are divided according to their common domains (Jakoby *et al.*, 2002). The bZIPs known to be MTFs are located in the ER membrane. In Arabidopsis, it is assumed that cleavage of AtbZIP28 (sensor for ER stress) by S1P and S2P proteases, allows this MTF to be released from the ER membrane, redistributing it to the nucleus where it fulfils its function (Seo *et al.*, 2008). However, although some AtS2P have been found, only one AtS1P is known, being responsible for the cleavage of AtbZIP17, which is processed and reallocated to the nucleus, resulting in an up-regulation of salt stress genes (Chen and Zhang, 2010).

In general, proteolytic cleavage appears to be a very important mechanism of regulation, in immediate processes, such as the processing of a signalling peptide, but also in the regulation of genes, by a controlled proteolytic cleavage of membrane-bound TFs. These MTFs are of great importance in response to stress conditions (Seo *et al.*, 2008). Specifically in the chloroplast, which has a major function in the plant cell, these mechanisms appear to be of utmost importance, as highlighted by the PSII repair cycle, or the series of proteases that reside in the chloroplast (Takahashi and Badger, 2011). However, the membrane system, namely the ER membrane, is also a focus of attention, not only due to the comparative studies in other eukaryotes, but also because it is in the ER that proteins enter the secretory pathway, suffer protein fold, and are assembled correctly. Besides, the ER's morphological organisation is affected by stress, turning it into a very

4.1. – Metalloproteases and their role in plant abiotic stress

stress-responsive organelle (Denecke, 2001). Even though it is known that abiotic stress affects proteolytic processes, more research is needed to specifically unveil how chloroplastidic proteases are regulated by abiotic stress and play an important role in the plant facing those stresses.

4.2.

"Functional characterisation of EGY3"

4.2.1. Rationale behind a bioinformatics and web-based data-mining strategy for the identification of new abiotic stress determinants

The *Arabidopsis thaliana* genome was the first plant genome to be sequenced, in a worldwide effort that projected plant research into a post-genomic era (The Arabidopsis Genome Initiative, 2000). Quoting directly from the MASC Report (2002): “the mission of the project is to identify all of the genes by using a functional biological approach leading to determination of the complete sequence of the Arabidopsis genome by the end of this century” in order “to understand the physiology, biochemistry, growth and developmental processes of a flowering plant at the molecular level, using Arabidopsis as an experimental model system”. The initial mission was to attribute a function to every gene in Arabidopsis by 2010 (MASC Report, 2002). Despite the fact that many of the initial objectives were accomplished within the following decade, the main objective of unravelling gene function remains largely undone. Still, in 2010 almost 96% of Arabidopsis unique genes (27543 of a total of 28691) comprised at least one sequenced insertion element, and 62% (17721 of a total of 28691) presented one or more confirmed homozygous insertions (MASC Report, 2010). Furthermore, 95% of the genes (27257 of a total of 28691) present confirmed expression, through different transcriptomics approaches, including small RNA sequencing projects (MASC Report, 2010). As a consequence, loss-of-function analysis of a gene-of-interest (GOI) is now fairly straightforward and the effort made in recent years has produced toolkits for the assessment of the biological function of thousands of annotated genes. As previously stated, the model plant *Arabidopsis thaliana* displays unarguable attributes for gene function discovery: a fast and compacted growth, small and compacted genome, simplicity of genetic manipulation, together with the efficient ability to be transformed by *Agrobacterium*. Moreover, a variety of resources for functional discovery (databases, analysis tools, cDNA, seed and mutant line collections) became available as a result of the previously mentioned post-genomic advancements in research (Feng and Mundy, 2006; Koornneef and Meinke, 2010).

Phenotype-centred approaches have become fundamental in functional discovery, namely through forward and reverse genetic strategies (Alonso and Ecker, 2006). Presently, forward genetics, in which a mutagenised population is screened for a phenotype of interest, has resulted in a near saturation of the most obvious phenotypes. Meanwhile, reverse genetics, which goes from gene selection to detection of a visible phenotype, is currently the most widespread methodology. This progress was due to the use of insertion mutants using biological vectors

(mainly T-DNA or transposons). In the presence of a sequenced genome, these vectors combine a low insertion-per-plant ratio to fast mutation mapping (Feng and Mundy, 2006; O'Malley and Ecker, 2010).

As mentioned before, abiotic stress has been the focus of intense research, mainly because of current climate changes that endanger worldwide agricultural yield production and produce annual losses of billions of dollars. The main abiotic stresses affecting plants have been thoroughly studied in *Arabidopsis*, including drought, salinity, heat, cold, chilling, high light intensity, all with a main common element that is water availability (Mittler and Blumwald, 2010). Nevertheless, knowledge on the capacity of plants to cope with all these stresses is still clearly insufficient, as many stress-related genes remain functionless in the whole-plant concept (Ahuja *et al.*, 2010). Research in *Arabidopsis*, with its high-throughput and omics-based approaches, has been key to understanding molecular processes and networks involved in stress tolerance, quickly translating this knowledge onto other plant species (Century *et al.*, 2008; MASC Report, 2010).

Microarray transcriptomics technology has become a standard method to analyse gene expression at the whole-genome level, permitting the detection of qualitative expression differences that result from the response to various stimuli in different plant species (Wullschleger and Difazio, 2003). In *Arabidopsis*, insight into stress responses has been significantly improved by expression profile analysis, through which several genes playing a role in wounding, cold, salt, drought and heat stresses have been identified (reviewed by Cheong *et al.*, 2002; Rizhsky *et al.*, 2004; Oono *et al.*, 2006; Sreenivasulu *et al.*, 2007; Swindell *et al.*, 2007). The huge amount of *in silico* data available for *Arabidopsis*, including systematic microarray analysis of fundamental plant processes or the profiling of various knockout mutants, help to assess relations between gene expression patterns and stress responses, allowing for a narrow search of putatively determinant genes (Kilian *et al.*, 2007). The *AtGenExpress* project, a systematic transcriptomics study in *Arabidopsis* conducted using the Affymetrix ATH1 microarray chip is a good example of systematic transcriptomics analysis, with a transcript profiling of ~24,000 *Arabidopsis* genes using the Affymetrix one-colour microarray gene expression technology (Redman *et al.*, 2004).

The present work initiated with a straightforward data-mining strategy (Silva-Correia *et al.*, unpublished data), with the purpose of identifying new abiotic stress determinants, namely previously unresolved heat stress-related genes. A reverse genetics-based strategy using gain- or loss-of-function mutant analysis was subsequently pursued. An overview and brief description of the gene discovery process is outlined in figure 4.7.

«Identification of novel, functionally unresolved chloroplastidial proteins that would be specifically involved in the heat stress response»

No. of genes	
~24000	1. Access to raw microarray data for process of interest
	◁ Total pool of genes in the ATH1 array
471	2. Identification of differentially expressed genes during 3h heat stress imposition
	◁ Define differential expression cut-off (<i>2x in shoots</i>); Define minimum expression value cut-off (<i>500 pixel count</i>);
147	Identify highly up-regulated genes; Cut-off genes that were singled out only once as differentially expressed;
	◁ Analyse hierarchical clustering of expression patterns;
	3. Analysis of functional state-of-the-art
36	◁ Cross-reference with 5208 plastid-predicted genes (<i>Chloroplast 2010</i>);
22	◁ Perform GO categorisation and literature analysis (<i>functionally unresolved genes</i>);
13	◁ Exclude HSPs or HSP-like and putative chaperonines;
1	◁ Identify a gene-of-interest based on potential functional involvement in the heat stress response (<i>EGY3; At1g17870</i>);

Figure 4.7. - Overview of the transcriptomics-based strategy behind the identification of a novel and unresolved gene-of-interest, putatively involved in the heat stress response.

4.2.2. Identification of *EGY3* as a putative heat-stress determinant

The initial objective was the identification of a novel, functionally unresolved chloroplastidial protein that would be specifically involved in the heat stress response. Bulk transcriptomics data from the *AtGenExpress: Abiotic stress series*, consisting of the global spatial-temporal gene expression pattern of the Arabidopsis response to heat stress was retrieved from the web (Kilian *et al.*, 2007). A straightforward raw data analysis was adopted, based on the identification of heat inducible genes that (1) were differentially expressed and (2) possessed significant expression levels on shoots of Arabidopsis during a three-hour period of heat stress imposition. Analysis was performed on standard spreadsheet-based software (e.g. Excel). Given the one-colour nature of the ATH1 microarray, absolute expression values for each gene/probe set were represented in the form of pixel count. Differentially expressed genes were selected based on a two-fold cut-off value in the expression ratios between stress and control experiments.

Subsequently, genes were isolated that presented >500 pixel count for heat stress experiments, as previously described (Goda *et al.*, 2002; Moseyko *et al.*, 2002). This analysis was performed for each sampling time (0.25, 0.5, 1 and 3 hours). Genes that evidenced expression lower than 500 pixel count were cut-off, thus avoiding the lack of sensitivity in the signal/pixel count of low-expressing genes, which would inflate differential expression ratios and compromise the analysis based on a two-fold cut-off value of relative expression ratios.

Analysis revealed 471 genes that were subsequently subjected to selection of up-regulated genes, disregarding genes down-regulated upon heat stress. Genes displaying up-regulation at only one time point were also disregarded. With these 147 genes, a hierarchical clustering was performed using Multiple Array Viewer (Figure 4.8-A). This analysis is relevant to permit at this point a global analysis of the expression pattern of the 147 genes, eyeing potentially interesting patterns in the expression of up-regulated genes. It is worthwhile to refer that important genes might have been cut off at this stage, but in this way (defining a cut-off of two-fold and a minimum of 500 pixel) it was ensured that only highly up-regulated genes were selected that responded in a consistent manner to heat stress.

At this stage, an analysis of the state-of-the-art was initiated (Figure 4.7). Since the initial objective was to have a chloroplastial protein, the 147 genes were cross-referenced with a list of 5208 Arabidopsis genes predicted *in silico* to code for plastid-targeted proteins (Chloroplast 2010; URL no.9) (Figure 4.8-B). Only 36 proteins were predicted to be targeted to the plastids. The corresponding 36 genes were then analysed with GO categorisation (Figure 4.8-C) and literature analysis, to select putatively unresolved genes with great potential for functional characterisation. The state-of-the-art for each heat stress-associated gene was assessed by literature search in TAIR (URL no.5) and PubMed (URL no.27), giving preference to genes with low/inexistent functional knowledge, or whose implication in the heat stress response had not been proposed up to that point. GO analysis was performed using Genome Ontology at TAIR (URL no.34), allowing for an overview of biological processes or molecular functions of the 36 genes. As demonstrated in figure 4.8-C, 43% respond to stress, namely abiotic and biotic stimulus, which is consistent with the selection strategy for heat stress-related genes. Regarding the molecular function, 22% presented unknown function which also was of interest since unresolved genes were the target of the search. Taking also in consideration the existing literature on the genes, the number descended to 22 unresolved genes-of-interest.

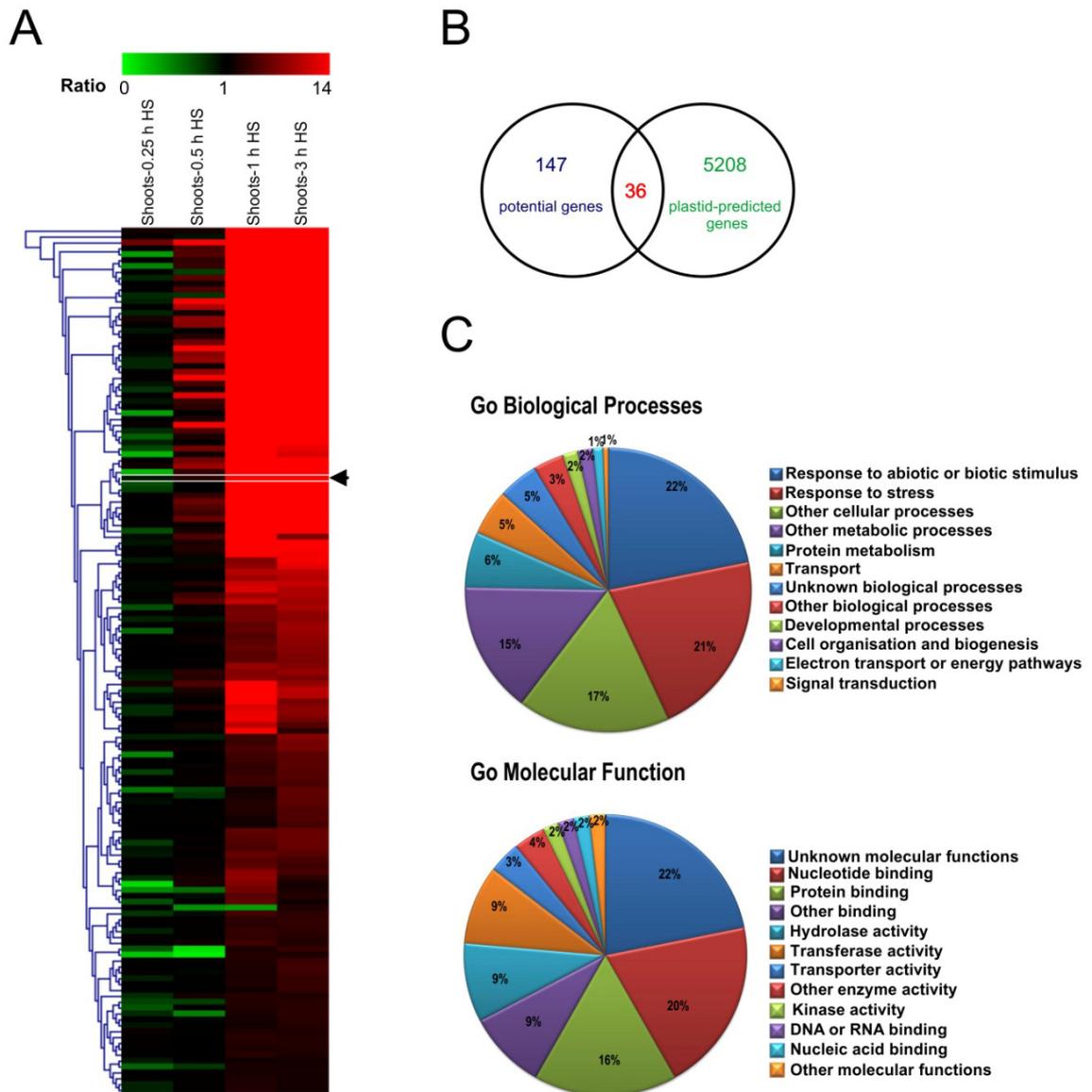


Figure 4.8. - *In silico* analysis of selected differentially expressed genes. A - Hierarchical clustering of 147 genes exhibiting up-regulation in shoots within a 3 h heat stress imposition period. The arrow indicates the gene singled out at the end of the strategy (*EGY3*). Values represent expression data from shoots in the *AtGenExpress Heat Series*, in the form of expression ratio in relation to control non-stressed plants. B - Venn diagram of cross-referenced plastid-predicted genes from Chloroplast 2010, with the 147 potential genes of the current analysis (see figure 4.7). C - GO categorization for biological processes and molecular function of the 36 unresolved genes, using TAIR (URL no.34).

For the subsequent selection, all HSP-, HSP-like- and chaperonin-coding genes were excluded, due to their immediate association with heat stress processes, that were more likely to already have or be the subject of ongoing functional characterisation. After this selection, 13 genes remained, which suffered a more thorough analysis of their expression pattern. Pixel count/relative expression values were analysed graphically in a time-line basis for each imposing stress time point. In the end a gene was chosen that evidenced a less constitutive expression and

demonstrated high inducibility, particularly during the heat exposure period, also evidencing a major relative ratio of induction; the selected gene was *EGY3* (At1g17870). To highlight the responsiveness of *EGY3* to abiotic stress, analysis in *Genevestigator - Electronic Northern (Abiotic stresses)* is depicted (Figure 4.9). *EGY3* was spotted as being induced very specifically to a few stress responses that included high light, drought, salt and osmotic. The expression pattern for some of these is shown in figure 4.10. It evidences how *EGY3* is very specifically induced during the 3 h heat stress imposition period, where it suffers a significant 70-fold plus induction in transcript levels (4.10-C). Meanwhile, during osmotic and salt stresses the transcriptional response seems to occur much later. This pattern highlights the potential of this gene in terms of importance to abiotic plant responses.

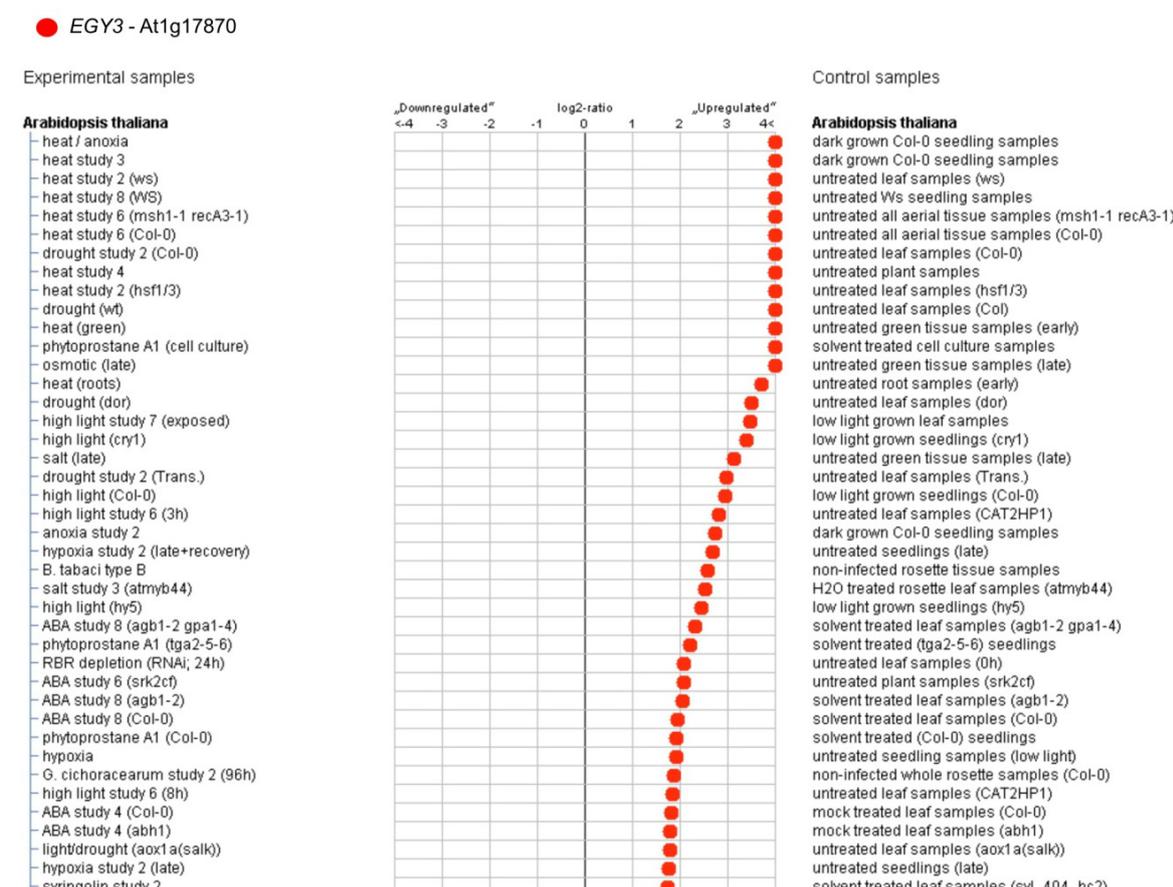


Figure 4.9. - Electronic Northern analysis through Genevestigator (URL no.11), of the *EGY3* transcriptional response to abiotic stresses in Arabidopsis.

In summary, the use of publicly available transcriptomics data coupled with a data-mining strategy that relied on the identification of differentially expressed genes under heat stress, followed by the gathering of functional information, lead us to the identification of gene, *EGY3*, coding a putative chloroplastial protein involved in heat stress. Using a reverse genetics approach,

functional characterisation of *EGY3* was subsequently initiated, to unravel/confirm the molecular, cellular and biological processes in which *EGY3* is involved.

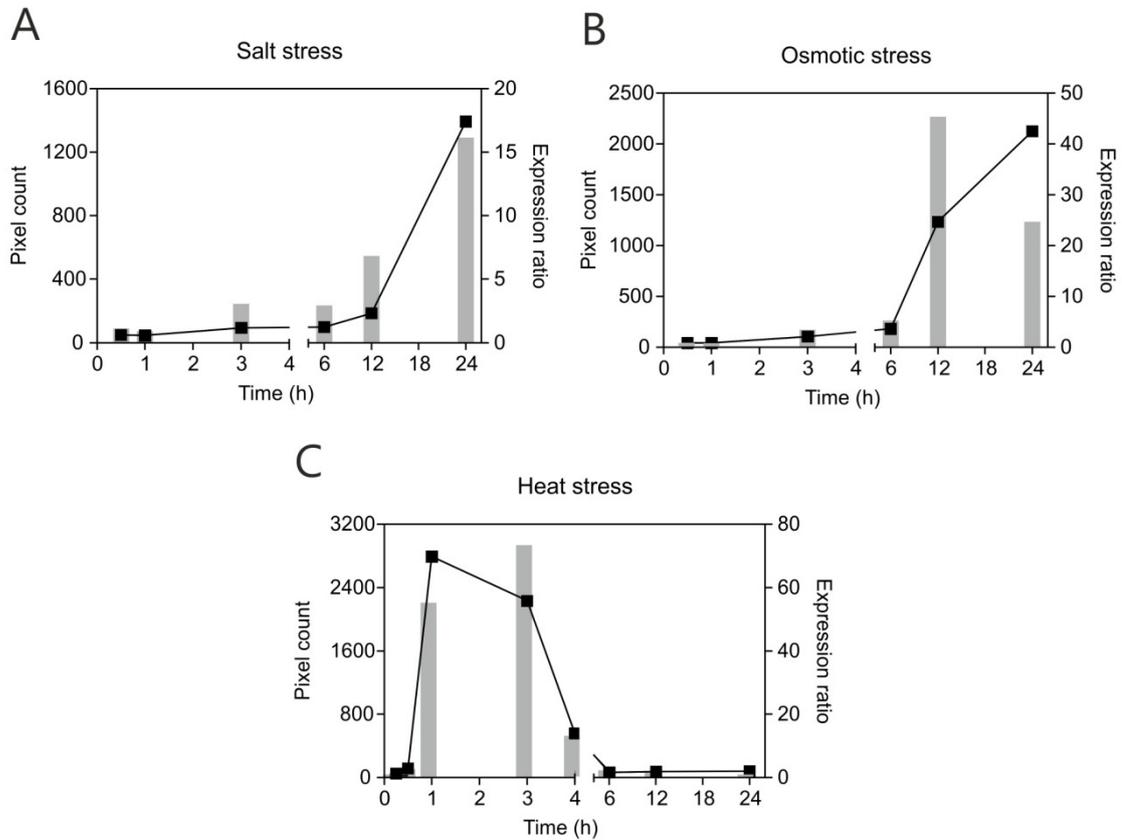


Figure 4.10. – *EGY3* expression analysis in shoots under stress conditions. Absolute expression levels (Bars; expressed in pixel count), and relative expression ratio (Closed squares; representing pixel count in the treatment/pixel count in the control) during a time-course exposure to salt stress (A), osmotic stress (B), as well as exposure to heat stress for 3 h followed by a 21 h recovery period (C). Data was retrieved from the *AtGenExpress Abiotic Series* (URL no.8).

4.2.3. Characterisation of the *EGY* gene family

Following the previously detailed gene selection, the *EGY* family in *Arabidopsis* (*EGY1* - At5g35220; *EGY2* - At5g05740 and *EGY3* - At1g17870) was studied more thoroughly *in silico*. A phylogenetic analysis was conducted, based on sequence homology using NCBI blastp against the *EGY1* a.a. sequence. After searching for *EGY*-like proteins, homologous sequences were retrieved, and a ClustalW alignment was performed. The corresponding phylogenetic tree is depicted in figure 4.11. Results show that the majority of the *EGY*-like M50 peptidases are mostly from the cyanobacteria phylum, being the rest from the kingdom Plantae (from Chorophyta to Gimospermic, Eudicots and Monocots, and even a Lycophyta) and only two sequences represent the kingdom Chromista.

4.2. – Functional characterisation of EGY3

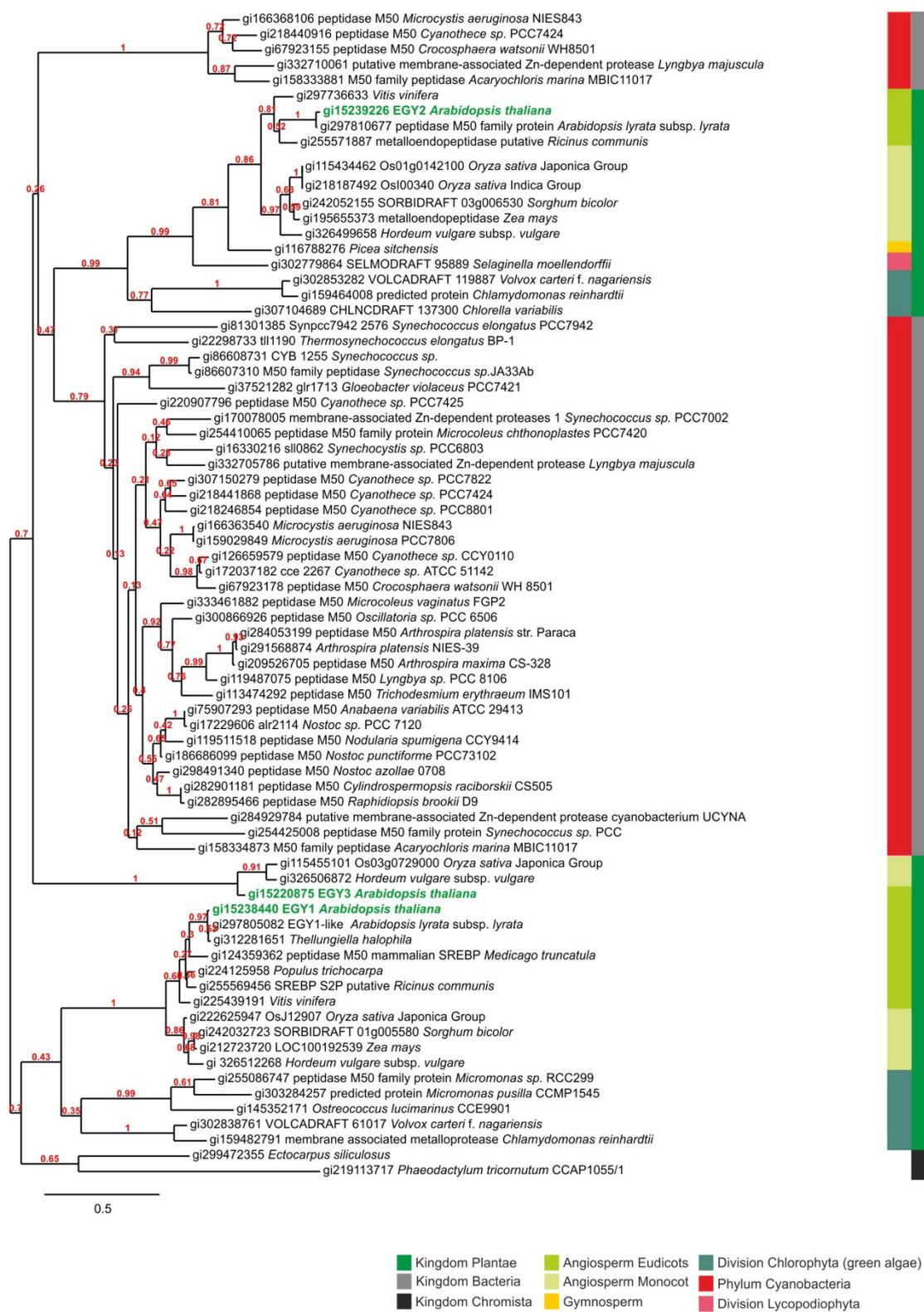


Figure 4.11. - Phylogenetic tree of EGY-like M50 peptidases. Sequences were retrieved following blastp homology search (NCBI, URL no.26) using the EGY1 a.a. sequence. The MEGA 5 software (Tamura *et al.*, 2011) was used to perform a ClustalW (BLOSUM) alignment of all sequences after exclusion of previously retrieved partial proteins. Phylogenetic reconstruction was performed using Maximum-likelihood with subsequent Bootstrap analysis (500 trees). The rate of aminoacid substitution was empirically calculated using the WAG model (URL no.29).

The kingdom Chromista was established as being distinct from Plantae and Protozoa because of the fact that their chloroplasts were acquired secondarily by enslavement of a red alga (Kingdom Plantae) and because of their unique membrane topology (Cavalier-Smith, 1981; Cavalier-Smith, 2010). Most organisms of this kingdom are algae containing chloroplasts with chlorophylls a and c, which are located within the lumen of the rough ER (in plants or in protozoan algae they are in the cytosol). It is said that all chromistan algae are evolutionary chimaeras between a eukaryotic host and a eukaryotic (probably red algal) symbiont (Cavalier-Smith, 1998). Regarding the Bacteria kingdom representatives, there seems to be a great diversity, which implicates a great importance of these proteins in this group of organisms, none withstanding protease importance in the remaining organisms.

The phylogenetic tree is subdivided primarily in five different clades. The most representative clade consists of EGY2-like proteins, which include, from the plant kingdom, Eudicotyledons, Monocotyledons, Gymnosperms, Lycophyta, and Chlorophytes, and the cyanobacteria from the Bacteria kingdom. A small clade of cyanobacteria-specific EGY-like proteins is phylogenetically related to EGY2-like proteins. Both clades seem to combine into the most ancestral branch of EGY-like proteins, as they are the only containing cyanobacteria, as well as ancestral plant organisms such as the Gymnosperms, Lycophyta, and Chlorophytes. This clearly suggests that EGY2 might be the *Arabidopsis* representative with the most ancient topology. The second most representative group is composed of EGY1-like proteins, which includes both Monocot and Eudicot plants and green algae (Chlorophyta). A small clade of the kingdom Chromista appears phylogenetically more related to EGY1-like proteins, however clearly more distant in terms of evolution, they seem to have conserved some EGY1-like proteins. EGY3-like proteins appear as a small but distinct clade of angiosperm-specific members, which are phylogenetically more related to EGY2-like proteins. Results suggest that this branch has been the product of recent evolution, which may result in a putative subfunctionalisation or neofunctionalisation of these proteins, after a recent duplication event from an ancestral EGY2-like gene.

From a topological point of view, members of the EGY-like gene family seem to vary in the number of transmembrane domains (TM). To exemplify, analysis was carried out in both *Arabidopsis* and *Oryza sativa* (rice) EGY family members, present in the three major subclades. For EGY1-like proteins the sequences presented seven TM in rice and eight in *Arabidopsis*, for EGY2-like proteins they had five TM in rice and three in *Arabidopsis*, and finally for the EGY3-like proteins the number of TM was of eight (rice) and six (*Arabidopsis*). EGY2 seem to have the lower number of TM within paralogs, but while EGY1 has a higher number of TM in *Arabidopsis*

compared to its rice paralog, the opposite takes place in EGY3-like paralogs. Nevertheless, TM number is based on *in silico* prediction, and *in vivo* the protein topology may be marginally different. None withstanding it could be said that topology is somewhat conserved.

EGY1 has an *Arabidopsis lyrata* homolog and is closer to other M50 peptidases in Eudicots and in some Monocots. EGY2 and EGY3 also have an *Arabidopsis lyrata* homolog, though for EGY3 it is not depicted in the phylogenetic tree because sequences were retrieved based on the EGY1 a.a. sequence. The fact that such a variety of kingdoms exists also reflects the conservation of these proteins and their function and likely importance in photosynthetic organisms. A most significant finding is the fact that EGY-like proteins seem to be restricted to organism displaying photosynthetic activity, which implicates these proteins as ancestral, but highly specific proteins, which are likely to be functionally connected to photosynthesis.

A more thorough topological analysis of Arabidopsis EGY family members was carried out, in order to characterise the proteins as to their putative location within the cell, number of transmembrane domains and motifs that could enlighten their function. Analysis was performed using InterProScan (for Pfam domains) (URL no.25), TMHMM server 2.0 (for detection of transmembrane domains) (URL no.20) and ChloroP 1.1 server (for chloroplast transit peptide detection) (URL no.22) (Figure 4.12). All three proteins evidence predicted transit peptides that direct them to the chloroplast, as described also by Chen *et al.* (2005). All proteins have transmembrane domains but differ significantly in their number, EGY1 has eight predicted transmembrane domains, while EGY2 has three, and EGY3 six. Only EGY1 and EGY2 have the predicted Pfam domain representing the motif of a peptidase of the M50 family (Rawlings, 1999). This Pfam domain (PF02163) is lacking in the EGY3 protein because of the absence of the HEXXH motif, characteristic of metalloproteases. The NPDG motif is also slightly different in EGY3, which presents an aminoacidic alteration (D to an E). This could mean no severe alteration since they are part of the same group of negatively charged nonpolar aminoacids. However, looking at the literature, studies on SpoIVFB, a S2P family member from *Bacillus subtilis*, demonstrated that a substitution on the aspartic acid (D) of the NPDG motif to a glutamic acid (E) impaired the processing of a transcription factor during the sporulation process (Rudner *et al.*, 1999).

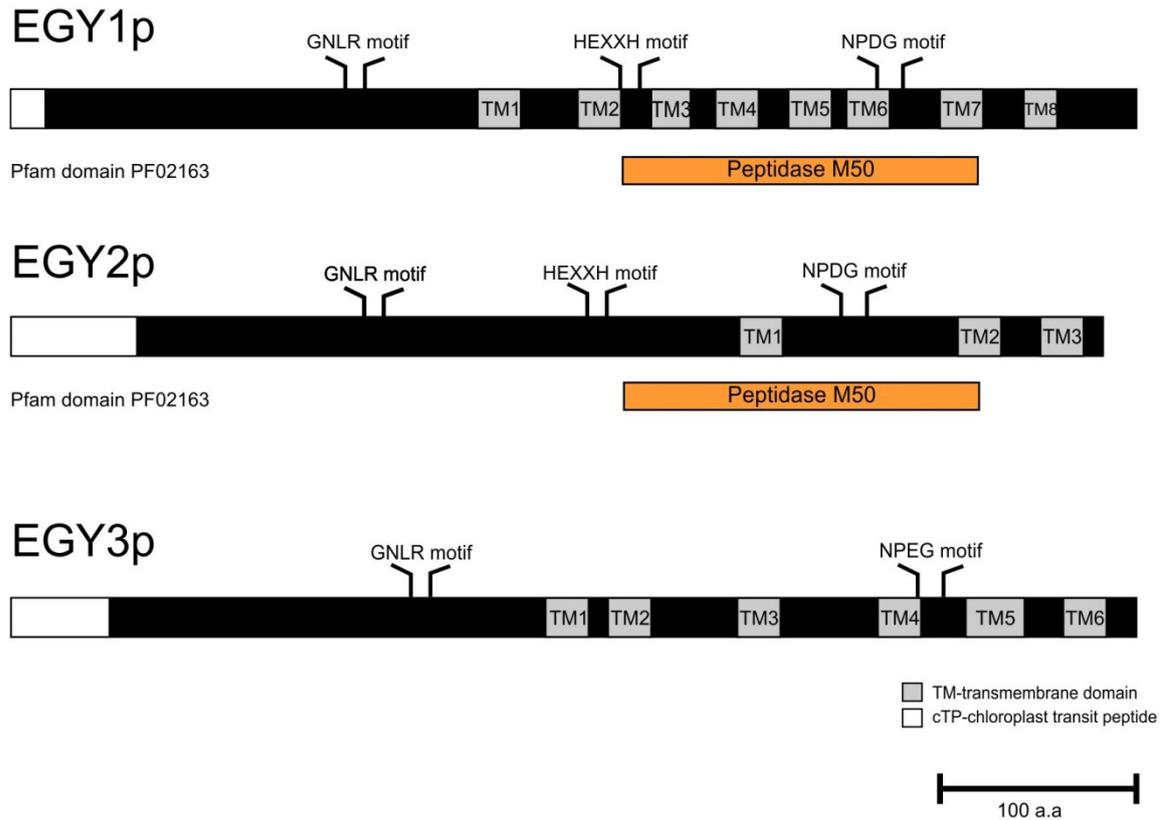


Figure 4.12. - *Arabidopsis thaliana* EGY protein topology. Pfam domains were determined by InterProScan (URL no.25), transmembrane domains were defined by TMHMM 2.0 server (URL no.20) and cTP (chloroplast transit peptide) by the ChloroP 1.1 Server (URL no.22).

Studies in the human S2P on the same motif reinforced the importance of the conserved aspartic acid, consistent with the possibility of it being the third metal ligand, present in the NPDG motif, whereas the other two metal ligands (the histidines; H) are present in the HEXXH motif (Rudner *et al.*, 1999; Zelenski *et al.*, 1999). With the lack of the HEXXH motif and the presence of glutamic acid instead of aspartic acid in EGY3, it is highly unlikely that EGY3 displays metalloprotease activity. Interestingly, neofunctionalisation of EGY3-like proteins might have occurred, since both monocot EGY3-like proteins lack the HEXXH and NPDG motif, having instead the SEIAT and NPEG motifs. The question arises then, of why these three EGY3-like proteins conserved such a high-impact change in protein topology. Finally, the GNLR motif is unique to all three EGY's and was previously designated a unique signature motif of EGY-like proteins, being present in higher plants and cyanobacteria (Chen *et al.*, 2005). The significance or function of this motif is yet to be identified.

4.2.4. EGY3 expression pattern analysis

In silico analysis based on microarray data allowed comparison between *EGY1*, *EGY2*, and *EGY3* expression data available in the Arabidopsis BAR database. Whole tissue expression analysis of these genes showed that *EGY1* is the most expressed in almost every tissue, followed subsequently by *EGY2* and *EGY3* (Figure 4.13). However, the values of absolute expression are not very high (below the 500 pixel count). Both *EGY1* and *EGY2* are more expressed in the young rosette, first stages of leaf development, and pedicels; while *EGY3* is more expressed during the formation of the embryo/seed, and in some flower stages, being most highly expressed in the dry seed. Concerning the potential involvement of EGY family members in abiotic stress responses, analysis of the BAR-eFP browser/e-Northern Expression and Genevestigator in response to various abiotic stresses confirmed that *EGY3* is stress-responsive, as previously stated. However, *EGY1* and *EGY2* showed no induction in transcript levels (data not shown). Though *EGY3* has lower expression in standard growth conditions in comparison to other family members (Figure 4.13), it seems that it is the EGY family member that responds to abiotic stresses, albeit the fact that others homologues might be more responsive to non-abiotic stresses. This could indicate some specialisation and a specific involvement of *EGY3* in the abiotic stress response, namely the heat stress response, for which it has higher induction ratios.

Functional characterisation of *EGY3* was initiated with the analyses of the spatial expression pattern of *EGY3* at the tissue level. Promoter::GUS constructs were generated by amplification from Arabidopsis gDNA of the putative promoter sequence defined by the AGRIS database (URL no.14), which was inserted into the pCAMBIA1303 plant expression vector (Figure 4.14-A,B). With this construct, stable transformants were obtained, homozygous lines were isolated after three generations, and GUS histochemical analysis in different tissues was performed in a line selected for strong GUS staining (Figure 4.14). Results showed a high level of expression for *EGY3*, mainly at the seedling stage. In 6-day-old seedlings, staining is evident in cotyledons, stomata, hypocotyl embryonary root, meristematic areas and internally in the root, but not in the root apical meristem (Figure 4.14-C). In 10-day-old seedlings, an identical pattern was observed, also with primordial leaves, trichomes, roots, and especially lateral root formation regions evidencing GUS staining (Figure 4.14-D). Adult plants (1-month-old) however, show no staining in both rosette and caulinar leaves, but they did present staining at the extremities of the siliques, and on flower sepal and carpel structures (Figure 4.14-E).

Arabidopsis eFP Browser-Developmental map

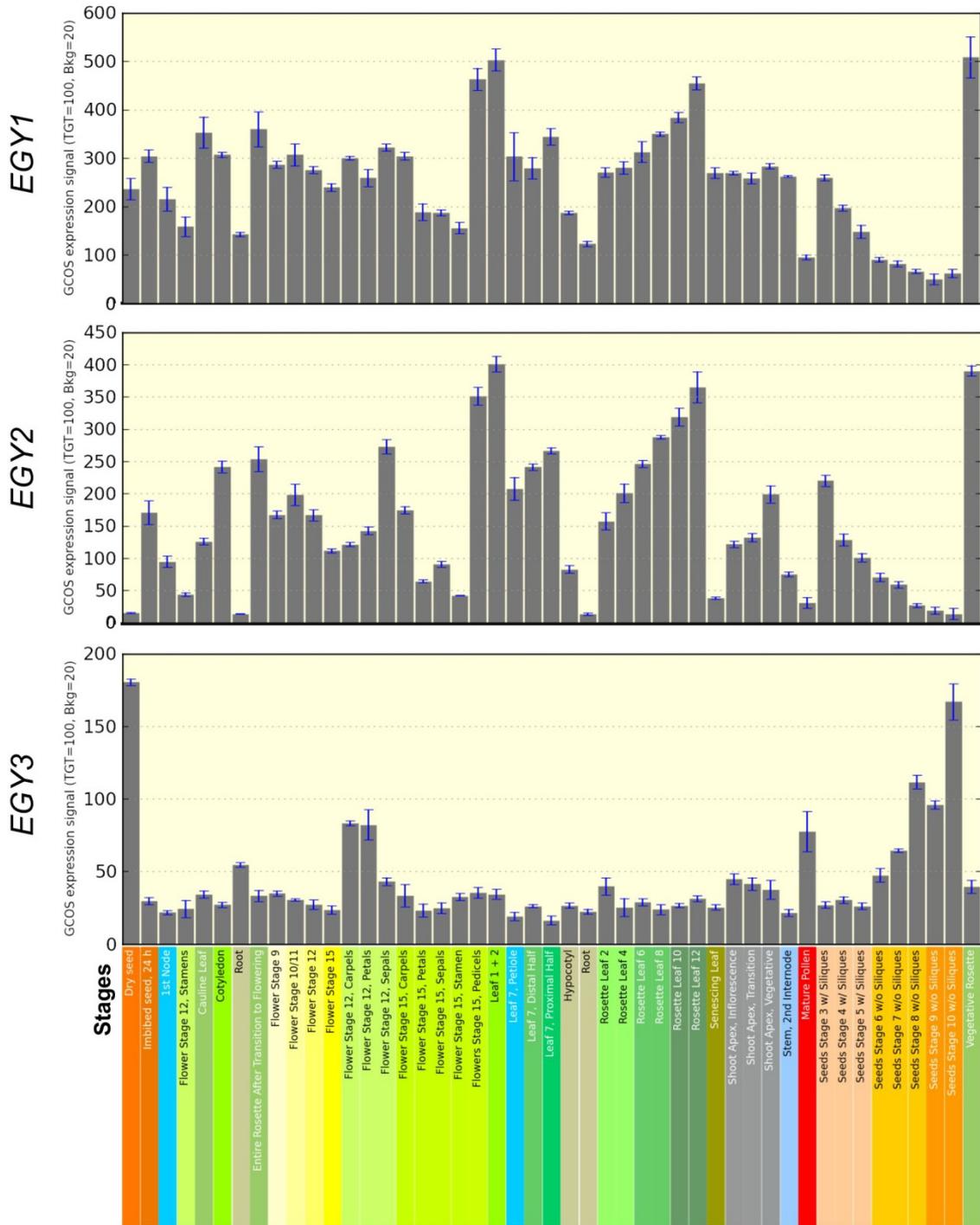


Figure 4.13. - Gene expression pattern of Arabidopsis EGY family members using eFP Browser from BAR (URL no.10) (Winter et al., 2007). Developmental maps of *EGY1*, *EGY2* and *EGY3* are depicted, with stages being described at the bottom of the chart. Total expression values and standard deviations are listed in table A.4 (*EGY1*), A.5 (*EGY2*), and A.6 (*EGY3*) of the Appendix I.

4.2. – Functional characterisation of EGY3

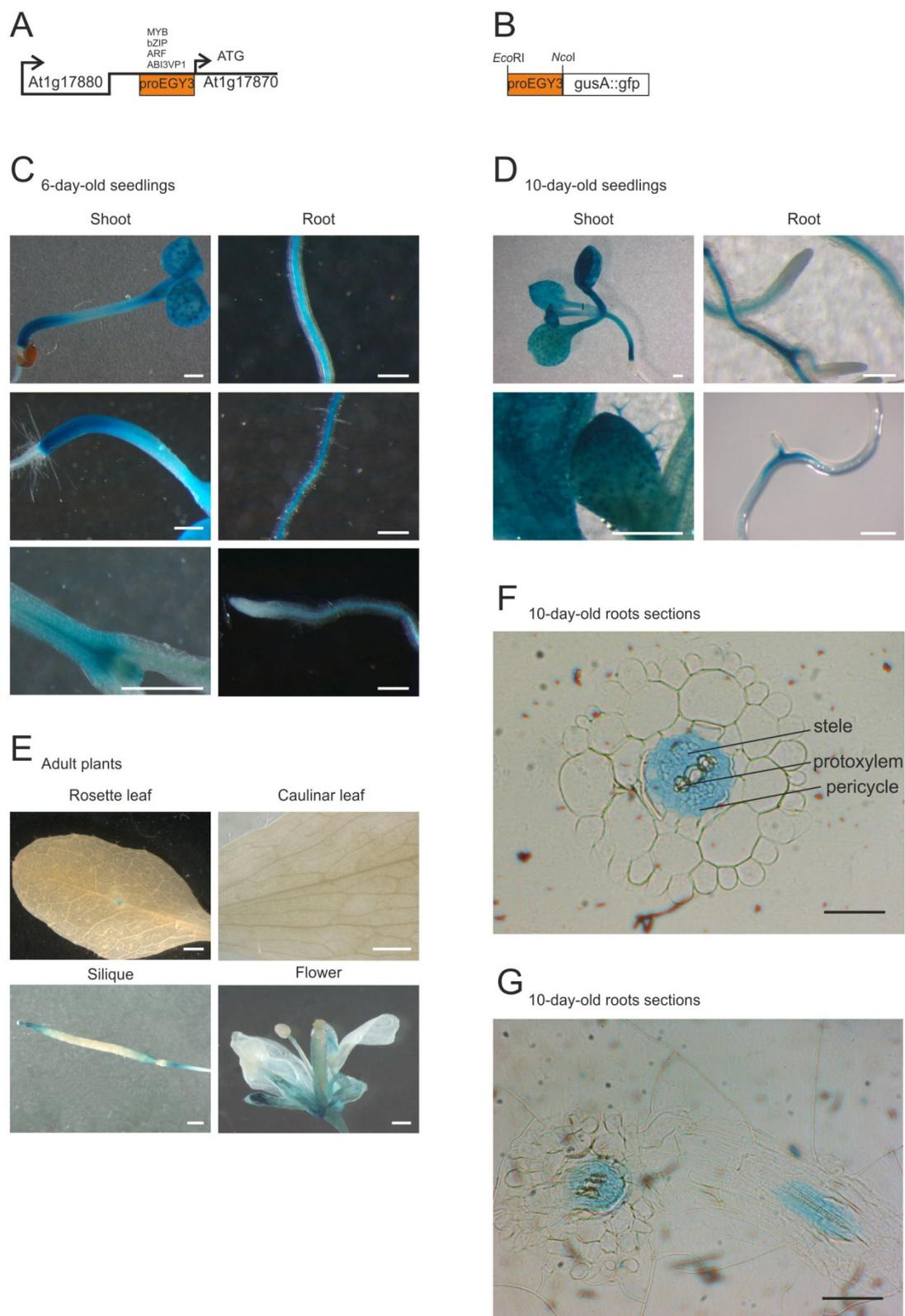


Figure 4.14. - Histochemical analyses of GUS activity in *proEGY3::GUS*. A – Schematic representation of the *EGY3* promoter and its *cis*-elements. B - Schematic representation of the *proEGY3::GUS* construct in pCAMBIA1303. C – Shoots and roots of 6-day-old seedlings. Scale bar represents 0.5 mm. D - Shoots and roots of 10-days-old seedlings. Scale bar represents 1 mm. E – Mature leaves (caulinar plus rosette), siliques and flowers of 1-month-old plants. Scale bar represents for leaves 1 mm. F and G - Root transverse section of 10-day-old seedlings. Scale bar represents 0.25 μ m (Dolan *et al.*, 1993; Taiz and Zeiger, 2003).

Results evidence the presence of EGY3 internally within the root, particularly the in the pericycle cells, stele, but not in the protoxylem, as seen in the transversal plane sectioning in figure 4.14-F. Resolution at the cellular level is important because lateral or branch roots arise from pericycle cells in the mature regions of the root; cells divisions in the pericycle establish secondary meristems that grow out through the cortex and epidermis to establish a new growth axis. The so called pericycle founder cells are primed xylem pole pericycle cells that undertake initiation by auxin activation, turning them into pericycle founder cells that undergo the first anticlinal and asymmetric divisions (Taiz and Zeiger, 2003; Péret *et al.*, 2009). In figure 4.14-F, it can be observed that only the main root is present, whereas in the 4.14-G there is some sectioning of the lateral root. Staining is present in the surrounding of the tracheal elements of the xylem, probably in the stele as well as in the primary root, despite the small portion stained. This and the fact that expression is high in lateral root formation in 4.14-D, might indicate a possible involvement of EGY3 in lateral root formation, especially in the beginning of the formation of the putative pericycle founder cells, and initial divisions. Staining in the stele may implicate some importance in vascular tissue formation or differentiation, or even in the maturity of the xylem until it becomes dead tissue, since EGY3 might be involved in protease activity (Taiz and Zeiger, 2003).

4.2.5. Isolation of *egy3* loss-of-function mutants

To functionally characterise EGY3, a reverse genetic approach was conducted, making use of the enormous collections of Arabidopsis T-DNA insertion mutants presently available (Alonso and Ecker, 2006). A search for mutant lines potentially interrupting EGY3 was carried out in SALK T-DNA Express: Arabidopsis Gene Mapping Tool (URL no.16). Various parameters were taken into consideration, namely the background ecotype (preferably Col-0) and the location of the T-DNA, avoiding promoter and intron regions in order to increase the probability of a knockout, and selecting for interruptions as upstream in the coding region as possible. In the case of EGY3, a SALK_042231 line was selected, introducing a T-DNA insertion in the 5th intron of EGY3. Schematic representation of EGY3 gene is depicted in figure 4.15-A, in which the selected T-DNA line and other aspects that are considered relevant to the experimental work and the discussion are also depicted. EGY3 contains six exons and five introns, as well as a small promoter of 204 bp, predicted by AGRIS (Davuluri *et al.*, 2003) to be the intergenic region and including 17 nucleotides of the 3'UTR of the adjacent gene At1g17880.

Mutant seeds were ordered from the public seed stock centre NASC (URL no.3), that provided seeds in the form of a heterogenous population that could contain insertions in none, one

or both alleles. Therefore, a diagnostic PCR was required in order to genotype homozygous plants for the insertion allele (Figure 4.15-B). The diagnostic PCR was performed with three simultaneous primers: LP and RP that anneal to the *EGY3* sequence plus the LbB1 primer of SALK lines, which anneals to the left border segment of the inserted T-DNA, amplifying outwards. In light of this, and knowing that LP-RP amplification consists of 1012 bp, it was estimated that the LbB1-RP amplification would be between 440-740 bp. As shown in figure 4.15-B, LbB1-RP amplification is above 500 bp, which is within the expectable size. In this case, the coordinate B7 is a wild-type plant, C1 a homozygous insertion mutant, and D1 a heterozygous insertion mutant. Homozygous mutant plants were thus obtained and subsequently designated *egy3-1*. In order to confirm the knockout mutant, semi-quantitative RT-PCR analysis was performed (Figure 4.15-C).

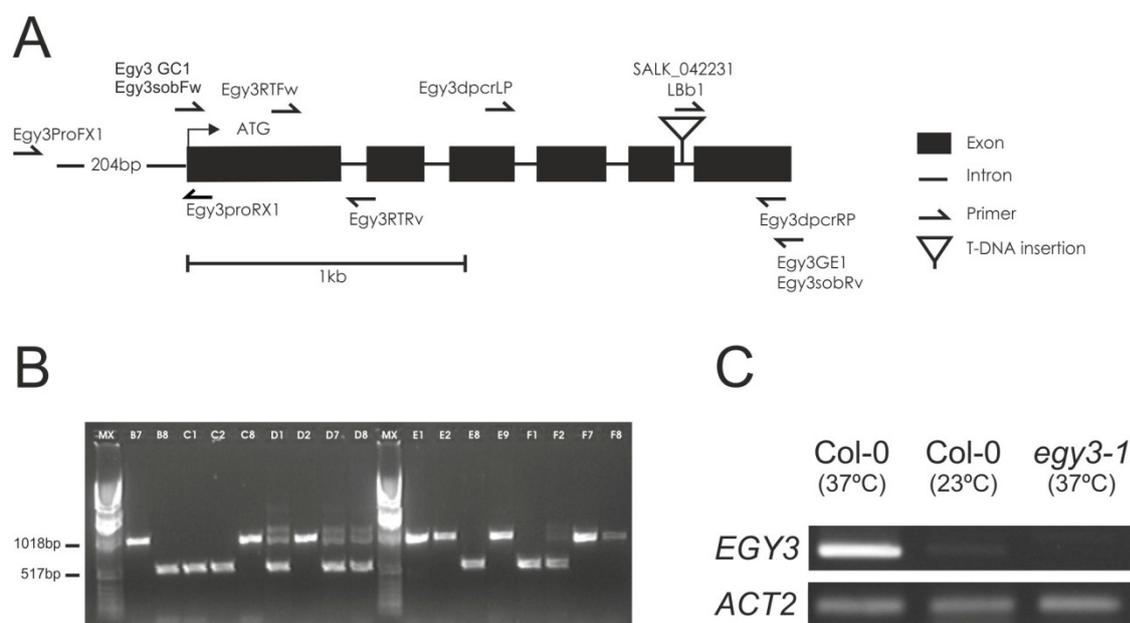


Figure 4.15. - Establishment of a loss-of-function *egy3-1* mutant. A – Schematic representation of *EGY3* (At1g17870). Disruption by the T-DNA occurs downstream of the ATG codon at 1737 bp. The positions of gene-specific PCR primers used for diagnostic-PCR genotyping of the T-DNA insertion (*EGY3* dpcr LP and RP) and RT-PCR analysis (*EGY3* RTRv and Rv), among other primers, are represented by arrows. Exons are represented by dark boxes, introns by dark lines and the T-DNA insertion by a triangle. Upstream of the ATG is the indication of the estimated promoter size, as predicted by the AGRIS database (URL no.14) (Davuluri *et al.*, 2003). B - Diagnostic PCR genotyping of *egy3-1* (SALK_042231) insertion mutant from a heterogeneous mutant population. Electroforetic analysis of *EGY3* using samples from a multiplex PCR using the following three primers: LP, RP (left and right primers, respectively) and LbB1 (left border specific for the SALK T-DNA lines). C - Semi-quantitative RT-PCR analysis of *EGY3* transcript levels in seedlings of Col-0 in both normal growth at 23°C and after heat stress imposition for 3 h at 37°C, as well as *egy3-1* mutant after heat stress imposition for 3 h at 37°C; *ACT2* (At3g18780) was used as constitutive gene to normalise gene expression. MW - Molecular Marker MassRuler DNA Ladder Mix.

The *egy3-1* primers were designed to amplify the region between the end of the 1st exon and the beginning of the 2nd exon, so as to amplify an intron in the case of gDNA contamination. In the present case, only the *EGY3* cDNA band of about 295 bp was present, and the mutant

displayed no transcript confirming the knockout of this allele. At the same time, the Col-0 14-day-old seedlings were used to check for heat stress-induced expression. The *in vitro*-grown seedlings were placed at 37°C for a period of 3 hours and then *EGY3* expression was analysed. Though this is only semi-quantitative and comparison is limited, an induction is clearly visible, which suggest that *EGY3* is indeed induced by heat, confirming the microarray data for which the gene was chosen.

4.2.6. *In vivo* developmental and heat stress phenotype characterisation of *egy3-1*

Plants from genotyped Col-0 and *egy3-1* backgrounds were grown to obtain a synchronised bulk seed stock. They were subsequently screened for a series of phenotypes both in plate (seedling stage) and in soil (rosette stage). An initial developmental phenotype analysis was conducted, based on Boyes *et al.* (2001). This paper described a method for the rapid discovery of gene function in development, by establishing a high-throughput phenotypic analysis based on a series of defined growth stages, which are developmental points generated for a morphological collection of data. This collection process was divided into two-stages based on growth and development, over the entire life-cycle of *Arabidopsis*. The first stage characterises the early seedling growth on vertical plates for a 2-week-period, while the second is based on extensive measurements from plants grown on soil for a period of ~2 months. With this methodology slight differences can be spotted that could be a result from genetic variation and/or environmental stress (Boyes *et al.*, 2001).

This method was simplified and used to analyse the role of long term heat stress on plant development in the *egy3-1* mutant, on plants grown on soil. It involved measurements of the rosette radius and the number of leaves per rosette in 3-week-old and 6-week-old plants. Also, a time-line analysis was performed to evaluate the first day of flowering of each plant, so as to observe differences that could indicate early or late flowering phenotypes. These parameters were measured in plants growing in long-day photoperiod at 23°C and at 28°C (Figure 4.16). As results demonstrate, there are no significant differences with 3-week-old plants either in the number of leaves per rosette (figure 4.16-A) or the rosette radius (figure 4.16-B). The *egy3-1* mutant did not display significant differences from its wild-type at 23°C or 28°C, and differences between 23°C and 28°C are not significant at this stage. The same does not occur after six weeks, when plants

had significant differences both in the number of rosette leaves (figure 4.16-C) and rosette radius (figure 4.16-D).

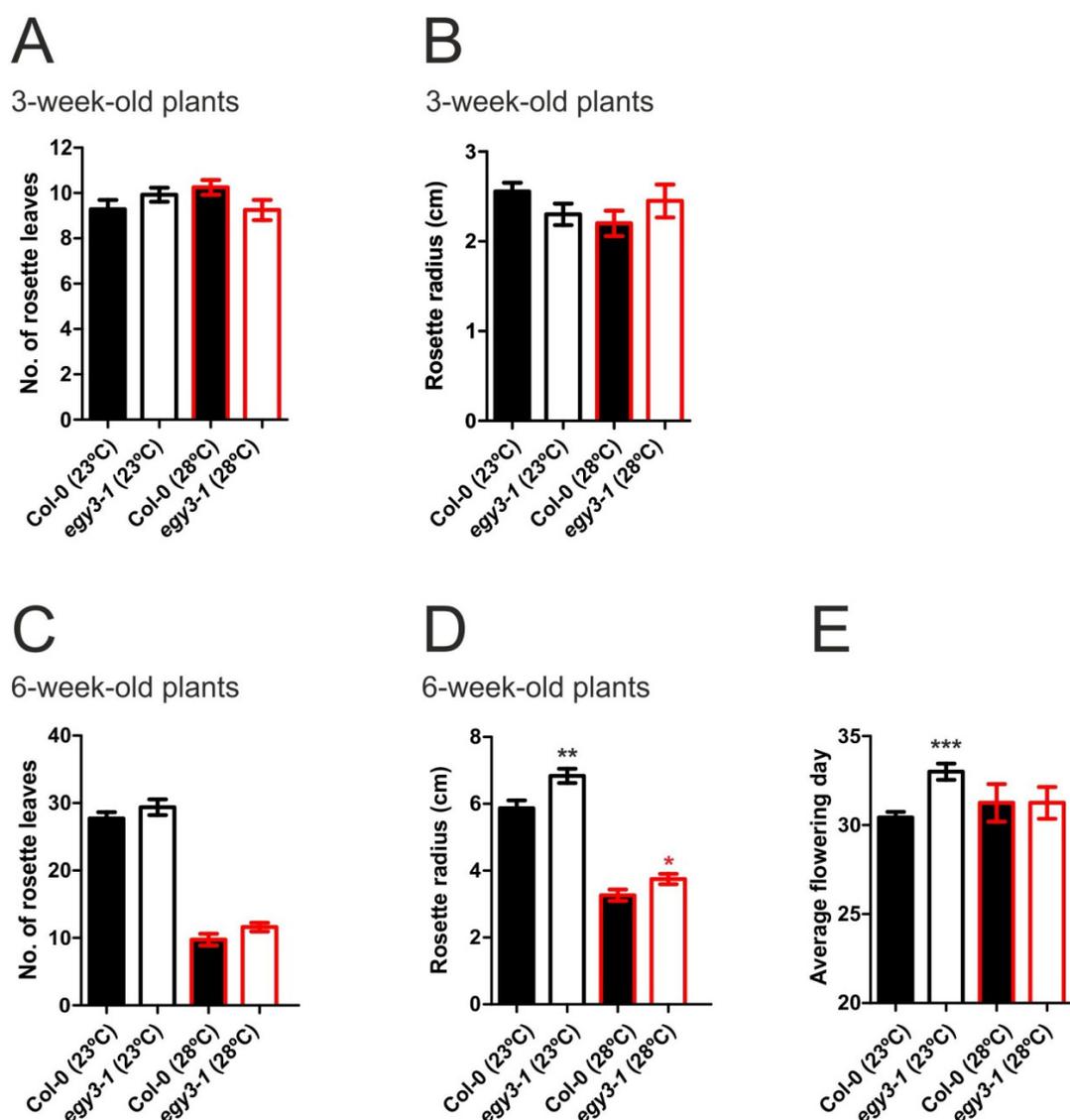


Figure 4.16. - Long term developmental phenotype analysis of Col-0 and *egy3-1* in long-day photoperiod. Number of rosette leaves in 3-week-old (A) and 6-week-old (C) plants. Rosette radius in 3-week-old (B) and 6-week-old (D) plants. Average flowering day (E). All measurements and data were collected in Col-0 and *egy3-1* plants growing at normal temperature conditions (23°C), and in plants subjected to long term heat stress (28°C). Error bars represent SEM (n ≥ 12). Asterisks represent significantly different levels between genotypes within the same growth conditions (student t-test: *, P < 0.05; **, P < 0.01; ***, P < 0.001).

As a consequence of stress, plants at 28°C are significantly weakened when compared to those at 23°C. The long term stress causes a pressure in development and plant growth, and the plant's effort is directed to the bolt and the production of flowers for survival (Balasubramanian *et al.*, 2006). This is corroborated by the fact that, although having a rosette with smaller and fewer leaves, the average day of flowering is very similar to the wild-type at 23°C (Figure 4.16-E). In this

experiment the *egy3-1* mutant presented phenotypic alterations, with 6-week-old mutant plants having a significantly higher radius than the wild-type (Figure 4.16-D), both at 23°C and 28°C, though differences are higher at 23°C. Also, the *egy3-1* mutant displayed a delayed flowering, since its average day of flowering was significantly higher than its wild-type at 23°C (Figure 4.16-E). Differences in flowering between the two genotypes were not observed at 28°C. When looking at plant morphology throughout their life-cycle, as is depicted in figure 4.17, one can see that at the seedling stage the *egy3-1* mutant has no differences in root size or development when compared to the wild-type (Figure 4.17-A). However, 3-week-old *egy3-1* mutants do appear smaller (Figure 4.17-B), though after a week (Figure 4.17-C) they gain the characteristics of the wild-type. As we can see in figure 4.17-D, Col-0 plants develop earlier bolting and flowering than *egy3-1*, which corroborates results from figure 4.16. In fact, by observing figure 4.17 it is clear that *egy3-1* appears to have more leaves or at least a bushier morphology than Col-0, which senesces earlier than *egy3-1*. EGY3 may play a relevant role in plant development, promoting a shorter length of the life cycle, as *egy3-1* plants have late flowering, bigger adult plants and generally a longer life cycle. Considering that heat stress accelerates the life cycle and hence the flowering time (Balasubramanian *et al.*, 2006), EGY3 may be involved in this mechanism during heat stress.

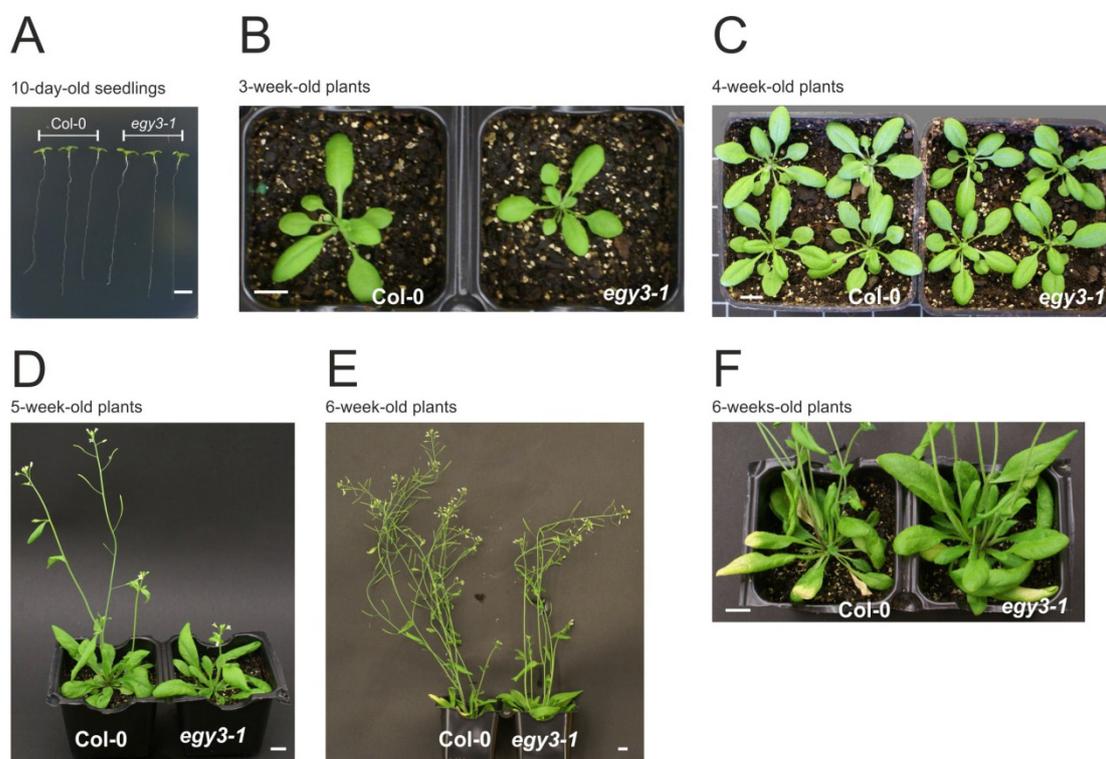


Figure 4.17. – The *egy3-1* mutant morphology throughout its life-cycle. Ten day-old seedlings (A), 3 week-old plants (B), 4-week-old plants (C), 5-week-old plants (E) and 6 week-old-plants (F) of *egy3-1* mutants and its wild-type Col-0. Scale bars represent 1 cm.

4.2.7. Heat stress-related phenotype analysis of *egy3-1*

As previously stated, *EGY3* was selected for its very specific heat-induced responsiveness, based on 14-day-old seedling expression analysis using microarray data. Albeit this, information on the abiotic stress transcriptional response of *EGY3*, for example using the BAR Database interface, also suggested that *EGY3* is not only highly induced in heat stress (70 fold), but also in osmotic (26 fold), and in salt stress (15 fold) (Figure 4.18). While in the heat stress response the gene expression is induced earlier in time, the osmotic and salt stresses induce a more delayed response comparatively to the heat response for *EGY3*. Both osmotic and salt stresses responses are probably due to the involvement of the osmotic component of the stress and not the osmoticum itself.

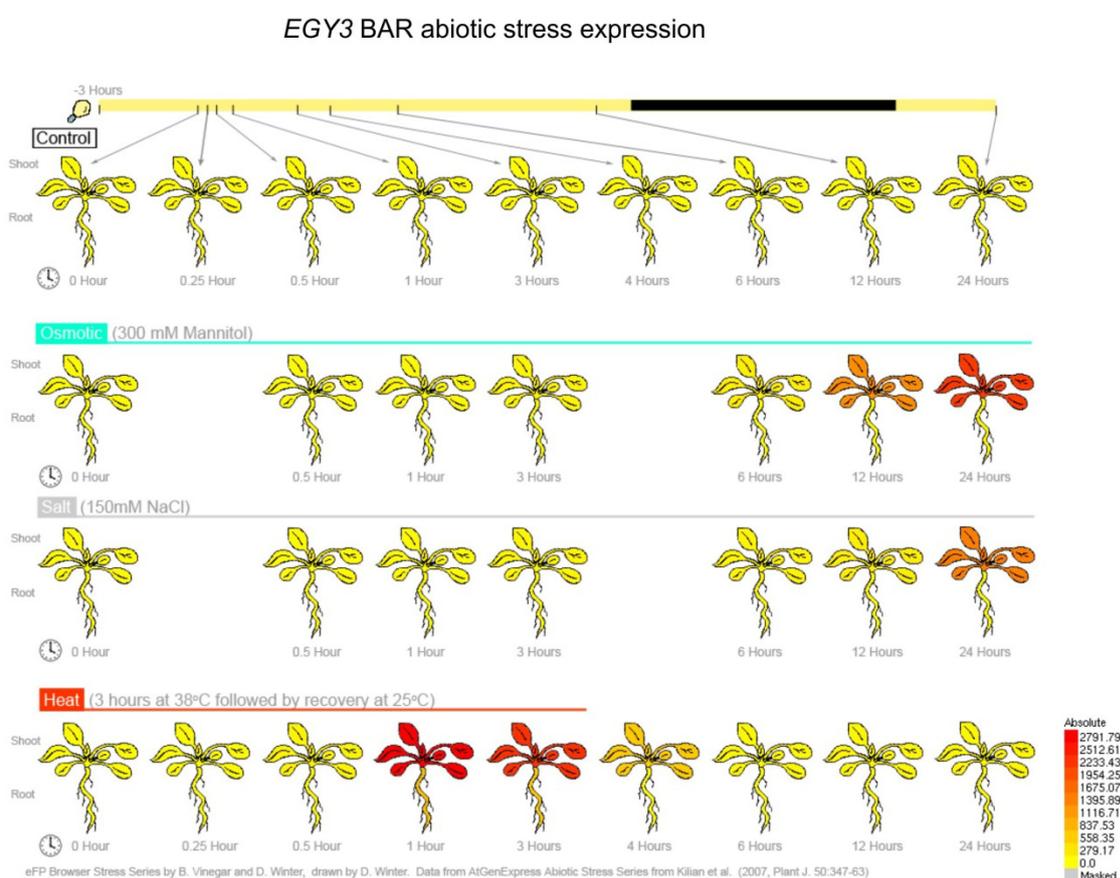


Figure 4.18. – BAR Arabidopsis eFP Browser (URL no.10) analysis of the abiotic stress response of *EGY3* (Winter *et al.*, 2007). Picture only evidences stresses that were shown to induce up-regulation of *EGY3* when compared to the control situation. Total expression values and standard deviations are listed in table A.7 of the Appendix I.

In order to characterise the transcriptional response of *EGY3* to heat, the promoter $EGY3::GUS$ reporter line was submitted to heat stress for three hours (Figure 4.19). As

previously shown in figure 4.14, at the seedling stage under standard 23°C, expression is already too strong, which lowers the resolving power of the GUS assay. Since adult plants, as figure 4.14 demonstrates, have a lower GUS signal, they were ideal to be subjected to a 37°C heat stress treatment using normal GUS staining conditions. They also allowed for more plant organs to be analysed. Rosette leaves, caulinar leaves, flowers, and siliques were analysed at a period between 0-3 h of exposure to 37°C. Caulinar leaves showed no relevant induction of GUS staining, but rosette leaves, that had no signal at 23°C (0 h) had a highly induced staining in vascular tissues that decreased after 3 h of stress. Regarding siliques, the GUS signal was intensified at the extremities, but it was in the flower that changes were more significant, passing from 0 hours, where sepals and carpel were stained, to an intensified staining in sepals, anthers, and the stigma tissue, important for the process of fertilisation. The overall conclusion is that heat stress induces *EGY3* expression in the vasculature of rosette leaves and in important structures of the flower. This might corroborate the theory that *EGY3* is involved in the regulation of flowering time.

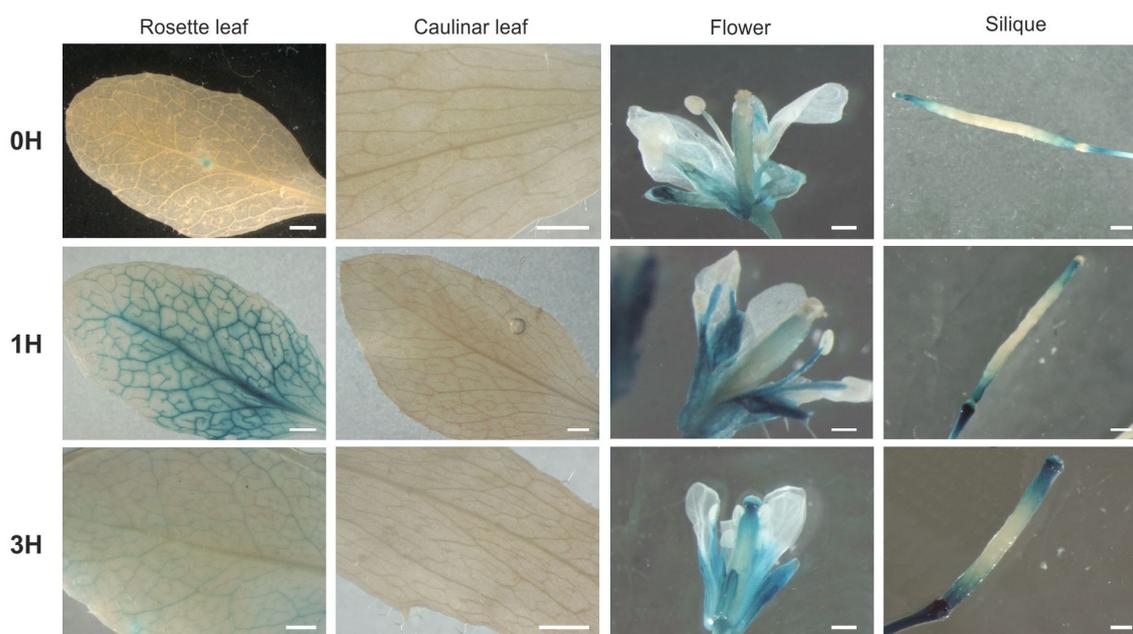


Figure 4.19. - Histochemical analyses of GUS activity in *proEGY3::GUS* expressing plants when facing heat stress. Mature leaves (caulinar and rosette), siliques and flowers of 5-week-old plants were imposed to 0-3 hours of 37°C. Scale bars represent 1 mm.

A heat stress-related phenotype was subsequently searched for *egy3-1* mutants. A series of heat stress experiments at the seedling and germination stages were performed in *egy3-1* and wild-type plants, taking into consideration that *EGY3* seems constitutively expressed in seeds and young tissues. A summary of the experiments is detailed in table 4.2.

Table 4.2. – Heat stress experimental conditions performed on *egy3-1* mutants to assess survival or root growth. The name of the test is according to figure 4.20.

Test	Time of exposure and Temperature	Phenotype assessment
Heat A	2 h at 23, 30, 34, 37, 40, 42, 44 and 45°C	Root measurement
Heat B	- 37°C (2 h) - 37°C (2 h) + 23°C (2 h) + 44°C (2 h) - 44°C (2 h)	Basal and acquired thermotolerance by root measurement
Heat C Heat D	60 min at 23, 38, 41, 44, 47, 50, 53 and 56°C 60 min at 23, 47, 49, 51, 53 and 55°C	Germination
Heat E	50°C for 0, 15, 30, 60, 120, 180, 240 and 300 min	Germination

Plants are able to survive to temperatures above their optimal growth temperature (basal thermotolerance), and also have the ability to gain tolerance to potentially lethal temperatures (acquired thermotolerance). This acquired thermotolerance is a result of a short period of acclimation at a moderately high yet survivable (sub-lethal) temperature (Larkindale *et al.*, 2005). In order to evaluate both responses, experiences were made following the experimental design of Larkindale *et al.* (2005) for testing basal and acquired thermotolerance. All seedling experiments consisted of 7-day-old vertically-grown *in vitro* plants, subjected to heat stress and analysed for growth after imposition of the stress. Test heat A (Figure 4.20-A) consisted in subjecting seedling plates to a range of temperatures (23°-45°C) for 2 hours, and subsequently measuring root growth. Differences even at 37°C presented no statistical significance (student t-test), likely due to the low number of replica. A new test was performed using the Larkindale *et al.* (2005) protocol, trying to verify the existence of differences in the basal or acquired thermotolerance in the mutant. More specifically, the heat B experiment consisted of exposing seedlings plates to (i) 2 hours at 37°C (control); (ii) 2 hours at 37°C plus 2 hours at 23°C plus 44°C (acquired thermotolerance test); (iii) 2 hours at 44°C (basal thermotolerance) (Figure 4.20-B). Results showed no difference in acquired or basal thermotolerance when compared to the wild-type. However, for the control a two hour period at 37°C suggested a difference, although it did not present statistical significance (student t-test). Moreover, it reversed the phenotypic tendency of the matching condition in the previous heat A test, which just shows the inconsistency of the experiment. *EGY3* expression in response to heat stress seems to be limited to the duration of the heat stress stimulus. This may explain the failure of the previous tests based on Larkindale *et al.* (2005) to produce significant differences, upon the fact that heat stress should be continuous to amount a significant difference between mutant and wild-type. A continuous heat stress should implicate a lower sub-lethal temperature.

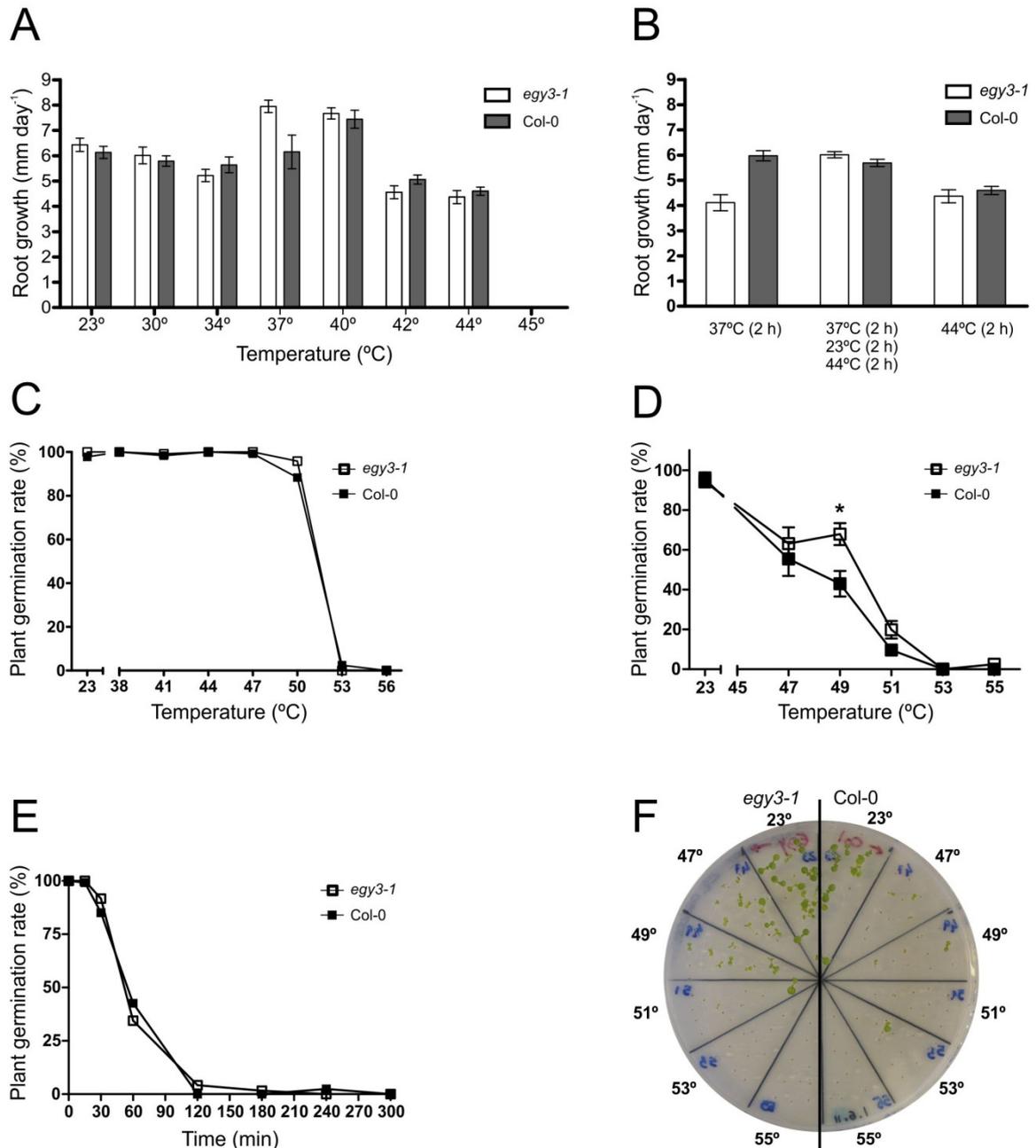


Figure 4.20. - Phenotype characterisation of heat stress-related phenotypes in Col-0 and the *egy3-1* mutant. A – 7-day-old seedlings were subjected to 23, 30, 34, 37, 40, 42, 44 and 45°C for a period of two hours. Measurement of the root was made for four days, every two days. N=5. B – 7-day-old seedlings were subjected to 37°C (2 h); 37°C (2 h) + 23°C (2h) + 44°C (2 h), and 44°C (2 h). Measurement of the root was made for six days, every two days. N=5. C – Plant germination rate after treatment for 1 h at 23, 38, 41, 44, 47, 50, 53 and 56°C. N=100 with 4 replica. Germination was assessed every day for a total of 9 days. D – Plant germination rate after treatment for 1 h at 23, 47, 49, 51, 53 and 55°C. E – Plant germination rate at 50°C for 0, 15, 30, 60, 120, 180, 240 and 300 min. N=100 with 4 replica. Germination was assessed every day for a total of 10 days. F – Photograph of the experimental result of the germination rate depicted in D. Asterisks represent significantly different levels between genotypes within the same conditions (student t-test: *, P < 0.05; **, P < 0.01; ***, P < 0.001).

As previously shown, the highest basal expression level predicted for *EGY3* takes place in dry seeds (Figure 4.13). At seed level, the germination rate was assessed by applying a heat shock before sowing. Two different experiments were performed, one with varying temperatures

during a one hour shock - heat test C and D (Figure 4.20-C,D), while the other - heat test E, provided a 50°C heat shock at various times (Figure 4.20-E). A slight heat resistance of the *egy3-1* mutant was shown in heat test C (Figure 4.20-C). Low resolution in the temperature range in the proximity of the phenotypical differences led to a second test with a better temperature resolution (heat test D). The number of replica was also doubled. The result (figure 4.20-D,F) confirmed that indeed statistically significant phenotypic differences were encountered at 49°C. A slight inconsistency with the dose response curve of control plants in both experiments can be explained by technical difficulties in controlling temperatures. When looking at heat test E (time-course analysis of seed germination) (Figure 4.20-E), no statistical differences were observed.

Results show that *EGY3* might be potentially important in the heat stress response at the seed level. The *egy3-1* mutant seed appears more resistant than the wild-type at 49°C. Heat stress affects enormously the chloroplast machinery, not only because it induces oxidative stress affecting photosynthesis, but also because it induces HSFs and recruits a great number of chloroplast-targeted HSPs and chaperones (Lee *et al.*, 2006; Kotak *et al.*, 2007). Loss-of-function mutants of heat-responsive effector genes will traditionally render germinating seeds more susceptible to heat stress. However heat-tolerant mutants have also been reported. For instance *tgr1* (*thermoinhibition-resistant germination 1*) and *tgr2* (*thermoinhibition-resistant germination 2*) are related to ABA, dormancy and inhibition of germination (Tamura *et al.*, 2006), while the double mutant *fad7/fad8* (mutants encoding chloroplastial proteins that are involved in the membrane fatty acid composition) showed increased tolerance to heat stress in leaves, due to a reduced trienoic fatty acid level that resulted in increased stability of photosynthesis (Larkindale *et al.*, 2005). Results suggest that loss-of-function mutants rendering tolerant germination seeds are not uncommon so *EGY3* might for instance act as a suppressor of one such heat effector protein, which is consistent with a potential role of *EGY3* in proteolytic activity or at least control of protein turnover of some target.

The major conclusion of the previous data is that *egy3-1* seems less susceptible to heat stress. Embryos contain photosynthetically active chloroplasts during seed development in some plants, which can improve seed storage products in the seed. Plastids are involved in essential metabolic activities, such as biosynthesis of nucleic acids, amino acids, and various lipids during all stages of plant development (Ruppel *et al.*, 2011). They develop early (48 h) after fertilisation in *Arabidopsis* embryos and persist for ~10 days. However, during seed maturation embryos turn colourless as chloroplasts dedifferentiate to nonphotosynthetic plastids by losing their chlorophyll, internal membrane structures, and starch. It is only upon germination that these dedifferentiated plastids (eoplasts) are converted into chloroplasts (shoots) and amyloplasts (hypocotyls and root

tips) assuring a proper seedling development (Ruppel *et al.*, 2011). Little is known about the molecular requirements for plastid development in embryo-derived cells and the real impact of chloroplasts in seeds, except that plastid-targeted proteins are important for metabolism, especially glycolysis in developing and germinating seeds, and that during stress the plastid biosynthetic activity is important for overcoming/tolerating the stress (Gutierrez *et al.*, 2007; Bentsinka and Koornneef, 2008; Ruppel *et al.*, 2011). During heat stress the chloroplast suffers structural changes, with high impact on the photosynthetic activity that takes place in thylakoid lamellae, the carbon metabolism in the stroma of the chloroplast, which has been suggested as the primary site of injury after high temperatures, and tissue dehydration that generates oxidative damage (Wahid *et al.*, 2007). However, a specific correlation during seed development and germination, between chloroplastidial proteins and heat stress remains elusive.

Collecting all data from the presence of GUS induction in important tissues, particularly in flowers, with the resistant phenotype of germinating seeds and the developmental phenotype of late flowering, it seems clear that EGY3 may play an important role in the reproductive and germination stages of development. Further experiments that corroborate this evidence and try to better analyse the seedling and whole-plant response to heat stress will be extremely valuable. Moreover, results from GUS staining analysis suggested that EGY3 is favourably expressed in the lateral roots primordia (Figure 4.14). Cross-sectioning analysis specifically showed that within the roots, the stele and the pericycle were stained. As already discussed, pericycle founder cells are involved in the formation of lateral root, as these are initiated when either individual or pairs of pericycle founder cells undertake several anticlinal divisions to produce a single layered primordium, formed of up to ten small cells of equal length. Then cells divide periclinally, forming an inner and an outer layer, after which more anticlinal and periclinal divisions create a dome-shaped primordium that emerge from the main root (Péret *et al.*, 2009).

Based on EGY3 expression data, phenotype characterisation involving a continuous heat stress could provide more information since the gene is rapidly induced and repressed. Considering this and the need to further understand the potential involvement of EGY3 in the root architecture during both development and the heat stress response, a more thorough analysis of the root architecture was performed. This analysis was conducted by growing plants *in vitro* for four days at 23°C and then at 23°C or 28°C for four more days (continuous stress). The EZRhizo software was used to measure a series of root parameters (Figure 4.21). In figure 4.21-A plants are seen at the last day of the experiment. Wild-type plants in normal conditions seem to have less lateral roots, and heat seems to induced root branching.

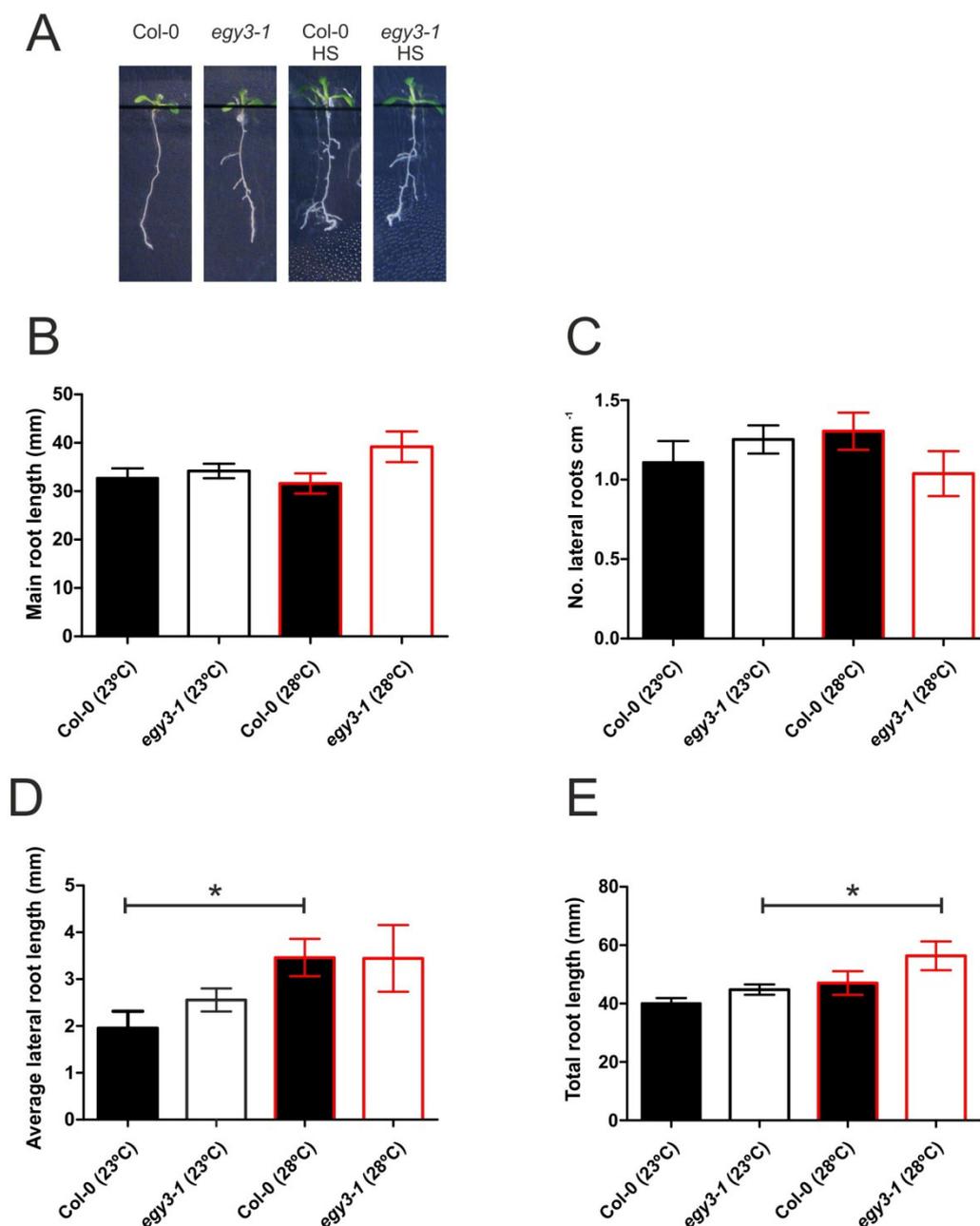


Figure 4.21. – Phenotype characterisation of root architecture in normal and heat stressed conditions in Col-0 and in the *egypt-1* mutant. A – Plant pictures from day 4 of the trial period of growth at 23°C or 28°C. Pictures were used for analysis in the EZRhizo software. B – Main root length. C – Number of lateral roots per cm of main root. D – Average lateral root length. E – Total root length. Error bars represent SEM. N≥16. Asterisks represent significantly different levels between the same genotype within different conditions (student t-test: *, P < 0.05; **, P < 0.01; ***, P < 0.001).

When analysing wild-type plants in response to heat, significantly different lateral root length was observed, which might be correlated with the heat stress response since they suffer heat while growing *in vitro* which causes more transpiration to occur, and since plates are filled with water a possible flooding of the roots takes place which is known to create adventitious roots (Potters *et al.*, 2006).

More significantly, the *egy3-1* mutants also display a difference that is statistically different in the total root length, which is higher at 28°C. In general, there are no differences in the main root length (Figure 4.21-B), despite a slightly higher tendency for *egy3-1* at 28°C. When comparing the number of lateral roots per cm, *egy3-1* at 23°C and Col-0 at 28°C present higher values (Figure 4.21-C). The average lateral root length is higher for *egy3-1* at 23°C when compared to the wild-type under identical conditions (Figure 4.21-D); however, wild-type values at 28°C are more elevated than at 23°C with a significant difference. As for the total root length (Figure 4.21-E), *egy3-1* at 28°C is significantly higher than the same genotype at 23°C.

Two important assumptions can be made after this analysis, first that temperature influences root architecture particularly in lateral root formation and therefore its length, thus influencing total root length. The search for water and nutrients is performed by the root, and its reorganisation can be due to stress conditions, either by enlarging the main root, or more efficiently by branching its architecture. Novel lateral roots promote an overall bigger surface for the absorption of nutrients, even more if we would consider root hairs (Osmont *et al.*, 2007). This is mainly an auxin-mediated process (Gray *et al.*, 1998). The second assumption is that *egy3-1* seems to possess a bigger total root length, at the cost of lateral roots. The apparent evidence that *egy3-1* has a bigger root surface than the wild-type may correlate with previous evidence that adult plants develop bigger rosette radius, are bushier and have late flowering (Figures 4.16 and 4.17). The better developed root system would thus promote growth through more resource availability, leading to a healthier and leafier plant, before starting to focus on the flowering and progeny stages (Osmont *et al.*, 2007).

It can be speculated that in seedlings and seeds, tissues where *EGY3* is putatively more expressed, the knockout plant *egy3-1* seems to possess a heat stress-like response. In standard temperatures of 23°C, the mutant root displays more branching. Meanwhile, seeds seem more prepared to survive a transient (1 h) heat shock treatment. Once again, a repressor or down-regulator role for *EGY3* seems to be suggested.

4.2.8. Analysis on the regulation of *EGY3* expression

Functional data involving *EGY3* is scarce. This gene was reported as one of the differentially expressed genes in the overexpression line of the transcription factor HsfA2, which suggested *EGY3* inclusion in a list of putative target genes of this TF (Nishizawa *et al.*, 2006). HsfA2 is part of the HSF family, a member of class A TFs, which contains AHA motifs. In eukaryotes, heat shock factors act as key components of signal transduction pathways, being

involved in the activation of genes in response to various stimuli (Nishizawa-Yokoi *et al.*, 2011). In *Arabidopsis*, 21 Hsf genes have been reported (Nover *et al.*, 2001). Plant Hsfs have an N-terminal DNA-binding domain (DBD), an adjacent domain with heptad hydrophobic repeats (HR-A/B) involved in oligomerisation, a cluster of basic amino acid residues essential for nuclear import (nuclear localization signal) and a C-terminal activation domain. In some cases, the latter can be also characterised by a nuclear export signal and short peptide motifs (AHA motifs) crucial for an activator function (Nover *et al.*, 2001; Nishizawa-Yokoi *et al.*, 2011). In plants there are three classes of Hsfs (A-C), that are defined by their singularities in the HR-A/B regions (Nover *et al.*, 2001). Hsfs bind to the highly conserved heat shock element (HSE), containing at least three 50-nGAAn-30 repeats in alternating orientations in the promoters of HS-inducible genes of all eukaryotes (Nishizawa-Yokoi *et al.*, 2011).

Arabidopsis heat shock transcription factor HsfA2 (At2g26150) was described by Nishizawa and co-workers reported (2006) as regulating the response to various stresses. These authors referred that an overexpression of the gene in *Arabidopsis* conferred resistance to a combination of heat shock (HS), high light (HL) and oxidative stress by Paraquat treatment (Pq). Paraquat (methyl viologen) is a redox-cycling herbicide that undertakes univalent reduction and transfers electrons to oxygen generating superoxide ion. In combination with high light, Pq at the chloroplast level will act as an acceptor of electrons from the PSI, guaranteeing a steady formation of superoxide ion and therefore inhibiting photosynthesis. A short and severe stress provoked by superoxide ion was proved to induce metabolic changes and de-regulation of essential genes as a short mechanism of survival in *Arabidopsis* plants (Maurino and Flügge, 2008). Regarding HsfA2, treatment with HS, HL and Pq resulted in the death of knockout and wild-type plants within 5 hours, but not of the overexpression line (Nishizawa *et al.*, 2006).

As stated, analysis revealed that *EGY3* was up-regulated in HsfA2 overexpressing plants, being therefore described as a target gene for this transcription factor in conditions of HL and HS. Also, *EGY3* protein was induced when facing HS and HL in two-week-old *Arabidopsis* plants, and induced in three-day-old *Arabidopsis* T87 cells, when treated with exogenous H₂O₂. Transcription levels were also assessed with both HL and HS treatments in *hsfa2* knockout plants, and results clearly showed a deregulation of *EGY3* in knockout plants, reinforcing that *EGY3* might be regulated by this transcription factor (Nishizawa *et al.*, 2006). Considering this, an experiment was designed to test for resistance/susceptibility of *egy3-1* knockout plants to high light, heat stress and oxidative stress, in conditions identical to those identified by Nishisawa *et al.* (2006) (Figure 4.22). During a period of 8 hours after treatment, plants were observed every hour for death symptoms.

Results of the time-course analysis indicated no resistance or susceptibility of *egy3-1* knockout mutants in comparison to its wild-type.

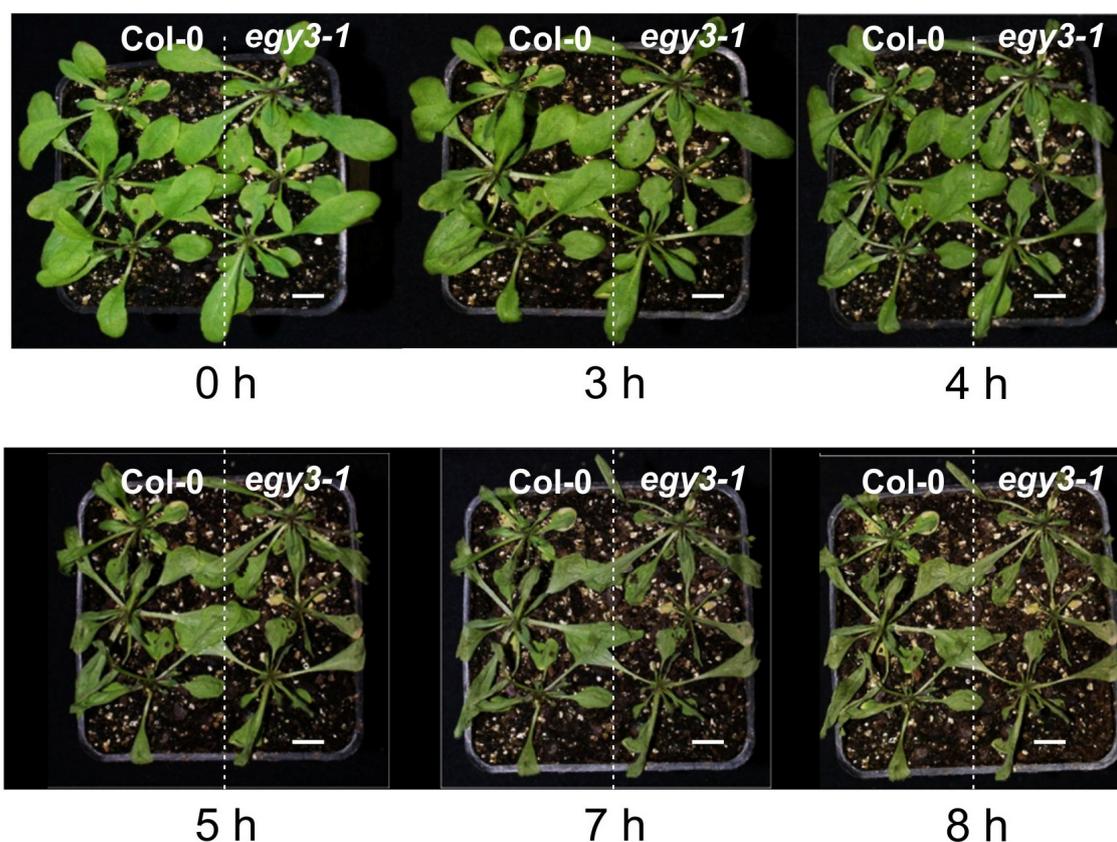


Figure 4.22. – High light, heat stress and Paraquat treatment in *egy3-1* mutants. One month-old plants were placed at 45°C with high light intensity (300 $\mu\text{mol PAR m}^{-2} \text{s}^{-1}$) and sprayed with a solution of 50 μM of Paraquat. Nine replicas were visually analysed on an hourly-basis for a total period of eight hours.

HsfA2 has been associated with processes other than oxidative stress caused by HL and HS (Nishizawa *et al.*, 2006), namely salt, osmotic and anoxic stress (Ogawa *et al.*, 2007; Banti *et al.*, 2010); inhibition of 26S proteasome function and/or Hsp90 activity in response to oxidative stress (Nishizawa-Yokoi *et al.*, 2010), and regulation of the expression of various genes related to defence against environmental stress (Nishizawa *et al.*, 2006). HsfA2 transcriptional activity in Arabidopsis was also proven to be regulated by sumoylation (Cohen-Peer *et al.*, 2010). This shows that HsfA2 is involved in major signalling pathways for various different environmental stresses. *Cis*-elements and *trans*-acting factors were identified for HsfA2, and HsfA1d and HsfA1e appear to mediate the induction of HsfA2 expression in response to oxidative stress caused by HL and HS, therefore functioning as both transcriptional regulators of HsfA2 and as key regulators of Hsf-mediated signalling in response to environmental stress, creating a new model for Hsf signalling (Nishizawa-Yokoi *et al.*, 2011).

A functional connection between *HsfA2* and *EGY3* was established using GeneMANIA. GeneMANIA is a tool to visualise functional networks, and based upon the query of genes it can give results of co-expression, physical interaction, co-localisation, pathway, and predicted functional relationships between genes, among others (Montejo *et al.*, 2010). Analysis of *EGY1*, *EGY2*, *EGY3* and *HsfA2* was performed in this web-tool (Figure 4.23-A). A significant aspect of the resulting network is that all EGY family members appear as predicted interactors of At1g78620, which codes for an unknown protein, with domains of unknown function. Another important aspect is the co-localisation and co-expression of *EGY3* with *HsfA2* which reinforces a putative connection. *EGY3* is also co-localised and co-expressed with an unknown protein (At5g35320) and with a DNAJ heat shock family protein (At2g20560) that is putatively involved in protein folding and in heat shock factor binding, reinforcing the *EGY3* role in heat stress. The majority of *EGY3* co-expressed genes are heat shock proteins like *HSP101*, *HSP70b*, *HSP17.6II*, *HSP17.6A*, and *HSP70T-2*. There is also one HSF (*HSFB2B*) and some genes of unknown protein function. Moreover, *EGY3* seems the only EGY family member associated with HSFs. As stated previously when referring to expression analysis of the EGY members, *EGY3* is the only gene up-regulated in abiotic stress conditions. Finally, *HSFB2b* appears in this analysis as co-expressed with *EGY3*, however, literature indicates that HsfB2b is not directly involved in the regulation of the onset of the heat shock response, but appears as a negative regulator of defensin gene expression and pathogen resistance (Kumara *et al.*, 2009).

Interestingly, knockout mutants for *HsfA2* revealed no visible phenotype for HL, HS and Pq and only overexpressing plants displayed this complex phenotype. This suggests that the *egy3-1* mutant may not display a dramatic or morphologically significant phenotype, much like its potential regulator, given also the complexity of the pathways in which *HsfA2* is involved. On the other hand, it is possible that, like in *HsfA2*, a phenotype will only be visible in the *EGY3* overexpression line (currently under development). Based on present evidence, a series of subsequent studies could be designed. A loss-of-function mutant for *HsfA2* was obtained and will be analysed for alterations in gene expression of *EGY3*, by monitoring *EGY3* transcript levels by qRT-PCR or by performing a cross with the pro*EGY3*::*GUS* line. For phenotyping purposes, crosses were already made in order to develop an *egy3-1/hsfa2* double mutant, with the F2 population presently being genotyped.

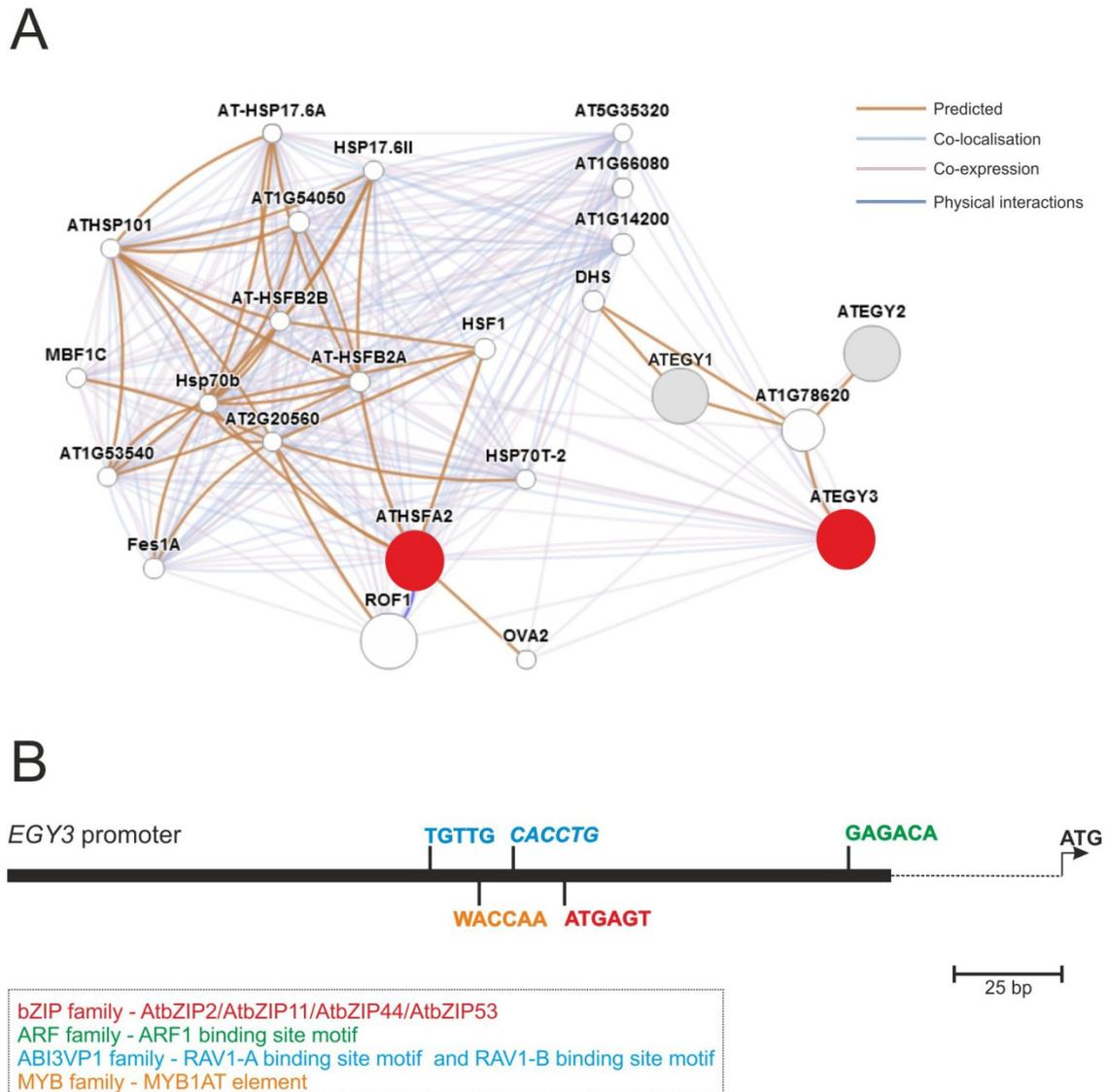


Figure 4.23. – Graphical depiction of the *EGY3* putative interaction and co-expression network and promoter *cis*-element presence. A – GeneMANIA (URL no.13) (Montjojo *et al.*, 2010) analysis of *EGY3* network containing *EGY2*, *EGY3* and *HsfA2*. B – *EGY3* promoter *cis*-elements, highlighted according to the family of binding transcription factors, based on AGRIS (URL no.14) (Davuluri *et al.*, 2003) and Athena (URL no.15) (O'Connor *et al.*, 2005). W stands for A/T.

As previously stated, it is most likely that *EGY3* is transcriptionally downstream of *HsfA2*. However, the promoter region of *EGY3* displays three different types of *cis*-elements that can be the binding site of four different TF families: bZIP, ARF, ABI3VP1 and MYB family (Figure 4.23-B). They do not include however HSE elements as far as the -3000 bp position, when analysed with the Athena promoter analysis web-tool (O'Connor *et al.*, 2005). Therefore, if *HsfA2* regulates *EGY3* expression it might be through some downstream component or TF of the *HsfA2* regulatory pathway that is directly regulating *EGY3* transcription.

The bZIP transcription factors are important in the regulation of processes such as pathogen defence, light and stress signalling, seed maturation and flower development. The

putative bZIP TF (AtbZIP2, AtbZIP11, AtbZIP44 and AtbZIP53) that might bind to the *EGY3* promoter sequences are all members of the group S, the largest group in Arabidopsis (Jakoby *et al.*, 2002). The AtbZIP11 is located in the chloroplast and is known to be repressed by sucrose through a translational inhibition mechanism. It is proposed that AtbZIP11 is a powerful regulator of carbohydrate metabolism that functions in a growth regulatory network that includes trehalose-6-phosphate and the sucrose non-fermenting-1 related protein kinase 1 (SnRK1) (Ma *et al.*, 2011) also being up-regulated by light (Jakoby *et al.*, 2002). Arabidopsis AKIN10/AKIN11 are protein kinases that are designated as SnRK1s, orthologs to the SnRF1 in yeast. They trigger changes in the expression of over 1000 genes that allow the re-establishment of homeostasis by repressing energy consuming processes and promoting catabolism when there is a stress-associated energy deficiency. SnRK1 signalling occurs via the S group of bZIP TFs, this regulation is complex as it appears that they can form heterodimers with members of the C group of bZIP TFs (Baena-González, 2010). AtbZIP2 is annotated as expressed in the pollen and pollen tube, and reported as a potential substrate of AKIN10, responsible for example for ASN1 (*ASPARAGINE SYNTHETASE1*) activation in response to darkness. The ASN1 synthesises asparagine and glutamate from aspartate and glutamine. Asparagine is important for nitrogen storage, transport compound, and it is synthesised at night and during low-sugar conditions (Hanson *et al.*, 2008). It is also thought that AtbZIP11 and AtbZIP53 function redundantly as ASN1 activators (Ufaz *et al.*, 2011). The bZIP53 is expressed mainly at the pollen, pollen tube and seeds, and has been reported as a transcriptional regulator of Arabidopsis seed maturation genes (Alonso *et al.*, 2009), and involved in amino acid metabolism during low energy stress (Dietrich *et al.*, 2011). To summarise, energy and sugar deprivation is related to all the referred TFs as they mediate AKIN10/11 signalling. This signalling converges reprogramming of transcription in response to apparently unrelated darkness, sugar and stress conditions. After deprivation of sugar and energy, signalling events mediated by AKIN10 target a wide range of genes to promote catabolism and suppress anabolism (Baena-González *et al.*, 2007). The fact that some of these TFs are particularly involved in seed maturation and expressed in the pollen, and hence in the process of fertilisation/reproduction should also be highlighted.

The ARF family is composed of auxin response factors with binding specificity to auxin response elements (AuxRes) in promoters of primary or early auxin-responsive genes. ARFs are characterised by an N-terminal DNA-binding domain; some also contain transcriptional activation domains, while others contain repression domains. In general, ARFs appear to be important to auxin-regulated gene expression (Guilfoyle *et al.*, 1998).

Concerning the ABI3VP1 family, the *EGY3* promoter includes both a RAV1-A and RAV1-B element. RAV1 proteins have two types of DNA binding domains, AP2 and VP1/B3. The AP2 domain was first identified as a DNA binding domain in a family of tobacco ethylene response element binding proteins, and in Arabidopsis APETALA2 (AP2), a transcriptional factor involved in flower development. VP1/B3 is conserved in a number of DNA binding proteins, such as VP1, ABI3 and ARF1, which have been shown to mediate the ABA and auxin responses (Hu *et al.*, 2004). Interestingly, it has been shown that overexpression of RAV1 in Arabidopsis results in retardation of lateral root and rosette leaf development, and underexpression results in earlier flowering. Also RAV1 may act as a negative growth regulator in the Brassinolide signalling pathway during growth and development (Hu *et al.*, 2004), and also as an inducer of senescence (Woo *et al.*, 2010).

The promoter analysis also showed the presence of a MYB1AT element, which acts as a MYB recognition site found in the promoters of the dehydration-responsive gene *RD22* and many other genes in Arabidopsis. MYB proteins are key factors in regulatory networks controlling development, metabolism, and responses to biotic and abiotic stresses, being particularly responsive to ABA (Dubos *et al.*, 2010). Presence of this *cis*-element might explain induction of *EGY3* by salt and osmotic stress.

On the whole, *in silico* analysis of *cis*-elements in the *EGY3* promoter provides a significant insight into *EGY3* function and asserts the phenotypes observed in the *egy3-1* mutant. Concerning developmental alterations, if RAV1 induces senescence it makes sense that the *egy3-1* mutant presents later flowering, bigger rosette leaves in adult 6-week-old plants, and basically a delayed senescence phenotype (Figure 4.16 and 4.17). It may also explain the putative *EGY3* involvement in lateral root formation, as the promoter has a RAV1 binding site and RAV1 promotes inhibition of lateral roots. It is also known that in Arabidopsis roots there is a periodic initiation of lateral root primordia based on an auxin oscillatory mechanism. In fact, local auxin accumulation in primed pericycle cells activates the auxin signalling cascade, which leads to the degradation of IAA14, derepressing ARFs (7 and 19), which will activate downstream gene expression (Péret *et al.*, 2009). The fact that *EGY3* possesses both ARF and RAV binding sites in its promoter supports the fact that *EGY3* seems to be expressed in lateral root primordia and the indication (though statistically non-significant) that *egy3-1* mutant plants have more tendency for the formation of lateral roots than the wild-type. Interesting experiments in the near future include the analysis of root morphology and architecture of *egy3-1* in the presence of IAA and expression pattern analysis of promoter $EGY3::GUS$ in the presence of IAA. A cross between *egy3-1* and *DR5::GUS* (a reporter system for IAA localization, using a synthetic IAA-responsive promoter regulating the *GUS* gene),

may also elucidate a possible deregulation in the auxin response that leads to lateral root formation. Finally, the regulation of S-bZIP TFs by AKIN10, a gene that is induced in nutrient deprivation, and the presence of these TFs binding sites in the *EGY3* promoter could explain the slower growth of the mutant and also the late flowering phenotype. For instance, the *EGY3* expression pattern during seed development (BAR, Figure 4.13) and resistance of *egy3-1* seeds to heat shock treatment (Figure 4.20) may correlate to AtbZIP53, a known regulator of seed maturation genes.

In order to provide additional functional clues, *EGY3* expression was also analysed in Genevestigator, exploring altered expressing in microarray experiments of known mutants/overexpressing lines (Figure 4.24). Some of the more significant results are now highlighted. *EGY3* appears to be highly expressed in CAT2HP1, which is a catalase deficient plant (Gadjev *et al.*, 2006) that was exposed to high light for 8h in this microarray experiment. *EGY3* is putatively responsive to oxidative stress. Double mutant *gun1/gun5* also displays high levels of *EGY3* transcripts; *gun1* and *gun5* mutants are defective in plastid retrograde signalling (Cheng *et al.*, 2011), and *gun1* showed also altered sensitivity to sucrose and abscisic acid, and alterations in early seedling development (Cottage *et al.*, 2010). As for the lines where *EGY3* seems to be down-regulated, the most relevant include a line overexpressing MYB44, the same mutant that suppresses jasmonate-responsive gene activation. MYB44 is induced by various hormones and abiotic stresses, and a MYB element was found in *EGY3* promoter (MYB1AT). Also the *hsf1/hsf3* double mutant, that is impaired in the HS response, seems to down-regulate *EGY3* (Lohmann *et al.*, 2004). *EGY3* transcription also seems to correlate to circadian clock- and photomorphogenesis-related mutants such as: *hy5* (*long hypocotyl 5*; mutant of a basic leucine zipper TF that functions downstream of multiple families of photoreceptors) (Zhang *et al.*, 2011) and *cop1-4* which represents the mutant of a gene that is a major repressor of photomorphogenesis in darkness, promotes HY5 degradation, contributes to day length perception and delays flowering under short-days (Jang *et al.*, 2008). *EGY3* is up-regulated in the *ein3/eil1* double mutant, which is involved in the ethylene signalling. EIN3/EIL1, ethylene-stabilised transcription factors are induced in the presence of ethylene, and seem to act downstream of COP1, regulating photomorphogenesis in Arabidopsis seedlings (Zhong *et al.*, 2009).

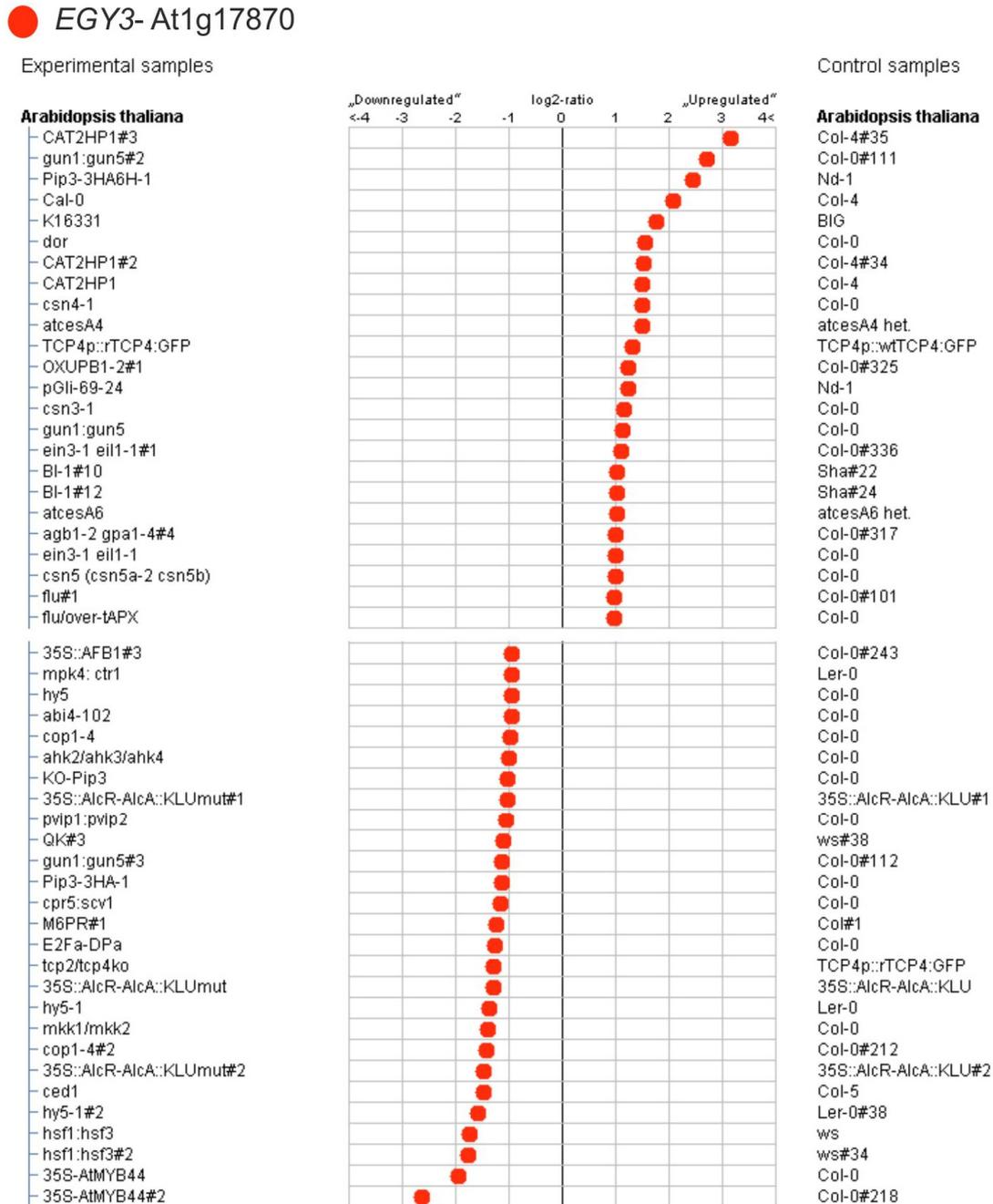


Figure 4.24. - Electronic Northern analysis of *EGY3* response in several loss- or gain-of-function *Arabidopsis* lines, performed through Genevestigator (URL no.11). # represents the number of the experiment.

In an overview of the results from this analysis, it is important to note that *EGY3* is up-regulated in a mutant involved in oxidative stress, which lacks the important enzyme catalase that catalyses the decomposition of hydrogen peroxide to water and oxygen. This is relevant since *EGY3* is also up-regulated by hydrogen peroxide, especially if in HS and HL conditions, as reported in Nishizawa et al. (2006). Other HSFs are also present, and putatively down-regulate *EGY3* expression, either directly or through HsfA2. Very important is the

observation of EGY3 up-regulation in the *ein3/eil1* double mutant (two insensitive ethylene mutants), and down-regulation in the *cop1-4* and *hy5* mutants. HY5 acts as an integrator that links various gene networks to coordinate plant development, being particularly important in promoting/repressing the transcription of photosynthesis associated nuclear genes (PhANGs), which is repressed by ubiquitination/degradation by COP1 (Larkin and Ruckle, 2008). Zhong and co-workers (2009) demonstrated that COP1 is a positive regulator of EIN3/EIL1 in the control of seedling greening and ethylene is important for the repression of the toxic accumulation of protochlorophyllide, an intermediate in the chlorophyll biosynthetic pathway. COP1 may indirectly or directly induce both EIN3/EIL1 and PIF1 (phytochrome interacting factor 1) (Zhong et al., 2009). COP1 is also involved in the photoperiodic flowering: in long-days the onset of flowering is made through phyA and cry photoreceptors that elevate the CONSTANS protein to induce flowering, while COP1 is inhibited during the day. However, at night CONSTANS is degraded by the 26S proteasome mediated by COP1. In short-days, flowering is delayed because COP1 forms a complex with suppressors of phytochrome A (SPAs) to degrade CONSTANS during the night period, and during the day no CONSTANS accumulation is induced (Henriques *et al.*, 2009).

4.3.7. Overexpression, complementation and heterologous expression studies of EGY3

Complementation studies are of the utmost importance when a phenotype is found in knockout plants, not only proving that the lack of phenotype in the complementation line is due to the gene in question but also serving as a control in experiments. The T-DNA insertion ratio is in average of about 1.5 per plant, so it is important to confirm the presence of only one T-DNA in the mutant, that this T-DNA is the one affecting the gene studied, and also that the phenotype is due to that interruption (Alonso and Ecker, 2006). To have additional insights into these phenotypes, an overexpression line would be important. For instance, in the case of the germination phenotype, where *egy3-1* displayed resistance to heat stress, tests should be made with the overexpressor in order to observe susceptibility rather than resistance. As part of the functional genomics strategy to characterise *EGY3*, overexpression lines were generated in order to allow gain-of-function studies. An *EGY3* overexpression line with a GFP tag fusion was also pursued using the Gateway cloning system. A successful BP entry clone was obtained and is currently being confirmed. More specifically, *EGY3* cDNA (Figure 4.25-A) was generated from extracted RNA through reverse transcription, and subcloned into the pGEM-T Easy vector. Following sequencing for confirmation

of the correct cDNA sequence, the cDNA was cloned by restriction in a pCAMBIA1303 vector, and placed under the regulation of the strong 35S plant promoter (Figure 4.25-B). Subsequently, the construct was used to transform plants by *Agrobacterium*-mediated floral dipping. Two genotypes were transformed: the *egy3-1* mutant, to generate a complementation line; wild-type Col-0, to obtain overexpression lines. Constructs were successfully transformed and T1 plants were already obtained for both constructs. Following the T2 (selection for one T-DNA insertion) and T3 plants (selection of homozygous plants), plants will be analysed for *EGY3* expression levels by semi-quantitative RT-PCR. This will be of importance to verify the quality of the lines, namely the strength of the overexpression and the adequacy for use in complementation of the *egy3-1* mutant. Once lines are confirmed, phenotype tests may be performed in synchronised seed populations.

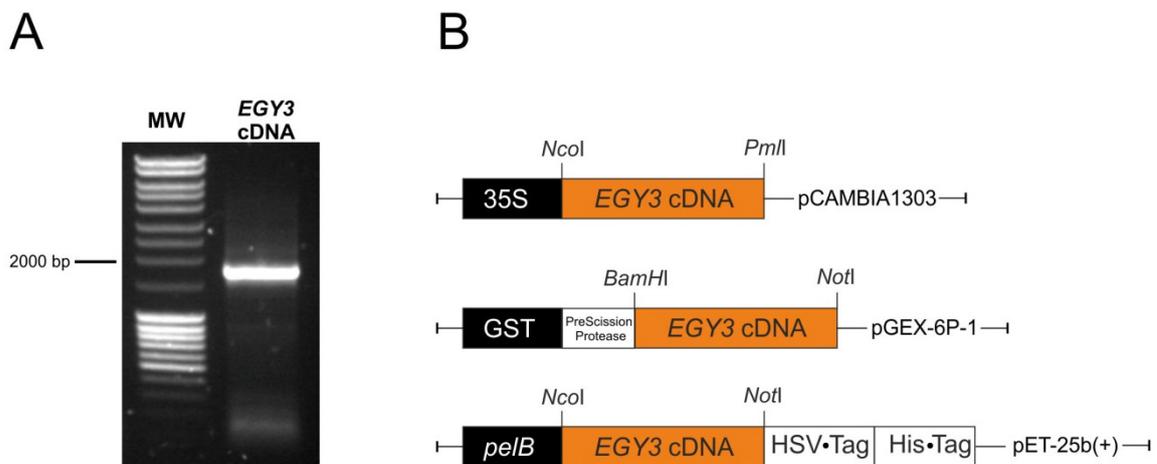


Figure 4.25. – Constructs with *EGY3* cDNA used for overexpression and protein expression studies. A - Electrophoretic analysis of *EGY3* cDNA amplification. MW - Molecular Marker MassRuler DNA Ladder Mix. B - Constructs made using *EGY3* cDNA to overexpress the gene *in planta* with pCAMBIA1303, and to express the protein in *E.coli* in fusion with a GST tag (pGEX-6P-1) or a His tag (pET-25b(+)). Schematics are not at scale. His - histidine. HSV - herpes simplex virus. *pelB* - *pelB* leader sequence that directs proteins to the periplasmic space.

Another pursued research line was the heterologous expression of *EGY3* in *E. coli*. The objective was to affirm or disclaim by *in vitro* analysis, the potential metalloprotease activity of *EGY3*. For this purpose *EGY3* cDNA was introduced into two vectors for heterologous protein expression. The pGEX-6P-1 vector was used to introduce an N-terminal GST tag, and also a PreScission protease cleavage site, to excise the GST if required. *EGY1* was similarly cloned in a pGEX vector in order to determine its ATP-independent metalloprotease activity (Chen *et al.*, 2005). The main objective was to use the GST tag to increase solubility of the fusion protein, given the fact that, as *EGY* family members and particularly *EGY3* possess a series of transmembrane domains, they are likely to interfere with protein solubility and processing in the *E. coli* environment. In a similar fashion, *EGY3* was cloned into pET-25b(+), in which the tags consist of six histidines (His tag) plus a HSV (herpes simplex virus) tag in the C-terminal, and an N-terminal

peIB leader that directs proteins to the periplasmic space. The GST tag, by having a high molecular weight and having a globular tertiary structure, could facilitate solubility and overexpression. Fusion proteins that possess the complete amino acid sequence of GST also demonstrate GST enzymatic activity and can undergo dimerisation similar to that observed in nature. The *peIB* leader in pET-25b(+) could also help move EGY3 to a more accessible place within the cell. With these two strategies, chances of getting correct expression were higher, increasing the probability of obtaining sufficient protein for the assay. In figure 4.25-B, a schematic of the construct was provided in order for the cloning strategy to be visually understandable. Restriction enzymes used for all constructs are depicted, as well as the corresponding vectors in which EGY3 cDNA was inserted.

As a preliminary assay for induction of EGY3 in *E. coli*, the protein expression vectors were used to transform the *E. coli* strain BL21(DE3)pLysE. BL21 strains are protease-deficient and designed to maximise expression of full-length fusion proteins, and BL21(DE3)pLysE is a high-stringency expression host, where the expression of the protein of interest is under the control of the *lac* operon. The high stringency is achieved by the presence of the pLysE vector that allows for the expression of basal T7 lysozyme that inhibits basal production of the T7 RNA polymerase. IPTG (0.4 or 1 mM) was added to exponentially-growing cells to induce expression of the construct. As shown in the figure 4.26-A, expression of pGEX-6P-1 with EGY3 protein was induced for a maximum of 3 hours, and then total protein extracts were resolved by SDS-PAGE. The M_r of the expected band was estimated to be of 89.55 kDa. However, the expression pattern did not differ from that of the control point at 0 hours, and no increase in intensity was observed in bands with probable M_r (arrows). Figure 4.26-B details expression of pET25b (+) with EGY3, which was induced with IPTG until a maximum of 4 hours. The expected protein M_r was 68.22 kDa. Once again the pattern of protein expression was not different from the 0 hour control, but there seemed to be a highlighted band at 1 hour. However, both gels for 0.4 and 1 mM of IPTG induction had relatively fewer bands than what would be expectable. Both assays should now be repeated, with different acrylamide concentrations, and expanding the optimisation criteria. The running gel's acrylamide concentration used was of 12%, but as in this case of high molecular weights, it should be decreased to at least 10%. Different *E. coli* protein expression strains, currently available in the laboratory, will be used. These include BL21(DE3)pLysS, which is more productive than pLysE because of a less effective T7 lysozyme; and Origami B, a strain that contains the *lacY1* mutation eliminating the active transport of lactose into cells via *lac* permease, consequently becoming less sensitive to lactose in the media. This strain allows for a more uniform entry of IPTG into all cells in

the population, and also promotes disulfide bond formation in the cytoplasm, which may yield maximum levels of soluble, active, properly folded target proteins.

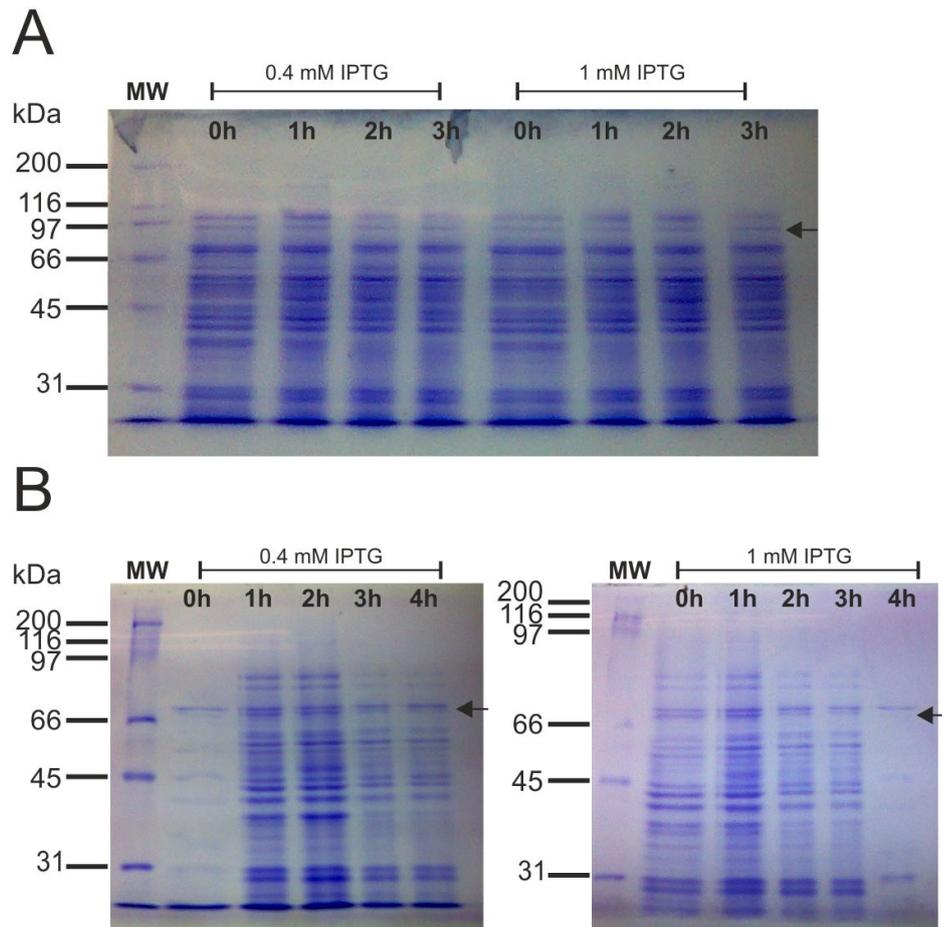


Figure 4.26. - Electrophoretic analysis of protein expression in BL21 (DE3) pLysE sbet strain. A - pGEX-6P-1 with *EGY3* cDNA expression for 3 h. B - pET25b(+) with *EGY3* cDNA expression for 4 h. Arrows highlight the putative size of the EGY3-Tag fusion protein, for each construct.

The medium could also be switched to a richer medium such as terrific broth for example. Finally, induction should also be tested using lactose rather than IPTG. As previously stated, new assays for protein overexpression are underway to test for different variables and obtain specific induction of the protein. The aim is to purify EGY3 in a soluble fraction and subsequently use it *in vitro* to perform a metalloprotease activity assay, proving that EGY3 has or most likely lacks metalloprotease activity.

4.3.8. Concluding remarks

In *Arabidopsis* a novel gene family was identified when Chen and co-workers (2005) first characterised the mutant *egy1* (for *ethylene-dependent gravitropism-deficient and yellow-green 1*), which presented reduced chlorophyll levels and an abnormal hypocotyl gravicurvature. Studies showed that EGY1 possessed *in vitro* metalloprotease activity and was important for chloroplast development, with a role in regulating endodermal plastid size and number (Chen *et al.*, 2005; Guo *et al.*, 2008). Following EGY1 discovery two functionally unresolved genes were annotated as being members of the same family, and were named EGY2 and EGY3 (Chen and co-workers (2005)). EGY3 became a focus of attention following an *in silico*-based strategy to encounter an unresolved chloroplastidial protein which was up-regulated in response to heat stress. It involved analysis of differential expression in microarray data and a series of targeted cut-offs, as described previously in this chapter. EGY3 appeared as a putative membrane and chloroplastidial protein that showed ~70 fold transcript increase when heat stress was imposed to 18-day-old seedlings, also evidencing up-regulation in the face of a few other abiotic stresses (namely osmotic and salt). The heat stress responsiveness was confirmed by semi-quantitative RT-PCR analysis. Also, a knockout mutant *egy3-1* was generated to allow for subsequent loss-of-function studies.

EGY3 is the only family member that seems to be involved in the abiotic stress response. After a phylogenetic analysis it was shown that EGY3 is fairly distant to EGY1- and EGY2-like proteins. In fact, EGY family members vary considerably in the number of transmembrane domains and particularly EGY3 lacks the canonical metalloprotease motif and has an aminoacidic change in its NPDG motif, which was also encountered in remaining EGY3-like proteins. This may represent an important neofunctionalisation of EGY3-like proteins. When comparing the expression pattern through the plant life cycle, *EGY3* is the least expressed, while *EGY2* and *EGY1* are more expressed, especially in the young rosette, first stages of leaf development and pedicels. None withstanding, *EGY3* is particularly expressed in the dry seed, formation of embryo/seed and some flower stages. Though generally low expressed when compared to the other family members, promoter-GUS studies with *EGY3* revealed that it is highly responsive to heat stress. Developmentally speaking they evidenced high expression of *EGY3* at the seedling stage, comprising both the shoot and the root. At this stage, GUS staining was evident in the cotyledons, stomata, hypocotyls embryonic root, meristematic areas in the shoot, primordial leaves, trichomes, regions of lateral root formation and internally in the root, except the meristematic zone. A transversal sectioning of the root revealed GUS presence in the stele and in the pericycle cells. As for adult tissues staining was present in the sepals and carpel structures and in the silique

extremity, but not in the adult leaves. Staining of adult structures exposed to heat stress showed induction in the vascular tissue in rosette leaves, and a high induction on flower structures such as sepals, anthers and stigma tissue. The presence of staining in areas such as the lateral root region and flower structures may suggest a role in both lateral root formation and flowering/fertilisation.

Characterisation of the phenotype of the *egy3-1* knockout mutant was performed based on a simplified methodology by Boyes *et al.* (2001), measuring specific developmental parameters throughout a life cycle at 23° and 28°C. Results showed that six-week-old *egy3-1* mutants had bigger rosettes (at 23° and 28°C) and possessed delayed flowering (at 23°C). Visually, *egy3-1* plants seemed initially smaller but grew into bushier plants, with late flowering and later senescence. This developmental phenotype could suggest a role in plant development, flowering and life-cycle. Heat stress tests were also performed in the seedling and seed/germinating stages. Seedling tests for acquired and basal thermotolerance (Larkindale *et al.*, 2005) revealed no relevant results. However, germination heat stress-survival tests showed a significant phenotype in which the mutant *egy3-1* seemed resistant to the stress. Given the putative relationship between EGY3, heat stress and lateral root formation, root architecture analysis was performed at both 23° and 28°C. The most significant result was the presence of a higher total root length for *egy3-1* grown at 28°C, when compared to the standard 23°C. This seemed to be the result of longer main and lateral roots rather than an increase in root branching. A very interesting conclusion is that the knockout of *EGY3* functioning seems to systematically produce a positive effect, namely by inducing growth of adult plants by extending the life cycle, by providing resistance to heat stress in seeds, and by promoting seedling root growth in response to high temperatures. All these results suggest that in *Arabidopsis* *EGY3* might act as a repressor. The proteasic nature of *EGY3* supports the idea that *EGY3* might act as a repressor by promoting the degradation of a target protein or being involved in protein turnover. Even considering that *EGY3* might not possess proteasic activity (given its topology and conserved residues), but acting as a possible chaperone, in a proteasic event.

Nishizawa *et al.* (2006) reported that *EGY3* was a target gene of the heat transcription factor HsfA2, and that *EGY3* was induced following oxidative stress (H_2O_2), and in a combination of HS and HL. The HsfA2 overexpression line displays a resistance phenotype to this combination of stresses. When the experiment was replicated in *egy3-1* no difference in phenotype was observed, but then again the same result is observed in the *hsfa2* mutant, and the experiment should be repeated with *EGY3* overexpression lines soon to be available. Still, *in silico* analysis supports a strong association between HsfA2 and other HSFs and *EGY3*, whether by co-expression, co-localisation, or deregulation of *EGY3* in *HSF* mutant lines, which reinforces the

present phenotypes displayed by *egy3-1*. Underway studies to confirm this hypothesis include the generation of an *egy3-1/hsfa2* double mutant, phenotype characterisation of *EGY3* overexpression lines, and qRT-PCR analysis of *EGY3* transcripts in the *hsfa2* mutant.

Surprisingly, no HSF-related HSE *cis*-elements were shown to be present in the promoter of *EGY3*, which suggests that HSF regulation of *EGY3* occurs higher in the regulatory cascade. Instead, four types of TF families were found to putatively bind the promoter: bZIP, ARF, ABI3VP1 and MYB. Most of the bZIP that putatively interact with the *EGY3* promoter are related to energy and sugar deprivation, as well as seed maturation. ARFs are related to auxin-regulated expression and hence plant and root development (Guilfoyle *et al.*, 1998). ABI3VP1, more precisely RAV1, is involved in lateral root and rosette development, acting as a negative growth regulator in the brassinosteroid signalling pathway during growth and development and as an inducer of senescence (Hu *et al.*, 2004; Woo *et al.*, 2010). The MYB TFs are well known for being key factors in the regulatory networks that control development, metabolism, and responses to abiotic and biotic stresses (Dubos *et al.*, 2010). Presence of these *cis*-elements in the *EGY3* promoter all support the developmental and abiotic stress-related phenotypes observed in *egy3-1*. Functional information was also provided by analysing *EGY3* deregulation in available microarray data of known mutants (either knockout or overexpressing lines).

EGY3 does not present the canonical metalloprotease motif so performing an *in vitro* catalytic assay is extremely relevant. For that reason two separate constructs were made to produce an *EGY3*-Tag fusion protein in *E. coli*, to later purify by affinity to the tag and use in the proteasic assay. Unfortunately although both constructs were obtained, no visible protein induction was observed, with both GST and His tag fusions. Given present results, optimisation of growth, medium, and induction conditions is required, including altering the strain used for heterologous expression of *EGY3*.

A 35S::*EGY3* construct was developed and transformed in both wild-type and the *egy3-1* backgrounds, to develop overexpression and complementation lines, respectively. The overexpression line can provide additional phenotypes that may be impossible to detect with the null mutant, as reported for HsfA2 (Nishizawa *et al.*, 2006). They may also corroborate phenotypes of the loss-of-function mutant. This is particularly relevant since the *egy3-1* mutant has a resistant phenotype, being less susceptible to heat stress, suggesting that the gene acts as a negative regulator. It will be important to observe whether overexpression of *EGY3* will provide phenotypes opposing those of *egy3-1*, namely one of susceptibility to heat stress. It will also be important to confirm present *egy3-1* phenotypes, by complementation but also by obtaining a second mutant allele.

At this stage it is highly speculative to infer the role of *EGY3* during development and how it correlates to heat stress. Taking for instance the observed *egy3-1* seed resistance to heat shock; the association of plastids with seed development and germination is scarce. Chloroplasts exist after 48 h of fertilisation (persist ~10 days) and then dedifferentiate to non-chlorophyllic structures and lose internal membranes. When germination occurs these eoplasts transform into the structures needed for each specific tissue, namely chloroplasts for the shoots and amyloplasts for the hypocotyls and root tips (Ruppel *et al.*, 2011). *EGY3* function in this process may range from a role in membrane organisation, to control of protein turnover of a heat-resistance determinant protein. The most suitable path seems to be the identification of interacting proteins or of potential targets, if *EGY3* does in fact maintain a role in proteasic activity. In this case, screening for proteins with a split ubiquitin system, more suited for membrane proteins, would be a feasible strategy.

5.

“Final remarks and future perspectives”

Despite extensive effort in the last few years, there are a large number of genes whose function remains unknown in the genome of *Arabidopsis thaliana*. In the present work we have investigated the role of several abiotic stresses. *SQE1* homologs, *SQE2* and *SQE3*, were studied because the characterisation of *SQE1* suggested an elusive role for sterols in the regulation of Reactive Oxygen Species (ROS) and displayed a drought phenotype in *dry2/sqe1-5* mutants (Posé *et al.*, 2009). Additionally, *EGY3* was selected through a web-based data-mining strategy that suggested a specific role in abiotic stress tolerance and particularly heat stress. Of the *EGY* gene family, *EGY3* seems to be the only member to be up-regulated by abiotic stress.

SQE2 and SQE3. Our analysis of SQE genes evidenced a complex regulation of sterol biosynthesis in plants. Despite squalene epoxidase 1 (*SQE1*) being foretold as the main enzyme in the conversion of squalene, our results showed that some tissue specificity exists within SQEs. In terms of biochemical analysis of the mutants, *SQE1* role is more specific of roots, while *SQE3* seems to have a major role in shoots. The mutant *sqe3-1* presented increased growth sensitivity in the presence of the squalene epoxidase inhibitor terbinafine, and also evidenced altered squalene and sterol profiles. Meanwhile, the double *dry2/sqe3-1* mutant was unviable, worsening the already acute pleiotropic phenotype of *dry2/sqe1-5* and highlighting a role for *SQE3* in sterol biosynthesis. Because *SQE3* cannot complement *SQE1* even driven by the same promoter, analysis were made to determine whether *SQE3* could be involved in an alternative sterol biosynthetic pathway, given the recent evidence that the lanosterol pathway could provide ~1.5% of the final sterol production in *Arabidopsis* (Ohyama *et al.*, 2009). Therefore, it was hypothesised that *SQE3* could be supplying 2,3-oxidosqualene to alternative OSCs that would therefore turn 2,3-oxidosqualene into other compounds (namely lanosterol), affecting the final sterol composition. This was supported by the finding that both, *LAS1* and *SQE3* were induced by biotic stress based on public transcriptomic data. However *sqe3-1* mutants did showed similar resistance to WT plants when infected with *Pto*, and further expression analysis did not show relevant differences in the mutant. However, this is still an open hypothesis, with double mutants being generated for both *dry2/sqe1-5* and *sqe3-1*, with *las1*. Subcellular localisations were a priority and both *SQE1* and *SQE3* seemed to be present in the ER. However, functional GFP fusion and stable transformations are under way to determine the localisation of both proteins and how this can influence the enzymatic activity and the timing of protein action. However preliminary results point to both proteins being indeed localised in the ER, which makes it difficult to explain the lack of complementation of *dry2/sqe1-5* by *SQE3* when driven by *SQE1* promoter. Topological analysis may provide an explanation since it predicts that the catalytic domains are faced in opposite directions of the membrane. Important assessments were

made in the seed, since *dry2/dry2* *SQE3/sqe3-1* presented an abortion rate very proximate to 25%, which may implicate these two genes in fertilisation or embryo/seed development. Therefore, the sterol biosynthetic pathway is apparently becoming less straightforward than previously predicted. Latest reports on the isoprenoid biosynthetic pathway showed that although *SQE1* and *SQE3* were considered more important on the overall sterol biosynthetic pathway, *SQE2* may too have an important role, previously disregarded because of its putative location in the mitochondria (Vranová *et al.*, 2011). This report shows a very complex network of sterol biosynthesis with *SQE2* located in the mitochondria together with at least two OSCs, suggesting that conversion of 2,3-oxidosqualene to other secondary metabolites, such as marneral, camelliol and arabidiol, has an importance yet to be established.

Although the basis have been established in this work, future experiments will be required to confirm the subcellular localisation, and also to fully understand mechanistically the role of *SQE3* in sterol biosynthesis. Seeds of *dry2/dry2* *SQE3/sqe3-1* should be analysed at earlier stages of development to understand in which tissue and at which stage of development the embryo arrests. A phenotypic analysis of a segregating *DRY2/dry2* *sqe3-1/sqe3-1* mutant should be interesting to analyse, as this could determine whether *SQE3* has a real function in seed coat development. Also interesting would be a co-localisation study of both *SQE1* and *SQE3* with *CAS1* and *LAS1* (the main OSC enzymes), advancing into *in vivo* assays of protein-protein interaction such as BiFc (Bimolecular fluorescence complementation), or co-immunoprecipitation. This would resolve whether *SQE3* also participates in the main sterol biosynthetic pathway through *CAS1*. Also, a split-ubiquitin assay (since we are in presence of a membrane protein) would be of value to verify putative interactors that could enlighten a more specific function for *SQE3* as well as the rest of the *SQEs* in the family. Furthermore, *sqe2-1/sqe3-1* double mutants should be analysed more thoroughly for phenotype analysis and sterol profiling. More focus should also be given to *SQE2*, since though low expressed its importance and putative localisation in the mitochondria could project some role not yet foreseen for sterols or other compounds in plants. Microarray analyses of *sqe2-1* and *sqe3-1* plants can provide valuable data, particularly considering that a *dry2/sqe1-5* microarray analysis was already performed. With the three data sets one could resolve the specificities of the three enzymes, the processes their involved in, and possibly identify the deregulation of key genes in the pathway, if that regulation occurred at a transcriptional level.

EGY3. The present work on *EGY3* highlighted its involvement in heat stress responses, previously suggested by transcription analysis in microarrays, later confirmed by both semi-quantitative RT-PCR and the resistance phenotype of the *egy3-1* mutant following a 49°C (1 h) shock treatment to

seeds. A developmental phenotype was also encountered in *egy3-1*, showing delayed flowering, bushier plants and delayed senescence. *EGY3* expression was shown by the reporter GUS system to be highly expressed in young tissues, and in sepals, carpels and in silique extremities. Root sectioning of young seedlings revealed that *EGY3* expression is present in the stele and in the pericycle cells. The presence in the lateral root and pericycle might suggest a role in lateral root formation, which is supported by the presence of ARFs and RAV TF binding sites in the promoter. The deregulation of *EGY3* expression in known mutants showed a possible involvement in oxidative stress, circadian rhythm and possible involvement in the plant developmental responses that direct the transcription of photosynthesis associated nuclear genes (PhANGs). This association with oxidative stress was also present when Nishizawa *et al.* (2006) reported that *EGY3* was a target gene of the TF HsfA2, involving a combined response to heat shock, high light and oxidative stress. These results further reinforce a putative role in heat and oxidative stresses. The ongoing effort to characterise *EGY3* is now diverted into various functional aspects, and has already involved the development of a series of molecular tools.

The overexpression line currently being produced will be also analysed relative to heat germination, developmental and flowering phenotypes, to determine whether this gene could be of biotechnological interest. We hypothesise to have an opposite effect *in planta* to the loss-of-function mutation, possibly allowing for additional phenotypes of interest. Root architecture analysis should also be repeated, namely with the overexpression line. Additional, more specific phenotype characterisation should also be pursued. For instance, developmental phenotypes should be analysed in short-day photoperiod, since there is a strong connection of *EGY3* with the plant circadian rhythm (de-regulation in *cop1* and *hy5* mutants). It will be relevant to monitor by qRT-PCR the expression of *EGY3* in different tissues, in response to heat or different genetic backgrounds such as *hsfa2*, and also monitor the expression of key genes in pathways related with the observed phenotypes. Ideally, microarray analysis of differentially expressed genes in *egy3-1* would contribute to the establishment of *EGY3* function. Since *EGY3* is a putative chloroplast-targeted protein, and knowing that the *egy1-1* mutant presented altered chloroplasts (Chen *et al.*, 2005; Guo *et al.*, 2008), analysis of the chloroplast ultra structure is currently underway. A construct is at the final stages of production, for an *EGY3*-GFP fusion protein, that will confirm *EGY3* targeting to the chloroplast. It is hypothesised that *EGY3* is regulated by HsfA2, yet it is not known if transcriptional control by this transcription factor occurs directly or indirectly. For that reason a double mutant is being generated and Electrophoretic Mobility Shift Assay (EMSA) using HsfA2 protein with the promoter of *EGY3* is suggested, thus evidencing if HsfA2 binds directly to the promoter of *EGY3*. To analyse whether the interaction between HsfA2 and *EGY3* occurs at protein level, BiFc can be

used. Experiments with other putative transcriptions factors, such as RAV1, ARF1, bZIPs (AtbZIP2, AtbZIP11, AtbZIP44 and AtbZIP53) and MYB1AT that are thought to bind the promoter sequence of *EGY3* would also be of interest. With these strategies the role of *EGY3* in processes such as lateral root formation, seed maturation, senescence and abiotic stress-related phenotypes could be better understood. The ongoing overexpression of the protein in *E. coli* will prove fundamental to determine the existence of metalloproteasic activity in *EGY3*. Finally, a screening for potential interactors could be important in order to establish the function and precise roles assigned to *EGY3*.

6.

"Bibliography"

- Ahuja I., de Vos R.C.H., Bones A.M. and Hall R.D.** (2010). Plant molecular stress responses face climate change. *TRENDS in Plant Science* **15**: 664-674.
- Ajjawi I., Lu Y., Savage L.J., Bell S.M. and Last R.L.** (2010). Large-Scale Reverse Genetics in Arabidopsis: Case Studies from the Chloroplast 2010 Project. *Plant Physiology* **152**: 529–540.
- Alba B.M., Leeds J.A., Onufryk C., Lu C.Z. and Gross C.A.** (2002). DegS and YaeL participate sequentially in the cleavage of RseA to activate the σ^E -dependent extracytoplasmic stress response. *Genes & Development* **16**: 2156–2168.
- Alonso J.M., Stepanova A.N., Leisse T.J., Kim C.J., Chen H., Shinn P., Stevenson D.K., Zimmerman J., Barajas P., Cheuk R., Gadrinab C., Heller C., Jeske A., Koesema E., Meyers C.C., Parker H., Prednis L., Ansari Y., Choy N., Deen H., Geralt M., Hazari N., Hom E., Karnes M., Mulholland C., Ndubaku R., Schmidt I., Guzman P., Aguilar-Henonin L., Schmid M., Weigel D., Carter D.E., Marchand T., Risseuw E., Brogden D., Zeko A., Crosby W.L., Berry C.C. and Ecker J.R.** (2003). Genome-Wide Insertional Mutagenesis of *Arabidopsis thaliana*. *Science* **301**: 653-657.
- Alonso J.M. and Ecker J.R.** (2006). Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. *Nature Reviews Genetics* **7**: 524-536.
- Alonso R., Oñate-Sánchez L., Weltmeier F., Ehler A., Diaz I., Dietrich K., Vicente-Carbajosa J. and Dröge-Laserb W.** (2009). A Pivotal Role of the Basic Leucine Zipper Transcription Factor bZIP53 in the Regulation of Arabidopsis Seed Maturation Gene Expression Based on Heterodimerization and Protein Complex Formation. *The Plant Cell* **21**: 1747–1761.
- Armengaud P., Zambaux K., Hills A., Sulpice R., Pattison R.J., Blatt M.R. and Amtmann A.** (2009). EZ-RHIZO: integrated software for the fast and accurate measurement of root system architecture. *The Plant Journal* **57**: 945–956.
- Azpiroz R., Wu Y., LoCascio J.C. and Feldmann K.A.** (1998). An Arabidopsis Brassinosteroid-Dependent Mutant Is Blocked in Cell Elongation. *The Plant Cell* **10**: 219–230.
- Baena-González E., Rolland F., Thevelein J.M. and Sheen J.** (2007). A central integrator of transcription networks in plant stress and energy signalling. *Nature* **448**: 938-943.
- Baena-González E.** (2010). Energy Signaling in the Regulation of Gene Expression during Stress. *Molecular Plant* **3**: 300–313.
- Balasubramanian S., Sureshkumar S., Lempe J. and Weigel D.** (2006). Potent Induction of *Arabidopsis thaliana* Flowering by Elevated Growth Temperature. *PLoS Genetics* **2**: e106.
- Banti V., Mafessoni F., Loreti E., Alpi A. and Perata P.** (2010). The Heat-Inducible Transcription Factor *HsfA2* Enhances Anoxia Tolerance in Arabidopsis. *Plant Physiology* **152**: 1471-1483
- Bassel G.W., Fung P., Chow T.-f.F., Foong J.A., Provart N.J. and Cutler S.R.** (2008). Elucidating the Germination Transcriptional Program Using Small Molecules. *Plant Physiology* **147**: 143–155.
- Bassham D.C., Brandizzi F., Otegui M.S. and Sanderfoot A.A.** (2008). The Secretory System of Arabidopsis. *The Arabidopsis Book*.
- Bentsinka L. and Koornneef M.** (2008). Seed Dormancy and Germination. *The Arabidopsis Book*.
- Benveniste P.** (2002). Sterol Metabolism. *The Arabidopsis Book*.

Benveniste P. (2004). Biosynthesis and Accumulation of Sterols. *Annual Review Plant Biology* **55**: 429–457.

Berg J.M., Tymoczko J.L. and Stryer L. (2002). The Biosynthesis of Membrane Lipids and Steroids. *Biochemistry*. New York, W. H. Freeman.

Berleth T. and Chatfield S. (2002). Embryogenesis: Pattern Formation from a Single Cell. *The Arabidopsis Book*.

Bernard P. and Couturier M. (1992). Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *Journal of Molecular Biology* **226**: 735-745

Bhushan S., Lefebvre B., Ståhl A., J.Wright S., Bruce B.D., Boutry M. and Glaser E. (2003). Dual targeting and function of a protease in mitochondria and chloroplasts. *EMBO reports* **4**: 1073-1078.

Bölter B., Nadaa A., Fulgosib H. and Soll J. (2006). A chloroplastic inner envelope membrane protease is essential for plant development. *FEBS Letters* **580**: 789–794.

Borsani O., Valpuesta V. and Botella M.A. (2001). Evidence for a Role of Salicylic Acid in the Oxidative Damage Generated by NaCl and Osmotic Stress in Arabidopsis Seedlings. *Plant Physiology* **126**: 1024–1030.

Bouché N. and Bouchez D. (2001). Arabidopsis gene knockout: phenotypes wanted. *Current Opinion in Plant Biology* **4**: 111–117.

Boutté Y. and Grebe M. (2009). Cellular processes relying on sterol function in plants. *Current Opinion in Plant Biology* **12**: 1-9.

Boyes D.C., Zayed A.M., Ascenzi R., McCaskill A.J., Hoffman N.E., Davis K.R. and Görlach J. (2001). Growth Stage–Based Phenotypic Analysis of Arabidopsis: A Model for High Throughput Functional Genomics in Plants. *The Plant Cell* **13**: 1499–1510.

Briggs G.C., Osmont K.S., Shindo C., Sibout R. and Hardtke C.S. (2006). Unequal genetic redundancies in *Arabidopsis* – a neglected phenomenon? *Trends in Plant Science* **11**: 492-498.

Brown M.S. and Goldstein J.L. (1997). The SREBP Pathway: Regulation of Cholesterol Metabolism by Proteolysis of a Membrane-Bound Transcription Factor. *Cell* **89**: 331–340.

Buchanan B.B., W Gruissem and Jones R.L. (2000). *Biochemistry & Molecular Biology of Plants*. 1st edition. Rockville, Maryland, American Society of Plant Physiologists.

Bullock W.O., Fernandez J.M. and Short J.M. (1987). XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. *BioTechniques* **5**: 376-379.

Carland F., Fujioka S. and Nelson T. (2010). The Sterol Methyltransferases SMT1, SMT2, and SMT3 Influence Arabidopsis Development through Nonbrassinosteroid Products. *Plant Physiology* **153**: 741–756.

Cavalier-Smith T. (1981). Eukaryote Kingdoms: Seven or Nine? *BioSystems* **14**: 461-481.

Cavalier-Smith T. (1998). A revised six-kingdom system of life. *Biological Reviews* **73**: 203-266.

Cavalier-Smith T. (2010). Kingdoms Protozoa and Chromista and the eozoan root of the eukaryotic tree. *Biology Letters* **6**: 342-345.

- Cederholm S.N., Bäckman H.G., Pesaresi P., Leister D. and Glaser E.** (2009). Deletion of an organellar peptidasome PreP affects early development in *Arabidopsis thaliana*. *Plant Molecular Biology* **71**: 497–508.
- Century K., Reuber T.L. and Ratcliffe O.J.** (2008). Regulating the Regulators: The Future Prospects for Transcription-Factor-Based Agricultural Biotechnology Products. *Plant Physiology* **147**: 20–29.
- Chaudhury A.M., Letham S., Craig S. and Dennis E.S.** (1993). *amp1* – a mutant with high cytokinin levels and altered embryonic pattern, faster vegetative growth, constitutive photomorphogenesis and precocious flowering. *The Plant Journal* **4**: 907-916.
- Chen G., Yu Rong Bi and Li N.** (2005). *EGY1* encodes a membrane-associated and ATP-independent metalloprotease that is required for chloroplast development. *The Plant Journal* **41**: 364–375.
- Chen G. and Zhang X.** (2010). New insights into S2P signaling cascades: Regulation, variation, and conservation. *Protein Science* **19**: 2015–2030.
- Chen J., Burke J.J., Velten J. and Xin Z.** (2006). FtsH11 protease plays a critical role in Arabidopsis thermotolerance. *The Plant Journal* **48**: 73–84.
- Cheng J., He C.-X., Zhang Z.-W., Xu F., Zhang D.-W., Wang X., Yuan S. and Lin H.-H.** (2011). Plastid Signals Confer Arabidopsis Tolerance to Water Stress. *Journal of Biosciences* **66**: 47-54.
- Cheong Y.H., Chang H.S., Gupta R., Wang X., Zhu T. and Luan S.** (2002). Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in Arabidopsis. *Plant Physiology* **129**: 661-677.
- Choe S., Dilkes B.P., Fujioka S., Takatsuto S., Sakurai A. and Feldmann K.A.** (1998). The *DWF4* Gene of Arabidopsis Encodes a Cytochrome P450 That Mediates Multiple 22 α -Hydroxylation Steps in Brassinosteroid Biosynthesis. *The Plant Cell* **10**: 231–243.
- Choe S., Noguchi T., Fujioka S., Takatsuto S., Tissier C.P., Gregory B.D., Ross A.S., Tanaka A., Yoshida S., Tax F.E. and Feldmann K.A.** (1999). The Arabidopsis *dwf7/ste1* Mutant Is Defective in the Δ^7 Sterol C-5 Desaturation Step Leading to Brassinosteroid Biosynthesis. *The Plant Cell* **11**: 207-221.
- Choe S., Tanaka A., Noguchi T., Fujioka S., Takatsuto S., Ross A.S., Tax F.E., Yoshida S. and Feldman K.A.** (2000). Lesion in the sterol Δ^7 reductase gene of Arabidopsis cause dwarfism due to a block in brassinosteroid biosynthesis. *The Plant Journal* **21**: 431-443.
- Chory J., Nagpal P. and Petob C.A.** (1991). Phenotypic and Genetic Analysis of *def2*, a New Mutant That Affects Light-Regulated Seedling Development in Arabidopsis. *The Plant Cell* **3**: 445-459.
- Clarke A.K.** (2005). Plant proteases – an appetite for destruction. *Physiologia Plantarum* **123**: 359–361.
- Clough S.J. and Bent A.F.** (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16**: 735–743.
- Clouse S.D., Langford M. and McMorris T.C.** (1996). A Brassinosteroid-Insensitive Mutant in *Arabidopsis thaliana* Exhibits Multiple Defects in Growth and Development. *Plant Physiology* **111**: 671-678.
- Clouse S.D.** (2002). Brassinosteroids. *The Arabidopsis Book*.
- Cohen-Peer R., Schuster S., Meiri D., Breiman A. and Avni A.** (2010). Sumoylation of Arabidopsis heat shock factor A2 (HsfA2) modifies its activity during acquired thermotolerance. *Plant Molecular Biology* **74**: 33–45.

- Collmera A., Lindeberg M., Petnicki-Ocwiejac T., Schneiderb D.J. and Alfano J.R.** (2002). Genomic mining type III secretion system effectors in *Pseudomonas syringae* yields new picks for all TTSS prospectors. *Trends in Microbiology* **10**: 462-469
- Conway L.J. and Poethig R.S.** (1997). Mutations of *Arabidopsis thaliana* that transform leaves into cotyledons. *Proceedings of the National Academy of Sciences* **94**: 10209–10214.
- Cottage A., Mott E.K., Kempster J.A. and Gray J.C.** (2010). The *Arabidopsis* plastid-signalling mutant *gun1* (*genomes uncoupled1*) shows altered sensitivity to sucrose and abscisic acid and alterations in early seedling development. *Journal of Experimental Botany* **61**: 3773-3786.
- Craigon D.J., James N., Okyere J., Higgins J., Jotham J. and May S.** (2004). NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. *Nucleic Acids Research* **32**: D575-D577.
- Cuppels D.A.** (1986). Generation and Characterization of Tn5 Insertion Mutations in *Pseudomonas syringae* pv. *tomato*. *Applied and Environmental Microbiology* **51**: 323-327.
- Curtis M.D. and Grossniklaus U.** (2003). A Gateway Cloning Vector Set for High-Throughput Functional Analysis of Genes in Planta. *Plant Physiology* **133**: 462–469.
- Davuluri R.V., Sun H., Palaniswamy S.K., Matthews N., Molina C., Kurtz M. and Grotewold E.** (2003). AGRIS: Arabidopsis Gene Regulatory Information Server, an information resource of Arabidopsis cis-regulatory elements and transcription factors. *BMC Bioinformatics* **4**: 1-11.
- Denecke J.** (2001). Plant Endoplasmic Reticulum. Encyclopedia of Life Sciences, John Wiley & Sons, Ltd.
- Dereeper A., Guignon V., Blanc G., Audic S., Buffet S., Chevenet F., Dufayard J.-F., Guindon S., Lefort V., Lescot M., Claverie J.-M. and Gascuel O.** (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research* **36**: W465–W469.
- Diener A.C., Li H., Zhou W.-x., Whoriskey W.J., Nes W.D. and Fink G.R.** (2000). *STEROL METHYLTRANSFERASE 1* Controls the Level of Cholesterol in Plants. *The Plant Cell* **12**: 853-870.
- Dietrich K., Weltmeier F., Ehlert A., Weiste C., Stahl M., Harter K. and Dröge-Lasera W.** (2011). Heterodimers of the Arabidopsis Transcription Factors bZIP1 and bZIP53 Reprogram Amino Acid Metabolism during Low Energy Stress. *The Plant Cell* **23**: 381–395.
- Dolan L., Janmaat K., Willemsen V., Linstead P., Poethig S., Roberts K. and Scheres B.** (1993). Cellular organisation of the *Arabidopsis thaliana* root. *Development* **119**: 71-84.
- Doyle J.J. and Doyle J.L.** (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry Bulletin* **19**: 11-15.
- Dubos C., Stracke R., Grotewold E., Weisshaar B., Martin C. and Lepiniec L.** (2010). MYB transcription factors in Arabidopsis. *Trends in Plant Science* **15**: 573-581
- Dufourc E.J.** (2008). Sterols and membrane dynamics. *Journal of Chemical Biology* **1**: 63-77.
- Edwards K., Johnstone C. and Thompson C.** (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research* **19**: 1349.
- Emanuelsson O., Nielsen H. and Heijne G.V.** (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Science* **8**: 978–984.

- Emanuelsson O., Brunak S., Heijne G.v. and Nielsen H.** (2007). Locating proteins in the cell using TargetP, SignalP, and related tools. *Nature Protocols* **2**: 953 - 971.
- Enjuto M., Balcells L., Campos N., Caelles C., Arr M. and Boronat A.** (1994). *Arabidopsis thaliana* contains two differentially expressed 3-hydroxy-3-methylglutaryl-CoA reductase genes, which encode microsomal forms of the enzyme. *Proceedings of the National Academy of Sciences* **91**: 927-931.
- Enjuto M., Lumbreras V., Marín C. and Boronat A.** (1995). Expression of the *Arabidopsis HMG2* Gene, Encoding 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase, Is Restricted to Meristematic and Floral Tissues. *The Plant Cell* **7**: 517-527.
- Feng C.-P. and Mundy J.** (2006). Gene Discovery and Functional Analyses in the Model Plant *Arabidopsis*. *Journal of Integrative Plant Biology* **48**: 5-14.
- Feuillet C., Leach J.E., Rogers J., Schnable P.S. and Eversole K.** (2011). Crop genome sequencing: lessons and rationales. *Trends in Plant Science* **16**: 77-88.
- Field B. and Osbourn A.E.** (2008). Metabolic Diversification-Independent Assembly of Operon-Like Gene Clusters in Different Plants. *Science* **320**: 543-547.
- Friedrichsen D.M., Joazeiro C.A.P., Li J., Hunter T. and Chory J.** (2000). Brassinosteroid-Insensitive-1 Is a Ubiquitously Expressed Leucine-Rich Repeat Receptor Serine/Threonine Kinase. *Plant Physiology* **123**: 1247–1255.
- Fujioka S., Li J., Choi Y.-H., Seto H., Takatsuto S., Noguchi T., Watanabe T., Kuriyama H., Yokota T., Chory J. and Sakurai A.** (1997). The *Arabidopsis deetiolated2* Mutant Is Blocked Early in Brassinosteroid Biosynthesis. *The Plant Cell* **9**: 1951-1962.
- Gadjev I., Vanderauwera S., Gechev T.S., Laloi C., Minkov I.N., Shulaev V., Apel K., Inzé D., Mittler R. and Breusegem F.V.** (2006). Transcriptomic Footprints Disclose Specificity of Reactive Oxygen Species Signaling in *Arabidopsis*. *Plant Physiology* **141**: 436–445.
- Gao H., Brandizzi F., Benning C. and Larkin R.M.** (2008). A membrane-tethered transcription factor defines a branch of the heat stress response in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* **105**: 16398–16403.
- Gapper C. and Dolan L.** (2006). Control of Plant Development by Reactive Oxygen Species. *Plant Physiology* **141**: 341–345.
- García-Lorenzo M., Sjödin A., Jansson S. and Funk C.** (2006). Protease gene families in *Populus* and *Arabidopsis*. *BMC Plant Biology* **6**.
- Glazebrook J.** (2005). Contrasting Mechanisms of Defense Against Biotrophic and Necrotrophic Pathogens. *Annual Review Phytopathology* **43**: 205–227.
- Goda H., Shimada Y., Asami T., Fujioka S. and Yoshida S.** (2002). Microarray Analysis of Brassinosteroid-Regulated Genes in *Arabidopsis*. *Plant Physiology* **130**: 1319–1334.
- Goda H., Sasaki E., Akiyama K., Maruyama-Nakashita A., Nakabayashi K., Li W., Ogawa M., Yamauchi Y., Preston J., Aoki K., Kiba T., Takatsuto S., Fujioka S., Asami T., Nakano T., Kato H., Mizuno T., Sakakibara H., Yamaguchi S., Nambara E., Kamiya Y., Takahashi H., Hirai M.Y., Sakurai T., Shinozaki K., Saito K., Yoshida S. and Shimada Y.** (2008). The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. *The Plant Journal* **55**: 526–542.

Gray W.M., Östin A., Sandberg G., Romano C.P. and Estelle M. (1998). High temperature promotes auxin-mediated hypocotyl elongation in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **95**: 7197–7202.

Griffin H.G. and Griffin A.M. (1994). *PCR Technology: Current Innovations*. 1st, Boca Raton: CRC Press.

Griffith M. and Gietz R.D. (2003). *Escherichia coli endA* deletion strain for use in two-hybrid shuttle vector selection. *BioTechniques* **35**: 272-278.

Guilfoyle T.J., Ulmasov T. and Hagen G. (1998). The ARF family of transcription factors and their role in plant hormone-responsive transcription. *Cellular and Molecular Life Sciences* **54**: 619–627.

Guo D., Gao X., Li H., Zhang T., Chen G., Huang P., An L. and Li N. (2008). *EGY1* plays a role in regulation of endodermal plastid size and number that are involved in ethylene-dependent gravitropism of light-grown *Arabidopsis* hypocotyls. *Plant Molecular Biology* **66**: 345-360.

Gutierrez L., Wuytswinkel O.V., Castelain M. and Bellini C. (2007). Combined networks regulating seed maturation. *Trends in Plant Science* **12**: 294-300.

Hanson J., Hanssen M., Wiese A., Hendriks M.M.W.B. and Smeekens S. (2008). The sucrose regulated transcription factor bZIP11 affects amino acid metabolism by regulating the expression of *ASPARAGINE SYNTHETASE1* and *PROLINE DEHYDROGENASE2*. *The Plant Journal* **53**: 935–949.

Haughn G. and Chaudhury A. (2005). Genetic analysis of seed coat development in *Arabidopsis*. *Trends in Plant Science* **10**: 472-477.

Heazlewood J.L., Verboom R.E., Tonti-Filippini J. I. S. and AH M. (2007). SUBA: the *Arabidopsis* Subcellular Database. *Nucleic Acids Research* **35**: 213–218.

Hellmann H. and Estelle M. (2002). Plant Development: Regulation by Protein Degradation. *Science* **297**: 793-797.

Henriques R., Jang I.-C. and Chua N.-H. (2009). Regulated proteolysis in light-related signaling pathways. *Current Opinion in Plant Biology* **12**: 49-56.

Hilson P., Small I. and Kuiper M.T. (2003). European consortia building integrated resources for *Arabidopsis* functional genomics. *Current Opinion in Plant Biology* **6**: 426–429.

Holmes D.S. and Quigley M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Analytical Biochemistry* **114**: 193-197

Hood E.E., Helmer G.L., Fraley R.T. and Chilton M.-D. (1986). The Hypervirulence of *Agrobacterium tumefaciens* A281 Is Encoded in a Region of pTiBo542 Outside of T-DNA. *Journal of Bacteriology* **168**: 1291-1301.

Hooper N.M. (1994). Families of zinc metalloproteases. *FEBS Letters* **354**: 1-6.

Hopkins W.G. and Huner N.P.A. (2008). *Introduction to Plant Physiology*. 4th edition. NJ, USA, John Wiley & Sons Inc.

Hruz T., Laule O., Szabo G., FransWessendorp, Bleuler S., Oertle L., PeterWidmayer, Gruissem W. and Zimmermann P. (2008) Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes. *Advances in Bioinformatics* vol. 2008, Article ID 420747, 5 pages, doi:10.1155/2008/420747

Hu Y.X., Wang Y.H., Liu X.F. and Li J.Y. (2004). *Arabidopsis* RAV1 is down-regulated by brassinosteroid and may act as a negative regulator during plant development. *Cell Research* **14**: 8-15.

- Huang B. and Xu C.** (2008). Identification and characterization of proteins associated with plant tolerance to heat stress. *Journal of Integrative Plant Biology* **50**: 1230–1237.
- Hunter S., Apweiler R., Attwood T.K., Bairoch A., Bateman A., Binns D., Bork P., Das U., Daugherty L., Duquenne L., Finn R.D., Gough J., Haft D., Hulo N., Kahn D., Kelly E., Laugraud A., Letunic I., Lonsdale D., Lopez R., Madera M., Maslen J., McAnulla C., McDowall J., Mistry J., Mitchell A., Mulder N., Natale D., Orengo C., Quinn A.F., Selengut J.D., Sigrist C.J.A., Thimma M., Thomas P.D., Valentin F., Wilson D., Wu C.H. and Yeats C.** (2009). InterPro: the integrative protein signature database. *Nucleic Acids Research* **37**: D211–D215.
- Husselstein-Muller T., Schaller H. and Benveniste P.** (2001). Molecular cloning and expression in yeast of 2,3-oxidosqualenetripenoid cyclases from *Arabidopsis thaliana*. *Plant Molecular Biology* **45**: 75–92.
- Inoue H., Nojima H. and Okayama H.** (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96**: 23-28.
- Jakoby M., Weisshaar B., Dröge-Laser W., Vicente-Carbajosa J., Tiedemann J., Kroj T. and Parcy F.** (2002). bZIP transcription factors in *Arabidopsis*. *Trends in Plant Science* **7**: 106-111.
- Jang S., Marchal V., Panigrahi K.C., Wenkel S., Soppe W., Deng X.-W., Valverde F. and Coupland G.** (2008). *Arabidopsis* COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic flowering response. *The EMBO Journal* **27**: 1277–1288.
- Jefferson R.A.** (1987). Assaying chimeric genes in plants: The GUS gene fusion system. *Plant Molecular Biology Reporter* **5**: 387-405.
- Johnston M., Foley J.A., Holloway T., Kucharik C. and Monfreda C.** (2009). Resetting global expectations from agricultural biofuels. *Environmental Research Letters* **4**: 014004.
- Kato Y. and Sakamoto W.** (2010). New insights into the types and function of proteases in plastids. *International Review of Cell and Molecular Biology*. KW Jeon. Burlington, Academic Press Elsevier. **280**: 185-218.
- Kersey P.J., Lawson D., Birney E., Derwent P.S., Haimel M., Herrero J., Keenan S., Kerhornou A., Koscielny G., Kähäri A., Kinsella R.J., Kulesha E., Maheswari U., Megy K., Nuhn M., Proctor G., Staines D., F. Valentin, Vilella A.J. and Yates A.** (2010). Ensembl Genomes: Extending Ensembl across the taxonomic space. *Nucleic Acids Research* **38**: D563–D569.
- Kilian J., Whitehead D., Horak J., Wanke D., Weinl S., Batistic O., D'Angelo C., Bornberg-Bauer E., Kudla J. and Harter K.** (2007). The AtGenExpress global stress expression data set: protocols, evaluation and model analysis of UV-B light, drought and cold stress responses. *The Plant Journal* **50**: 347-363.
- Kim H.B., Kwon M., Ryu H., Fujioka S., Takatsuto S., Yoshida S., An C.S., Lee I., Hwang I. and Choe S.** (2006). The Regulation of *DWARF4* Expression Is Likely a Critical Mechanism in Maintaining the Homeostasis of Bioactive Brassinosteroids in *Arabidopsis*. *Plant Physiology* **140**: 548–557.
- Kim T.-W. and Wang Z.-Y.** (2010). Brassinosteroid Signal Transduction from Receptor Kinases to Transcription Factors. *Annual Review Plant Biology* **61**: 681–704.
- Klahre U., Noguchi T., Fujioka S., Takatsuto S., Yokota T., Nomura T., Yoshida S. and Chuua N.-H.** (1998). The *Arabidopsis* *DIMINUTO/DWARF1* Gene Encodes a Protein Involved in Steroid Synthesis. *The Plant Cell* **10**: 1677–1690.
- Koiwa H., Bressan R.A. and Hasegawa P.M.** (2006). Identification of plant stress-responsive determinants in *Arabidopsis* by large-scale forward genetic screens. *Journal of Experimental Botany* **57**: 1119–1128.

Kolesnikova M.D., Wilson W.K., Lynch D.A., Obermeyer A.C. and Matsuda S.P.T. (2007). *Arabidopsis* Camelliol C Synthase Evolved from Enzymes That Make Pentacycles. *Organic Letters* **9**: 5223-5226.

Koncz C. and Schell J. (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Molecular and General Genetics* **204**: 383-396.

Koornneef M. and Meinke D. (2010). The development of *Arabidopsis* as a model plant. *The Plant Journal* **61**: 909–921.

Kotak S., Larkindale J., Lee U., Koskull-Döring P.v., Vierling E. and Scharf K.-D. (2007). Complexity of the heat stress response in plants. *Current Opinion in Plant Biology* **10**: 310–316.

Krogh A., Larsson B., Heijne G.v. and Sonnhammer E.L.L. (2001). Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. *Journal of Molecular Biology* **305**: 567-580.

Kumara M., Busch W., Birke H., Kemmerling B., Nürnberger T. and Schöffl F. (2009). Heat Shock Factors HsfB1 and HsfB2b Are Involved in the Regulation of *Pdf1.2* Expression and Pathogen Resistance in *Arabidopsis*. *Molecular Plant* **2**: 152–165.

Laemmli U.K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227**: 680-685.

Landl K.M., Klösch B. and F Turnowsky (1996). *ERG1*, encoding squalene epoxidase, is located on the right arm of chromosome VII of *Saccharomyces cerevisiae*. *Yeast* **12**: 609-613.

Larkin R.M. and Ruckle M.E. (2008). Integration of light and plastid signals. *Current Opinion in Plant Biology* **11**: 593-599.

Larkindale J., Hall J.D., Knight M.R. and Vierling E. (2005). Heat Stress Phenotypes of *Arabidopsis* Mutants Implicate Multiple Signaling Pathways in the Acquisition of Thermotolerance. *Plant Physiology* **138**: 882–897.

Lee U., Rioflorido I., Hong S.-W., Larkindale J., Waters E.R. and Vierling E. (2006). The *Arabidopsis* ClpB/Hsp100 family of proteins: chaperones for stress and chloroplast development. *The Plant Journal* **49**: 115–127.

Leonelli S. (2006). *Arabidopsis*, the botanical *Drosophila*: from mouse cress to model organism. *Endeavour* **31**: 34-38.

Lewis A.P. and Thomas P.J. (1999). A novel clan of zinc metallopeptidases with possible intramembrane cleavage properties. *Protein Science* **8**: 439-442.

Li J. and Chory J. (1997). A Putative Leucine-Rich Repeat Receptor Kinase Involved in Brassinosteroid Signal Transduction. *Cell* **90**: 929–938.

Li J. and Jin H. (2006). Regulation of brassinosteroid signaling. *Trends in Plant Science* **12**: 37-41.

Lindsey K., Pullen M.L. and Topping J.F. (2003). Importance of plant sterols in pattern formation and hormone signalling. *Trends in Plant Science* **8**: 521-525.

Liu J.-X., Srivastava R., Che P. and Howell S.H. (2007). Salt stress responses in *Arabidopsis* utilize a signal transduction pathway related to endoplasmic reticulum stress signaling. *The Plant Journal* **51**: 897–909.

Lodeiro S., Xiong Q., Wilson W.K., Kolesnikova M.D., Onak C.S. and Matsuda S.P.T. (2007). An Oxidosqualene Cyclase Makes Numerous Products by Diverse Mechanisms: A Challenge to Prevailing Concepts of Triterpene Biosynthesis. *Journal of the American Chemical Society* **129**: 11213-11222.

Lohmann C., Eggers-Schumacher G., Wunderlich M. and Schöffl F. (2004). Two different heat shock transcription factors regulate immediate early expression of stress genes in Arabidopsis. *Molecular Genetics and Genomics* **271**: 11-21.

Lovato M.A., Hart E.A., Segura M.J.R., Giner J.-L. and Matsuda S.P.T. (2000). Functional Cloning of an *Arabidopsis thaliana* cDNA Encoding Cycloecalenol Cycloisomerase. *The Journal of Biological Chemistry* **18**: 13394–13397.

Ma J., Hanssen M., Lundgren K., Hernández L., Delatte T., Ehlert A., Liu C.-M., Schluempmann H., Dröge-Laser W., Moritz T., Smeekens S. and Hanson J. (2011). The sucrose-regulated Arabidopsis transcription factor bZIP11 reprograms metabolism and regulates trehalose metabolism. *New Phytologist* **191**: 733–745.

Macho A.P., Guevara C.M., Tornero P., Ruiz-Albert J. and Beuzón C.R. (2010). The *Pseudomonas syringae* effector protein HopZ1a suppresses effector-triggered immunity. *New Phytologist* **187**: 1018–1033.

Marcus L., Hartnett J. and Storts D.R. (1996). The pGEM®-T and pGEM®-T Easy Vector Systems. *Promega Notes Magazine* **58**: 36-38.

Marion J., Bach L., Bellec Y., Meyer C., Gissot L. and Faure J.-D. (2008). Systematic analysis of protein subcellular localization and interaction using high-throughput transient transformation of Arabidopsis seedlings. *The Plant Journal* **56**: 169–179.

Martínez-Zapater J.M. (1993). Genetic Analysis of Variegated Mutants in *Arabidopsis*. *Journal of Heredity* **84**: 138-140.

MASC Report (2002). Annual Report. [The Multinational Coordinated Arabidopsis thaliana Functional Genomics Project.](#)

MASC Report (2010). Annual Report. [The Multinational Coordinated Arabidopsis thaliana Functional Genomics Project.](#)

MASC Report (2011). Annual Report. [The Multinational Coordinated Arabidopsis thaliana Functional Genomics Project.](#)

Mathur J., Molnár G., Fujioka S., Takatsuto S., Sakurai A., Yokota T., Adam G., Voigt B., Nagy F., Maas C., Schell J., Koncz C. and Szekeres M. (1998). Transcription of the *Arabidopsis* CPD gene, encoding a steroidogenic cytochrome P450, is negatively controlled by brassinosteroids. *The Plant Journal* **14**: 593–602.

Maurino V.G. and Flügge U.-I. (2008). Experimental systems to assess the effects of reactive oxygen species in plant tissues. *Plant Signaling & Behavior* **3**: 923-928.

Meinke D.W. and Sussex I.M. (1979). Embryo-lethal Mutants of *Arabidopsis thaliana*. *Developmental Biology* **72**: 50-61.

Meinke D., Muralla R., Sweeney C. and Dickerman A. (2008). Identifying essential genes in *Arabidopsis thaliana*. *Trends in Plant Science* **13**: 483-491.

Meyer T.S. and Lamberts B.L. (1965). Use of Coomassie brilliant blue R250 for the electrophoresis of microgram quantities of parotid saliva proteins on acrylamide-gel strips. *Biochimica et Biophysica Acta* **107**: 144-145.

Miko I. (2008). Phenotype variability: penetrance and expressivity. *Nature Education* **1**.

Mittler R. (2006). Abiotic stress, the field environment and stress combination. *Trends in Plant Science* **11**: 15-19.

Mittler R. and Blumwald E. (2010). Genetic Engineering for Modern Agriculture: Challenges and Perspectives. *Annual Review of Plant Biology* **61**: 443–462.

Moffatt B.A. and Studier F.W. (1987). T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* **49**: 221-227.

Montejo J., Zuberi K., Rodriguez H., Kazi F., Wright G., Donaldson S.L., Morris Q. and Bader G.D. (2010). GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics* **26**: 2927–2928.

Mordhorst A.P., Voerman K.J., Hartog M.V., Meijer E.A., Went J.v., Koornneef M. and Vries S.C.d. (1998). Somatic Embryogenesis in *Arabidopsis thaliana* Is Facilitated by Mutations in Genes Repressing Meristematic Cell Divisions. *Genetics* **149**: 549–563.

Moseyko N., Zhu T., Chang H.S., Wang X. and Feldman L.J. (2002). Transcription profiling of the early gravitropic response in *Arabidopsis* using high-density oligonucleotide probe microarrays. *Plant Physiology* **130**: 720-728.

Murashige T. and Skoog F. (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* **15**: 473-497.

Nagai M., Sakakibara J., Wakui K., Fukushima Y., Igarashi S., Tsuji S., Arakawa M. and Ono T. (1997). Localization of the Squalene Epoxidase Gene (SQLE) to Human Chromosome Region 8q24.1. *Genomics* **44**: 141-143

Nelson B.K., Cai X. and Nebenführ A. (2007). A multicolored set of in vivo organelle markers for co-localization studies in *Arabidopsis* and other plants. *The Plant Journal* **51**: 1126–1136.

Nelson D.L. and Cox M.M. (2000). Lehninger Principles of Biochemistry. 3rd edition. New York, Worth Publishers.

Neta-Sharir I., Isaacson T., Lurie S. and Weiss D. (2005). Dual Role for Tomato Heat Shock Protein 21: Protecting Photosystem II from Oxidative Stress and Promoting Color Changes during Fruit Maturation. *The Plant Cell* **17**: 1829–1838.

Nieto B., Forés O., Arró M. and Ferrer A. (2009). *Arabidopsis* 3-hydroxy-3-methylglutaryl-CoA reductase is regulated at the post-translational level in response to alterations of the sphingolipid and the sterol biosynthetic pathways. *Phytochemistry* **70**: 53-59.

Nishizawa-Yokoi A., Tainaka H., Yoshida E., Tamoi M., Yabuta Y. and Shigeoka S. (2010). The 26S Proteasome Function and Hsp90 Activity Involved in the Regulation of HsfA2 Expression in Response to Oxidative Stress. *Plant & Cell Physiology* **51**: 486-496.

Nishizawa-Yokoi A., Nosaka R., Hayashi H., Tainaka H., Maruta T., Tamoi M., Ikeda M., Ohme-Takagi M., Yoshimura K., Yabuta Y. and Shigeoka S. (2011). HsfA1d and HsfA1e Involved in the Transcriptional

Regulation of HsfA2 Function as Key Regulators for the Hsf Signaling Network in Response to Environmental Stress. *Plant & Cell Physiology* **52**: 933–945.

Nishizawa A., Yabuta Y., Yoshida E., Maruta T., Yoshimura K. and Shigeoka S. (2006). Arabidopsis heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *The Plant Journal* **48**: 535–547.

Nover L., Bharti K., Döring P., Mishra S.K., Ganguli A. and Scharf K.-D. (2001). Arabidopsis and the heat stress transcription factor world: how many heat stress transcription factors do we need? *Cell Stress & Chaperones* **6**: 177–189.

O'Connor T.R., Dyreson C. and Wyrick J.J. (2005). Athena: a resource for rapid visualization and systematic analysis of Arabidopsis promoter sequences. *Bioinformatics* **21**: 4411–4413.

O'Malley R.C. and Ecker J.R. (2010). Linking genotype to phenotype using the Arabidopsis unimutant collection. *The Plant Journal* **61**: 928–940.

Obayashi T., Kinoshita K., Nakai K., Shibaoka M., Hayashi S., Saeki M., Shibata D., Saito K. and Ohta H. (2007). ATTED-II: a database of co-expressed genes and cis elements for identifying co-regulated gene groups in Arabidopsis. *Nucleic Acids Research* **35**: D863-D869.

Obayashi T., Hayashi S., Saeki M., Ohta H. and Kinoshita K. (2009). ATTED-II provides coexpressed gene networks for Arabidopsis. *Nucleic Acids Research* **37**: D987-991.

Obayashi T., Nishida K., Kasahara K. and Kinoshita K. (2011). ATTED-II updates: condition-specific gene coexpression to extend coexpression analyses and applications to a broad range of flowering plants. *Plant & Cell Physiology* **52**: 213-219

Ogawa D., Yamaguchi K. and Nishiuch T. (2007). High-level overexpression of the Arabidopsis HsfA2 gene confers not only increased thermotolerance but also salt/osmotic stress tolerance and enhanced callus growth. *Journal of Experimental Botany* **58**: 3373-3383.

Oh M.-H., Ray W.K., Huber S.C., Asara J.M., Gage D.A. and Clouse S.D. (2000). Recombinant Brassinosteroid Insensitive 1 Receptor-Like Kinase Autophosphorylates on Serine and Threonine Residues and Phosphorylates a Conserved Peptide Motif *in Vitro*. *Plant Physiology* **124**.

Ohyama K., Suzuki M., Masuda K., Yoshida S. and Muranaka T. (2007). Chemical Phenotypes of the *hmg1* and *hmg2* Mutants of Arabidopsis Demonstrate the *In-planta* Role of HMG-CoA Reductase in Triterpene Biosynthesis. *Chemical & Pharmaceutical Bulletin* **55**: 1518-1521.

Ohyama K., Suzuki M., Kikuchi J., Saito K. and Muranaka T. (2009). Dual biosynthetic pathways to phytosterol via cycloartenol and lanosterol in Arabidopsis. *Proceedings of the National Academy of Sciences* **106**: 725–730.

Olinares P.D.B., Kim J. and Wijk K.J.v. (2010). The Clp protease system: a central component of the chloroplast protease network. *Biochimica et Biophysica Acta - Bioenergetics*.

Oono Y., Seki M., Satou M., Iida K., Akiyama K., Sakurai T., Fujita M., Yamaguchi-Shinozaki K. and Shinozaki K. (2006). Monitoring expression profiles of Arabidopsis genes during cold acclimation and deacclimation using DNA microarrays. *Functional & Integrative Genomics* **6**: 212-234.

Osmont K.S., Sibout R. and Hardtke C.S. (2007). Hidden Branches: Developments in Root System Architecture. *Annual Review of Plant Biology* **58**: 93-113.

- Palma J.M., Sandalio L.M., Corpas F.J., Romero-Puertas M.C., McCarthy I. and Rio L.A.d.** (2002). Plant proteases, protein degradation, and oxidative stress: role of peroxisomes. *Plant Physiology Biochemistry* **40**: 521-530.
- Pearson A., Budin M. and Brocks J.J.** (2003). Phylogenetic and biochemical evidence for sterol synthesis in the bacterium *Gemmata obscuriglobus*. *Proceedings of the National Academy of Sciences* **100**: 15352–15357.
- Péret B., Bert De Rybel, Casimiro I., Benková E., Swarup R., Laplaze L., Beeckman T. and Bennett M.J.** (2009). Arabidopsis lateral root development: an emerging story. *Trends in Plant Science* **14**: 399-408.
- Phillips D.R., Rasbery J.M., Bartel B. and Matsuda S.P.** (2006). Biosynthetic diversity in plant triterpene cyclization. *Current Opinion in Plant Biology* **9**: 305–314.
- Posé D. and Botella M.A.** (2009). Analysis of the arabidopsis *dry2/sqe1-5* mutant suggests a role for sterols in signaling. *Plant Signaling & Behavior* **4**: 873-874.
- Posé D., Castanedo I., Borsani O., Nieto B., Rosado A., Taconnat L., Ferrer A., Dolan L., Valpuesta V. and Botella M.A.** (2009). Identification of the Arabidopsis *dry2/sqe1-5* mutant reveals a central role for sterols in drought tolerance and regulation of reactive oxygen species. *The Plant Journal* **59**: 63–76.
- Potters G., Pasternak T.P., Guisez Y., Palme K.J. and Jansen M.A.K.** (2006). Stress-induced morphogenic responses: growing out of trouble? *Trends in Plant Science* **12**.
- Pullen M., Clark N., Zarinkamar F., Topping J. and Lindsey K.** (2010). Analysis of Vascular Development in the *hydra* Sterol Biosynthetic Mutants of Arabidopsis. *PLoS ONE* **5**: e12227.
- Rasbery J.M., Shan H., LeClair R.J., Norman M., Matsuda S.P.T. and Bartel B.** (2007). Arabidopsis thaliana Squalene Epoxidase Is Essential for Root and Seed Development. *The Journal of Biological Chemistry* **282**: 17002–17013.
- Rawlings N.D.** (1999). MEROPS: the peptidase database. *Nucleic Acids Research* **27**: 325-331.
- Rawson R.B., Zelenski N.G., Nijhawan D., Ye J., Sakai J., Hasan M.T., Chang T.Y., Michael S. Brown and Goldstein J.L.** (1997). Complementation Cloning of S2P, a Gene Encoding a Putative Metalloprotease Required for Intramembrane Cleavage of SREBPs. *Molecular Cell* **1**: 47–57.
- Redman J.C., Haas B.J., Tanimoto G. and Town C.D.** (2004). Development and evaluation of an Arabidopsis whole genome Affymetrix probe array. *The Plant Journal* **38-561**: 545.
- Richter S., Zhong R. and Lamppa G.** (2005). Function of the stromal processing peptidase in the chloroplast import pathway. *Physiologia Plantarum* **123**: 362–368.
- Rizhsky L., Liang H., Shuman J., Shulaev V., Davletova S. and Mittler R.** (2004). When defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress. *Plant Physiology* **134**: 1683-1696.
- Roine E., Wei W., Yuan J., Nurmiaho-Lassila E.-L., Kalkkinen N., Romantschuk M. and He S.Y.** (1997). Hrp pilus: An hrp-dependent bacterial surface appendage produced by *Pseudomonas syringae* pv. *tomato* DC3000. *Proceedings of the National Academy of Sciences* **94**: 3459–3464.
- Rubio-Somoza I. and Weigel D.** (2011). MicroRNA networks and developmental plasticity in plants. *Trends in Plant Science* **16**: 258-264.

- Rudner D.Z., Fawcett P. and Losick R.** (1999). A family of membrane-embedded metalloproteases involved in regulated proteolysis of membrane-associated transcription factors. *Proceedings of the National Academy of Sciences* **96**: 14765–14770.
- Ruppel N.J., Logsdon C.A., Whippo C.W., Inoue K. and Hangarter R.P.** (2011). A Mutation in *Arabidopsis* *SEEDLING PLASTID DEVELOPMENT1* Affects Plastid Differentiation in Embryo-Derived Tissues during Seedling Growth. *Plant Physiology* **155**: 342–353.
- Rus E.C.** (2009). Rutas alternativas de biosíntesis de ácido L-ascórbico (vitamina C) en plantas. PhD thesis. Biología Molecular y Bioquímica. Universidad de Málaga. Málaga.
- Ryder N.S.** (1992). Terbinafine: Mode of action and properties of the squalene epoxidase inhibition. *British Journal of Dermatology* **126**: 2-7.
- Sakamoto W., Tamura T., Hanba-Tomita Y., Sodmergen and Murata M.** (2002). The *VAR1* locus of *Arabidopsis* encodes a chloroplastic FtsH and is responsible for leaf variegation in the mutant alleles. *Genes to Cells* **7**: 769–780.
- Sakamoto W., Zaltsman A., Adam Z. and Takahashi Y.** (2003). Coordinated Regulation and Complex Formation of YELLOW VARIEGATED1 and YELLOW VARIEGATED2, Chloroplastic FtsH Metalloproteases Involved in the Repair Cycle of Photosystem II in *Arabidopsis* Thylakoid Membranes. *The Plant Cell* **15**: 2843–2855.
- Sambrook J. and Russel D.W.** (2001). Molecular Cloning. A laboratory manual. 3rd edition. New York, CSHL Press.
- Sánchez-Morán E., Jones G.H., Franklin C.H. and Santos J.L.** (2004). A Puromycin-Sensitive Aminopeptidase Is Essential for Meiosis in *Arabidopsis thaliana*. *The Plant Cell* **16**: 2895–2909.
- Schaller A.** (2004). A cut above the rest: the regulatory function of plant proteases. *Planta* **220**: 183-197.
- Schrick K., Mayer U., Horrichs A., Kuhnt C., Bellini C., Dangl J., Schmidt J. and Jürgens G.** (2000). FACKEL is a sterol C-14 reductase required for organized cell division and expansion in *Arabidopsis* embryogenesis. *Genes & Development* **14**: 1471-1484.
- Seo P.J., Kim S.-G. and Park C.-M.** (2008). Membrane-bound transcription factors in plants. *Trends in Plant Science* **3**: 550-556
- Shibuya M., Katsube Y., Otsuka M., Zhang H., Tansakul P., Xiang T. and Ebizuka Y.** (2009). Identification of a product specific β -amyrin synthase from *Arabidopsis thaliana*. *Plant Physiology and Biochemistry* **47**: 26-30.
- Somerville C. and Koornneef M.** (2001). A fortunate choice: the history of *Arabidopsis* as a model plant. *Nature Reviews Genetics* **3**: 883-889.
- Souter M., Topping J., Pullen M., Friml J., Palme K., Hackett R., Grierson D. and Lindsey K.** (2002). *hydra* Mutants of *Arabidopsis* Are Defective in Sterol Profiles and Auxin and Ethylene Signaling. *The Plant Cell* **14**: 1017-1031.
- Sreenivasulu N., Sopory S.K. and Kishor P.B.K.** (2007). Deciphering the regulatory mechanisms of abiotic stress tolerance in plants by genomic approaches. *Gene* **388**: 1-13
- Ståhl A., Moberg P., Ytterberg J., Panfilov O., Löwenhielm H.B.v., Nilsson F. and Glaser E.** (2002). Isolation and Identification of a Novel Mitochondrial Metalloprotease (PreP) That Degrades Targeting Presequences in Plants. *The Journal of Biological Chemistry* **277**: 41931–41939.

Suzuki M., Kamide Y., Nagata N., Seki H., Ohyama K., Kato H., Masuda K., Sato S., Kato T., Tabata S., Yoshida S. and Muranaka T. (2004). Loss of function of *3-hydroxy-3-methylglutaryl coenzyme A reductase 1 (HMG1)* in *Arabidopsis* leads to dwarfing, early senescence and male sterility, and reduced sterol levels. *The Plant Journal* **37**: 750–761.

Suzuki M., Nakagawa S., Kamide Y., Kobayashi K., Ohyama K., Hashinokuchi H., Kiuchi R., Saito K., Muranaka T. and Nagata N. (2009). Complete blockage of the mevalonate pathway results in male gametophyte lethality. *Journal of Experimental Botany* **60**: 2055–2064.

Swindell W.R., Huebner M. and Weber A.P. (2007). Transcriptional profiling of *Arabidopsis* heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways. *BMC Genomics* **8**: 125.

Szekeres M., Németh K., Koncz-Kálmán Z., Mathur J., Kauschmann A., Altmann T., Rédei G.P., Nagy F., Schell J. and Koncz C. (1996). Brassinosteroids Rescue the Deficiency of CYP90, a Cytochrome P450, Controlling Cell Elongation and De-etiolation in *Arabidopsis*. *Cell* **85**: 171-182

Taiz L. and Zeiger E. (2003). *Plant Physiology*. 3rd edition. Sunderland.

Takahashi S. and Badger M.R. (2011). Photoprotection in plants: a new light on photosystem II damage. *Trends in Plant Science* **16**: 53-58.

Takahashi T., Gasch A., Nishizawa N. and Chua N.-H. (1995). The *DIMINUTO* gene of *Arabidopsis* is involved in regulating cell elongation. *Genes & Development* **9**: 97-107.

Takechi K., Sodmergen, Murata M., Motoyoshi F. and Sakamoto W. (2000). The *YELLOW VARIEGATED (VAR2)* Locus Encodes a Homologue of FtsH, an ATP-Dependent Protease in *Arabidopsis*. *Plant & Cell Physiology* **41**: 1334-1346.

Takeda S., Gapper C., Kaya H., Bell E., Kuchitsu K. and Dolan L. (2008). Local Positive Feedback Regulation Determines Cell Shape in Root Hair Cells. *Science* **319**: 1241-1244

Tamura K., Peterson D., Peterson N., Stecher G., Nei M. and Kumar S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*, doi:10.1093/molbev/msr121

Tamura N., Yoshida T., Tanaka A., Sasaki R., Bando A., Toh S., Lepiniec L. and Kawakami N. (2006). Isolation and Characterization of High Temperature-Resistant Germination Mutants of *Arabidopsis thaliana*. *Plant & Cell Physiology* **47**: 1081–1094.

The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 796-815.

The UniProt Consortium (2011). Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Research* **39**: D214–D219.

Tomoyasu T., Yuki T., Morimura S., Mori H., Yamanaka K., Niki H., Hiraga S. and Ogura T. (1993). The *Escherichia coli* FtsH Protein Is a Prokaryotic Member of a Protein Family of Putative ATPases Involved in Membrane Functions, Cell Cycle Control, and Gene Expression. *The Journal of Bacteriology* **175**: 1344-1351.

Toufighi K., Brady S.M., Austin R., Ly E. and Provart N.J. (2005). The Botany Array Resource: e-Northerns, Expression Angling, and promoter analyses. *The Plant Journal* **43**: 153–163.

- Turnbull C.G.N., Booker J.P. and Leyser H.M.O.** (2002). Micrografting techniques for testing long-distance signalling in Arabidopsis. *The Plant Journal* **32**: 255-262.
- Turnbull C.G.N.** (2010). Grafting as a Research Tool. *Plant Developmental Biology - Methods in Molecular Biology*. L Hennig and C Köhler. New York, Humana Press - Springer Protocols. **655**: 11-26.
- Ufaz S., Shukla V., Soloveichik Y., Golan Y., Breuer F., Koncz Z., Galili G., Koncz C. and Zilberstein A.** (2011). Transcriptional control of aspartate kinase expression during darkness and sugar depletion in Arabidopsis: involvement of bZIP transcription factors. *Planta* **233**: 1025–1040.
- van der Hoorn R.A.L.** (2008). Plant Proteases: From Phenotypes to Molecular Mechanisms. *Annual Review Plant Biology* **59**: 191-223.
- Vinocur B. and Altman A.** (2005). Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. *Current Opinion in Biotechnology* **16**: 123–132.
- Voinnet O., Rivas S., Mestre P. and Baulcombe D.** (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *The Plant Journal* **33**: 949–956.
- Vranová E., Hirsch-Hoffmann M. and Gruissem W.** (2011). AtIPD: A Curated Database of Arabidopsis Isoprenoid Pathway Models and Genes for Isoprenoid Network Analysis. *Plant Physiology* **156**: 1655–1660.
- Wahid A., Gelani S., Ashraf M. and Foolad M.R.** (2007). Heat tolerance in plants: An overview. *Environmental and Experimental Botany* **61**: 199–223.
- Wan J., Bringloe D. and Lamppa G.K.** (1998). Disruption of chloroplast biogenesis and plant development upon down-regulation of a chloroplast processing enzyme involved in the import pathway. *The Plant Journal* **15**: 459-468.
- Weigel D., Ahn J.H., Blázquez M.A., Borevitz J.O., Christensen S.K., Fankhauser C., Ferrándiz C., Kardailsky I., Malancharuvil E.J., Neff M.M., Nguyen J.T., Sato S., Wang Z.-Y., Xia Y., Dixon R.A., Harrison M.J., Lamb C.J., Yanofsky M.F. and Chory J.** (2000). Activation Tagging in Arabidopsis. *Plant Physiology* **122**: 1003–1013.
- Wentzinger L.F., Bach T.J. and Hartmann M.-A.** (2002). Inhibition of Squalene Synthase and Squalene Epoxidase in Tobacco Cells Triggers an Up-Regulation of 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase. *Plant Physiology* **130**: 334–346.
- Wienkoop S., Baginsky S. and Weckwerth W.** (2010). *Arabidopsis thaliana* as a model organism for plant proteome research. *Journal of Proteomics* **73**: 2 2 3 9 – 2 2 4 8.
- Winter D., Vinegar B., Nahal H., Ammar R., Wilson G.V. and Provart N.J.** (2007). An “Electronic Fluorescent Pictograph” Browser for Exploring and Analyzing Large-Scale Biological Data Sets. *PLoS ONE* **2**: e718.
- Woo H.R., Kim J.H., Kim J., Kim J., Lee U., Song I.-J., Kim J.-H., Lee H.-Y., Nam H.G. and Lim P.O.** (2010). The RAV1 transcription factor positively regulates leaf senescence in Arabidopsis. *Journal of Experimental Botany* **61**: 3947–3957.
- Wullschlegel S.D. and Difazio S.P.** (2003). Emerging use of gene expression microarrays in plant physiology. *Comparative and Functional Genomics* **4**: 216-224.

Xiong Q., Wilson W.K. and Matsuda S.P.T. (2006). An *Arabidopsis* Oxidosqualene Cyclase Catalyzes Iridal Skeleton Formation by Grob Fragmentation. *Angewandte Chemie International Edition* **45**: 1285 – 1288.

Yin Y., Vafeados D., Tao Y., Yoshida S., Asami T. and Chory J. (2005). A New Class of Transcription Factors Mediates Brassinosteroid-Regulated Gene Expression in *Arabidopsis*. *Cell* **120**: 249–259.

Zelenski N.G., Rawson R.B., Brown M.S. and Goldstein J.L. (1999). Membrane Topology of S2P, a Protein Required for Intramembranous Cleavage of Sterol Regulatory Element-binding Proteins. *The Journal of Biological Chemistry* **274**: 21973–21980.

Zhang H., He H., Wang X., Wang X., Yang X., Li L. and Deng X.W. (2011). Genome-wide mapping of the *HY5*-mediated gene networks in *Arabidopsis* that involve both transcriptional and post-transcriptional regulation. *The Plant Journal* **65**: 346–358.

Zhang J.Z., Creelman R.A. and Zhu J.-K. (2004). From Laboratory to Field. Using Information from *Arabidopsis* to Engineer Salt, Cold, and Drought Tolerance in Crops. *Plant Physiology* **135**: 615–621.

Zhang X., Wu Z. and Huang C. (2008). Effects of gibberellin mutations on *in vitro* shoot bud regeneration of *Arabidopsis*. *African Journal of Biotechnology* **7**: 4159-4163.

Zhong R., Wan J., Jin R. and Lamppa G. (2003). A pea antisense gene for the chloroplast stromal processing peptidase yields seedling lethals in *Arabidopsis*: survivors show defective GFP import *in vivo*. *The Plant Journal* **34**: 802-812.

Zhong S., Zhao M., Shi T., Shi H., An F., Zhao Q. and Guo H. (2009). EIN3/EIL1 cooperate with PIF1 to prevent photo-oxidation and to promote greening of *Arabidopsis* seedlings. *Proceedings of the National Academy of Sciences* **106**: 21431–21436.

Zhu J.-K. (2002). Salt and Drought Stress Signal Transduction in Plants. *Annual Review Plant Biology* **53**: 247–273.

- URL no.1 – http://scienceblogs.com/pharyngula/2006/11/mads_boxes_flower_development.php
- URL no.2 – <http://www.arabidopsis.org/portals/education/aboutarabidopsis.jsp>
- URL no.3 – <http://arabidopsis.info/>
- URL no.4 – <http://www.cambia.org/daisy/cambia/585.html>
- URL no.5 – <http://www.arabidopsis.org/>
- URL no.6 – <http://oligo.net/>
- URL no.7 – <http://frodo.wi.mit.edu/primer3/>
- URL no.8 – <http://affy.arabidopsis.info/narrays/experimentbrowse.pl>
- URL no.9 – <http://www.plastid.msu.edu/>
- URL no.10 – <http://bar.utoronto.ca/welcome.htm>
- URL no.11 – <https://www.genevestigator.com/gv/>
- URL no.12 – <http://atted.jp/>
- URL no.13 – <http://genemania.org/>
- URL no.14 – <http://arabidopsis.med.ohio-state.edu/>
- URL no.15 – <http://www.bioinformatics2.wsu.edu/cgi-bin/Athena/cgi/home.pl>
- URL no.16 – <http://signal.salk.edu/cgi-bin/tdnaexpress>
- URL no.17 – <http://signal.salk.edu/tdnaprimers.2.html>
- URL no.18 – <http://suba.plantenergy.uwa.edu.au/>
- URL no.19 – <http://plants.ensembl.org/index.html>
- URL no.20 – <http://www.cbs.dtu.dk/services/TMHMM/>
- URL no.21 – <http://www.cbs.dtu.dk/services/SignalP/>
- URL no.22 – <http://www.cbs.dtu.dk/services/ChloroP/>
- URL no.23 – <http://www.uniprot.org/>
- URL no.24 – <http://ihg.gsf.de/ihg/mitoprot.html>
- URL no.25 – <http://www.ebi.ac.uk/Tools/pfa/iprscan/>
- URL no.26 – <http://www.ncbi.nlm.nih.gov/>
- URL no.27 – <http://www.ncbi.nlm.nih.gov/pubmed>
- URL no.28 – <http://blast.ncbi.nlm.nih.gov/Blast.cgi>
- URL no.29 – <http://www.phylogeny.fr/>
- URL no.30 – http://plants.ensembl.org/Arabidopsis_thaliana/Info/Index
- URL no.31 – <http://pfam.sanger.ac.uk/>
- URL no.32 – <http://www.chem.qmul.ac.uk/iubmb/enzyme/>
- URL no.33 – <http://www.seedgenes.org/>
- URL no.34 – <http://www.arabidopsis.org/tools/bulk/go/>

Appendix I – Expression raw data

Table A.1. – Expression values of the Arabidopsis eFP Browser from BAR database relative to *SQE1* (Figure 3.10, 3.11 and 3.31 in Chapter 3.2. is based on these values).

Tissue	Expression Level	Standard Deviation
Dry seed	51.18	4.7
Imbibed seed, 24 h	101.65	4.37
1st Node	173.55	14.0
Flower Stage 12, Stamens	137.53	9.21
Cauline Leaf	93.13	5.63
Cotyledon	148.85	4.18
Root	142.26	10.03
Entire Rosette After Transition to Flowering	126.6	12.43
Flower Stage 9	160.81	9.7
Flower Stage 10/11	146.98	3.82
Flower Stage 12	133.45	1.1
Flower Stage 15	70.78	4.18
Flower Stage 12, Carpels	75.61	4.49
Flower Stage 12, Petals	247.26	11.67
Flower Stage 12, Sepals	74.45	3.44
Flower Stage 15, Carpels	73.51	6.12
Flower Stage 15, Petals	113.55	8.51
Flower Stage 15, Sepals	48.41	3.73
Flower Stage 15, Stamen	111.21	8.82
Flowers Stage 15, Pedicels	140.61	5.25
Leaf 1 + 2	177.6	14.39
Leaf 7, Petiole	163.66	16.67
Leaf 7, Distal Half	170.8	8.22
Leaf 7, Proximal Half	176.5	13.24
Hypocotyl	118.53	6.65
Root	159.28	3.11
Rosette Leaf 2	93.7	6.53
Rosette Leaf 4	131.45	10.62
Rosette Leaf 6	157.35	7.83
Rosette Leaf 8	166.96	1.54
Rosette Leaf 10	143.96	17.48
Rosette Leaf 12	128.58	5.95
Senescing Leaf	52.91	3.55
Shoot Apex, Inflorescence	155.54	12.15
Shoot Apex, Transition	139.03	3.2
Shoot Apex, Vegetative	103.78	4.63
Stem, 2nd Internode	131.53	7.14
Mature Pollen	111.93	12.9

Seeds Stage 3 w/ Siliques	83.08	1.42
Seeds Stage 4 w/ Siliques	172.81	15.7
Seeds Stage 5 w/ Siliques	225.69	8.5
Seeds Stage 6 w/o Siliques	277.83	11.67
Seeds Stage 7 w/o Siliques	253.96	17.52
Seeds Stage 8 w/o Siliques	95.05	9.13
Seeds Stage 9 w/o Siliques	85.21	9.99
Seeds Stage 10 w/o Siliques	80.38	10.37
Vegetative Rosette	115.76	4.97
Root Stage III Stele	158.59	0.0
Root Stage III Endodermis	141.47	0.0
Root Stage III Cortex + Endodermis	221.1	0.0
Root Stage III Epidermal Artrichoblasts	138.37	0.0
Root Stage III Lateral Root Cap	194.87	0.0
Root Stage II Stele	163.25	0.0
Root Stage II Endodermis	145.63	0.0
Root Stage II Cortex + Endodermis	227.6	0.0
Root Stage II Epidermal Artrichoblasts	142.44	0.0
Root Stage II Lateral Root Cap	200.6	0.0
Root Stage I Stele	79.98	0.0
Root Stage I Endodermis	71.35	0.0
Root Stage I Cortex + Endodermis	111.52	0.0
Root Stage I Epidermal Artrichoblasts	69.79	0.0
Root Stage I Lateral Root Cap	98.28	0.0
Root Quiescent Center	287.82	19.47
Uninucleate Microphore	93.15	4.55
Bicellular Pollen	98.2	24.79
Tricellular Pollen	147.94	9.75
Mature Pollen Grain	134.5	0.0
Globular - Apical	206.93	108.35
Globular - Basal	105.2	11.39
Heart - Cotyledon	94.53	37.2
Heart - Root	118.36	81.62
Torpedo - Cotyledon	73.73	36.58
Torpedo - Root	91.5	76.8
Torpedo - Meristem	178.75	67.59
Torpedo - Apical	171.67	52.92
Torpedo - Basal	213.41	22.63
Xylem Col-0	75.13	5.05
Cork Col-0	59.23	5.27
Xylem MYB61 knockout	82.06	5.2
Cork MYB61 knockout	62.36	15.24
Xylem MYB50 knockout	72.46	6.28
Cork MYB50 knockout	65.1	4.82
Hypocotyl Col-0	57.2	4.81

Hypocotyl Ler	65.36	7.49
Hypocotyl abi1	65.53	4.55
Hypocotyl aba1	56.96	1.02
Hypocotyl max4	63.26	7.8
Hypocotyl axr1	57.63	1.71
Mesophyll cells, no ABA	159.32	73.41
Guard cells, no ABA	189.81	82.19
Mesophyll cells, with 100 uM ABA	168.03	69.55
Guard cells, with 100 uM ABA	192.42	61.4
Guard cells, no ABA, no cordycepin nor actinomycin	272.0	0.0
Guard cells, no ABA, cordycepin and actinomycin added during protoplasting	107.62	0.0
Guard cells, with 100uM ABA, no cordycepin nor actinomycin	253.83	0.0
Guard cells, with 100uM ABA, cordycepin and actinomycin added during protoplasting	131.01	0.0
Mesophyll cells, no ABA, no cordycepin nor actinomycin	232.73	0.0
Mesophyll cells, no ABA, cordycepin and actinomycin added during protoplasting	85.9	0.0
Mesophyll cells, with 100uM ABA, no cordycepin nor actinomycin	237.59	0.0
Mesophyll cells, with 100uM ABA, cordycepin and actinomycin added during protoplasting	98.48	0.0
Stem epidermis, top of stem	184.85	20.47
Stem epidermis, bottom of stem	107.05	6.26
Whole stem, top of stem	222.61	30.09
Whole stem, bottom of stem	174.88	8.78
Stigma tissue	88.55	3.18
Ovary tissue	166.21	13.61
Dry pollen	23.37	3.15
Pollen, germinated in vitro for 30 minutes	22.1	2.78
Pollen, germinated in vitro for 4 hours	43.15	9.53
Pollen tubes harvested after growth through pistil explants	681.89	40.57
Rib Meristem	473.3	46.45
Peripheral Zone	288.84	37.6
Central Zone	346.46	13.47
embryo pre-globular stage	69.82	5.5
micropylar endosperm pre-globular stage	27.75	1.84
peripheral endosperm pre-globular stage	37.03	6.69
chalazal endosperm pre-globular stage	47.7	4.52
chalazal seed coat pre-globular stage	25.46	24.67
general seed coat pre-globular stage	66.25	20.78
embryo globular stage	33.92	4.79
suspensor globular stage	326.68	84.26
micropylar endosperm globular stage	31.86	11.3
peripheral endosperm globular stage	24.11	0.3
chalazal endosperm globular stage	95.34	37.64
chalazal seed coat globular stage	10.03	0.92
general seed coat globular stage	123.5	17.93

embryo proper heart stage	47.22	1.7
micropylar endosperm heart stage	95.49	28.98
peripheral endosperm heart stage	60.96	6.45
chalazal endosperm heart stage	20.72	6.12
chalazal seed coat heart stage	21.66	7.87
seed coat heart stage	81.65	14.61
embryo proper linear-cotyledon stage	96.95	23.29
cellularized endosperm linear-cotyledon stage	196.28	0.71
chalazal endosperm linear-cotyledon stage	51.49	6.02
chalazal seed coat linear-cotyledon stage	34.84	29.22
general seed coat linear-cotyledon stage	142.51	1.0
embryo proper mature green stage	216.51	69.84
micropylar endosperm mature green stage	143.85	20.59
peripheral endosperm mature green stage	143.98	16.83
chalazal endosperm mature green stage	132.12	34.27
chalazal seed coat mature green stage	42.84	12.46

Table A.2. – Expression values of the Arabidopsis eFP Browser from BAR database relative to SQE2 (Figure 3.10, 3.11, and 3.34 in Chapter 3.2. is based on these values).

Tissue	Expression Level	Standard Deviation
Dry seed	5.01	0.23
Imbibed seed, 24 h	12.7	0.55
1st Node	33.4	2.26
Flower Stage 12, Stamens	49.71	7.96
Cauline Leaf	38.55	3.4
Cotyledon	50.66	9.49
Root	28.68	4.32
Entire Rosette After Transition to Flowering	36.86	4.98
Flower Stage 9	40.71	1.2
Flower Stage 10/11	65.21	4.79
Flower Stage 12	55.9	3.16
Flower Stage 15	41.65	7.6
Flower Stage 12, Carpels	47.31	1.63
Flower Stage 12, Petals	36.35	8.34
Flower Stage 12, Sepals	29.41	5.61
Flower Stage 15, Carpels	48.58	2.12
Flower Stage 15, Petals	57.01	8.13
Flower Stage 15, Sepals	30.85	2.57
Flower Stage 15, Stamen	38.35	3.36
Flowers Stage 15, Pedicels	35.83	4.56
Leaf 1 + 2	36.08	5.58
Leaf 7, Petiole	34.05	2.57
Leaf 7, Distal Half	42.73	3.26

Leaf 7, Proximal Half	34.93	2.04
Hypocotyl	33.19	3.98
Root	32.76	3.15
Rosette Leaf 2	37.1	8.57
Rosette Leaf 4	35.45	2.48
Rosette Leaf 6	39.63	2.5
Rosette Leaf 8	34.28	4.72
Rosette Leaf 10	37.93	1.31
Rosette Leaf 12	27.33	2.63
Senescing Leaf	32.91	3.33
Shoot Apex, Inflorescence	45.6	1.71
Shoot Apex, Transition	49.03	3.6
Shoot Apex, Vegetative	35.86	6.38
Stem, 2nd Internode	28.3	3.86
Mature Pollen	37.46	9.79
Seeds Stage 3 w/ Siliques	42.16	4.35
Seeds Stage 4 w/ Siliques	41.66	6.23
Seeds Stage 5 w/ Siliques	37.63	3.49
Seeds Stage 6 w/o Siliques	39.61	2.86
Seeds Stage 7 w/o Siliques	35.13	0.87
Seeds Stage 8 w/o Siliques	19.06	11.65
Seeds Stage 9 w/o Siliques	32.56	3.37
Seeds Stage 10 w/o Siliques	22.4	5.46
Vegetative Rosette	48.93	0.96
Vegetative Rosette	345.31	34.56
Root Stage III Stele	35.21	0.0
Root Stage III Endodermis	27.71	0.0
Root Stage III Cortex + Endodermis	29.57	0.0
Root Stage III Epidermal Artrichoblasts	31.42	0.0
Root Stage III Lateral Root Cap	38.44	0.0
Root Stage II Stele	43.83	0.0
Root Stage II Endodermis	34.5	0.0
Root Stage II Cortex + Endodermis	36.82	0.0
Root Stage II Epidermal Artrichoblasts	39.12	0.0
Root Stage II Lateral Root Cap	47.86	0.0
Root Stage I Stele	35.04	0.0
Root Stage I Endodermis	27.58	0.0
Root Stage I Cortex + Endodermis	29.43	0.0
Root Stage I Epidermal Artrichoblasts	31.28	0.0
Root Stage I Lateral Root Cap	38.26	0.0
Root Quiescent Center	103.06	33.88
Uninucleate Microphore	54.75	2.85
Bicellular Pollen	55.55	1.34
Tricellular Pollen	40.5	4.79
Mature Pollen Grain	22.2	0.0

Globular - Apical	56.5	41.3
Globular - Basal	39.63	30.99
Heart - Cotyledon	21.46	8.94
Heart - Root	34.69	33.41
Torpedo - Cotyledon	31.7	14.67
Torpedo - Root	127.19	130.9
Torpedo - Meristem	36.34	24.91
Torpedo - Apical	27.44	7.79
Torpedo - Basal	20.87	7.59
Xylem Col-0	13.56	3.08
Cork Col-0	28.93	5.72
Xylem MYB61 knockout	11.4	2.38
Cork MYB61 knockout	24.33	4.99
Xylem MYB50 knockout	20.6	3.25
Cork MYB50 knockout	41.5	2.2
Hypocotyl Col-0	18.36	2.87
Hypocotyl Ler	13.26	3.8
Hypocotyl abi1	20.6	2.3
Hypocotyl aba1	22.56	5.17
Hypocotyl max4	20.93	5.28
Hypocotyl axr1	24.66	2.78
Mesophyll cells, no ABA	19.41	12.36
Guard cells, no ABA	95.91	9.02
Mesophyll cells, with 100 uM ABA	39.57	29.27
Guard cells, with 100 uM ABA	76.35	0.37
Guard cells, no ABA, no cordycepin nor actinomycin	104.94	0.0
Guard cells, no ABA, cordycepin and actinomycin added during protoplasting	86.89	0.0
Guard cells, with 100uM ABA, no cordycepin nor actinomycin	76.73	0.0
Guard cells, with 100uM ABA, cordycepin and actinomycin added during protoplasting	75.98	0.0
Mesophyll cells, no ABA, no cordycepin nor actinomycin	31.78	0.0
Mesophyll cells, no ABA, cordycepin and actinomycin added during protoplasting	7.05	0.0
Mesophyll cells, with 100uM ABA, no cordycepin nor actinomycin	68.84	0.0
Mesophyll cells, with 100uM ABA, cordycepin and actinomycin added during protoplasting	10.3	0.0
Stem epidermis, top of stem	132.53	1.33
Stem epidermis, bottom of stem	70.93	5.34
Whole stem, top of stem	40.63	2.19
Whole stem, bottom of stem	36.7	0.24
Stigma tissue	97.77	9.59
Ovary tissue	131.55	13.64
Dry pollen	3.59	2.56
Pollen, germinated in vitro for 30 minutes	7.59	1.2
Pollen, germinated in vitro for 4 hours	4.04	1.69

Pollen tubes harvested after growth through pistil explants	7.6	3.17
Rib Meristem	138.53	2.84
Peripheral Zone	111.05	14.94
Central Zone	63.11	10.4
seed coat mature green stage	21.85	1.49
Seedling control at 30 Minutes	38.89	6.27
Seedling ABA Treated at 30 Minutes	27.51	3.77
Seedling Control at 1 Hour	18.53	5.44
Seedling ABA Treated at 1 Hour	14.5	0.93
Seedling control at 3 Hours	28.34	6.49
Seedling ABA Treated at 3 Hours	5.01	0.23
Seed no Treatment	12.7	0.55
Seed treated with Water	11.61	4.25
Seed treated with 3 μ M ABA	10.85	0.82
Seed treated with 30 μ M ABA	21.85	1.49

Table A.3. – Expression values of the Arabidopsis eFP Browser from BAR database relative to SQE3 (Figure 3.10, 3.11, 3.31 and 3.34 in Chapter 3.2. is based on these values).

Tissue	Expression Level	Standard Deviation
Dry seed	408.65	25.49
Imbibed seed, 24 h	119.92	11.1
1st Node	194.41	2.9
Flower Stage 12, Stamens	487.11	14.37
Cauline Leaf	305.31	10.56
Cotyledon	407.96	22.37
Root	12.03	1.29
Entire Rosette After Transition to Flowering	264.71	22.09
Flower Stage 9	169.8	12.17
Flower Stage 10/11	290.16	27.51
Flower Stage 12	306.64	20.15
Flower Stage 15	641.68	29.69
Flower Stage 12, Carpels	184.38	2.75
Flower Stage 12, Petals	344.51	6.89
Flower Stage 12, Sepals	569.98	30.1
Flower Stage 15, Carpels	226.31	9.42
Flower Stage 15, Petals	1184.11	24.84
Flower Stage 15, Sepals	692.88	32.07
Flower Stage 15, Stamen	1289.63	21.15
Flowers Stage 15, Pedicels	362.78	4.61
Leaf 1 + 2	376.76	35.39
Leaf 7, Petiole	209.01	23.68
Leaf 7, Distal Half	220.23	3.93
Leaf 7, Proximal Half	232.25	19.91

Hypocotyl	198.43	17.68
Root	14.35	1.59
Rosette Leaf 2	292.36	14.16
Rosette Leaf 4	258.26	17.72
Rosette Leaf 6	252.03	18.32
Rosette Leaf 8	201.41	13.84
Rosette Leaf 10	180.66	6.77
Rosette Leaf 12	191.75	4.18
Senescing Leaf	171.13	12.27
Shoot Apex, Inflorescence	104.98	6.09
Shoot Apex, Transition	86.26	1.01
Shoot Apex, Vegetative	165.48	13.94
Stem, 2nd Internode	233.4	15.42
Mature Pollen	2965.5	275.58
Seeds Stage 3 w/ Siliques	253.63	7.25
Seeds Stage 4 w/ Siliques	466.2	6.49
Seeds Stage 5 w/ Siliques	449.26	14.39
Seeds Stage 6 w/o Siliques	252.16	10.54
Seeds Stage 7 w/o Siliques	271.11	21.76
Seeds Stage 8 w/o Siliques	381.81	14.84
Seeds Stage 9 w/o Siliques	476.46	47.81
Seeds Stage 10 w/o Siliques	428.91	23.72
Vegetative Rosette	345.31	34.56
Root Stage III Stele	38.49	0.0
Root Stage III Endodermis	46.29	0.0
Root Stage III Cortex + Endodermis	28.74	0.0
Root Stage III Epidermal Artrichoblasts	27.62	0.0
Root Stage III Lateral Root Cap	75.22	0.0
Root Stage II Stele	13.75	0.0
Root Stage II Endodermis	16.53	0.0
Root Stage II Cortex + Endodermis	10.26	0.0
Root Stage II Epidermal Artrichoblasts	9.86	0.0
Root Stage II Lateral Root Cap	26.87	0.0
Root Stage I Stele	10.91	0.0
Root Stage I Endodermis	13.13	0.0
Root Stage I Cortex + Endodermis	8.15	0.0
Root Stage I Epidermal Artrichoblasts	7.83	0.0
Root Stage I Lateral Root Cap	21.33	0.0
Root Quiescent Center	28.35	3.62
Uninucleate Microphore	141.25	10.34
Bicellular Pollen	151.65	17.15
Tricellular Pollen	706.8	46.9
Mature Pollen Grain	1010.6	0.0
Globular - Apical	133.06	35.36
Globular - Basal	101.63	30.22

Heart - Cotyledon	59.66	30.26
Heart - Root	93.33	36.09
Torpedo - Cotyledon	135.43	20.23
Torpedo - Root	160.36	51.74
Torpedo - Meristem	102.48	62.79
Torpedo - Apical	44.38	13.9
Torpedo - Basal	89.6	57.02
Xylem Col-0	15.76	3.05
Cork Col-0	19.53	1.4
Xylem MYB61 knockout	10.4	4.0
Cork MYB61 knockout	16.26	1.04
Xylem MYB50 knockout	16.59	0.85
Cork MYB50 knockout	36.73	2.76
Hypocotyl Col-0	25.03	4.11
Hypocotyl Ler	35.76	5.04
Hypocotyl abi1	24.36	3.7
Hypocotyl aba1	24.4	1.8
Hypocotyl max4	20.2	2.24
Hypocotyl axr1	19.26	3.77
Mesophyll cells, no ABA	256.78	133.67
Guard cells, no ABA	2092.67	400.48
Mesophyll cells, with 100 uM ABA	250.28	137.38
Guard cells, with 100 uM ABA	2380.66	715.39
Guard cells, no ABA, no cordycepin nor actinomycin	1692.19	0.0
Guard cells, no ABA, cordycepin and actinomycin added during protoplasting	2493.16	0.0
Guard cells, with 100uM ABA, no cordycepin nor actinomycin	1665.27	0.0
Guard cells, with 100uM ABA, cordycepin and actinomycin added during protoplasting	3096.06	0.0
Mesophyll cells, no ABA, no cordycepin nor actinomycin	390.46	0.0
Mesophyll cells, no ABA, cordycepin and actinomycin added during protoplasting	123.11	0.0
Mesophyll cells, with 100uM ABA, no cordycepin nor actinomycin	387.66	0.0
Mesophyll cells, with 100uM ABA, cordycepin and actinomycin added during protoplasting	112.9	0.0
Stem epidermis, top of stem	313.68	29.88
Stem epidermis, bottom of stem	327.01	26.84
Whole stem, top of stem	145.64	17.65
Whole stem, bottom of stem	791.42	37.79
Stigma tissue	1141.08	278.34
Ovary tissue	768.14	167.37
Dry pollen	1004.79	74.94
Pollen, germinated in vitro for 30 minutes	1005.16	77.63
Pollen, germinated in vitro for 4 hours	1007.53	60.83
Pollen tubes harvested after growth through pistil explants	678.26	46.98
Rib Meristem	592.57	18.61
Peripheral Zone	307.06	24.07

Central Zone	415.81	32.63
embryo pre-globular stage	15.02	4.81
micropylar endosperm pre-globular stage	59.4	4.73
peripheral endosperm pre-globular stage	71.75	1.72
chalazal endosperm pre-globular stage	61.83	27.9
chalazal seed coat pre-globular stage	3.67	2.47
general seed coat pre-globular stage	6.23	5.67
embryo globular stage	7.93	6.95
suspensor globular stage	91.74	3.42
micropylar endosperm globular stage	53.79	9.09
peripheral endosperm globular stage	66.71	32.88
chalazal endosperm globular stage	213.38	58.13
chalazal seed coat globular stage	9.92	2.76
general seed coat globular stage	11.12	3.02
embryo proper heart stage	16.93	6.62
micropylar endosperm heart stage	25.27	12.79
peripheral endosperm heart stage	22.77	5.6
chalazal endosperm heart stage	261.71	97.48
chalazal seed coat heart stage	18.74	9.16
seed coat heart stage	11.6	2.3
embryo proper linear-cotyledon stage	5.03	3.81
cellularized endosperm linear-cotyledon stage	38.48	3.27
chalazal endosperm linear-cotyledon stage	142.5	4.01
chalazal seed coat linear-cotyledon stage	37.78	7.43
general seed coat linear-cotyledon stage	429.64	62.6
embryo proper mature green stage	51.7	7.63
micropylar endosperm mature green stage	94.99	12.27
peripheral endosperm mature green stage	125.04	20.33
chalazal endosperm mature green stage	105.28	33.82
chalazal seed coat mature green stage	104.63	9.67
seed coat mature green stage	355.49	66.56
Seedling control at 30 Minutes	102.82	23.11
Seedling ABA Treated at 30 Minutes	104.18	3.77
Seedling Control at 1 Hour	108.54	6.65
Seedling ABA Treated at 1 Hour	129.83	15.44
Seedling control at 3 Hours	134.45	0.17
Seedling ABA Treated at 3 Hours	361.48	34.03
Seed no Treatment	408.65	25.49
Seed treated with Water	119.92	11.1
Seed treated with 3 μ M ABA	108.75	17.72
Seed treated with 30 μ M ABA	109.99	3.6

Table A.4. – Expression values of the Arabidopsis eFP Browser from BAR database relative to *EGY1* (Figure 4.13 in Chapter 4.2. is based on these values).

Tissue	Expression Level	Standard Deviation
Dry seed	236.03	22.02
Imbibed seed, 24 h	304.19	12.98
1st Node	215.75	24.56
Flower Stage 12, Stamens	158.9	20.21
Cauline Leaf	353.23	32.0
Cotyledon	307.06	5.3
Root	142.73	4.75
Entire Rosette After Transition to Flowering	359.88	36.0
Flower Stage 9	286.78	6.86
Flower Stage 10/11	307.05	22.78
Flower Stage 12	275.83	6.97
Flower Stage 15	239.75	7.86
Flower Stage 12, Carpels	300.2	4.03
Flower Stage 12, Petals	259.21	17.26
Flower Stage 12, Sepals	322.21	7.33
Flower Stage 15, Carpels	304.26	7.98
Flower Stage 15, Petals	189.1	17.27
Flower Stage 15, Sepals	187.1	6.44
Flower Stage 15, Stamen	155.73	11.54
Flowers Stage 15, Pedicels	463.23	22.66
Leaf 1 + 2	503.05	22.53
Leaf 7, Petiole	303.64	49.47
Leaf 7, Distal Half	279.84	22.18
Leaf 7, Proximal Half	344.43	17.05
Hypocotyl	187.13	3.8
Root	123.38	5.59
Rosette Leaf 2	270.73	9.33
Rosette Leaf 4	280.26	12.87
Rosette Leaf 6	312.98	21.51
Rosette Leaf 8	350.15	4.55
Rosette Leaf 10	384.06	10.63
Rosette Leaf 12	454.3	13.38
Senescing Leaf	269.51	10.9
Shoot Apex, Inflorescence	269.56	3.65
Shoot Apex, Transition	258.46	11.44
Shoot Apex, Vegetative	283.48	5.83
Stem, 2nd Internode	262.53	1.91
Mature Pollen	95.33	5.34
Seeds Stage 3 w/ Siliques	259.78	6.39

Seeds Stage 4 w/ Siliques	197.43	5.99
Seeds Stage 5 w/ Siliques	148.28	13.8
Seeds Stage 6 w/o Siliques	90.3	4.92
Seeds Stage 7 w/o Siliques	82.31	6.26
Seeds Stage 8 w/o Siliques	66.5	3.95
Seeds Stage 9 w/o Siliques	50.41	11.03
Seeds Stage 10 w/o Siliques	62.11	8.36
Vegetative Rosette	508.16	42.76

Table A.5. – Expression values of the Arabidopsis eFP Browser from BAR database relative to *EGY2* (Figure 4.13 in Chapter 4.2. is based on these values).

Tissue	Expression Level	Standard Deviation
Dry seed	15.72	0.79
Imbibed seed, 24 h	170.49	18.36
1st Node	94.96	8.67
Flower Stage 12, Stamens	44.2	2.76
Cauline Leaf	126.23	5.18
Cotyledon	241.8	8.86
Root	13.86	0.54
Entire Rosette After Transition to Flowering	253.63	18.97
Flower Stage 9	167.5	6.22
Flower Stage 10/11	198.4	16.4
Flower Stage 12	166.76	8.5
Flower Stage 15	111.58	3.23
Flower Stage 12, Carpels	121.6	3.13
Flower Stage 12, Petals	142.73	6.01
Flower Stage 12, Sepals	272.71	11.16
Flower Stage 15, Carpels	174.71	5.75
Flower Stage 15, Petals	64.43	2.28
Flower Stage 15, Sepals	91.06	4.5
Flower Stage 15, Stamen	42.51	0.15
Flowers Stage 15, Pedicels	350.95	13.83
Leaf 1 + 2	400.86	12.3
Leaf 7, Petiole	207.88	17.61
Leaf 7, Distal Half	241.33	4.94
Leaf 7, Proximal Half	266.7	4.91
Hypocotyl	83.01	6.14
Root	13.6	2.04
Rosette Leaf 2	157.56	13.3
Rosette Leaf 4	201.1	14.36
Rosette Leaf 6	246.08	5.49
Rosette Leaf 8	287.8	2.71

Rosette Leaf 10	318.75	13.85
Rosette Leaf 12	365.05	24.22
Senescing Leaf	38.38	1.81
Shoot Apex, Inflorescence	121.76	4.69
Shoot Apex, Transition	132.16	6.47
Shoot Apex, Vegetative	199.28	13.24
Stem, 2nd Internode	75.18	3.94
Mature Pollen	30.86	8.27
Seeds Stage 3 w/ Siliques	220.18	8.36
Seeds Stage 4 w/ Siliques	128.75	9.28
Seeds Stage 5 w/ Siliques	101.43	6.46
Seeds Stage 6 w/o Siliques	70.78	6.54
Seeds Stage 7 w/o Siliques	59.16	5.04
Seeds Stage 8 w/o Siliques	27.16	3.13
Seeds Stage 9 w/o Siliques	19.16	5.52
Seeds Stage 10 w/o Siliques	13.85	8.69
Vegetative Rosette	390.25	8.21

Table A.6. – Expression values of the Arabidopsis eFP Browser from BAR database relative to *EGY3* (Figure 4.13 in Chapter 4.2. is based on these values).

Tissue	Expression Level	Standard Deviation
Dry seed	180.54	2.22
Imbibed seed, 24 h	29.64	2.5
1st Node	21.81	1.42
Flower Stage 12, Stamens	24.36	5.93
Cauline Leaf	34.25	2.59
Cotyledon	27.13	1.75
Root	54.71	1.51
Entire Rosette After Transition to Flowering	33.36	3.81
Flower Stage 9	34.98	1.75
Flower Stage 10/11	30.46	0.85
Flower Stage 12	27.25	3.28
Flower Stage 15	23.5	2.91
Flower Stage 12, Carpels	83.38	1.72
Flower Stage 12, Petals	82.18	10.36
Flower Stage 12, Sepals	43.21	2.63
Flower Stage 15, Carpels	33.48	7.61
Flower Stage 15, Petals	23.36	4.49
Flower Stage 15, Sepals	24.88	3.76
Flower Stage 15, Stamen	32.56	2.49
Flowers Stage 15, Pedicels	35.49	3.68

Leaf 1 + 2	34.46	3.31
Leaf 7, Petiole	19.25	2.83
Leaf 7, Distal Half	26.26	1.2
Leaf 7, Proximal Half	16.38	3.03
Hypocotyl	26.63	1.79
Root	22.5	1.73
Rosette Leaf 2	39.81	5.93
Rosette Leaf 4	25.35	6.22
Rosette Leaf 6	28.8	2.43
Rosette Leaf 8	23.86	3.34
Rosette Leaf 10	26.68	1.54
Rosette Leaf 12	31.28	2.01
Senescing Leaf	25.45	1.74
Shoot Apex, Inflorescence	44.85	3.68
Shoot Apex, Transition	41.48	4.36
Shoot Apex, Vegetative	37.61	6.42
Stem, 2nd Internode	21.66	2.37
Mature Pollen	77.66	13.81
Seeds Stage 3 w/ Siliques	26.98	2.32
Seeds Stage 4 w/ Siliques	30.43	2.37
Seeds Stage 5 w/ Siliques	26.16	2.28
Seeds Stage 6 w/o Siliques	47.53	4.55
Seeds Stage 7 w/o Siliques	64.51	1.28
Seeds Stage 8 w/o Siliques	111.69	4.54
Seeds Stage 9 w/o Siliques	95.86	2.92
Seeds Stage 10 w/o Siliques	166.93	12.44
Vegetative Rosette	39.66	4.46

Table A.7. – Expression values of the Arabidopsis eFP Browser from BAR database relative to *EGY3* abiotic stress data (Figure 4.18 in Chapter 4.2. is based on these values).

Tissue	Expression Level	Standard Deviation
Control Shoot 0 Hour	44.53	4.6
Osmotic Shoot 0 Hour	44.53	4.6
Salt Shoot 0 Hour	44.53	4.6
Heat Shoot 0 Hour	44.53	4.6
Control Root 0 Hour	24.22	0.95
Osmotic Root 0 Hour	24.22	0.95
Salt Root 0 Hour	24.22	0.95
Heat Root 0 Hour	24.22	0.95
Control Shoot After 15 Minutes	51.68	5.81
Heat Shoot After 15 Minutes	51.91	0.9
Control Root After 15 Minutes	23.77	2.98

Heat Root After 15 Minutes	23.57	3.57
Control Shoot After 30 Minutes	45.02	2.97
Osmotic Shoot After 30 Minutes	41.18	1.49
Salt Shoot After 30 Minutes	49.72	8.77
Heat Shoot After 30 Minutes	116.23	43.73
Control Root After 30 Minutes	19.08	3.11
Osmotic Root After 30 Minutes	44.63	1.08
Salt Root After 30 Minutes	43.25	9.53
Heat Root After 30 Minutes	27.86	6.49
Control Shoot After 1 Hour	50.67	2.77
Osmotic Shoot After 1 Hour	41.47	0.57
Salt Shoot After 1 Hour	46.02	2.58
Heat Shoot After 1 Hour	2791.79	93.59
Control Root After 1 Hour	22.45	0.76
Osmotic Root After 1 Hour	37.79	8.89
Salt Root After 1 Hour	25.68	6.41
Heat Root After 1 Hour	683.05	112.33
Control Shoot After 3 Hours	30.47	0.79
Osmotic Shoot After 3 Hours	104.32	9.77
Salt Shoot After 3 Hours	93.5	20.38
Heat Shoot After 3 Hours	2234.49	90.47
Control Root After 3 Hours	25.36	4.01
Osmotic Root After 3 Hours	35.31	1.84
Salt Root After 3 Hours	35.9	0.1
Heat Root After 3 Hours	558.55	7.98
Control Shoot After 4 Hours	42.08	1.32
Heat Shoot After 4 Hours	560.01	29.04
Control Root After 4 Hours	26.86	0.02
Heat Root After 4 Hours	50.36	3.66
Control Shoot After 6 Hours	34.25	7.18
Osmotic Shoot After 6 Hours	179.21	43.28
Salt Shoot After 6 Hours	99.67	13.41
Heat Shoot After 6 Hours	69.44	5.2
Control Root After 6 Hours	26.73	1.64
Osmotic Root After 6 Hours	46.88	12.58
Salt Root After 6 Hours	50.91	2.84
Heat Root After 6 Hours	35.65	3.05
Control Shoot After 12 Hours	27.19	4.07
Osmotic Shoot After 12 Hours	1233.64	320.06
Salt Shoot After 12 Hours	185.37	34.87
Heat Shoot After 12 Hours	76.75	5.79
Control Root After 12 Hours	26.32	1.92
Osmotic Root After 12 Hours	57.2	3.41

Salt Root After 12 Hours	51.86	2.06
Heat Root After 12 Hours	69.07	5.7
Control Shoot After 24 Hours	86.31	5.03
Osmotic Shoot After 24 Hours	2125.79	329.19
Salt Shoot After 24 Hours	1394.06	117.85
Heat Shoot After 24 Hours	81.02	5.15
Control Root After 24 Hours	22.35	0.47
Osmotic Root After 24 Hours	63.8	5.54
Salt Root After 24 Hours	51.21	3.17
Heat Root After 24 Hours	28.71	0.59

Appendix II – Squalene and sterol analysis raw data

Table A.8. – Data from the quantification of sterols in 1-month-old leaves of *dry2/sqe1-5*, *sqe2-1*, *sqe3-1* and their ecotypes. Values are given in $\mu\text{g/g}$ dry weight \pm SE (N = 3).tr-traces. (Figure 3.18-A in chapter 3.2.).

Sterols	Ler	<i>dry2/sqe1-5</i>	Col-0	<i>sqe2-1</i>	<i>sqe3-1</i>
Sitosterol	1436 \pm 114	1447 \pm 181	1434 \pm 61	1522 \pm 164	1297 \pm 74
Stigmasterol	tr	tr	tr	tr	tr
Stigmastanol	285 \pm 62	201 \pm 17	181 \pm 21	256 \pm 27	177 \pm 18
Campesterol	183 \pm 18	125 \pm 8	158 \pm 13	173 \pm 16	152 \pm 23
Cholesterol	42 \pm 7	25 \pm 7	36 \pm 6	40 \pm 7	32 \pm 6

Table A.9. – Data from the quantification of squalene and other sterols in 14-day-old root seedlings of *sqe3-1* and Col-0. Values are given in $\mu\text{g/g}$ dry weight \pm SE (N = 3). (Figure 3.18-B in chapter 3.2.).

Root tissue compounds	Col-0	<i>sqe3-1</i>
Squalene	16 \pm 3	30 \pm 7
Cycloartenol	117 \pm 21	112 \pm 17
Isofucosterol	114 \pm 13	134 \pm 14
Sitosterol	1640 \pm 90	1683 \pm 74
Stigmasterol	1192 \pm 31	1149 \pm 28
Stigmastanol	284 \pm 48	309 \pm 31
Campesterol	431 \pm 15	441 \pm 16
Cholesterol	57 \pm 9	72 \pm 8

Table A.10. – Data from the quantification of squalene and other sterols in 14-day-old shoot seedlings of *sqe3-1* and Col-0. Values are given in $\mu\text{g/g}$ dry weight \pm SE (N = 3 for *sqe3-1* and N=2 for Col-0). tr-traces. (Figure 3.18-C in chapter 3.2.).

Aerial part compounds	Col-0	<i>sqe3-1</i>
Squalene	14 \pm 5	37 \pm 2
Cycloartenol	178 \pm 6	112 \pm 19
Isofucosterol	40 \pm 6	48 \pm 8

Sitosterol	2229 ± 59	2665 ± 385
Stigmasterol	tr	tr
Stigmastanol	446 ± 135	307 ± 32
Campesterol	590 ± 15	592 ± 18
Cholesterol	59 ± 14	56 ± 2

Appendix III – Vector maps

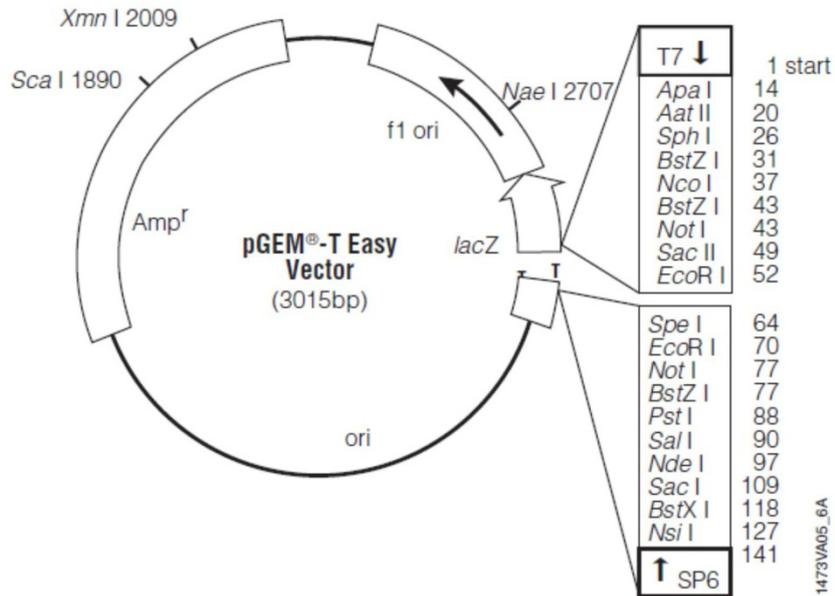


Figure A.1. – pGEM-T subcloning vector circle map.

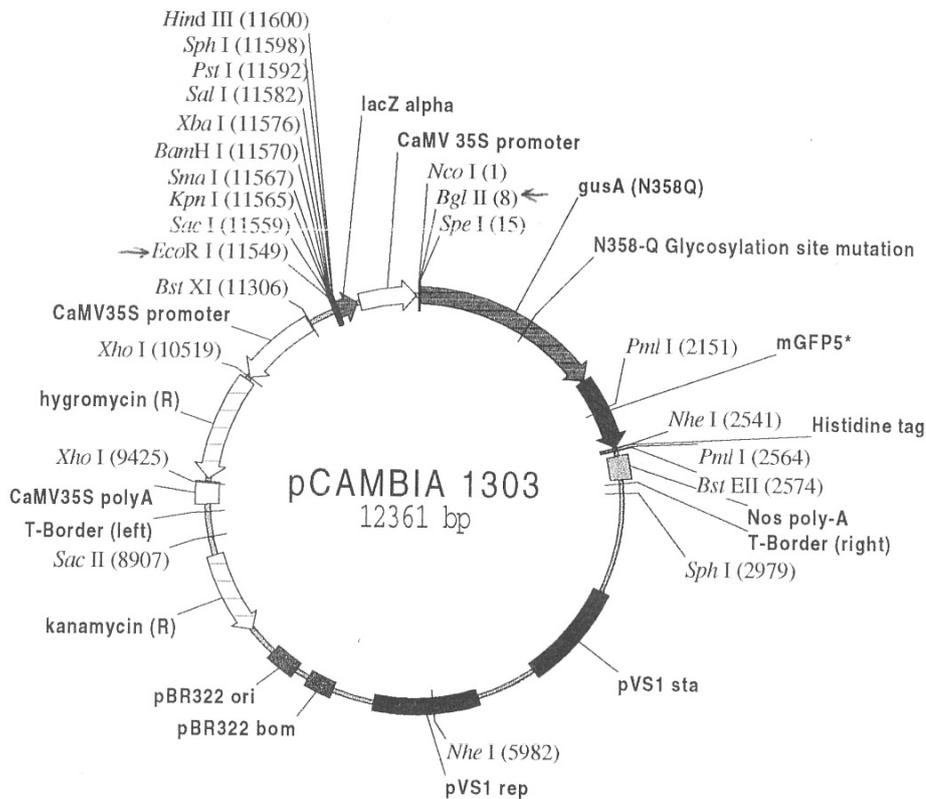
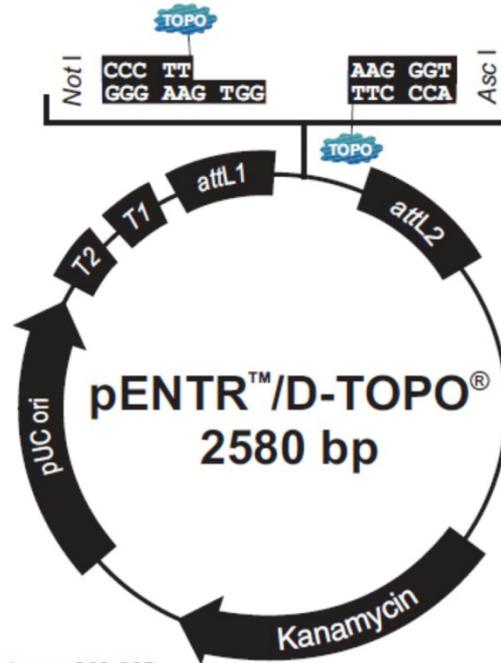


Figure A.2. – pCAMBIA1303 vector circle map (URL no.4). It is a 35S::GUS-GFP vector used for expression *in planta*.

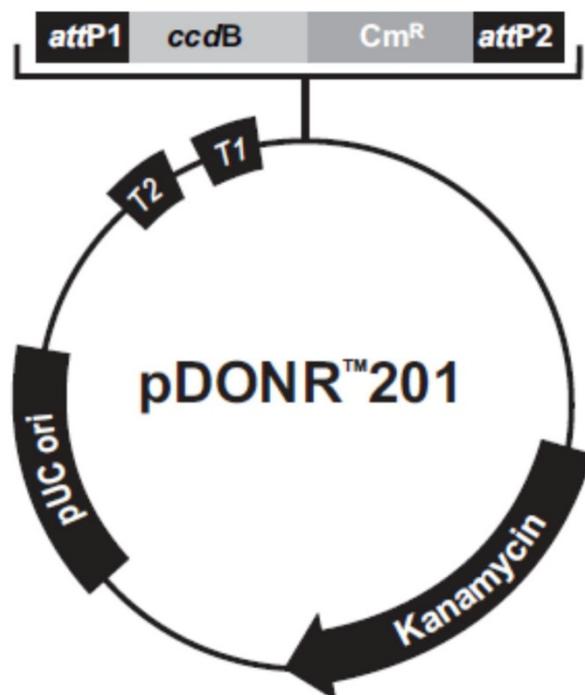


**Comments for pENTR™/D-TOPO®
2580 nucleotides**

rrmB T2 transcription termination sequence: bases 268-295
rrmB T1 transcription termination sequence: bases 427-470
M13 forward (-20) priming site: bases 537-552
attL1: bases 569-668 (c)
TOPO® recognition site 1: bases 680-684
Overhang: bases 685-688
TOPO® recognition site 2: bases 689-693
attL2: bases 705-804
T7 Promoter/priming site: bases 821-840 (c)
M13 reverse priming site: bases 845-861
Kanamycin resistance gene: bases 974-1783
pUC origin: bases 1904-2577

(c) = complementary sequence

Figure A.3. – pENTR entry Gateway vector circle map.



Comments for:

pDONRTM201
4470 nucleotides

<i>rmB</i> T2 transcription termination sequence (c):	73-100
<i>rrnB</i> T1 transcription termination sequence (c):	232-275
Recommended forward priming site:	300-324
<i>attP1</i> :	332-563
<i>ccdB</i> gene (c):	959-1264
Chloramphenicol resistance gene (c):	1606-2265
<i>attP2</i> (c):	2513-2744
Recommended reverse priming site:	2769-2792
Kanamycin resistance gene:	2868-3677
pUC origin:	3794-4467
(c) = complementary strand	

Figure A.4. – pDONR201 entry Gateway vector circle map.

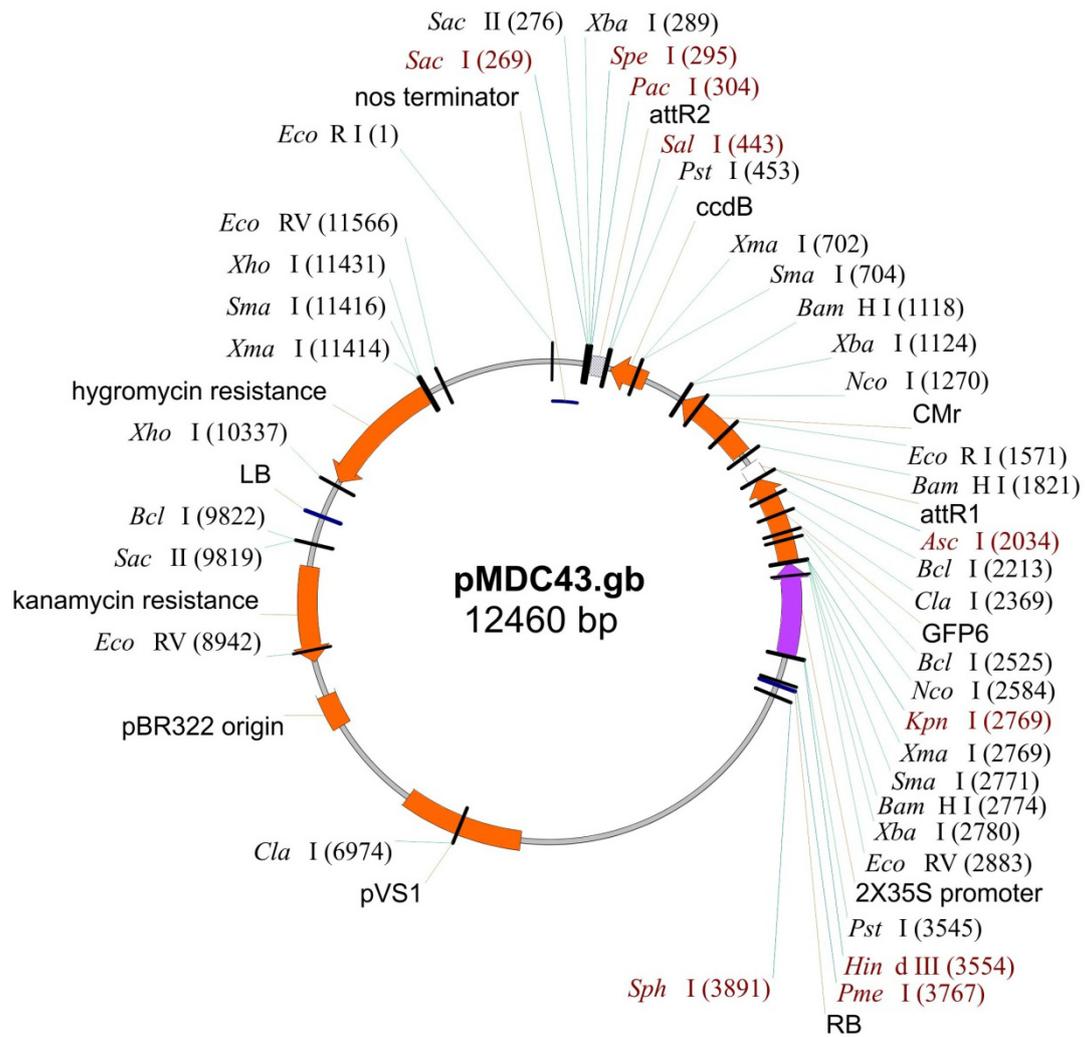


Figure A.5. – pMDC43 destination gateway plant vector circle map for 35S::GFP-GENE constructs (Curtis and Grossniklaus, 2003).

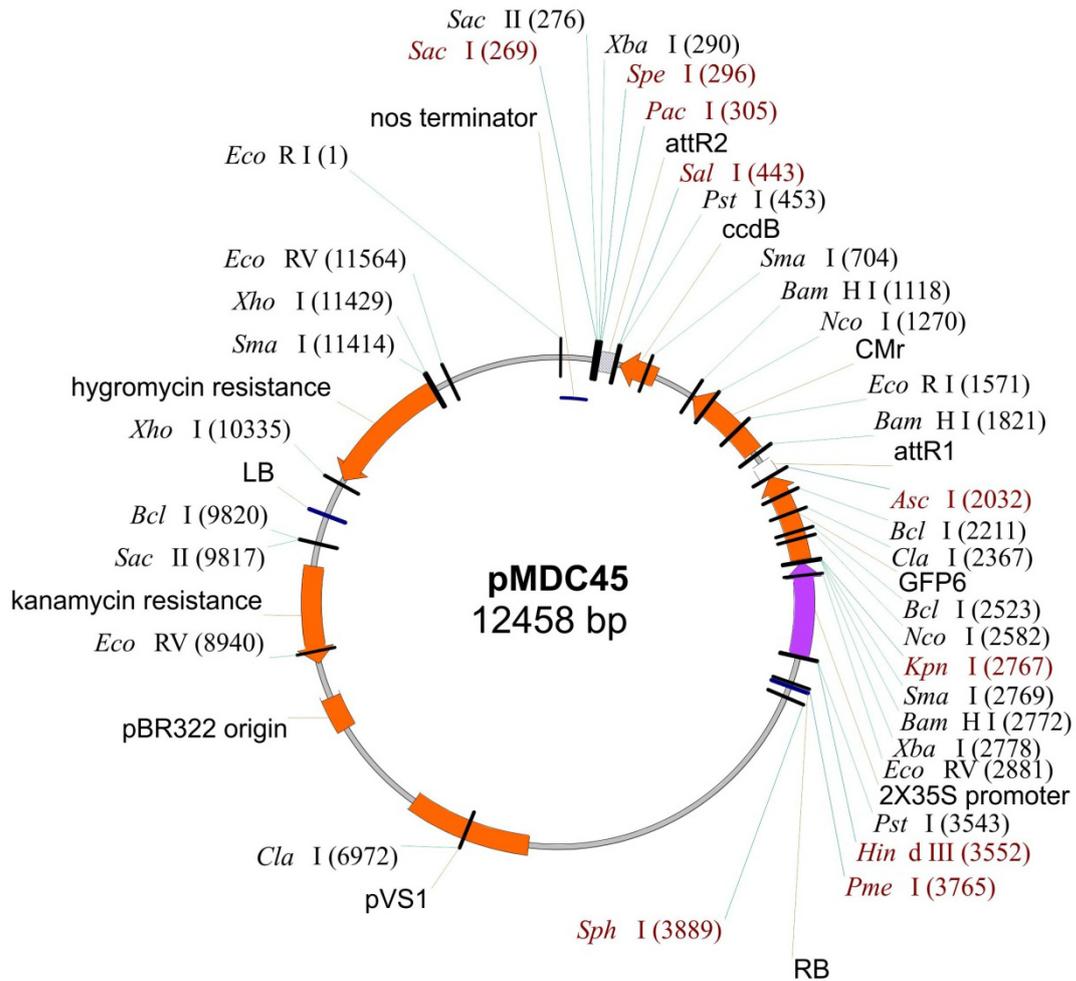


Figure A.6. – pMDC45 destination gateway plant vector circle map for 35S::GFP-GENE constructs (Curtis and Grossniklaus, 2003)

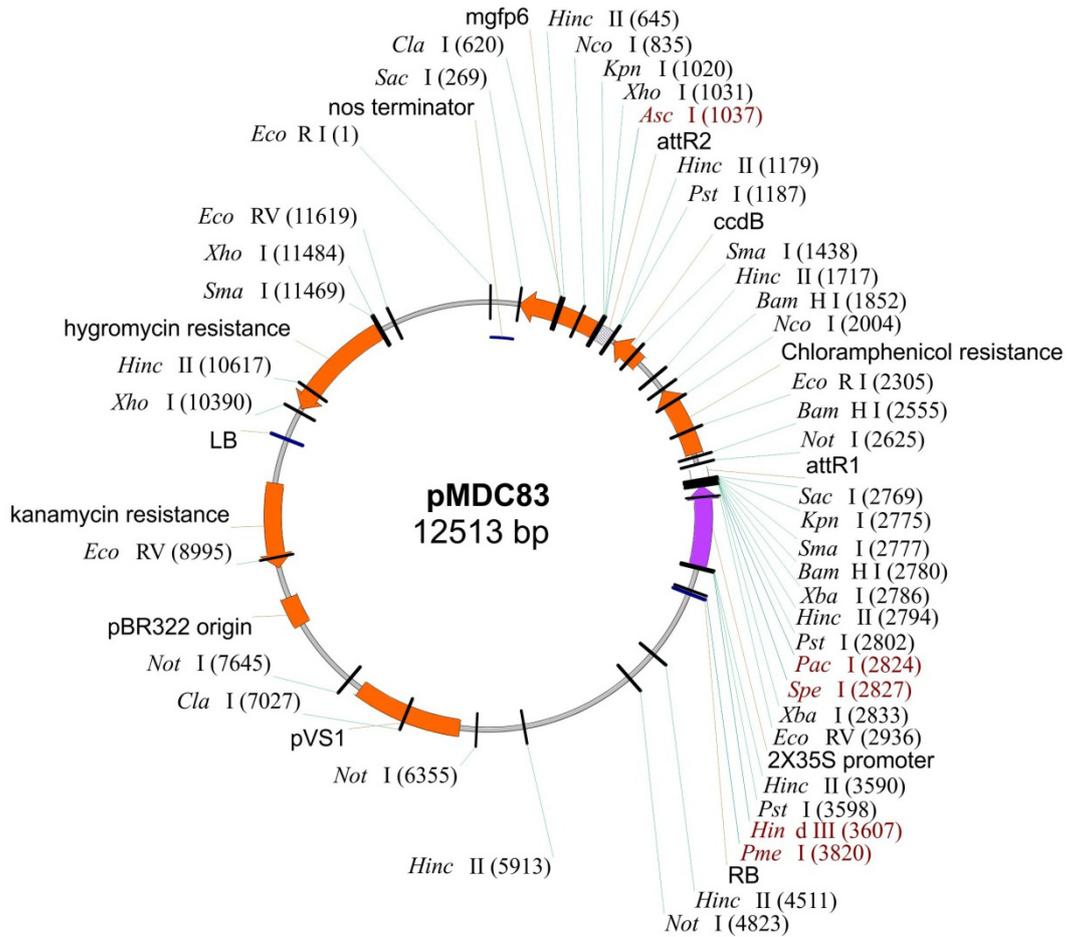


Figure A.7. – pMDC83 destination gateway plant vector circle map for 35S::GENE-GFP constructs (Curtis and Grossniklaus, 2003).

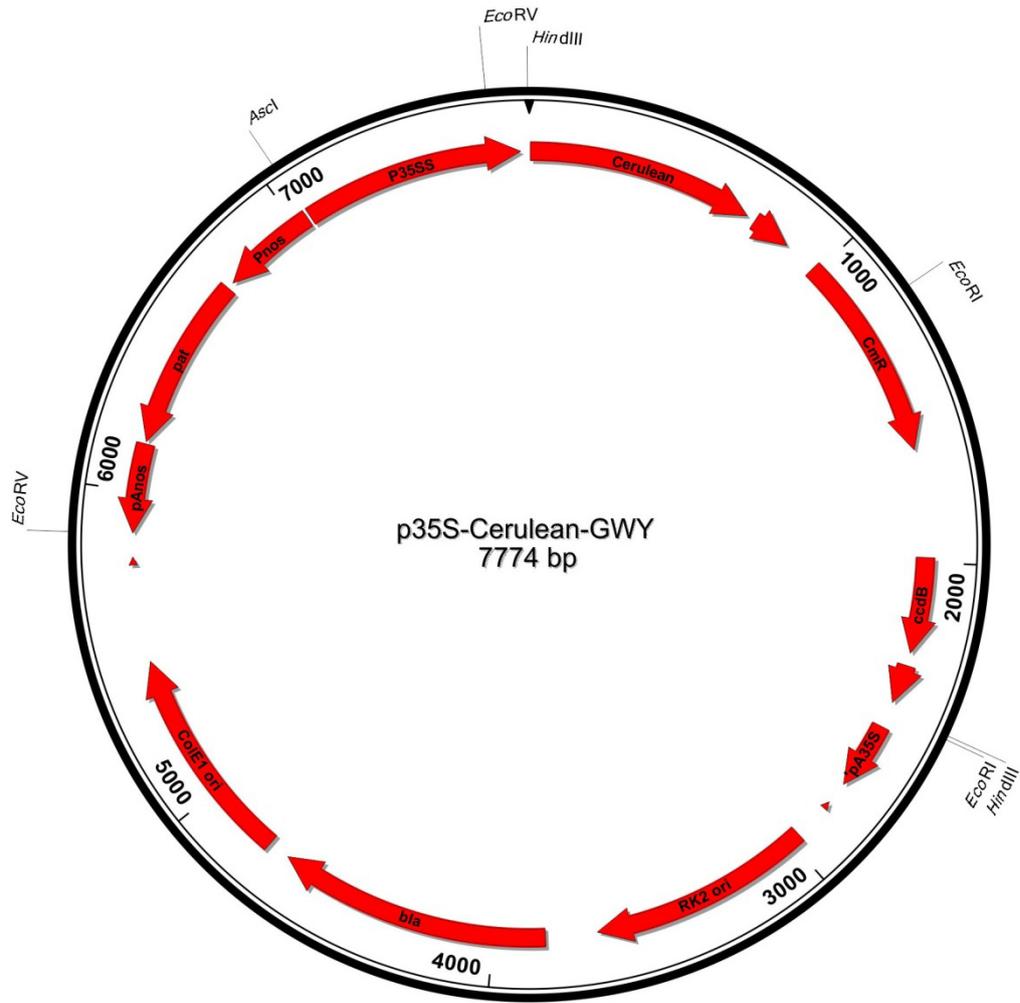


Figure A.8. – Gateway destination plant vector circle map for 35S::Cerulean-Gene constructs

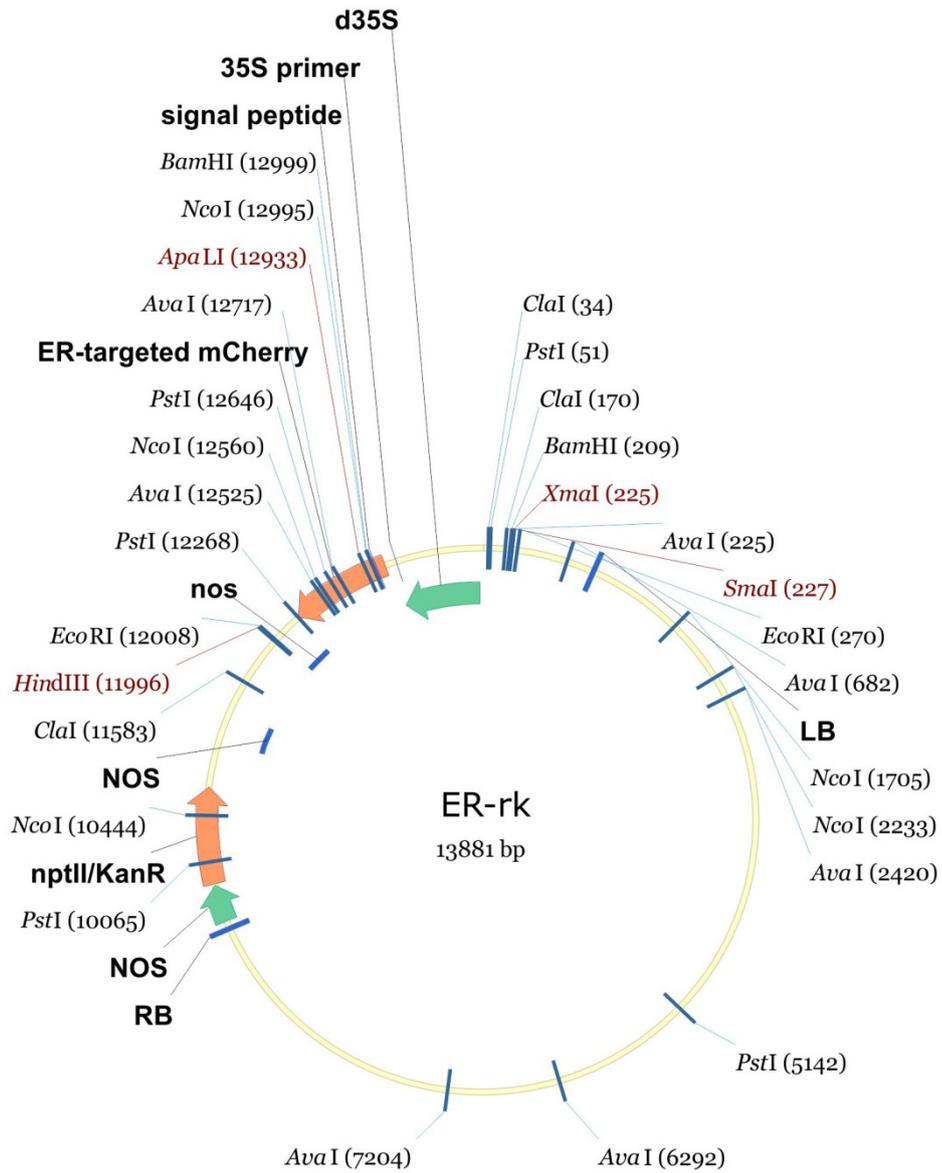


Figure A.9. – Er-rb vector circle map - vector with fluorescent endoplasmic reticulum marker. A sequence tagging to the ER fused with mcherry fluorescent CDS (Nelson *et al.*, 2007).

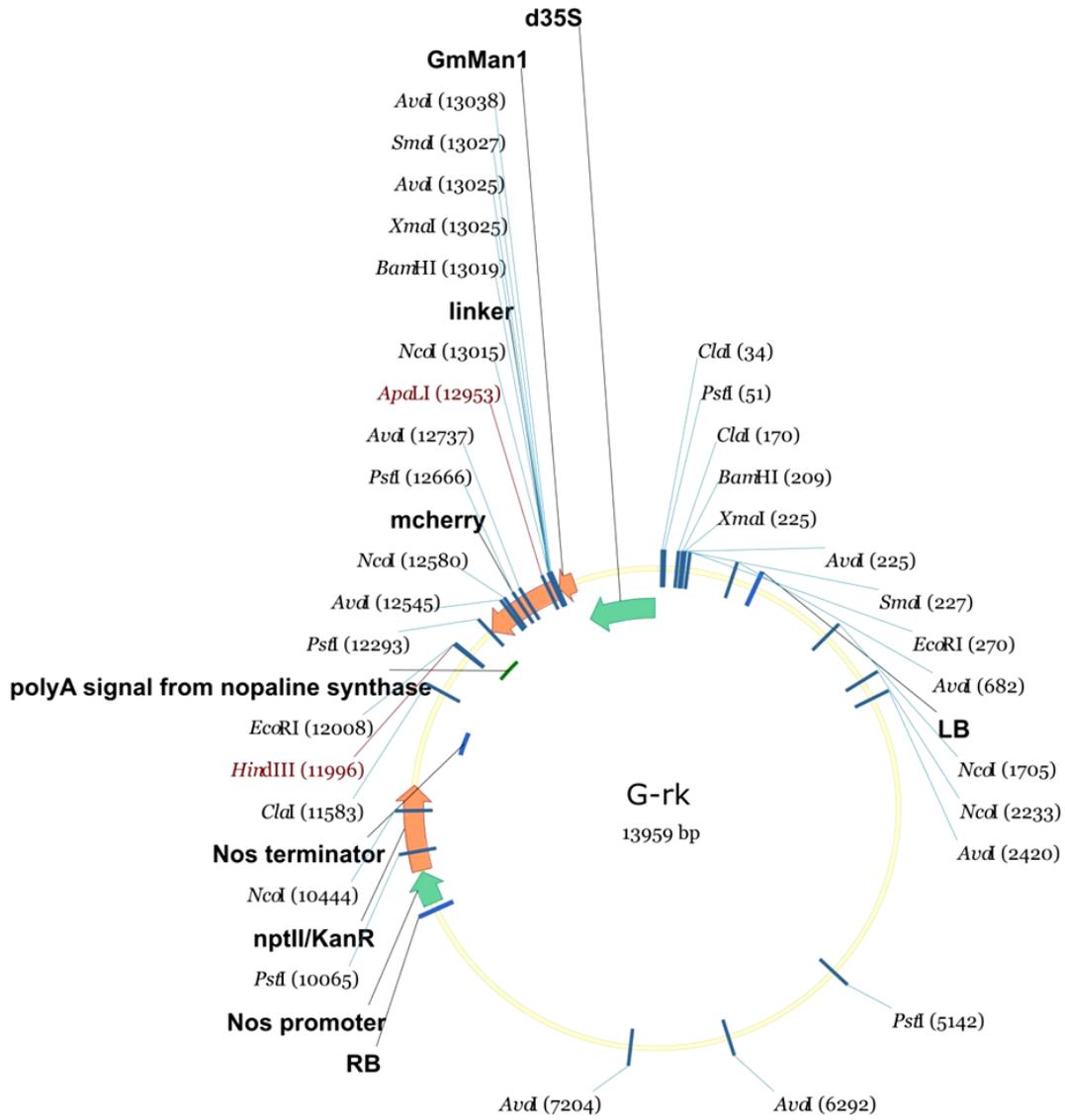


Figure A.10. – G-rb vector circle map - vector with fluorescent golgi marker. A sequence tagging to the Golgi fused with mcherry fluorescent CDS (Nelson *et al.*, 2007).

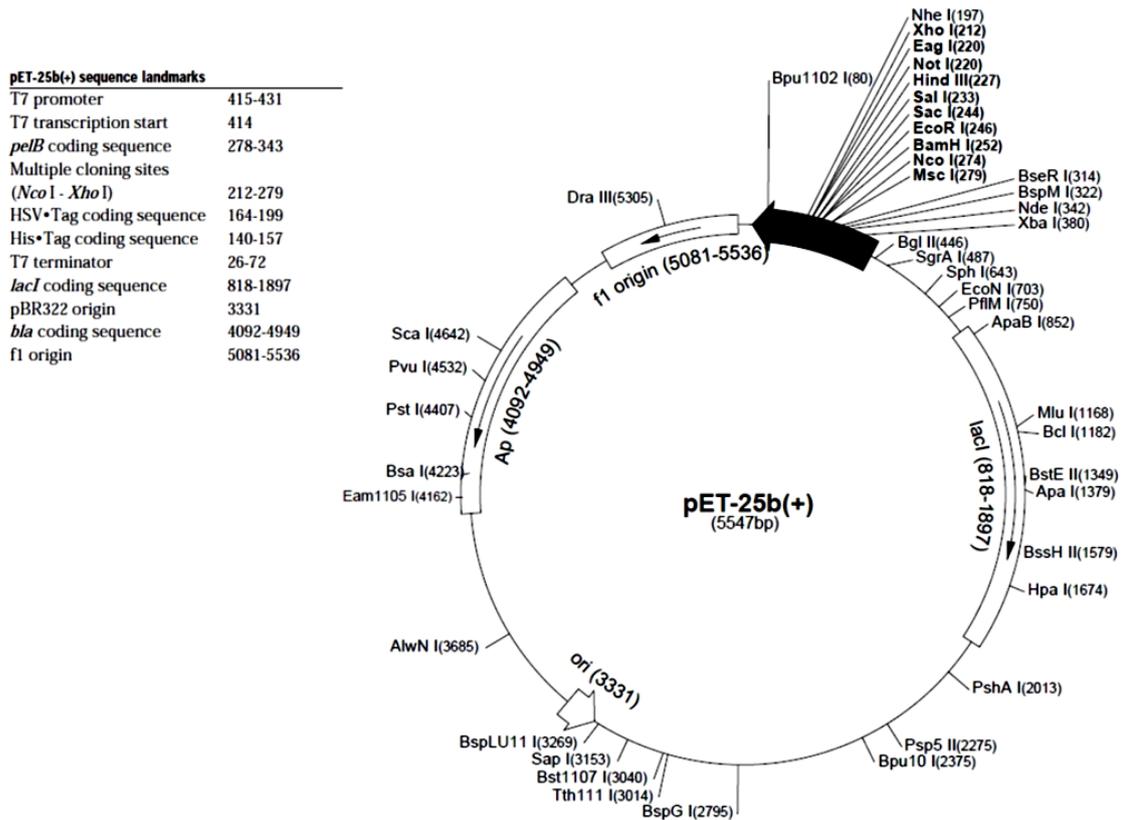


Figure A.11. – pET-25b(+)G-rb protein expression vector circle map with Histidine and HSV tag at the C-terminal.

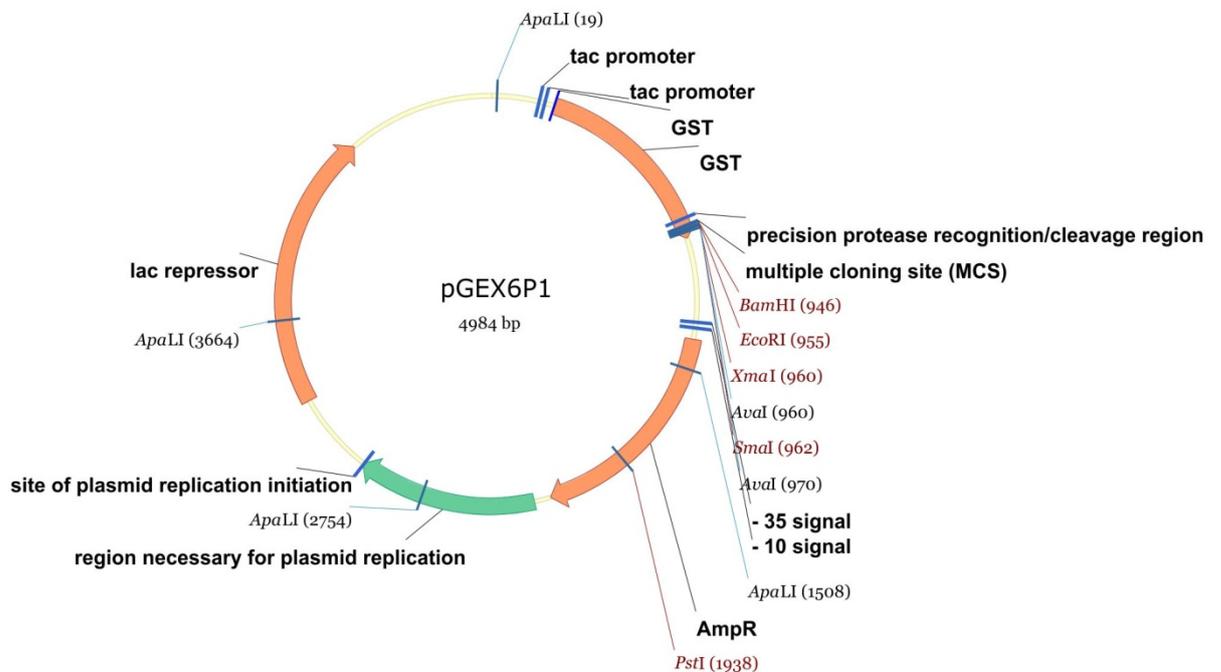


Figure A.12. – pGEX-6P-1 protein expression vector circle map with GST tag at the N-terminal.

