



**HAL**  
open science

**Flavonoid oxidation in Arabidopsis thaliana seeds.  
Functional analysis of the TRANSPARENT TESTA 10  
gene encoding a polyphenoloxidase of the laccase type.**

Przemyslaw P. Bidzinski

► **To cite this version:**

Przemyslaw P. Bidzinski. Flavonoid oxidation in Arabidopsis thaliana seeds. Functional analysis of the TRANSPARENT TESTA 10 gene encoding a polyphenoloxidase of the laccase type.. Molecular biology. Universite Paris XI Orsay, 2009. English. NNT: . tel-02177497

**HAL Id: tel-02177497**

**<https://theses.hal.science/tel-02177497>**

Submitted on 9 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE DE PARIS-SUD  
U.F.R. SCIENTIFIQUE D'ORSAY

**THESE**

présentée  
pour obtenir le grade de

**DOCTEUR EN SCIENCES  
DE L'UNIVERSITE PARIS 11, ORSAY**

*Discipline: Sciences de la Vie*

par

**Przemyslaw BIDZINSKI**

**Oxydation des flavonoïdes dans les graines d'*Arabidopsis thaliana*.  
Analyse fonctionnelle du gène *TRANSPARENT TESTA 10* codant  
une polyphénoloxydase de type laccase**

Soutenue le 7 Juillet 2009 devant la commission d'examen composée de :

Graham NOCTOR	Professeur, Université Paris 11, Orsay	Président
Alain GOOSSENS	Chercheur, Université Gand/VIB	Rapporteur
Wim SOPPE	Chercheur, MPI Cologne	Rapporteur
Lise JOUANIN	Directrice de Recherche CNRS	Examinatrice
Nancy TERRIER	Chargée de Recherche INRA	Examinatrice
Isabelle DEBEAUJON	Chargée de Recherche INRA	Directrice de Thèse

*Life, sometimes so wearying  
Is worth its weight in gold  
The experience of traveling  
Lends a wisdom that is old  
Beyond our 'living memory'  
A softly spoken prayer:*

*"It's the journey that's important,  
Not the getting there!"*

[...]

By John McLeod

## Acknowledgments

After more than three years of the journey through the PhD, the time has come to look back and acknowledge all the people I have shared this part of the the life with.

First, I would like to thank my supervisor Isabelle Debeaujon. Thank you for all your help, discussions, critics and advice. Thank you for patiently answering all my questions (some of them more than once) and sharing your research experience with me.

I would like to thank Annie Marion-Poll and all the members of the Seed Biology Laboratory for excellent work environment.

I would like to acknowledge Adeline Berger for enormous help with the cytology, Anne Frey for guiding and help for the MUG experiment, Nathalie Berger for sharing the secrets of 'boiling' and 'dirty cloning', Bertrand Dubreucq for explaining the qRT-PCR and all the members of the laboratory for sharing their experience.

I would like to thank the FLAVO team: Loïc Lepiniec for leading the group and all the support I got from him during my PhD. Lucille Pourcel who shared her first hand experience on work with *TT10* and gave the idea to run Paris-Versailles. Jean-Marc Routaboul for all the metabolomic analyses, discussions and comments as well as the best cassoulet one can imagine. Special thanks to Christian Dubos for all the advice and help, especially when it was needed the most.

Special thanks to Martine Miquel for support, encouragement and explaining how to make stroudele dough very thin.

Managing the French administration would not be possible without Thomas Goujon, who also spread good mood everywhere he was passing and shared expert advice on rollerblading.

Thanks to Marie Lacruz and Stéphane Raude who took care of the administrative part during the whole PhD.

One must not forget Michaël Anjuere who was looking after the plants in the greenhouse.

Big thanks to Anne Plessis who explained me how to use LyX and Joseph Tran who helped to put this manuscript in the final shape.

I would like to thank the members of the jury for reading and evaluating this manuscript as well as members of my thesis committees, Lise Jouanin, Olivier Loudet and Alain Goossens who shared their time and experience. Thanks to Gregory Mouille for being the reporter at my first Monday seminar in the Versailles amphitheater.

Thanks to Wim Soppe for hosting me in his lab and Christina Philip for technical assistance in BAC library screening.

I would like to thank all my friends who were patiently listening when I was sharing my last experiments with them. Big thanks to all the VERT students. Special thanks to Laura who shared not only passion for research but also long running kilometers and a climbing rope. Thank you Theodoros for all long discussions, Greek dances and all great moments. Big thanks to Tino. Thanks to all great people I have met living in France and especially in Grignon for the multi-cultural exchange.

Thank you Loren for all your support, encouragement and patience. Thank you more than the universe.

Thanks to my family and especially my sister Joanna, who was supporting me from far away. Thank you for always being there.

I dedicate this thesis to my parents Jadwiga and Leon.

## Abstract

### Title:

Flavonoid oxidation in *Arabidopsis thaliana* seeds. Functional analysis of the *TRANSPARENT TESTA 10* gene encoding a polyphenoloxidase of the laccase type.

**Abstract:** Arabidopsis seeds accumulate flavonoids (proanthocyanidins and flavonols) during their development. A previous study has shown that a laccase (AtLAC15) encoded by the *TRANSPARENT TESTA 10* (*TT10*) gene could trigger flavonoid oxidation in the seed coat. If both proanthocyanidins (PAs) and flavonols appear to be TT10 protein substrates, only PA oxidation leads to brown pigments responsible for the mature seed coat color. An important consequence of TT10 activity on seed flavonoid metabolism is an increased ratio of insoluble to soluble PAs. The physiological functions of TT10 are still unknown, however defense against biotic and abiotic stresses, either constitutive or induced, may be predicted on the basis of present knowledge on polyphenoloxidases.

The purpose of this thesis was to perform a functional characterization of the *TT10* gene. A part of the work was devoted to the analysis of the regulatory mechanisms controlling the developmental pattern of *TT10* gene expression in seeds and vegetative plant parts. The functional 5'-dissection of a 2.0-kb promoter realized with the *uidA* reporter gene encoding  $\beta$ -glucuronidase (GUS) was performed to identify regions responsible for activation in seed and other plant organs. *TT10* promoter happens to be activated exclusively in seed coat and siliques. Directed mutagenesis was undertaken to precise the regulatory role of in silico-detected cis-acting regulatory elements (CAREs) located in a 194-bp region necessary for expression in seed coat. *TT10* gene expression assessed in different tissues at various stages of development using qRT-PCR matched promoter activity pattern. Natural variation for TT10 expression among Arabidopsis accessions was also detected, with the levels of *TT10* mRNA in Cvi, Ler and Sha being strongly reduced compared to the ones in Ws, Col and Bay. The impact of this molecular polymorphism on seed flavonoid composition, as analyzed on mature seeds with LC-MS, is discussed. *In silico* analysis of the *TT10* promoter revealed the presence of putative CAREs potentially involved in signaling and response to biotic and abiotic stresses. However histochemical analysis of GUS activity in transgenic Arabidopsis plantlets expressing *pTT10<sub>2.0-kb</sub> : GUS* failed to detect any ectopic activity when submitted to a variety of stresses. This result suggests that transcriptional response to environmental stimuli is highly constrained by developmental parameters. *TT10* appeared to be the only member of the laccase gene family to be strongly expressed in seeds. TT10 function may have evolved towards flavonoid oxidation by co-localization with these substrates, which is ensured by tissue-specific gene expression.

**Key words:** Arabidopsis, flavonoids, gene, laccase, mutant, promoter, seed, transcriptional regulation

### Resident laboratory:

Seed Biology Laboratory – INRA, Jean-Pierre Bourgin Institute  
UMR 204 INRA-AgroParisTech  
Route de Saint-Cyr, 78026 Versailles Cedex, France

## Résumé

### Titre:

Oxydation des flavonoïdes dans les graines d'*Arabidopsis thaliana*. Analyse fonctionnelle du gène *TRANSPARENT TESTA 10* codant une polyphénoloxydase de type laccase.

### Résumé:

Les graines d'*Arabidopsis* accumulent des flavonoïdes (proanthocyanidines et flavonols) durant leur développement. Une étude précédente a montré qu'une laccase (AtLAC15) codée par le gène *TRANSPARENT TESTA 10* (*TT10*) induisait l'oxydation des flavonoïdes au niveau des téguments. Si les proanthocyanidines (PAs) et les flavonols sont des substrats de la protéine TT10, seule l'oxydation des PAs conduit à la formation de pigments bruns responsables de la couleur de la graine mature. Une importante conséquence de l'activité de TT10 sur les flavonoïdes de la graine est l'augmentation du rapport PAs insolubles / PAs solubles. Les fonctions physiologiques de TT10 sont toujours inconnues, mais un rôle dans la défense contre des stress biotiques et abiotiques constitutifs ou induits peut être prédit sur la base des connaissances actuelles sur les polyphénoloxydases.

L'objectif de cette thèse était de réaliser une analyse fonctionnelle du gène *TT10*. Une partie de l'étude était consacrée à l'analyse des mécanismes de régulation contrôlant le pattern développemental d'expression du gène *TT10* dans les graines et les parties végétatives. Une dissection fonctionnelle en 5' d'un promoteur de 2.0-kb réalisée à l'aide du gène rapporteur *uidA* codant la  $\beta$ -glucuronidase (GUS) a été réalisée pour identifier les régions responsables de l'activation du promoteur dans les graines et les autres parties de la plante. Le promoteur de *TT10* est activé exclusivement dans les téguments de la graine et les siliques. Une mutagenèse dirigée a été réalisée dans une région de 194 pb nécessaire à l'expression dans les téguments, pour préciser la fonctionnalité des éléments régulateurs agissant en cis (ERACs) identifiés *in silico*. L'expression du gène *TT10* mesurée par qRT-PCR dans différents tissus et à différents stades de développement est en accord avec le pattern d'activité du promoteur. Une variation naturelle pour l'expression de *TT10* a aussi été détectée parmi plusieurs accessions d'*Arabidopsis*, avec des niveaux d'ARNm mesurés chez Cvi, Ler et Sha fortement réduits par rapport à ceux relevés chez Ws, Col et Bay. L'impact de ce polymorphisme moléculaire sur la composition en flavonoïdes de la graine, analysée sur graines matures par LC-MS, est discuté. L'analyse *in silico* du promoteur de *TT10* a révélé la présence d'ERACs potentiellement impliqués dans la signalisation et la réponse aux stress biotiques et abiotiques. Cependant l'analyse histochimique de l'activité GUS de plantes transgéniques exprimant *pTT10<sub>2,0-kb</sub> : GUS* n'a pas permis de détecter d'activité ectopique en présence de stress variés. Ce résultat suggère que la réponse transcriptionnelle aux stimuli environnementaux est fortement conditionnée par les paramètres développementaux. *TT10* est le seul membre de la famille des laccases à être exprimé fortement dans les graines. La fonction de *TT10* a probablement évolué vers l'oxydation des flavonoïdes par la co-localisation avec ces substrats, qui est assurée par la spécificité tissulaire d'expression du gène.

**Mots clés:** *Arabidopsis*, flavonoïdes, gène, graine, laccase, mutant, promoteur, régulation transcriptionnelle

### Laboratoire d'accueil:

Laboratoire de Biologie des Semences – INRA, Institut Jean-Pierre Bourgin

UMR 204 INRA-AgroParisTech

Route de Saint-Cyr, 78026 Versailles Cedex, France

# Contents

Acknowledgments . . . . .	i
Abstract . . . . .	iii
Résumé . . . . .	iv
Table of contents . . . . .	v
List of Figures . . . . .	ix
List of Tables . . . . .	xii
Abbreviations . . . . .	xv
<b>1 Introduction</b>	<b>1</b>
1.1 Seed biology . . . . .	1
1.1.1 Seed coat development in <i>Arabidopsis</i> . . . . .	1
1.2 Flavonoids in <i>Arabidopsis</i> seed . . . . .	2
1.2.1 Biosynthesis . . . . .	3
1.2.2 Biological roles of flavonoids . . . . .	5
1.2.3 Oxidation of end-products . . . . .	5
1.2.3.1 Laccases . . . . .	7
1.2.3.2 Catechol oxidases . . . . .	8
1.3 Regulation of gene expression . . . . .	8
1.3.1 Regulation of <i>BAN</i> and <i>MYBL2</i> expression . . . . .	8
1.3.2 Two-cell layer expression pattern of <i>TT10</i> . . . . .	9
1.3.3 One layer, two layers, which layers? . . . . .	9
1.4 Objectives of the thesis . . . . .	9
<b>2 Results</b>	<b>11</b>
2.1 Transcriptional regulation of <i>TT10</i> expression . . . . .	11
2.1.1 <i>TT10</i> transcript accumulation during plant development . . . . .	11
2.1.1.1 Transcript accumulation - RT-PCR . . . . .	11
2.1.1.2 Transcript accumulation - Transcriptomics . . . . .	12
2.1.1.3 Analysis of publicly available transcriptomic data for <i>TT10</i> expression during plant development - Summary . . . . .	14
2.1.2 <i>In silico</i> analysis of <i>TT10</i> promoter sequence . . . . .	15
2.1.2.1 Putative <i>cis</i> -acting regulatory elements found in the <i>TT10</i> promoter . . . . .	16
2.1.2.2 Putative CAREs associated with seed development, stress response and/or protection against various biotic and abiotic stresses . . . . .	17
2.1.2.3 CAREs found in promoters of flavonoid biosynthetic genes . . . . .	18



2.1.2.4	Additional comments about known transcription factors involved in flavonoid biosynthesis . . . . .	19
2.1.2.5	Other CAREs . . . . .	19
2.1.2.6	TATA box . . . . .	20
2.1.2.7	<i>In silico</i> analysis of <i>TT10</i> promoter sequence - Summary . . . . .	20
2.1.3	Studies of <i>TT10</i> promoter activity during plant development . . . . .	22
2.1.3.1	5' dissection of the <i>TT10</i> promoter . . . . .	22
2.1.3.2	<i>In planta</i> analysis of GUS activity . . . . .	23
2.1.3.3	Quantitative GUS assay - MUG . . . . .	25
2.1.3.4	Sub-dissection and site-directed mutagenesis - introduction . . . . .	25
2.1.3.5	Patterns of GUS activities after <i>TT10</i> promoter sub-dissection and site-directed mutagenesis . . . . .	28
2.1.4	Promoter GUS activity in protoplasts . . . . .	30
2.2	Finding transcription factors . . . . .	31
2.2.1	<i>TT10</i> promoter activity in candidate regulatory mutant backgrounds . . . . .	31
2.2.2	Confirmation of <i>TTG2</i> promoter activity in developing seeds . . . . .	32
2.2.3	Analysis of public micro-array data - searching for co-expressed genes . . . . .	33
2.2.3.1	<i>TT10</i> expression in candidate regulatory mutant backgrounds obtained from co-expression analysis - seed set . . . . .	33
2.2.3.2	Results for regulatory mutants chosen from co-expression analysis . . . . .	35
2.2.3.3	Co-expression analysis - developmental set . . . . .	37
2.2.4	Analysis of <i>TT10</i> expression in mutants involved in flavonoid biosynthesis . . . . .	38
2.2.5	Candidate genes related to stress response which could be involved in the regulation of <i>TT10</i> expression . . . . .	39
2.2.6	Yeast one hybrid screening - fishing out transcription factors binding <i>TT10</i> promoter . . . . .	42
2.3	Insights from natural variation . . . . .	44
2.3.1	Differences in <i>TT10</i> transcript accumulation . . . . .	44
2.3.2	Analysis of promoter sequences . . . . .	45
2.3.2.1	Focus on the proximal part of the promoter . . . . .	46
2.3.2.2	SNP <i>vs.</i> <i>TT10</i> transcript accumulation . . . . .	48
2.3.2.3	Analysis of long promoter fragments . . . . .	48
2.3.2.4	BAC library screening . . . . .	48
2.3.2.5	Going further in the promoter: amplification of longer fragments . . . . .	48
2.3.3	Flavonoid composition in mature seeds of various <i>Arabidopsis</i> accessions . . . . .	49
2.3.4	Functional complementation of <i>Ler</i> accession with Col-0 <i>TT10</i> genomic region . . . . .	52
2.3.5	Functional analysis of the Columbia promoter in other backgrounds . . . . .	53
2.4	Stress induction of the <i>TT10</i> promoter activity . . . . .	54
2.4.1	Abiotic stresses . . . . .	54
2.4.2	Wounding and biotic stresses . . . . .	55
2.5	Identification of laccases expressed in seeds . . . . .	57
2.5.1	Expression data . . . . .	57
2.5.2	Flavonoid composition of mature seeds of the laccase mutants . . . . .	58
2.5.3	Promoter GUS activities . . . . .	59
2.5.3.1	LAC3 . . . . .	60
2.5.3.2	LAC5 . . . . .	60

2.5.3.3	LAC12	60
2.5.3.4	Summary	60
2.6	Cell suspension as a tool to study flavonoid biosynthesis	65
<b>3</b>	<b>Materials and Methods</b>	<b>67</b>
3.1	Materials	67
3.1.1	Plant Material	67
3.1.1.1	<i>Arabidopsis thaliana</i> - wild type accessions	67
3.1.1.2	<i>Arabidopsis thaliana</i> - mutants	67
3.1.1.3	<i>Arabidopsis thaliana</i> - transgenic plants	71
3.1.1.4	<i>Arabidopsis thaliana</i> - cell suspension	72
3.1.2	Bacterial Strains	72
3.1.2.1	<i>Escherichia coli</i> , strain DH10B	72
3.1.2.2	<i>Escherichia coli</i> , One shot® TOP 10	72
3.1.2.3	<i>Escherichia coli</i> , DB3.1	72
3.1.2.4	<i>Agrobacterium tumefaciens</i> , strain 58C1pMP90	72
3.1.3	Yeast strain	72
3.1.3.1	<i>Saccharomyces cerevisiae</i> , strain EGY48	72
3.1.3.2	Yeast strains in the REGIA transcription factor library	72
3.1.4	Plasmids and expression vectors	73
3.1.5	Cvi and <i>Ler</i> BAC library at MPI, Cologne	74
3.1.6	Transcription factor library	75
3.2	Methods	75
3.2.1	Handling of bacteria	75
3.2.1.1	Culture	75
3.2.1.2	Preparation and transformation of electro-competent bacteria	75
3.2.2	Handling of yeast	75
3.2.2.1	Culture	75
3.2.2.2	Preparation and transformation of yeast	76
3.2.3	Handling of <i>Arabidopsis thaliana</i>	76
3.2.3.1	Seed sterilization	76
3.2.3.2	Plants growth conditions <i>in vitro</i>	76
3.2.3.3	Plants growth conditions in greenhouse	76
3.2.3.4	Crosses	76
3.2.3.5	Plant transformation by the floral dip method	77
3.2.3.6	Maintenance of cell suspensions	77
3.2.3.7	Genotyping	77
3.2.3.8	Transient expression of promoter:GUS constructs in <i>Arabidopsis</i> protoplasts	77
3.2.4	DNA methods	77
3.2.4.1	Isolation of plasmid DNA from <i>E. coli</i>	77
3.2.4.2	DNA fragments purification	77
3.2.4.3	Isolation of genomic DNA from plants	77
3.2.4.4	Rapid DNA extraction for genotyping	78
3.2.4.5	Polymerase chain reaction (PCR)	78
3.2.4.6	PCR fragments purification	78

3.2.4.7	Sequencing . . . . .	79
3.2.4.8	Site-directed mutagenesis . . . . .	79
3.2.5	RNA methods . . . . .	79
3.2.5.1	RNA isolation and cDNA synthesis . . . . .	79
3.2.5.2	RT-PCR . . . . .	79
3.2.5.3	Quantitative real-time RT-PCR (qRT-PCR) . . . . .	79
3.2.5.4	Normalization controls for qRT-PCR . . . . .	80
3.2.5.5	Primer efficiency . . . . .	80
3.2.6	Histochemical methods for GUS detection and quantification . . . . .	82
3.2.6.1	Histochemical detection of GUS activity . . . . .	82
3.2.6.2	Quantification of GUS activity (MUG assay) . . . . .	82
3.2.7	Analysis of seed flavonoids . . . . .	82
3.2.8	Routine techniques . . . . .	83
3.2.9	Internet resources and services . . . . .	83
3.2.9.1	The Arabidopsis Information Resource - TAIR . . . . .	83
3.2.9.2	The Arabidopsis Thaliana Integrated Database - ATIDB . . . . .	83
3.2.9.3	Integrative database around plant genomes - FLAGdb++ v.3.9 . . . . .	83
3.2.9.4	The Complete Arabidopsis Transcriptome Micro Array database - CATMA Database . . . . .	83
3.2.9.5	AtcisDB at AGRIS . . . . .	83
3.2.9.6	The Genevestigator - expression database and meta-analysis system . . . . .	84
3.2.9.7	The Bio-Array Resource for Arabidopsis Functional Genomics - BAR . . . . .	84
3.2.9.8	GeneCAT - Gene Co-expression Analysis Toolbox . . . . .	84
3.2.9.9	DIURNAL . . . . .	84
3.2.9.10	PLACE and Signal Scan - A Database of Plant <i>Cis</i> -acting Regulatory DNA Elements . . . . .	84
3.2.9.11	Plant CARE . . . . .	84
3.2.9.12	Plant PAN - Plant Promoter Analysis Navigator . . . . .	85
3.2.9.13	ATHENA . . . . .	85
3.2.9.14	Basic Local Alignment Search Tool - BLAST . . . . .	85
3.2.9.15	Multiple Sequence Alignment . . . . .	85
3.2.9.16	PrimerX - Primer design for the site directed mutagenesis . . . . .	85
<b>4</b>	<b>Discussion, Conclusions and Perspectives</b>	<b>87</b>
4.1	Regulation of <i>TT10</i> gene expression . . . . .	87
4.1.1	Transcript accumulation . . . . .	87
4.1.2	Developmental vs. stress regulated expression . . . . .	88
4.1.3	Regulatory components . . . . .	89
4.1.3.1	Promoter and <i>Cis</i> -acting regulatory elements . . . . .	90
4.1.3.2	Candidate transcription factors for developmental regulation . . . . .	92
4.1.3.3	Candidate transcription factors for stress regulation . . . . .	93
4.1.3.4	What have we learned from natural variation? . . . . .	94
4.1.3.5	Transcription factor library screening . . . . .	95
4.1.3.6	Co-expressed genes . . . . .	95
4.1.3.7	Epigenetics? . . . . .	96

<i>CONTENTS</i>	ix
4.2 <i>TT10</i> and flavonoid metabolism in seeds . . . . .	97
4.2.1 <i>TT10</i> and flavonoid metabolism . . . . .	97
4.2.2 Other functions? . . . . .	97
4.3 Regulation in the multi-gene family . . . . .	98
4.4 Final remarks . . . . .	98
<b>5 Appendices</b>	<b>99</b>
5.1 <i>In silico</i> analysis of the <i>TT10</i> promoter . . . . .	99
5.1.1 Putative <i>cis</i> -acting regulatory DNA elements found in PLACE database using the Signal Scan . . . . .	99
5.1.2 CAREs found in <i>AtcisDB</i> . . . . .	103
5.2 Primers . . . . .	105
5.3 Seeds colors . . . . .	110
5.4 Map of the methylation in the <i>TT10</i> gene . . . . .	112
5.5 Transposon in <i>TT10</i> promoter . . . . .	113
<b>Bibliography</b>	<b>114</b>



# List of Figures

1.1	Testa structure and flavonoid localization in <i>Arabidopsis</i> seed. . . . .	2
1.2	The structure of the flavonol quercetin. . . . .	3
1.3	Flavonoid biosynthesis pathway in <i>Arabidopsis</i> . . . . .	4
1.4	Enzymatic reactions for polyphenoloxidases (laccase, catechol oxidase). . . . .	6
1.5	Seed coat pigmentation in <i>Arabidopsis</i> : illustration of a browning process. . . . .	6
1.6	Seed coat sections showing the co-localization of <i>TT10</i> promoter activity, mRNA expression and protein activity with flavonoid substrates (5-daf seeds). . . . .	7
1.7	Spatial pattern of <i>TT10</i> promoter activity in WT seeds. Comparison with <i>BAN</i> and <i>MYBL2</i> promoter activities. . . . .	10
2.1	Expression pattern of the <i>TT10</i> gene in developing siliques and seeds of wild-type Col-0 plants. . . . .	12
2.2	<i>TT10</i> expression visualized with eFP Browser. . . . .	13
2.3	Comparison of <i>TT10</i> and <i>BAN</i> expression visualized with GeneCAT expression profiling tool. . . . .	14
2.4	Comparison of <i>TT10</i> and <i>CHS</i> expression visualized with DIURNAL. . . . .	15
2.5	<i>TT10</i> genomic region and neighboring genes. . . . .	16
2.6	<i>In silico</i> analysis of proximal (500bp) <i>TT10</i> promoter showing putative <i>cis</i> -acting regulatory DNA elements. . . . .	21
2.7	Pattern of <i>TT10</i> promoter activity in wild-type plants. . . . .	24
2.8	Comparison of GUS activity of 5'-dissected <i>TT10</i> promoter <i>in planta</i> . . . . .	26
2.9	GUS activity driven by the the <i>TT10</i> promoter in <i>pCAMBIA</i> binary vector. . . . .	27
2.10	Pattern of <i>TT10</i> promoter activity in seeds of the candidate regulatory mutant backgrounds. . . . .	32
2.11	Pattern of <i>TTG2</i> promoter activity in wild-type seeds. . . . .	33
2.12	Expression levels of candidate regulators of <i>TT10</i> in developing seeds. . . . .	35
2.13	<i>TT10</i> transcript accumulation in candidate regulatory mutant background - co-expressed candidates. . . . .	36
2.14	Flavonoid composition of mature seeds in the mutants for genes selected from co-expression analysis. . . . .	37
2.15	<i>TT10</i> transcript accumulation in candidate regulatory mutant backgrounds. . . . .	39
2.16	<i>TT10</i> transcript accumulation in candidate regulatory mutant backgrounds - stress related candidates. . . . .	41
2.17	Flavonoid composition of mature seeds of the mutants - stress related candidates. . . . .	41
2.18	Summary of yeast one-hybrid experiment. . . . .	43
2.19	Assessment of <i>TT10</i> expression levels in various <i>Arabidopsis</i> accessions. . . . .	45
2.20	Promoter sequence alignment and comparison of the putative CAREs present in various accessions. . . . .	48

2.21	Polymorphism for length of <i>TT10</i> promoter amplification. . . . .	49
2.22	Flavonoid composition of mature seeds of various accessions and <i>tt10</i> alleles. . . . .	51
2.23	<i>TT10</i> transcript accumulation in <i>Ler</i> plants transformed with the Col-0 wild-type <i>TT10</i> genomic region. . . . .	52
2.24	Flavonoid composition of mature seeds from <i>Ler</i> lines complemented with the Col-0 wild-type genomic region. . . . .	53
2.25	Stress experiments - experimental design. . . . .	54
2.26	Plantlets subjected to progressive dehydration. . . . .	56
2.27	Ectopic GUS activity of <i>TT10</i> promoter. . . . .	56
2.28	Comparison of the expression pattern of the laccase multigene family in seed and silique. . . . .	58
2.29	Flavonoid composition of mature seeds of the laccase mutants. . . . .	59
2.30	Pattern of <i>LAC3</i> promoter activity in wild-type plants. . . . .	62
2.31	Pattern of <i>LAC5</i> promoter activity in wild-type plants. . . . .	63
2.32	Pattern of <i>LAC12</i> promoter activity in wild-type plants. . . . .	64
2.33	Total RNA and flavonoid pathway genes expressed in the cell suspension. . . . .	65
3.1	Maps of the vectors used for cloning. . . . .	74
3.2	Flower marking and silique sample collection. . . . .	80
3.3	Expression of <i>APT1</i> and <i>TUBULINE</i> in siliques 7DAF. . . . .	81
4.1	Organization of the regulatory elements in the <i>TT10</i> Promoter. . . . .	91
5.1	Visualization of <i>cis</i> -acting regulatory elements found with <i>AtcisDB</i> search engine. . . . .	103
5.2	Seed colors of various Arabidopsis accessions and <i>tt10</i> mutant alleles, <b>(A)</b> . Seed color of the plants used in the <i>Ler</i> 'complementation' experiment, <b>(B)</b> . . . . .	110
5.3	Seed color of various mutants used in the study - Part 1. . . . .	111
5.4	Seed color of various mutants used in the study - Part 2. . . . .	111
5.5	Seed color of mutants in laccase gene. . . . .	111
5.6	Methylation map of the <i>TT10</i> gene region. . . . .	112
5.7	Alignment of transposon sequence. . . . .	113

# List of Tables

2.1	Summary of <i>TT10</i> promoter activity pattern <i>in planta</i> . . . . .	25
2.2	Patterns of GUS activity observed in the seed coat after sub-dissection and site-directed mutagenesis of <i>TT10</i> promoter. . . . .	30
2.3	List of selected genes co-expressed with <i>TT10</i> - Seed Set. . . . .	34
2.4	List of selected genes co-expressed with <i>TT10</i> - developmental set. . . . .	38
2.5	Concentration of hormones, salts and other treatments in stress experiments. . . . .	55
2.6	Comparison of laccase promoter activities. . . . .	61
3.1	Wild type accessions of <i>Arabidopsis thaliana</i> used in this work. . . . .	68
3.2	Alleles of the <i>tt10</i> mutant studied in this work. . . . .	68
3.3	<i>Arabidopsis thaliana</i> mutant lines used in this study. . . . .	68
3.4	Transgenic lines generated during this work. . . . .	71
3.5	List of vectors used in this study. . . . .	73
3.6	Antibiotics . . . . .	75
3.7	Efficiency of the 3Fw/3Rw set of primers. . . . .	81
5.1	Putative <i>cis</i> -acting regulatory elements found in <i>TT10</i> promoter with Signal Scan. . . . .	103
5.2	Putative <i>cis</i> -acting regulatory elements found in <i>TT10</i> promoter with <i>AtcisDB</i> . . . . .	105
5.3	Primers used for promoter dissection to clone fragments in <i>pBS-GUS</i> vector. . . . .	105
5.4	Primers used for amplification of long promoter fragments in <i>Ler</i> and <i>Cvi</i> accessions. . . . .	105
5.5	Primers used to amplify promoter fragments for transcription factor library screening with <i>pHISi</i> vector. . . . .	106
5.6	Primers used for site-directed mutagenesis on <i>pTT10_3</i> promoter fragment in <i>pBS-GUS</i> vector. . . . .	106
5.7	Primers used for probes preparation for BAC library screening. . . . .	106
5.8	Primers used for promoter dissection - GATEWAY® cloning. . . . .	106
5.9	Primers used to study laccase family genes expression. . . . .	107
5.10	Primers used for laccase promoter cloning. . . . .	107
5.11	Primers used in quantitative RT-PCR. . . . .	108
5.12	Primers used for genotyping of SALK T-DNA mutants. . . . .	108
5.13	Other primers used in the study. . . . .	108
5.14	CATMA GST primers used to study expression of flavonoid biosynthesis genes. . . . .	109





## Abbreviations

#	number, clone
A	adenosine
ATP	adenosine-5'-triphosphate
aa	amino acid
ABA	abscisic acid
AGI	arabidopsis genome initiative
<i>attB1</i> , <i>attB2</i>	recombination sites in GATEWAY® cloning
BAC	Bacterial Artificial Chromosome
BAR	Bio-Array Resource
Bay	Bayreuth
bp	base pairs
BSA	Bovine Serum Albumin
C	cytidine
CARE(s)	<i>Cis</i> -Acting Regulatory Element(s)
cDNA	complementary DNA
CDS	coding sequence
Col	Columbia
Cvi	Cape Verde Islands
DAF	Days After Flowering
DMF	N,N-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide triphosphate
°C	degrees Celsius
DTT	dithiothreitol
EC	epicatechin
EDTA	ethylene diamine tetraacetic acid
<i>EF1</i>	<i>ELONGATION FACTOR 1aA4</i>
eFP	electronic fluorescent pictograph
e.g.	<i>exempli gratia</i> (for example)
EMS	ethane methyl sulfonate
EST	expressed sequence tag
EtOH	ethanol
g	gram
G	guanosine
GA	gibberellic acid
gcRMA	GeneChip Robust Multiarray Averaging
gDNA	genomic DNA
GFP	green fluorescent protein

<i>GL</i>	<i>GLABRA</i>
GUS	$\beta$ -glucuronidase
h	hours
HIS	histidine
ii	inner integument
INRA	French National Institute for Agricultural Research ( <i>fr.</i> l'Institut National de la Recherche Agronomique)
kb	kilo base
l	liter
LAC	laccase
LB	Luria-Bertani medium
LC-MS	liquid chromatography coupled with mass spectrometry detector
<i>Ler</i>	Landsberg <i>erecta</i>
$\mu$	micro
m	meter
M	mega
M	molar concentration (mole/liter)
MCS	multi cloning site
mm	millimeter
MPI(K)	Max Planck Institute (Cologne)
mRNA	messenger ribonucleic acid
MS	mass spectrometry
n	nano
OD	optical density
oi	outer integument
o/n	over night
PA(s)	proanthocyanidins (condensed tannins)
p	promoter
PCR	polymerase chain reaction
pH	negative logarithm of the proton concentration
qRT-PCR	quantitative real time - PCR
RIL(s)	recombinant inbred line(s)
RMA	Robust Multiarray Averaging
RNA	ribonucleic acid
RNase H	ribonuclease H
RT	reverse transcription
RT	room temperature
s	seconds
SD	minimal synthetic defined media
SDS	sodium dodecyl sulphate
Sha	Shakdara
T	thymidine
T-DNA	transferred DNA

TF(s)	transcription factor(s)
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
TT/tt	TRANSPARENT TESTA
<i>uidA</i>	$\beta$ -glucuronidase
UTR	untranslated region
UV	ultraviolet
v/v	volume per volume
Ws	Wassilewskija
w/v	weight per volume
w/w	weight per weight
WT	wild-type
$x \cdot g$	gravity force
X-gluc	5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucuronide
YPD	yeast peptone dextrose medium

Nomenclature:

TT10	written in capital letters for the protein
<i>TT10</i>	written in capital letters, italics for the gene
<i>tt10</i>	written in small letters, italics for the mutant in the gene
<i>pTT10:GUS</i>	transcriptional fusion of the <i>TT10</i> promoter and reporter gene <i>uidA</i>
T1	primary transformants
T2	second generation of the T-DNA transformed plants

In the degenerated motifs, letters stand for nucleotides: **N**: A, C, G or T; **R**: A or G; **W**: A or T; **Y**: C or T



# Chapter 1

## Introduction

### 1.1 Seed biology

Formation of the seed is crucial in the life cycle of higher plants. Seeds serve several functions, among which the keys are nourishment of the embryo, dispersal to a new location, and dormancy to survive during unfavorable conditions. A typical seed of angiosperms, including *Arabidopsis* is constituted by the embryo enclosed in the endosperm, which supplies nutrients and a seed coat (integuments, testa) which provides protection (see Fig. 1.1 A.).

#### 1.1.1 Seed coat development in *Arabidopsis*

In spermatophytes, the ovule is formed by the embryo sac which is surrounded by the nucellus and integument(s). After fertilization, the ovule develops into a seed containing a newly formed embryo immersed in the endosperm and enclosed in the testa. The seed coat is a maternal tissue derived from the differentiation of the integuments, whereas endosperm and embryo are of both maternal and paternal origin because of the double fertilization. The endosperm in *Arabidopsis* mature seed is reduced to one cell layer and is tightly associated with the seed coat.

The *Arabidopsis* testa includes two integuments in which micropyle pore is formed allowing entrance of the pollen tube, and the chalazal tissues. Three cell layers constitute the inner integument (ii), whereas the outer integument is two-layered. The five cell layers follow one of four distinct fates, e.g. subepidermal oil layer is accumulating flavonols, whereas the innermost cell layer (ii1, endothelium) specializes in proanthocyanidin (PA) biosynthesis. PAs are also accumulating in some cells in the chalazal region (pigment strand) creating a continuum of the tannin-producing cells. The vascular tissues of the maternal funiculus are connected with the seed in the chalazal region, which after seed detachment forms a scar called hilum. Formation of the testa involves cellular differentiation as well as programmed cell death. The cellular organization of the integuments is easily distinguishable at the heart stage of embryo development (see Fig. 1.1 A.), however further modifications are leading to the formation of the mature seed coat (see Fig. 1.1 B.). At the late stages of seed development, structure of the epidermal cells which were accumulating starch and forming mucilage is preserved. Other layers are crushed together by that time, PAs are released from the endothelial cells and impregnate the adjacent cell layers. The mature seed coat protects the embryo, contributes to seed dormancy by creation of a physicochemical barrier and helps seed dispersal. Many studies demonstrated that phenolic compounds, especially flavonoids play an important role in plant protection against environmental constraints (e.g. resistance to pathogens and herbivores, UV radiation) and also contribute to the germination-inhibiting

effects. Seed coat development has been recently reviewed in Haughn and Chaudhury (2005), and Debeaujon et al. (2007).

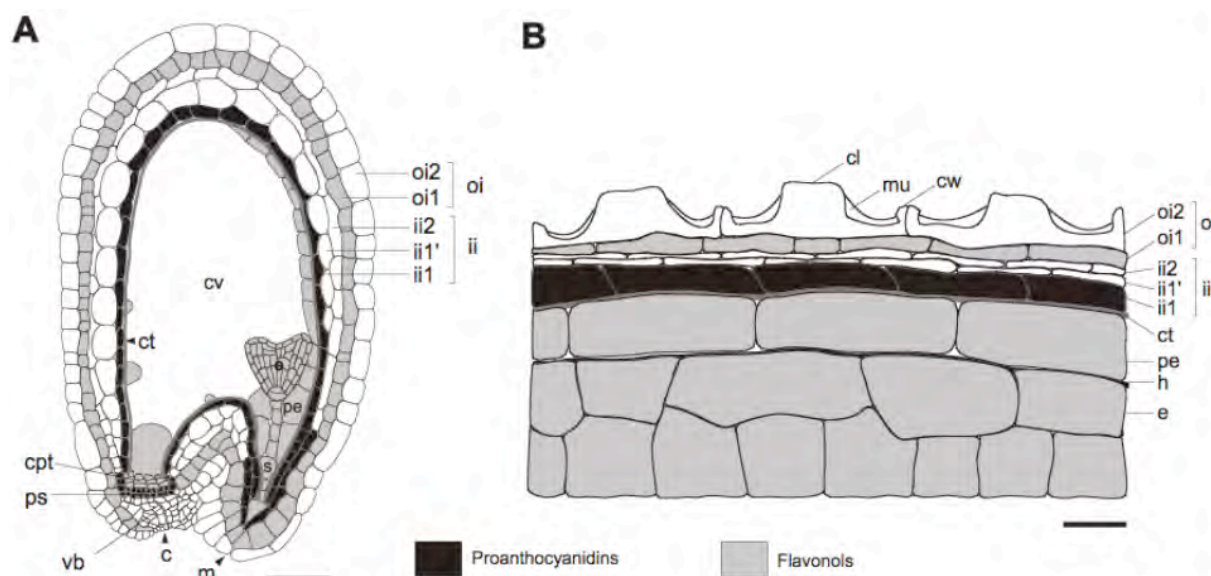


Figure 1.1: Testa structure and flavonoid localization in *Arabidopsis* seed.

(A) Anatomy of a developing seed at the heart stage of embryo development (longitudinal section). Cells accumulating either proanthocyanidins or flavonols are highlighted in black or gray, respectively. The endothelium corresponds to the ii1 layer. (B) Cross section of a mature testa. c, chalaza; cl, columella; cpt, chalazal proliferating tissue (nucellus); ct, cuticle; cv, central vacuole; cw, cell wall; e, embryo; h, hyaline layer; ii, inner integument; m, micropyle; mu, mucilage; oi, outer integument; pe, peripheral endosperm (aleurone layer); ps, pigment strand; s, suspensor; vb, vascular bundle.

Bar = 40  $\mu\text{m}$  in A and 7  $\mu\text{m}$  in B.

(Adapted from Debeaujon et al., 2007).

## 1.2 Flavonoids in *Arabidopsis* seed

Flavonoids are plants secondary metabolites derived from the phenylpropanoid pathway. Most of the polyphenolic compounds, namely hydroxycinnamates, coumarins, lignins and flavonoids are derived from phenylalanine. Flavonoid structure is characterized by two aromatic cycles (A- and B- rings) linked by a heterocycle (C-ring) (see Fig. 1.2). At this moment there are more than 6500 molecules isolated and grouped in several classes according to the oxidation degree of the C-ring. These classes include flavonols, anthocyanins and flavan-3-ols. Additional modifications, including hydroxylations, methylations, glycosylations, acylations or prenylations account for the diversity within a compound class (Pourcel, 2006; Routaboul et al., 2006; Winkel-Shirley, 2006). The condensation of flavan-3-ols results in the formation of proanthocyanidins (Dixon et al., 2002; Marles et al., 2003; Pourcel et al., 2005).

Two classes of flavonoids, namely flavonols and proanthocyanidins are synthesized and accumulate in a tissue-specific manner in *Arabidopsis* seeds (Routaboul et al., 2006). Before ovule fertilization, cells of the endothelium layer are characterized by a dense cytoplasm. Rapid vacuolization occurs after fertilization and colorless tannin precursors can be detected (Devic et al., 1999). PA biosynthesis starts around 1-2 days after flowering (DAF). First, they accumulate in vacuoles of the endothelium cells as colorless compounds until 6-7 DAF (Debeaujon et al., 2003). At the beginning of seed desiccation (10 DAF) PAs stored in the vacuoles migrate to the cell walls, where oxidation occurs, leading to formation of brown pigments

contributing to the color of wild-type *Arabidopsis* seeds. Flavonols are found in the endosperm and embryo, however in the testa they are essentially accumulated in the oil layer. In mature seeds, flavonols are present mainly as glycoside derivatives with quercetin-3-*O*-rhamnoside being the most abundant (Pourcel et al., 2005; Routaboul et al., 2006). Interestingly, natural variation in the quantity of PAs and flavonols occurs among *Arabidopsis* accessions (Lepiniec et al., 2006; Routaboul et al., 2006), which might reflect an important trait of plant adaptation to different habitats. Flavonoid composition of mature *Arabidopsis* seeds has been recently described in details by Routaboul et al., 2006.



Figure 1.2: The structure of the flavonol quercetin.

Quercetin is given as an example of carbon numbering. Important features influencing antioxidant potential are the di-hydroxylated B-ring, unsaturation and a 4-oxo function at the C-ring (see Williams et al., 2004). (Adapted from Pourcel et al., 2007).

### 1.2.1 Biosynthesis

A majority of the genes involved in flavonoid biosynthesis were identified through visual screenings of various collections of *Arabidopsis* mutants for altered seed coat pigmentation (Koornneef, 1990; Shirley et al., 1995; Lepiniec et al., 2006). Many mutants have been called *transparent testa* (*glabra*) for the altered pigmentation of the seed coat. One of the mutants, *banyuls* (*ban*) is peculiar, because in place of PAs it accumulates anthocyanins. The *transparent testa 10* (*tt10*) mutant is also unique in the fact that it does not affect the biosynthesis of PAs but their subsequent oxidative browning (Pourcel et al., 2005). It is important to note that six of the *tt* mutants are encoding transcription factors which are involved in the flavonoid biosynthetic genes regulation as well as they are required for normal seed development and cell differentiation. Flavonoid metabolism in seeds and mutants in genes affecting this metabolism have been recently reviewed in Lepiniec et al. (2006) and are summarized in Figure. 1.3.

The *TT10/LAC15* gene encodes a polyphenol oxidase which belongs to a laccase multigene family of 17 members (Pourcel et al., 2005). *TT10/LAC15* has been shown to be involved in the formation of epicatechin quinones that spontaneously polymerize into brown derivatives. Moreover, it may catalyze the oxidative browning of colorless PAs, which is consistent with the fact that the *tt10* mutant accumulates wild-type PA levels but yellow seeds at harvest. *TT10/LAC15* may also be involved in the formation of biflavonols from quercetin-3-*O*-rhamnoside (see Fig. 1.3 and Fig. 1.5). Interestingly, natural variation in *TT10* transcript accumulation occurs between *Arabidopsis* accessions (Pourcel et al., 2005).



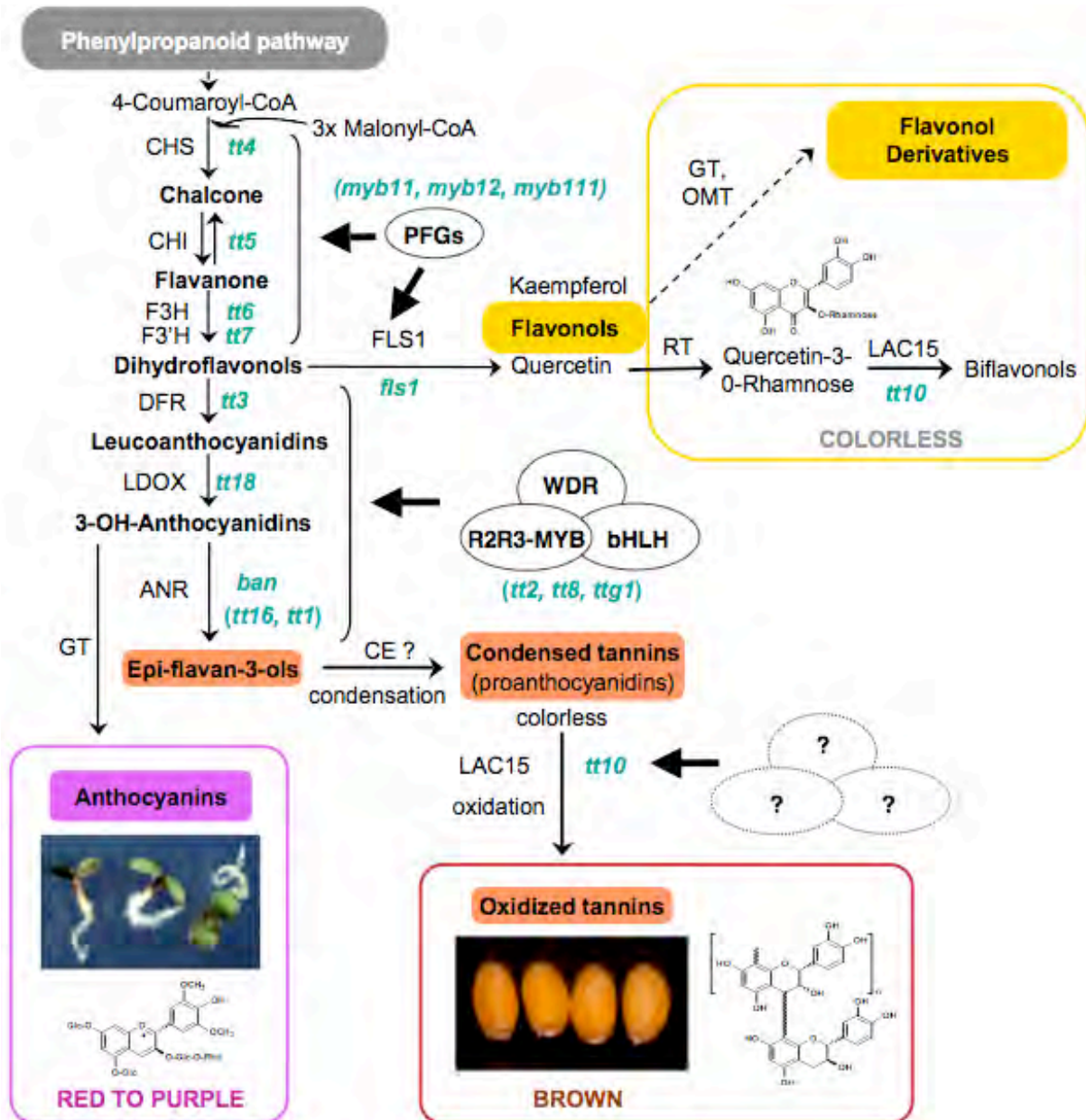


Figure 1.3: Flavonoid biosynthesis pathway in Arabidopsis.

Three major flavonoid end-products are formed: colorless proanthocyanidins (condensed tannins), which brownish upon oxidation by the TT10 / LAC15 laccase and are seed-coat specific, flavonol derivatives (colorless) which are ubiquitous and can be dimerized by TT10 to form biflavonols, and anthocyanins (purple), which are only in vegetative parts. Enzymes are presented in capital letters. Regulatory factors are in bold capital letters and circled. Corresponding mutants are in italics green. Abbreviations: ANR, anthocyanidin reductase; CE, condensing enzyme; CHI, chalcone isomerase; CHS, chalcone synthase; DFR, dihydroflavonol reductase; F3H, flavonol 3-hydroxylase; F3'H, flavonol 3'-hydroxylase; FLS, flavonol synthase; GT, glycosyltransferase; LAC, laccase; LDOX, leucoanthocyanidin dioxygenase; OMT, O-methyltransferase; PFG, production of flavonol glycosides; RT, rhamnosyltransferase; *tt*, transparent testa; *ttg*, *tt glabra*. (Adapted from Lepiniec et al., 2006; Pourcel, 2006 and Stracke et al., 2007).

### 1.2.2 Biological roles of flavonoids

Flavonoids are important in many aspects of plant life in relation to interactions with environment. It has been associated with resistance to pathogens and herbivores (Dixon and Paiva, 1995; Dixon et al., 2002). It seems that flavonoids can also contribute to seed longevity by providing protection against various biotic and abiotic stresses (Debeaujon et al., 2000; Winkel-Shirley, 2002; Rajjou et al., 2008). They can also reinforce testa-imposed dormancy by decreasing seed coat permeability (Debeaujon et al., 2000). Flavonols have been shown to act as signaling molecules for bacteria forming nitrogen-fixing nodules. They are also required for pollen tube germination (Taylor and Grotewold, 2005). Among other functions of flavonoids it is important to point their antioxidant and antimicrobial properties which are beneficial not only for plants but also for human and animal health (Scalbert et al., 2005; Selmi et al., 2006; Aron and Kennedy, 2008). However, some can also have a negative impact on the industrial use of some crop seeds (e.g rapeseed), or on protein digestibility and absorption from seed meal (Winkel-Shirley, 2001a; Marles et al., 2003; Dixon et al., 2005).

### 1.2.3 Oxidation of end-products

The protective function of flavonoids against biotic and abiotic stresses can be related to their cytotoxicity and antioxidant abilities. It seems that flavonoid oxidation contributes to these physiological and chemical properties. The oxidation of colorless flavan-3-ols and proanthocyanidins reinforces their cross-linking to the cell wall and leads to the formation of brown seed coat pigments, whereas oxidation of flavonols results in formation of antifungal agents. The topic of flavonoid oxidation in plants has been recently reviewed in Pourcel et al. (2007).

Flavonoid oxidation *in planta* is mainly mediated enzymatically by polyphenoloxidases (PPOs) and peroxidases, however it can also occur spontaneously (see Fig. 1.4 and Fig. 1.5). PPOs and peroxidases have been shown to be expressed during normal plant development, however they can be also induced by various environmental causes, like pathogen attacks, wounding or desiccation (Mayer and Staples, 2002; Thipyapong and Steffens, 1997; Thipyapong et al., 2004; Mayer, 2006). Physiological role of those genes seems to be regulated by differential subcellular localization of enzymes and substrates as well as by transcriptional and post-transcriptional events during gene expression.

Browning of *Arabidopsis* seed coat is mainly due to the oxidation of epicatechin and soluble PAs by *TT10* laccase (Pourcel et al., 2005). The *tt10* mutant displays a delay in seed coat browning, an increase in soluble (non-oxidized) PAs and a default in the formation of biflavonols from quercetin-3-*O*-rhamnoside. The pattern of *TT10* promoter activity analyzed with the GUS (see Fig. 1.6 A.) and GFP reporter genes as well as *in planta* TT10 activity (see Fig. 1.6 D.) showed that *TT10* colocalizes with tannin- and flavonol-accumulating cell layers in the seed coat (see Fig. 1.6 and Pourcel et al. (2005)). Recent results confirmed that *TT10* mRNA and also protein were localized in endothelium and outer integument 1 cell layer (see Fig. 1.6 B. and C.).

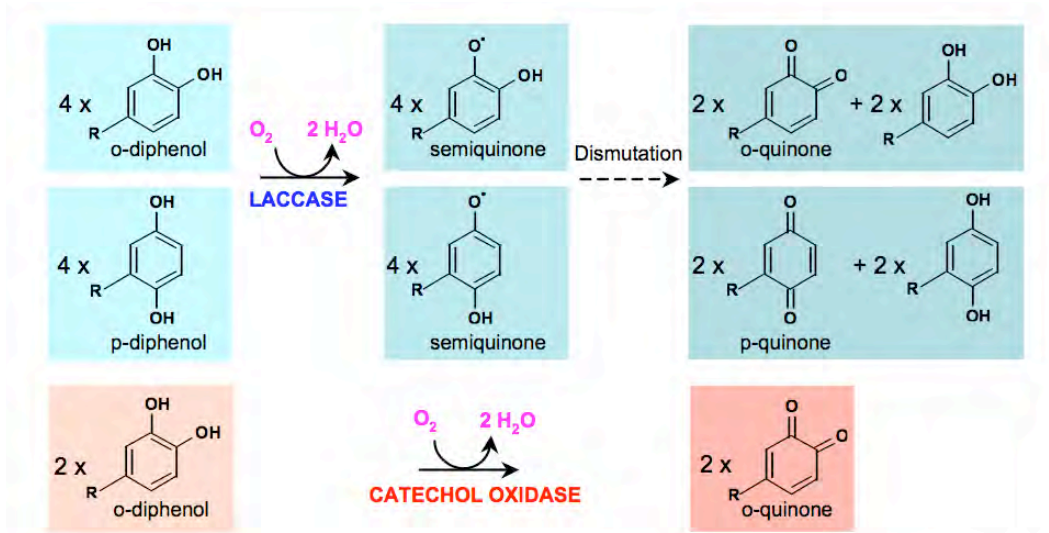


Figure 1.4: Enzymatic reactions for polyphenoloxidases (laccase, catechol oxidase). Polyphenoloxidases (PPOs) are oxidoreductase enzymes acting on diphenols in presence of molecular oxygen ( $O_2$ ). Laccase (LAC) is able to oxidize both ortho- and para-diphenols. Catechol oxidase (CO) is able to oxidize ortho-diphenols. (Adapted from Pourcel et al., 2007).

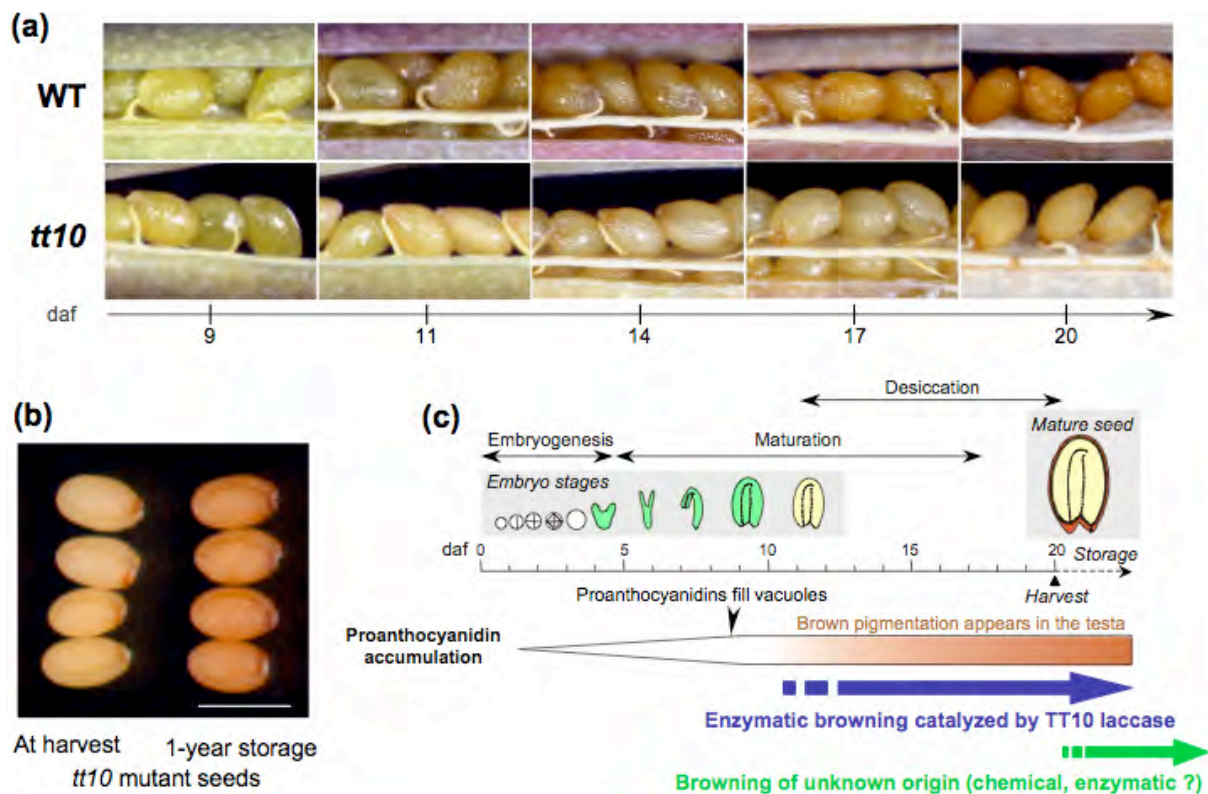


Figure 1.5: Seed coat pigmentation in *Arabidopsis*: illustration of a browning process. (a) Pictures showing the apparition of a brown pigmentation in the testa of the wild-type genotype (WT) during seed desiccation. The brown pigment is absent from the *transparent testa 10* (*tt10*) mutant defective for a laccase enzyme. (b) Mutant *tt10* seeds slowly get brown after harvest until resembling WT seeds. Bar = 550  $\mu\text{m}$ . (c) Schematical drawing pointing to the occurrence of brown pigmentation during Arabidopsis seed development. daf, days after flowering. (Adapted from Pourcel et al., 2007).

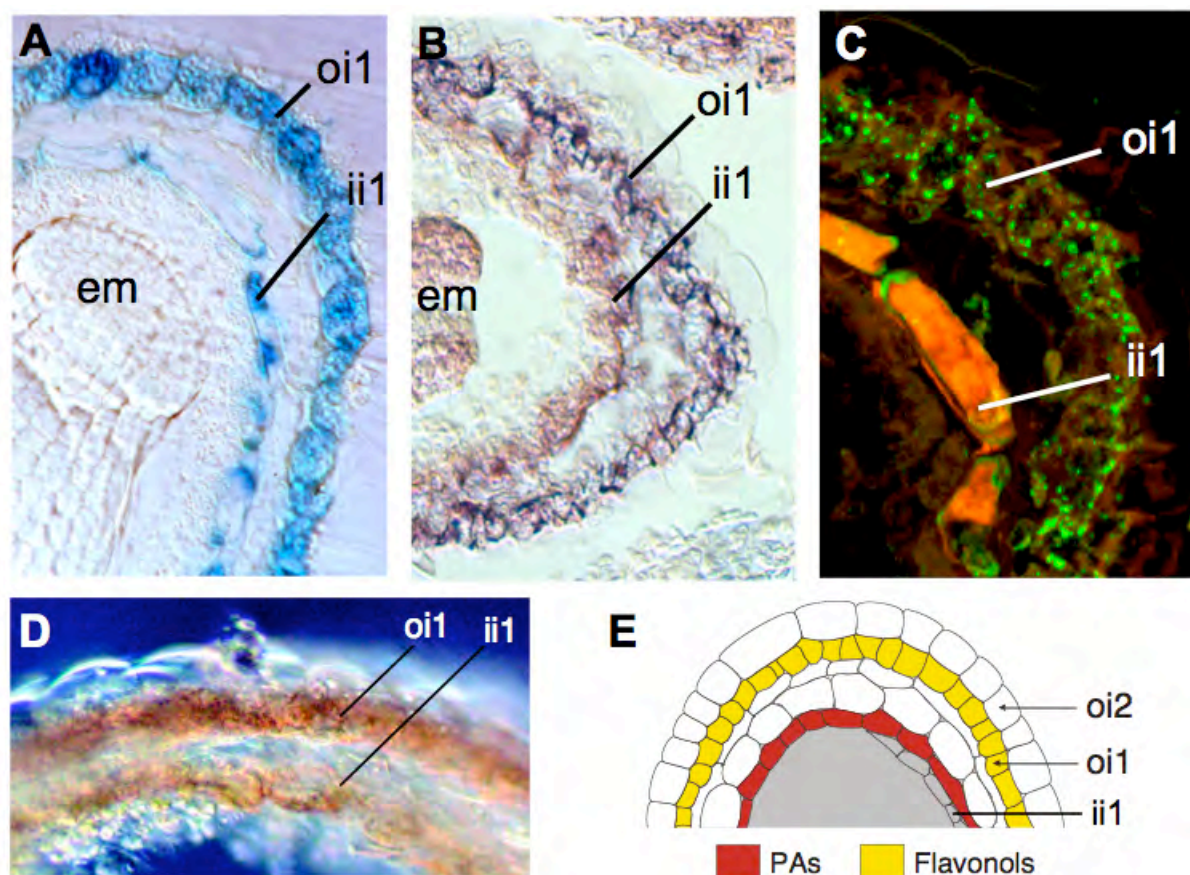


Figure 1.6: Seed coat sections showing the co-localization of *TT10* promoter activity, mRNA expression and protein activity with flavonoid substrates (5-daf seeds).

(A) Promoter activity (*pTT10:uidA*); (B) In situ hybridization with *TT10* cDNA probe; (C) Immunolocalization of TT10 protein with anti-HA antibody (*pTT10:TT10:3xHA*); (D) Protein oxidative browning activity in presence of epicatechin substrate; (E) Schematic representation of the integumentary structure showing flavonoid localization.

B. and C. are unpublished data from Lucille Pourcel. A., D. and E. are from Pourcel et al., 2005.

Abbreviations: em, embryo; ii, inner integument; oi, outer integument; PAs, proanthocyanidins; HA - hemagglutinin.

### 1.2.3.1 Laccases

Laccases (LAC), i.e. o- and p-diphenol:dioxygen oxidoreductases (EC 1.10.3.2) belong to a larger group of enzymes called multicopper or blue copper oxidases, which includes ascorbic acid oxidase and ceruloplasmin (reviewed in Mayer and Staples, 2002). These glycoproteins are characterized by four histidin-rich copper binding domains and ability to oxidize phenolic substrates in the presence of molecular oxygen (see Fig. 1.4). Also polyphenol oxidases of the catechol oxidase type (EC 1.10.3.1) require molecular oxygen for their activity, what makes them distinct from peroxidases, which require hydrogen peroxide (Pourcel et al., 2007).

Laccases have been found in eukaryotes (fungi, plants and insects) as well as in prokaryotes (Claus, 2004; Riva, 2006). The most studied are the fungal laccases, which are involved in spore pigmentation, virulence and delignification (Mayer and Staples, 2002; Pourcel et al., 2007). Plant laccases are less studied and until now were mainly associated with lignification, even if a direct activity of the enzyme on monolignol polymerization within the cell matrix has not yet been demonstrated (Mayer and Staples, 2002). Plant laccases were also shown to be involved in wound healing, iron metabolism detoxification and browning

reactions (Pourcel et al., 2007 and references therein).

### 1.2.3.2 Catechol oxidases

Catechol oxidases (CO) are enzymes catalyzing the oxidation of monophenols to o-diphenols as well as they can oxidize o-diphenols to the corresponding o-quinones (Marusek et al., 2006) (see Fig. 1.4). Like laccases, COs are glycosylated copper-binding proteins, that require molecular oxygen for their activity. They were found in eukaryotes and prokaryotes (Marusek et al., 2006; Mayer, 2006). Interestingly, the physiological roles of CO seem to be very close to those reported for laccases. It is not surprising as both groups of enzymes are involved in oxidative browning which in consequence can contribute to protection against UV radiations, and also senescence, wound healing and seed coat hardening (Pourcel et al., 2007 and references therein).

Interestingly, it seems that the *Arabidopsis* genome does not contain any typical catechol oxidase (Pourcel et al., 2005). Their lack could have been filled by laccases, in the sense of a specialization toward o-diphenol oxidation.

## 1.3 Regulation of gene expression

Regulation of gene expression includes all the processes leading to formation of the gene product, RNA and protein. All the steps of gene expression may be modulated, from gene transcription to the post-translational modifications of a protein. However, expression of many genes is regulated at the transcriptional level and depends on the combinatorial contribution of transcription factors, which can act as activators or repressors. Gene expression regulated at the transcriptional level results in precise spatio-temporal manner activity of their promoter. Below, expression patterns of flavonoid-related genes are given as an example.

### 1.3.1 Regulation of *BAN* and *MYBL2* expression

Regulation of *BAN* (Debeaujon et al., 2003) and *MYBL2* (Dubos et al., 2008) differs at the level of the spatial activities of their promoter in the seed coat: *BAN* is present in one integumentary cell layer (ii1) and *MYBL2* is present in two cell layers (ii1 and oi1).

The entry step into the proanthocyanidin branch of flavonoid biosynthesis is formation of epi-flavan-3-ols. This reaction is metabolized by anthocyanidin reductase encoded by *BANYULS* (*BAN*) (see Fig. 1.3). *BAN* promoter activity was detected specifically in PA-accumulating cells in the the inner integument 1 layer (endothelium) of the seed coat and pigment strand in the chalaza zone (see Fig. 1.7 G. and Debeaujon et al. (2003). The *BAN* promoter deletion and gain-of-function experiments led to identification of 86-bp promoter region which functions as an enhancer specific for PA-accumulating cells. Studies of the *BAN* promoter activity in regulatory mutant background revealed that its activity was abolished in *tt2*, *tt8* and *ttg1* (Debeaujon et al., 2003). The spatial pattern of *BAN* promoter activity was also modified in *tt1* and *tt16*, but not in *ttg2* (Debeaujon et al., 2003). Further studies revealed that TT2/TT8/TTG1 regulatory complex binding to the enhancer element in *BAN* promoter is required for its expression (Baudry et al., 2004).

Recently, a MYB family transcription factor, MYBL2 has been shown to regulate the transcription of genes involved in anthocyanin biosynthesis (Dubos et al., 2008). Although *MYBL2* promoter activity is detected in vegetative tissues, in seeds its activity is restricted to the endothelium and epidermis (see Fig. 1.7 H. and Dubos et al., 2008). *Cis*-acting regulatory elements in promoter region and transcription factors required for MYBL2 expression are yet unknown.

### 1.3.2 Two-cell layer expression pattern of *TT10*

*TT10* promoter activity colocalizes first with PA- and flavonol-producing cells of the testa. Activity of *TT10* promoter starts at early stages of seed development. First it is detected in the endothelium and in the pigment strand at the chalaza zone (see Fig. 1.7 A., C., and D.). Later, the activity increases and spreads to the outer integument 1 cell layer of the seed coat (see Fig. 1.7 E., and F.). The *uidA* gene was also strongly expressed in early aborted seeds (see Fig. 1.7 B.). Any difference in GUS activity compared with the wild type was observed in three regulatory mutants *tt2*, *tt8*, and *ttg1* (Pourcel et al., 2005).

### 1.3.3 One layer, two layers, which layers?

It seems that regulation of gene expression in the seed coat is very complex. Genes can be expressed in one or more layers of the testa. Moreover they can be expressed at various developmental stages of the seed coat development. Identification of *cis*-acting regulatory elements required for gene expression in each of the seed coat cell layers or combination of layers would allow to target expression of other genes for crop improvement.

## 1.4 Objectives of the thesis

The main focus of this thesis was the functional analysis of the *TRANSPARENT TESTA 10* gene. The work was following two axes, where developmental and possible stress regulation of *TT10* expression was studied.

The first part of the work dealing with the developmental regulation of *TT10* expression required detailed characterization of promoter activity throughout the plant to complete existing data. Functional analysis of *TT10* promoter was aiming at identification of the promoter region and *cis*-acting regulatory elements required for seed coat specific activity. Identification of the CAREs allowing to drive gene expression separately in endothelium (PA specific) or outer integument 1 (flavonol specific) cell layer, were one of the goals.

In parallel, existing natural variation for *TT10* expression was exploited as it could provide an additional information about regulation of gene expression.

Several complementary approaches, namely studies of *TT10* promoter activity and gene expression level in candidate regulatory mutant background as well as transcription factor library screening were aiming at identification of transcription factors which could directly regulate *TT10* expression.

Second part of the work dealing with possible stress regulation was aiming at identification of the stress conditions in which *TT10* would be expressed. The tools developed in the first part were planned to be used for identification of the promoter region and *cis*-acting regulatory elements involved in the stress regulation of *TT10* expression.

The secondary objectives of this thesis were aiming at identification of other laccases expressed in seeds and manipulations of *Arabidopsis* cell suspensions to study flavonoid metabolism in simplified system.

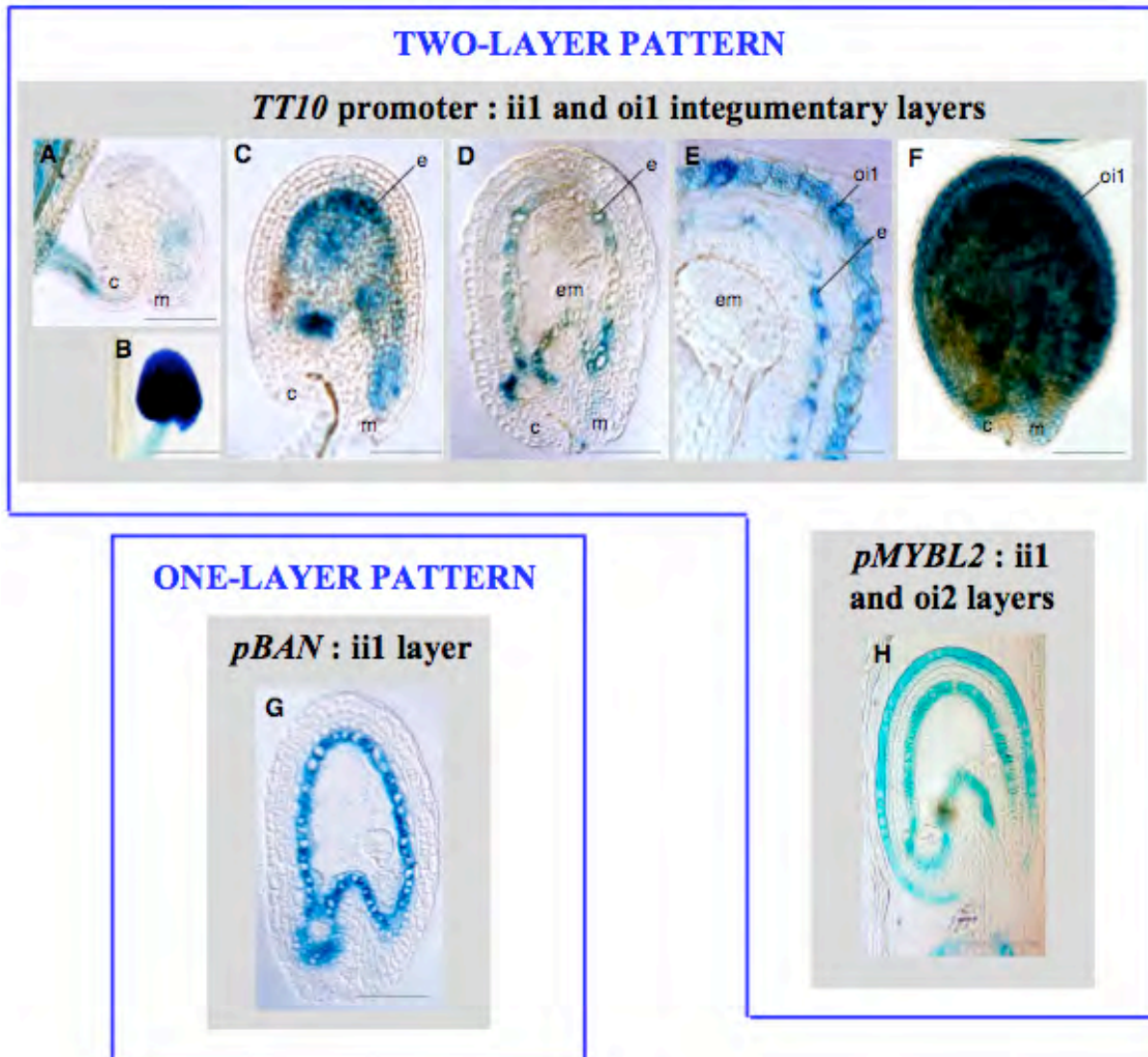


Figure 1.7: Spatial pattern of *TT10* promoter activity in WT seeds. Comparison with *BAN* and *MYBL2* promoter activities.

(A) to (F) Expression of the Pro*TT10*:uidA cassette in developing seeds at 1 DAF A, 3 DAF C and D, 8 DAF E and F, and in an aborted seed B. GUS activity is observed with Nomarski optics on whole mounts for A to C and F, and on sections for D and E.

(G) Pattern of *pBAN* activity in developing seed at 4 DAF.

(H) Pattern of *pMYBL2* activity in developing seed at 4 DAF.

Abbreviations: c, chalaza; em, embryo; e, endothelium; ii, inner integument; m, micropyle; oi, outer integument. Bars = 24  $\mu\text{m}$  for F and K; 70  $\mu\text{m}$  for A to D, G and H.

Adapted from Pourcel et al., 2005: A to F; Debeaujon et al., 2003: G and Dubos et al., 2008: H.

# Chapter 2

## Results

The main objective of my work was to study the regulation of *TT10* gene expression. The results of experiments, techniques and approaches, concerning different aspects of gene expression, which have been used to shed light on the mechanism of *TT10* expression, are described in details in following sections: *TT10* expression and promoter studies (section: 2.1), techniques aiming at finding transcription factors involved (section: 2.2), insights from natural variation occurring for *TT10* expression (section: 2.3) and potential stress response (section: 2.4). Each section consists of a short introduction, presentation of the results and technical discussion of the approach used.

Two secondary objectives of this work were: I) to identify other laccases expressed in seeds (section: 2.5) and II) to establish conditions and optimize techniques for the manipulation of *Arabidopsis* cell suspensions as a tool to study the flavonoid metabolism (section: 2.6).

### 2.1 Transcriptional regulation of *TT10* expression

This section is focusing on *TT10* expression during plant development and detailed characterization of *TT10* promoter. At first, our analysis of *TT10* transcript accumulation is presented and compared to publicly available expression data (2.1.1). Then the promoter activity is studied during plant development, using *uidA* reporter gene (2.1.3). *In silico* analysis (2.1.2) of the promoter sequence was used as a basis to design the 5' dissection of the promoter (2.1.3.1), and its site-directed mutagenesis (2.1.3.4). *In silico* analysis of the promoter sequence is supported by literature search, documenting on the functionality of *cis*-acting regulatory elements (CARE).

#### 2.1.1 *TT10* transcript accumulation during plant development

##### 2.1.1.1 Transcript accumulation - RT-PCR

The expression pattern of *TT10* was previously investigated by RT-PCR (McCaig et al., 2005; Cai et al., 2006) and semiquantitative RT-PCR (Pourcel et al., 2005). In those studies *TT10* transcript was detected to be: i) in all the plant organs except stem (McCaig et al., 2005), ii) predominantly in developing silique, at low level in stem, seedling and flower but not in root (Pourcel et al., 2005), iii) in silique, root, flower and stem but not in leaf (Cai et al., 2006). Here, we wanted to revisit and complete previously known and sometimes contradictory data using the very sensitive quantitative RT-PCR method, with special attention to silique development. More robust and precise data about *TT10* expression were required for the analysis of its regulation. Quantitative RT-PCR had also been chosen to study differences in *TT10* transcript



accumulation between accessions (see: 2.3.1) as well as in candidate regulatory mutant backgrounds (see: 2.2.3, 2.2.4 and 2.2.5).

The expression pattern was investigated in various tissues of 6-week-old Col-0 plants grown in long day conditions in the greenhouse (Fig. 2.1 A.). The *TT10* mRNA was detectable as early as the open-flower stage. Detailed analysis of transcript accumulation in developing siliques, revealed a huge increase between 5 and 7 days after flowering (DAF) followed by a slight decrease at 10 DAF, and a second peak of transcript accumulation around 15 DAF. In older siliques the amount of *TT10* mRNA decreases and is close to 0% of *ELONGATION FACTOR 1aA4* (*EF1*) control at 21 DAF, but is still detectable. *TT10* is highly expressed in developing seeds, reaching up to 350 %EF at 5 DAF and 250 %EF at 10 DAF (Fig. 2.1 B.). *TT10* transcript was not detected in dry seeds, cauline and rosette leaves, stem at the lower part and terminal node.

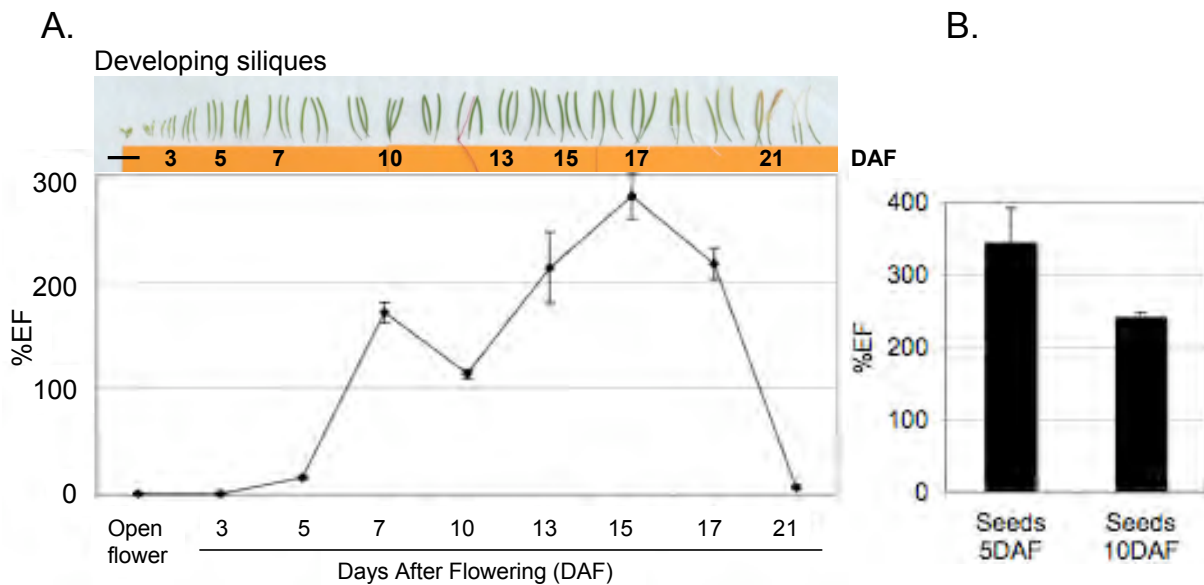


Figure 2.1: Expression pattern of the *TT10* gene in developing siliques and seeds of wild-type Col-0 plants. *TT10* gene expression in wild-type Col-0 plants was detected by quantitative RT-PCR and is presented as a percentage of the expression of the reference gene *ELONGATION FACTOR 1aA4* (*EF1*). For details see 3.2.5.

(A) Picture illustrates siliques development and samples used for RNA extraction. Graph is representing one out of two biological repeats of the experiment.

(B) *TT10* gene expression in extracted seeds.

Values represent averages  $\pm$ SE of three technical replicates. Bar = 2 cm.

### 2.1.1.2 Transcript accumulation - Transcriptomics

Publicly available transcriptomic data are a constantly growing resource, therefore we checked (last update 7.04.2009), if new combinations of tissue/developmental stage/treatment have revealed new information about *TT10* expression. Genevestigator, Bio-Array Resource (BAR), Gene Co-expression Analysis Toolbox (GeneCAT), DIURNAL and several other analysis and visualization tools were used (see section 3.2.9). For many functions those web-based tools are redundant, but for some analysis one is more adapted than the other. Below (Fig. 2.2) as an example, analysis of *TT10* expression with the eFP Browser (at BAR) is presented. This tool is well adapted to gene expression analysis thanks to easily readable color code and graphic representation of plant organs/tissues for which the data are presented. Fast switch between data sets (e.g. developmental/stress), standard deviation and threshold filtering as well as easily accessible expression

values for more detailed analysis are making eFP Browser a very convenient tool for gene expression studies. The *TT10* expression analyzed and visualized with the eFP Browser confirms that it is mainly expressed in developing seed/silique but also shows expression in pollen grain (Fig. 2.2).

- This probe set reaches its maximum expression level (expression potential) of **3258.06** in the **Developmental Map or Tissue** : source.

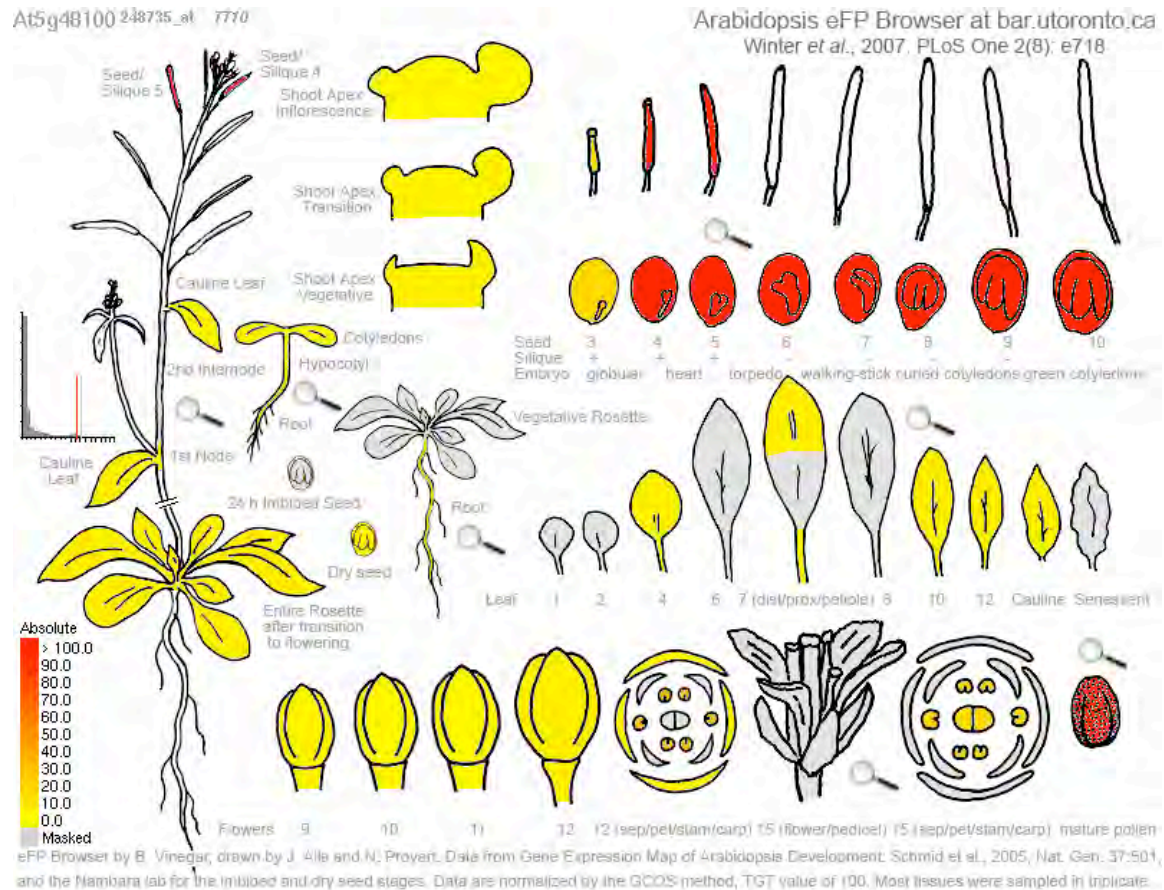


Figure 2.2: *TT10* expression visualized with eFP Browser.

eFP Browser is presenting gene expression as a heat map placed on the graphic representation of the *Arabidopsis* organs - Developmental Map - AtGenExpress Developmental Set (Schmid et al., 2005). Standard deviation filtering has been used to mask samples with the deviation greater than half of their expression value (grey). Signal threshold filtering has been set at 100 (expression potential) to eliminate samples where expression is considered to be below the detection level (yellow). Data are not available for separate siliques (white) from torpedo stage of embryo development onwards, where only dissected seeds were analyzed. Developing seed, silique and pollen where *TT10* gene expression potential is > 100 (detectable) are painted in red. For details see: <http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi> and Winter et al. (2007).

Figure 2.3 is presenting comparison of *TT10* and *BAN* expression, which is nicely visualized thanks to Expression Profiling Tool available at GeneCAT. This tool allows comparing the expression of multiple genes in different plant tissues. *BAN* is a core enzyme of the subpathway of flavonoid biosynthesis leading to proanthocyanidins, which are later oxidized by *TT10* (Debeaujon et al., 2003; Pourcel et al., 2005). The GeneCAT analysis shows two peaks of *TT10* expression in developing seeds/siliques as well as weak signal in pollen and even weaker in flower and root. From the GeneCAT analysis *BAN* expression seems to be seed/silique specific, with a peak of expression at an earlier stages than *TT10*, what corresponds to the RT-PCR results presented in Debeaujon et al. (2003).

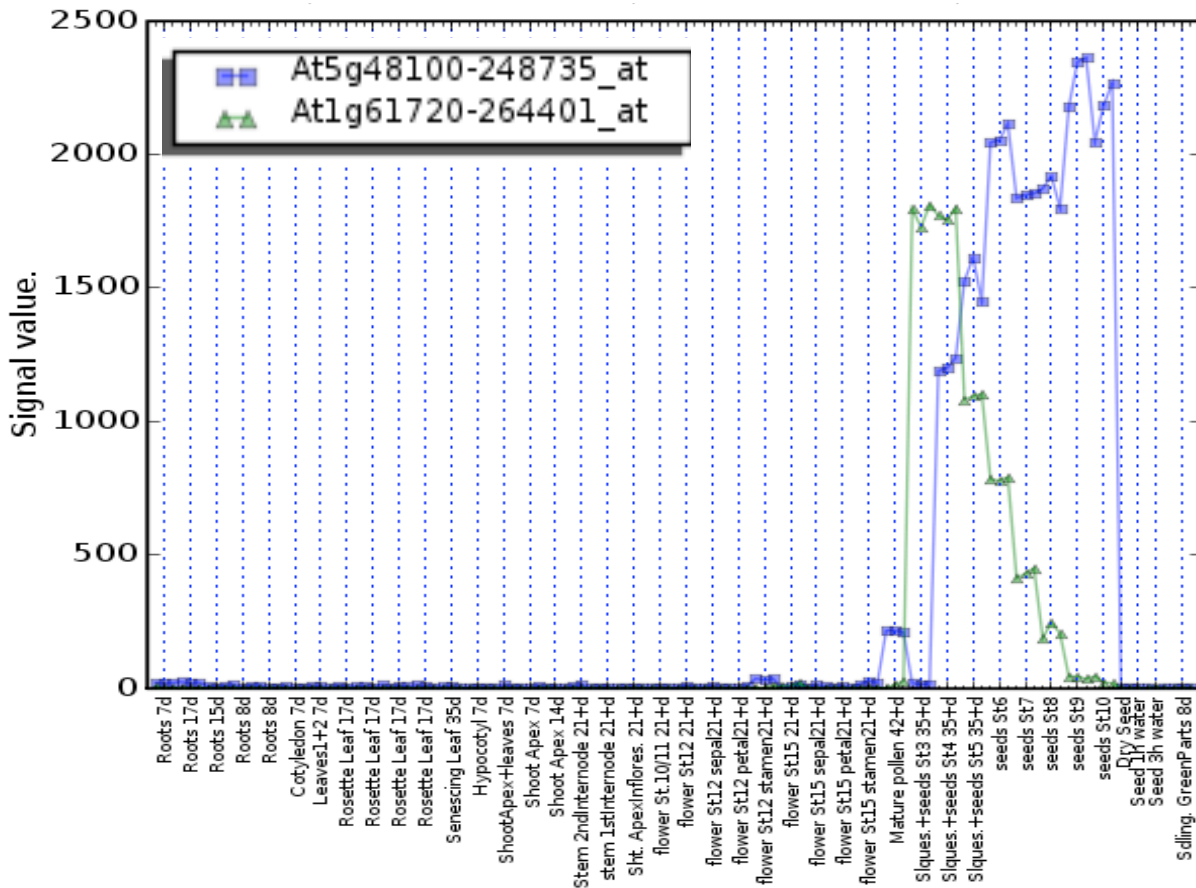


Figure 2.3: Comparison of *TT10* and *BAN* expression visualized with GeneCAT expression profiling tool. Graph presents expression level and tissues in which *TT10* (closed squares - blue) and *BAN* (triangles - green) are expressed. Expression profiling tool uses RMA normalized data from 121 ATH1 microarrays generated during AtGenExpress project. For details see: <http://genecat.mpg.de/cgi-bin/Ainitiator.py> and Mutwil et al. (2008).

Figure 2.4 is presenting the DIURNAL tool which was used to analyze if *TT10* expression could be affected by the circadian rhythm/photoperiod. Interestingly, expression of the key enzyme of flavonoid biosynthesis, encoded by *CHALCONE SYNTHASE* (*CHS*) is regulated by light and is presented as an example of gene expression affected by the circadian rhythm (Thain et al., 2002). We can clearly see that its transcript is accumulating during the day, whereas it is much lower during the night (Fig. 2.4). The signal values for *TT10* in those experiments are very low, therefore it is not possible to conclude if photoperiod is affecting its expression.

### 2.1.1.3 Analysis of publicly available transcriptomic data for *TT10* expression during plant development - Summary

Taken together all the transcriptomic data analyzed, it seems that *TT10* accumulation is significant only in developing silique and seed. In some experiments, *TT10* is on limit of detection level (pollen and root), however there are limited repetitions available and the standard variation is high for those data which should therefore be treated with caution. *TT10* expression is below the detection level in the circadian/photoperiod experiments.

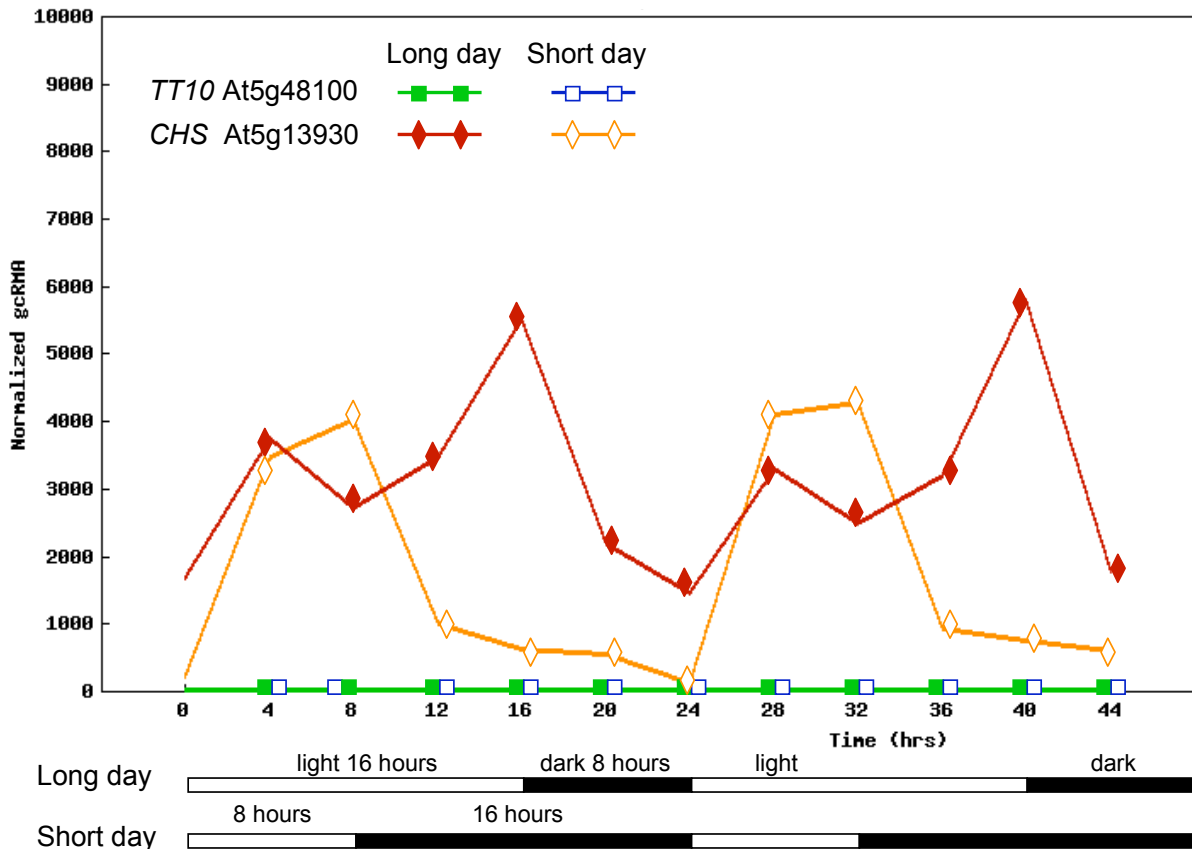


Figure 2.4: Comparison of *TT10* and *CHS* expression visualized with DIURNAL.

*TT10* and *CHS* expression in long and short day is presented as a graphical output adapted from the DIURNAL gene expression analysis tool. *TT10*, closed green and open blue squares for long and short days respectively. *CHS*, closed red and open orange diamond/rhombus for long and short days respectively. Samples, 7-day-old seedlings were grown in MS agar medium supplemented with 3% sucrose. For details see: [http://diurnal.cgrb.oregonstate.edu/diurnal\\_details.html](http://diurnal.cgrb.oregonstate.edu/diurnal_details.html) and Mockler et al. (2007).

### 2.1.2 *In silico* analysis of *TT10* promoter sequence

Several web-based computer programs (see 3.2.9) recognizing known *Cis*-Acting Regulatory Elements (CAREs) have been used for the *in silico* analysis of the 2.0-kb *TT10* promoter sequence Col-0, including the 5' UTR (for promoter sequence used in the study see Appendix, 5.1). Detection of CAREs in the promoter by so-called 'search by signal' methods are prone to give false positive results (Rombauts et al., 2003). For that reason the results obtained should be treated as putative *cis*-acting regulatory elements and their functionality should be tested *in vivo*. In order to select the best candidates for further analysis, literature mining accompanied *in silico* analysis to find out the context and organization of the motifs in the promoter.

PLACE SignalScan server was the preferred tool as it is searching for the presence of motifs identical or similar to the previously reported *cis*-acting regulatory elements in the PLACE database, which presently seems to be the richest with 469 entries of CAREs deposited (PLACE 30.0, 469 entries, Jan.8, 2007). The full list of putative *cis*-acting regulatory elements found in the *TT10* promoter is presented in the Appendix, Table 5.1. Similar results were obtained when querying PlantCARE, AtcisDB at Agris (see Appendix Tab. 5.2), ATHENA and PlantPan. In addition, *in silico* analysis revealed no tandem repeats regions, nor CpNpG islands in -1000bp - + 500bp of the At5g48100 locus (PlantPan) and no over-represented TF binding site has been found (ATHENA).

It is important to note that *TT10* is not sharing its promoter sequence with any known or predicted gene and that the preceding gene (*At5g48110*) is transcribed from the same strand. The region considered as *TT10* promoter is the 3' region for that gene (Fig. 2.5). In such a case any changes arising in the promoter region on the course of the evolution would probably affect essentially *TT10* expression.

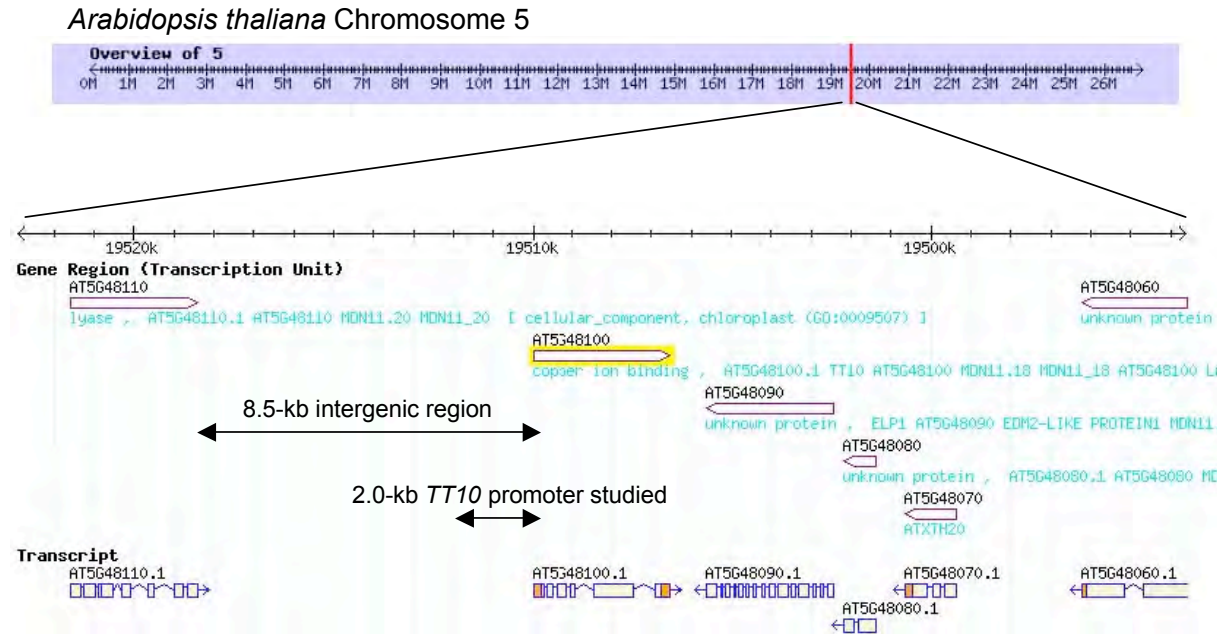


Figure 2.5: *TT10* genomic region and neighboring genes.

The organization of the *TT10* genomic region on chromosome 5 (top, blue rectangle) is presented. Transcription units are represented with arrows indicating the orientation of the transcripts. Bottom part is showing the gene structure where rectangles are representing exons and lines introns. UTR's, when present are shaded orange. Figure adapted from ATIDB.

### 2.1.2.1 Putative *cis*-acting regulatory elements found in the *TT10* promoter

Studies on yeast promoters suggest that regulatory elements are commonly present in the 500bp upstream region of the transcription initiation site (Caselle et al., 2002). Many plant genes have been shown to contain regulatory elements required for the tissue specific gene expression in the proximal part of their promoters (Abe et al., 2003; Debeaujon et al., 2003; Nakashima et al., 2006; Zhou et al., 2009). For that reason special attention has been taken to analyze this part of the *TT10* promoter in details as well as to focus on the dissection of the proximal part of the promoter (see 2.1.3.1). Selected putative *cis*-regulatory elements found in that part of the promoter are presented in Figure 2.6 and described below. Several criterions have been used to limit the list of putative CAREs considered: 1.) tandem repeats and length of the motif (longer regarded as more significant), 2.) literature evidence for being involved in seed development and/or defense/stress response (associated with putative function of laccases and flavonoids), and/or 3.) motifs similar to the one previously reported to be involved in flavonoid biosynthetic genes regulation.

Motifs are briefly described, moreover examples of transcription factors are given and briefly discussed in the context of the putative *cis*-acting regulatory elements.

### 2.1.2.2 Putative CAREs associated with seed development, stress response and/or protection against various biotic and abiotic stresses

In *Arabidopsis* seeds, flavonoids are accumulated and they are thought to provide protection against various environmental challenges and ensure embryo survival and seed germinability (Rajjou and Debeaujon, 2008). In particular, flavonoid oxidation triggered by *TT10* could be required for protection of the embryo against environmental challenges (Pourcel et al., 2005, 2007). Functions associated to laccases in plants are multiple. Apart from lignification, they are also thought to be involved in wound healing, plant defense against predators, herbivores and invasion by bacteria and fungi (Mayer and Staples, 2002). Plant hormones e.g. ABA, GA, ethylene, salicylic and jasmonic acids (SA and JA respectively) are involved not only in stress-induced gene expression, but also in plant and seed development and in cross-talk between those processes (Fujita et al., 2006; Chung et al., 2008; Pauwels et al., 2008, 2009; Santner et al., 2009). Taking in account that *TT10* encodes a laccase which is involved in flavonoid oxidation, *cis*-acting regulatory elements could be common to genes involved in both pathways and/or have multiple functions, i.e. motifs found in ABA and dehydration responsive genes and motifs in genes involved in seed development where their expression is mediated by ABA (Simpson et al., 2003; Borisjuk et al., 2004; Nakashima et al., 2006; Cao et al., 2007; Chinnusamy et al., 2008).

The putative CAREs for *TT10* regulation listed below were previously reported to be involved either in seed development or stress response.

- **W-box** - is a binding motif for WRKY transcription factors with the core sequence (T)TGAC(Y). Most of the studied WRKY TF appear to have a binding preference for the W-box, although the specificity is partly dependent on the sequence flanking the core motif (Ciolkowski et al., 2008). This group of TF seems to be unique to plants and it is involved in various physiological responses including pathogen defense, senescence and development (Eulgem et al., 2000). Recently WRKY TF have been also found in the protist *Giardia lamblia* and the slime mold *Dictyostelium discoideum*, which could indicate their earlier origin (Ulker and Somssich, 2004; Pandey and Somssich, 2009). It seems that they can act both as transcription activators and repressors. In rice, WRKY transcription factors have been shown to be transcriptional repressors of the gibberellin signaling pathway (Zhang et al., 2004), positive and negative regulators of ABA signaling (Xie et al., 2005), whereas in *Arabidopsis* they have been shown to be involved in salicylic acid signaling and *NONEXPRESSER OF PR GENES 1 (NPR1)* gene expression (Yu et al., 2001). W-boxes are present in promoters of many pathogen-related genes, and some of them are having as many as eleven W-boxes (Chen and Chen, 2002; Eulgem and Somssich, 2007). In the *TT10* promoter, several W-boxes have been found, but as the sequence outside of the core motif is also important, it is difficult to conclude about their functionality. It should be noted that one of the *TT* genes namely *TTG2*, is a WRKY transcription factor (AtWRKY44) involved in seed coat development and trichome formation in *Arabidopsis* (Johnson et al., 2002). It would be interesting to test if *TT10* could be a direct target of TTG2. Therefore we have tested the expression of *TT10* in *ttg2* mutant background (see section 2.2.2).
- **ABRE** (ABA Responsive Element) and **ABRE-like** - there are several overlapping motifs with the core ACGT sequence. They are binding sites for basic leucine zipper (bZIP)-type binding proteins, which are very often transcriptional activators (Foster et al., 1994). In rice, the bZIP transcriptional activator *RITA-1* is highly expressed during seed development (Izawa et al., 1994). ABRE elements act together with dehydration-responsive elements in the *RD29A* gene expression in response to ABA (Nakashima et al., 2006). ABRELATERD1 - ABRE-like (motif: ACGTG) and ACGTATERD1 (motif: ACGT) motifs are required for etiolation-induced expression of *ERD1* gene (Early Response to Dehyd-

ration1) (Simpson et al., 2003). The ABRE motifs in the *TT10* promoter are considered as important, because *TT10* expression could be regulated by ABA during seed development. Indeed when at the last stages of development the seed is desiccating, this could resemble the response to dehydration.

- **TCA1MOTIF** - (motif: TCATCTTCTT) - TCA-1 (Tobacco Nuclear Protein-1) binding site is related to salicylic acid-inducible expression of many genes. It has been found in more than 30 different plant genes which are known to be stress-induced. It was reported as a functional motif in barley, when repeated four times within 200bp (Goldsbrough et al., 1993). In *TT10* promoter it is present only once in the 2.0-kb promoter analyzed, but as it is 10bp long motif, it could be meaningful.

#### Laccase-related motifs:

- **MYB** - Recently it has been shown that AtMYB58 and AtMYB63 are regulating expression of genes involved in lignin biosynthesis by binding to the AC element in their promoter (AC-I: ACCTACC, AC-II: ACCAACC, AC-III: ACCTAAC). Moreover it has been demonstrated that they can directly upregulate the expression of *AtLAC4* but not *AtLAC17* (Zhou et al., 2009). It seems that the MYB family TF and therefore MYB binding motifs are highly important elements for gene expression in secondary metabolism. Perfect AC elements were not found in *TT10* promoter sequence, although in proximal 500bp, AC-I is present once with two mismatches, AC-II - once with one mismatch and AC-III twice with one mismatch.

#### 2.1.2.3 CAREs found in promoters of flavonoid biosynthetic genes

Several classes/families of transcription factors (TF) have been demonstrated to be involved in the regulation of the flavonoid biosynthetic genes, among them are MYB TF: AtMYB11, AtMYB12, AtMYB111 (flavonols; Stracke et al. (2007), AtMYB75/PAP1, AtMYB90/PAP2, AtMYB113, AtMYB114 (Anthocyanins; Borevitz et al. (2000)), AtMYB123/TT2 (Proanthocyanidins; Nesi et al. (2001) and the negative regulators AtMYB4 (Jin et al., 2000) and AtMYBL2 (Dubos et al., 2008). Other classes are represented by the bHLH family, with AtbHLH001/GL3, AtbHLH002/EGL3 (anthocyanins, Zhang et al. (2003) and AtbHLH042/TT8 (anthocyanins and proanthocyanidins; Nesi et al. (2000); Zhang et al. (2003). These bHLH factors are forming protein complexes with already mentioned AtMYBs involved in anthocyanin and proanthocyanidins biosynthesis, and the WDR regulatory protein TTG1 (Walker et al., 1999; Baud et al., 2004). The involvement of the other transcription factors AtWRKY44/TTG2 (Johnson et al., 2002), WIP-type Zn-Finger/TT1 (Sagasser et al., 2002) and MADS AtAGL32 - TT16 (Nesi et al., 2002) is less studied and the direct targets are not yet identified for them (reviewed in Lepiniec et al., 2006).

The motifs found in the *TT10* promoter which could bind transcription factors of the families mentioned are:

- **MYB2** - motif recognized by Signal Scan as MYB1AT (binding motif: WAACCA). MYB recognition site in the promoters of dehydration responsive genes. Originally discovered in the promoter of the *rd22* gene, which is up-regulated in the AtMYB2 / AtMYC2 (see below MYC2) over expressing plants in response to abscisic acid (ABA) under drought stress (Abe et al., 2003).
- **MYC2** - recognized by Signal Scan as MYCATRD22 (motif: CACATG) and MYCCONSENSUSAT (motif: CANNTG). Also described as E-box (binding site: CANNTG) and sometimes overlapping with ABRE (Stalberg et al., 1996). MYC2 is a binding motif for the AtMYC2/rd22BP1 TF belonging to bHLH family (Solano et al., 1995; Abe et al., 1997; Busk and Pagès, 1998). Involvement in cold stress response has also been reported (Chinnusamy et al., 2003).

Those two CAREs were reported to act together in many ABA and dehydration responsive genes (Abe et al., 2003). MYB-bHLH complex had been reported to be involved in the light-regulated activation of the phenylpropanoid genes (Hartmann et al., 2005). MYB2 and MYC2 sites in *TT10* promoter are located approximately at -400bp and -250bp, positions which were found in many genes upregulated in plants overexpressing AtMYC2/AtMYB2 (Abe et al., 2003). It would be interesting to evaluate their functionality.

Several other MYB and MYB-like motifs have been found:

- **GARE MYB** - MYBBAHV (motif: TAACAAA). This motif is an element of the gibberellin response complex of the alpha-amylase gene in barley (Gubler et al., 1995). GARE MYB is overlapping with AMYBOX1 (motif: TAACARA), which is a conserved sequence in the promoters of the alpha-amylase gene of barley (*Hordeum vulgare*), rice (*Oryza sativa*) and wheat (*Triticum aestivum*) (Huang et al., 1990). This motif is also similar to GAREAT (motif: TAACAAR) which has been reported as gibberellin(GA)-responsive element present in 20% of GA inducible genes during seed germination (Ogawa et al., 2003).
- **MYB** - MYBCOREATCYCB1 (motif: AACGG). Motif found in promoter of the *CycB1;1* gene, which is involved in cell cycle regulation. It had been found to be necessary for the binding of the proteins (e.g. putative MYB TF, At2g03470) and able to drive GUS expression in plants overexpressing hypothetical protein, At2g13640 (Planchais et al., 2002).

#### 2.1.2.4 Additional comments about known transcription factors involved in flavonoid biosynthesis

- Pourcel et al. (2005) have shown that AtMYB123/TT2, AtbHLH042/TT8 and WDR/TTG1 regulatory protein, are not required for the activity of the 2.0-kb *TT10* promoter in transgenic plants.
- If the situation is clear for above mentioned TT2, TT8 and TTG1 transcription regulators, it seems to be much more complex for WIP-type Zn-Finger/TT1, MADS AtAGL32/TT16 and AtWRKY44/TTG2. Those genes are not only involved in flavonoid biosynthesis but also in endothelium development. It is in the scope of this study to investigate the involvement of TT1, TT16 and TTG2 in *TT10* expression (discussed in more detail in section 2.2.1; see also: W-box above).

#### 2.1.2.5 Other CAREs

Motifs of lower abundance in *TT10* promoter, shorter binding site or which could be indirectly associated with *TT10* gene function are listed here.

- **DOF** - (motif: AAAG) - is a short motif binding zinc finger for plant specific DOF TF, which are associated with expression of multiple genes involved in carbon metabolism (Yanagisawa, 2000). There are several DOF motifs in proximal part of the *TT10* promoter and surprisingly, most of them (five) are spanned within 150bp region.
- **ARR1** - (motif: NGATT) - ARR1-binding element is a motif recognized by ARR1 and ARR2 response regulator proteins which faintly resemble mammalian MYB TF (Sakai et al., 2000).
- **ACTCAT** - motif recognized by Signal Scan as PREATPRODHD (motif: ACTCAT), which stands for Pro- or hypoosmolarity element. ACTCAT element has been show to bind several bZIP class TF and is necessary for the efficient expression of proline dehydrogenase gene in response to L-Proline and hypoosmolarity (Satoh et al., 2002, 2004; Weltmeier et al., 2006).



- **I-box** - motif recognized by Signal Scan as IBOXCORE (motif: GATAA). This motif is a conserved sequence in promoters of light-regulated genes in both monocots and dicots (Terzaghi and Cashmore, 2003).
- **-10 PE** - recognized by the Signal Scan as -10PEHVPSBD (binding motif: TATTCT). Motif originally found in barley as involved in light regulation of chloroplast *psbD* gene (Thum et al., 2001).
- **SORLREP3AT** - (motif: TGTATATAT) is one of the computationally identified Sequences Over-Represented in Light-Repressed Promoters (SORLREP) in *Arabidopsis* (Hudson and Quail, 2003).

Considering the protective role of flavonoids against UV radiation (Winkel-Shirley, 2002) and reported involvement of light in regulation of flavonoid biosynthetic genes (Thain et al., 2002), I-box, -10PE and SORLREP3AT were selected as candidate motifs.

#### Seed related CAREs:

- **E-box** - EBOXBNNAPA (motif: CANNTG). Motif is overlapping with MYC and ABRE. E-box has been shown as required for high expression of *napA* storage protein in transgenic *Brassica napus* seeds (Stalberg et al., 1996).
- **DPBF** - (motif: ACACNNG) - Element found by yeast one hybrid screen in a carrot Dc3, *lea* class gene (Kim et al., 1997). Motif is for binding of DPBF (Dc3 Promoter-Binding Factor, bZIP class TF), which shares up to 96% sequence similarity with ABI5 in the DNA binding and potential dimerization domains (Finkelstein and Lynch, 2000).

#### 2.1.2.6 TATA box

Predicted TATA box is in the 42bp distance from the beginning of the 5' UTR identified by RACE PCR (Pourcel et al., 2005). Interestingly at the beginning of the transcript, W-box and GT1CONSENSUS (GT-1) sequences co-localize. GT-1 binding site is present in many promoters of light-regulated genes, where it can stabilize the TFIIA-TBP-DNA (TATA box) complex (Gourrierc et al., 1999). Moreover binding of the GT-1 like factors to the *PATHOGENESIS-RELATED GENE 1a* (*PR-1a*) promoter influences the level of SA-inducible gene expression (Buchel et al., 1999). We could hypothesize that W-box could be part of a negative regulation, whereas GT-1 would act for activation.

#### 2.1.2.7 *In silico* analysis of *TT10* promoter sequence - Summary

The *in silico* analysis of the *TT10* promoter sequence, accompanied by literature mining is a base for the understanding of gene regulation. To sum up the information presented in this section, several hypotheses could be proposed. If *TT10* is regulated as one of the flavonoid biosynthetic genes, there is probability that MYC2 and MYB2 motifs are functional and that they are binding sites for bHLH and MYB transcription factors respectively. Second hypothesis is built on the assumption that *TT10* could be regulated as a pathogen related gene. In this case we could also expect that the MYB sites would be involved and perhaps accompanied by W-boxes and/or ABRE. Another hypothesis is that *TT10* is regulated in a completely novel manner. It is possible, because it is not involved directly in flavonoid biosynthesis (like early and late biosynthetic genes), but in oxidative modification of end products. The 5' promoter dissection experiment described in the next section was based on the *in silico* information gathered. CAREs with the functional evidences were separated in various promoter constructs to validate which ones are functional *in vivo*.

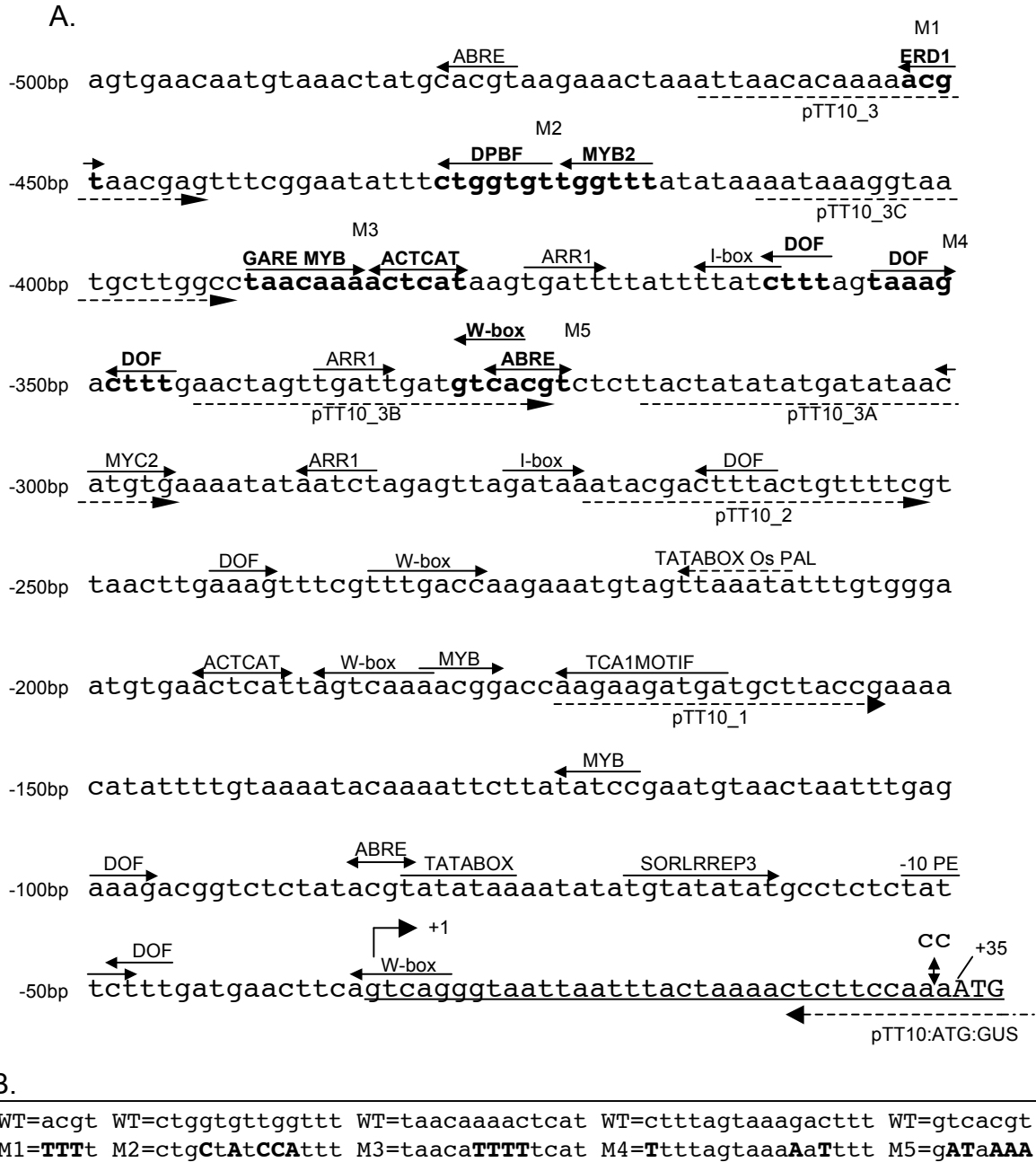


Figure 2.6: *In silico* analysis of proximal (500bp) *TT10* promoter showing putative *cis*-acting regulatory DNA elements.

(A) Wild-type Col-0 nucleotide sequence and selected putative *cis*-regulatory DNA elements found in *TT10* promoter. In bold are motifs considered for site directed mutagenesis (details of the modifications B, and description in section 2.1.3.4). Solid arrows are covering the corresponding motifs indicating their orientation. Dashed arrows are representing primers and cover the 'binding region'. 5' UTR is underlined. Vertical arrow is pointing nucleotide substitution before ATG, which are due to creation of *NcoI* restriction site required for transcriptional fusion with the reporter gene.

(B) Wild-type sequence of the putative *cis*-regulatory element analyzed in site-directed mutagenesis experiment. Nucleotides modified in the constructs are in bold uppercase and the resulting sequences are named M1 to M5.

### 2.1.3 Studies of *TT10* promoter activity during plant development

Pourcel et al. (2005) have investigated the regulation of *TT10* at the transcriptional level with a 2.0-kb *TT10* promoter from Columbia (Col-0) translationally fused to the *uidA* reporter gene encoding  $\beta$ -glucuronidase (GUS) protein transferred in Ws-4 background (previously named Ws-2). The promoter activity at the early stages of embryo morphogenesis (1 - 3 DAF) was detected in the endothelium layer of the seed coat and in the pigment strand at the chalaza zone. Later, the activity increased and was also observed in the outer integument (mostly in the oil layer of the seed coat). The GUS activity was also reported in early aborted seeds and in the transmitting tissue of the silique. Similar results were reported by Liang et al. (2006a), where activity of 1.5-kb *TT10* promoter activity was studied.

The present study intended to analyze the activity of the *TT10* promoter in its native background and extend it to the whole plant as well as to develop a tool which could be used for the analysis of stress response (see: 2.4).

#### 2.1.3.1 5' dissection of the *TT10* promoter

To investigate further the regulation of *TT10* and to define the domains of the promoter that might be important for the transcriptional regulation of tissue specificity, 5'-deletions were generated and translational fusions to the *uidA* gene were made.

#### Promoter fragments studied and cloning strategy

To limit the number of constructs and to maximize the output of the information about the promoter region which might be important, *in silico* analysis and putative CAREs were considered to plan promoter dissection, focusing on the proximal part of the promoter (as described in section 2.1.2). The promoter fragments studied were numbered, where: 1. stands for the shortest fragment of 173bp, 2. 271bp, 3. 465bp, 4. 1.0-kb, 5. 1.5-kb and 6. for 2.0-kb. Constructs were named as *pTT10\_X:GUS/pBIB-Hyg*, where X stands for the promoter fragment considered.

The shortest (1.) promoter fragment includes a predicted TATA box located 42bp upstream from the 5'UTR and putative CAREs: ABRE, MYB and TCA1MOTIF. Next fragment (2.) was planned to include two W-boxes which could be functional together. Fragment 3., includes MYC2, MYB2, ABRE and several other putative CAREs (see Fig. 2.6). Two other constructs have been planned to divide the upstream remaining sequence in equal parts, resulting in fragments 4. and 5.

The cloning strategy was aiming to conserve the native promoter composition (only two nucleotides changed to create *NcoI* restriction site) and is based on cloning of 2.0-kb *TT10* promoter described in Pourcel et al. (2005). Briefly, promoter fragments were PCR-amplified with the primers (Appendix, Tab. 5.3) including *XhoI* and *NcoI* restriction sites for translational fusion with the *uidA* (GUS) gene and sub-cloned into *pCR®-Blunt II-Topo®*. The *XhoI-NcoI* digested fragments were cloned into the *pBS-GUS* vector (Debeaujon et al., 2003), which was subsequently digested *SmaI-KpnI* to recover the corresponding promoter:GUS cassettes. The introduction of the constructs into the *pBIB-Hyg* binary vector was a limiting step of the cloning and it appeared to be of very low efficiency. It is important to note that the vector had to be prepared in two-step digestion, first with *SmaI* and then *KpnI*, because those sites are overlapping and reverse order of digestion results in loss of the *SmaI* site. Constructs were verified by restriction and sequencing. To fully appreciate the tissue specificity of the promoter, *Arabidopsis* Col-0 and Ws-4 wild-type plants were transformed with the six constructs described above. The results are illustrated on Figure 2.7, described below and summarized in table 2.1.

### 2.1.3.2 *In planta* analysis of GUS activity

The consistent GUS pattern established for each of the promoter fragments studied is based on analysis of 24 independent transformants lines (T1) (except *pTT10\_1:GUS/pBiB-Hyg*, 12 lines were studied). All the transgenic plants studied were genotyped for each construct and T2 plants were hygromycin selected to chose those segregating 3:1. Characteristic patterns of each promoter fragment activity observed in stable transformants lines (T2 and T3 generations; hygromycin resistant) was consistent with the one observed in primary transformants.

The same spatio-temporal GUS activity patterns were observed for *pTT10\_3*, *pTT10\_4*, *pTT10\_5* and *pTT10\_6* promoter fragments (Fig. 2.7). For those constructs, GUS activity was observed as early as the closed flower bud stage, however no activity was detected in mature pollen (Fig. 2.7 B.). It seems that pollination was not required for *TT10* promoter activity, because GUS activity was detected in siliques 7-days after flower castration (data not shown). On early stages of silique development (Fig. 2.7 C.) staining was present mainly in the transmitting tissue of the silique, and observed first at the top and then progressing to the bottom of the silique. Around 4-5 DAF (Fig. 2.7 D.) promoter activity was detectable in the funiculus as well as in the seed testa (endothelium layer). 2-3 days later (Fig. 2.7 E.), staining becomes very strong and the whole seeds were stained, which could be due to expression in outer integument 1 layer of the seed coat. Figure 2.7 F. is showing a silique at 8 DAF, which exhibits strong staining in aborted seeds and transmitting tissue. Stars at figure Fig. 2.7 C., E. and H. are highlighting the abscission zone of the sepals and petals. This staining is detectable shortly after petals are desiccating and sometimes is detectable as late as 14 DAF (Fig. 2.7 H.). At the later stages, GUS staining is no longer detectable in seeds which could be due to impermeability of the seed coat (Fig. 2.7 G. and H.), because in mechanically opened silique and scalpel wounded seeds, strong GUS staining is observed (Fig. 2.7 I.). In the mature siliques (Fig. 2.7 G. to I.) intensive GUS staining is visible in the dehiscence zone of the replum, but not in the valve. GUS activity is not detectable in any other tissues in plants grown in normal conditions in the greenhouse (Fig. 2.7 J. to M.). Plants grown *in vitro*, normally do not exhibit detectable GUS activity. Figure 2.7 N. is showing a seedling 10 days after transfer to the growth chamber, grown in the liquid 0,5x B5 medium supplemented with 0,5% sucrose. GUS activity is also not detectable in seedlings grown *in vitro* on vertical plates up to 14 days in long day or in constant light as well as in 10 day-old dark grown seedlings (data not shown). Other growth conditions and corresponding GUS patterns are described in section 2.4.

The *pTT10\_1* and *pTT10\_2* fragments did not exhibit detectable GUS activity *in planta*, except faint blue staining of the funiculus and abscission zone of the seed in *pTT10\_2*. The 194bp (between *pTT10\_2* and *pTT10\_3*) and 98bp region (between *pTT10\_1* and *pTT10\_2*) are required for seed coat and funiculus promoter activity, respectively. Moreover, it seems that the region upstream from -465bp (*pTT10\_3*) contains a sequence required for high gene expression, because much stronger GUS staining was observed for *pTT10\_4*, *pTT10\_5* and *pTT10\_6*, than *pTT10\_3*. For those constructs, much higher concentrations of the potassium ferricyanide and potassium ferrocyanide (FF) had to be used to eliminate GUS diffusion and artefactual staining of other tissues. No FF was used and material was ethanol bleached in the organs with weak *uidA* activity. Similar results were obtained for the constructs studied in Ws-4 background (data not shown).

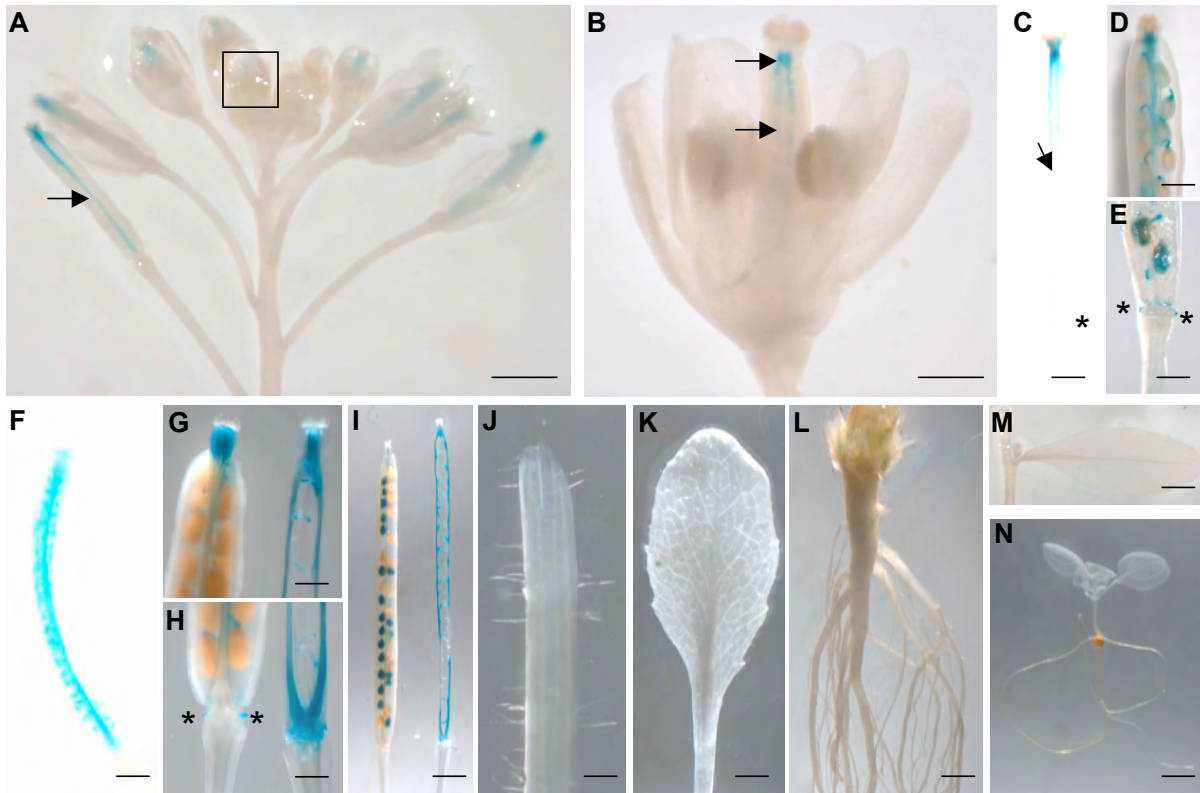


Figure 2.7: Pattern of *TT10* promoter activity in wild-type plants.

(A) to (H) Expression of the *pTT10\_3:GUS/pBIB-Hyg* cassette in wild-type Col-0 plants. (A) Inflorescence of 5-week-old plant; (B) dissected flower bud squared in (A); (C) to (H) various stages of developing siliques, (C) 4 DAF, (D) 5 DAF, (E) and (F) 8 DAF: (E) normal silique and (F) 'abnormal' silique with aborted seeds, (G) top and (H) bottom part of a 14-DAF silique (left) and a 18-DAF silique (right). Arrows are pointing week GUS staining. Stars highlight GUS staining in abscission zone of the petals.

(I) *pTT10\_5:GUS/pBIB-Hyg*, siliques  $\pm 14$  DAF (left) and  $\pm 18$  DAF (right).

(J) to (N) *pTT10\_6:GUS/pBIB-Hyg*; (J) to (M) various organs of 5-weeks-old plant, (J) stem, (K) rosette leaf, (L) root, (M) cauline leaf; (N) 10 days old seedling.

Different concentrations of potassium ferricyanide and potassium ferrocyanide were used to prevent GUS diffusion: A to H 1mM, I 2mM, J to N 0mM. Bars = 170 $\mu$ m B, 500 $\mu$ m C, 600 $\mu$ m D to F, 700 $\mu$ m G and H, 1mm J and N, 2mm A, I and K, 3mm L and M.

DAF - days after flowering

<b><i>AtTT10</i> promoter length</b>		<i>pTT10_6</i>	<i>pTT10_5</i>	<i>pTT10_4</i>	<i>pTT10_3</i>	<i>pTT10_2</i>	<i>pTT10_1</i>
<b>Organ/tissue</b>		2000bp	1451bp	1000bp	465bp	271bp	173bp
Seed	Integuments	+++++	++++	++++	+	-	-
	Micropyle	+++++	++++	++++	+	-	-
	Chalaza	+++++	++++	++++	+	-	-
	Funiculus	+++++	++++	++++	+	+	-
	Abscission zone	+++++	++++	++++	+	+	-
Silique - Flower	Flower bud	+++	+++	+++	+	-	-
	Stigma/style - Transmitting tissue	+++	+++	+++	+	-	-
	Abscission zone	+	+	+	+	-	-
	Receptacle	-	-	-	-	-	-
	Peduncle	-	-	-	-	-	-
	Stamen	-	-	-	-	-	-
	Pollen	-	-	-	-	-	-
Other organs	Cauline/Rosette leaf, trichomes	-	-	-	-	-	-
	Stem	-	-	-	-	-	-
	Root	-	-	-	-	-	-
Seedling	Cotyledons	-	-	-	-	-	-
	Hypocotyl	-	-	-	-	-	-
	Root	-	-	-	-	-	-

Table 2.1: Summary of *TT10* promoter activity pattern *in planta*.

This table is summarizing promoter dissection experiment and is presenting reproducible patterns observed for each of the constructs studied in wild-type plants. Relative promoter strength is represented by number of plus (+) signs and is based on the various concentrations of potassium ferricyanide and potassium ferrocyanide used to eliminate GUS diffusion (for quantitative results see Fig. 2.8 below).

### 2.1.3.3 Quantitative GUS assay - MUG

To precise the different strengths of the dissected promoter fragments, quantitative measurement of the GUS activity was realized. It is possible to measure precisely GUS activity, using fluorometric assays with very small amounts of transformed plant tissue (Jefferson et al., 1987). To confirm the differences in GUS staining observed for different promoter fragments, two or three lines for each construct have been selected. Because the amount of GUS protein could depend on the number of T-DNA inserts in transgenic plants, only the lines segregating 3:1 for antibiotic resistance : sensitivity were selected (good indication for single locus of the T-DNA insertion). Considering the pattern of *TT10* transcript accumulation (Fig. 2.1) and the promoter tissue specificity, 7-day-old siliques have been selected as samples to compare GUS activity.

There was a clear difference in GUS activity between the promoter fragments analysed in MUG experiment. The longest constructs *pTT10\_6*, *pTT10\_5* and *pTT10\_4*, have comparable, very high GUS activity, whereas it is 25 - 50 fold lower for *pTT10\_3*. It is still possible to detect weak GUS activity for *pTT10\_2*, but not for *pTT10\_1*. Similar results were obtained for the constructs studied in Ws-4 background (data not shown).

### 2.1.3.4 Sub-dissection and site-directed mutagenesis - introduction

To continue, sub-dissection of the 194bp region (between *pTT10\_2* and *pTT10\_3*), site-directed mutagenesis experiments were carried out to narrow down the region required for promoter activity in seeds. The ultimate goal of this part of the work was to find out which *cis*-regulatory elements in the promoter are functional *in*

*vivo*. The dissection experiment (2.1.3.1) defined a 194bp promoter region, which contains several putative CAREs defined in *in silico* analysis. A sub-dissection of this region was planned to further limit the promoter region and number of putative CAREs considered for site-directed mutagenesis.

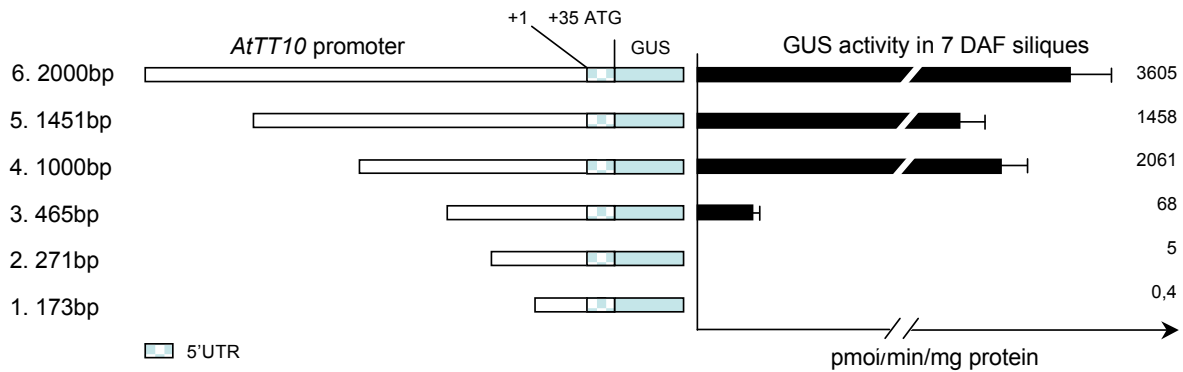


Figure 2.8: Comparison of GUS activity of 5'-dissected *TT10* promoter *in planta*.

The differences in a GUS activity, which was measured in siliques (7DAF) of T3 transformant lines in Col-0 background are shown. Promoter fragments studied and their lengths are represented by the open rectangles. Values represent averages  $\pm$ SE of six technical replicates.

### Sub-dissection

Three constructs have been carried out to test putative CAREs found in the 194bp region. Looking from the 5'-end of the pTT10\_3 promoter fragment, construct pTT10\_3C - 412bp removes ABRE-like, DPBF and MYC2 motifs. Next fragment, pTT10\_3B - 344bp, was removing GARE MYB, ACTCAT, ARR1, I-box and three DOF motifs. Finally pTT10\_3A - 318bp was planned to keep the MYC2 motif and not include W-box and ABRE which are just upstream (see Fig. 2.6 for details).

### Technical difficulties encountered

The cloning strategy was planned to follow the same steps used for the promoter dissection construct preparation (2.1.3.1). Nevertheless, in this case the transfer of the promoter:GUS cassette into the *pBIB-Hyg* binary vector did not result in the appropriate constructs. To optimize cloning conditions various ratios of the insert and vector were tested. We have tried to purify the DNA fragments with several commercially available kits (see 3.2.4.2). Restriction analysis of *pBIB-Hyg* binary vector with *SmaI* and *KpnI* restriction enzymes revealed that both sites were cut with high efficiency. However the religation after single cutting was of very low efficiency with different batches of ligase. In parallel, alternative approaches have been proposed and are described below.

### Cloning in *pCAMBIA 1300*

*pCAMBIA 1300* (*pCAMBIA*) ([www.cambia.org](http://www.cambia.org)) is a binary vector, which is routinely used in the Seed Biology Laboratory. The analysis of the restriction sites to take out the promoter:GUS cassette from the *pBS-GUS* and those in the multi-cloning site of the *pCAMBIA*, revealed that *KpnI-PstI* are the only sites available. Briefly, the cloning protocol was as follows: i) plasmid preparation was omitting gel purification after the restriction; ii) ligation was followed by *SalI* digestion, to eliminate false positives from not completely digested binary vector; iii) after restriction and ligation the reaction mix was desalted; iv) bacterial colonies were screened for positive clones by restriction analysis (strategy based on advises from Nathalie Berger). The integrity of all the constructs had been checked by sequencing and plants were transformed.

Unexpectedly the study of the GUS pattern in T1 plants, revealed very strong blue staining in all the plant tissues tested for all the constructs. One of the hypotheses to explain such pattern of GUS activity was first association of the possible induced activity with heavy aphid/greenfly invasion in the greenhouse. However, this GUS pattern was also observed in T2, antibiotic-resistant plants after the anti-aphid treatment. Another possible explanation is the effect of the strong promoter CaMV35S, used for *HptII* gene expression in *pCAMBIA* (Fig. 2.9). This promoter, used to ensure the antibiotic resistance could also interact with the *TT10* promoter fragments, which were inserted side-by-side, to drive ectopic GUS expression through its enhancers. Indeed, the pattern of GUS activity resembles the one driven by *35Sdual:GUS/pBIB-Hyg* (Fig. 2.9).

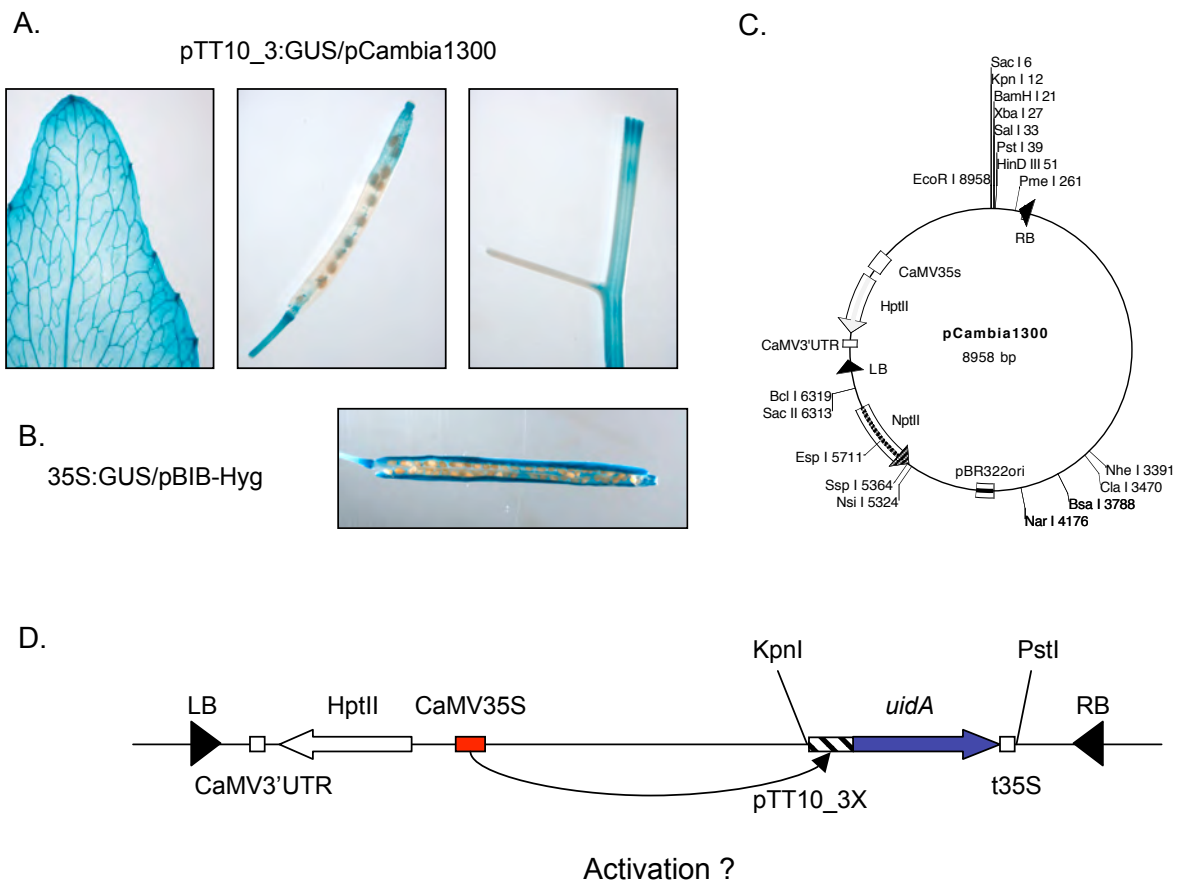


Figure 2.9: GUS activity driven by the the *TT10* promoter in *pCAMBIA* binary vector.

(A) Ectopic expression of the GUS gene driven by various *TT10* promoter fragments illustrated with *pTT10\_3:GUS/pCAMBIA1300*.

(B) GUS activity driven by CaMV35S dual promoter in siliques

(C) Map of the *pCAMBIA 1300* vector.

(D) Map of the T-DNA cassette showing the *HptII* - Hygromycin resistance gene and the CaMV35S promoter (red box). A *KpnI-PstI* insertion of the promoter:GUS constructs realized. LB/RB - left and right borders of the T-DNA insert. Arrow is pointing hypothesized involvement of CaMV35S enhancers in ectopic expression of the *uidA* gene.



### Cloning in *pBI101GUS-GTW*

Considering the unexpected results obtained with the promoter studies in the *pCAMBIA* vector, another alternative approach has been used to carry out the sub-dissection and site-directed mutagenesis experiments. Because of the time constraints of the thesis, the GATEWAY® cloning has been used, which allowed very fast vector preparation (Earley et al., 2006). On the other hand, a compromise had to be made and the 5'UTR sequence had to be modified to introduce the GATEWAY® recombination sites by adding 29bp at both ends of the promoter. The *pBI101GUS-GTW* and the cloning strategy are described elsewhere (Baudry et al., 2006; Dubos et al., 2008). Briefly, the promoter fragments have been PCR-amplified with the primers containing *attB1* and *attB2* GATEWAY® recombination sites (Appendix, Tab. 5.8) and introduced into the *pDONR207* vector by BP recombination. The next step was the LR recombination to transfer the promoter fragment into the binary vector *pBI101GUS-GTW*. The integrity of all the constructs was checked by sequencing and plants were transformed. Results obtained with sub-dissection are described below in the section 2.1.3.5, together with the results of the site-directed mutagenesis.

### Site-directed mutagenesis

A complementary approach to deletion analyses to check the functionality of the putative *cis*-acting regulatory elements found in the 194bp region (between *pTT10\_2* and *pTT10\_3*), which could be required for the promoter activity in seed coat, is site-directed mutagenesis. The motifs for the analysis were chosen based on the *in silico* analysis described in section 2.1.2. Some of the putative CAREs are located very close from each other or even overlapping, therefore it was decided to analyze those together. The constructs have been named - M - for mutagenesis, and numbered 1 to 5 referring to the respective construct. The nucleotide substitution in the constructs prepared is based on the one which had been found in the literature describing motif identification. We checked that the mutagenised CAREs considered in the experiment, were not recognized by *in silico* analysis (PLACE, Signal Scan). Figure 2.6 is illustrating the location of the motifs present in the five constructs studied, with the details of the modified sequence given in the box. The mutagenesis was prepared on the full *pTT10\_3* promoter fragment because elements downstream -271bp (*pTT10\_2*) could be necessary for mutual activation of GUS activity.

### Construct preparation

The point mutations have been introduced in the promoter sequence using QuickChange® Site-Directed Mutagenesis Kit (Stratagene®). It is a PCR based technique which allows efficient site-specific mutation in a double-stranded plasmid ([www.stratagene.com](http://www.stratagene.com) and product manual). Primers have been designed using PrimerX. Nucleotide substitutions have been confirmed by sequencing. The following, cloning steps and difficulties encountered were as described for sub-dissection construct preparation. Multiple attempts of cloning into the *pBIB-Hyg* did not result in the proper construct and those in *pCAMBIA 1300* gave ectopic GUS activity. Finally, the constructs were prepared with the GATEWAY® technique and the preliminary results of GUS pattern in T1 plants are discussed below.

#### 2.1.3.5 Patterns of GUS activities after *TT10* promoter sub-dissection and site-directed mutagenesis

The result of the *TT10* promoter sub-dissection and site-directed mutagenesis are summarized in Table 2.2. The analysis focused on seeds, where it was expected to observe blue staining for the minimal promoter, *pTT10\_3* - positive control, or no GUS activity in *pTT10\_2* - negative control. Those two constructs have been the reference of *TT10* promoter activity and their comparison was required to validate the data obtained in the new vector. For the other constructs it was expected that, if the promoter fragment/motif

required for GUS expression was eliminated, the blue staining would not be observed anymore. It would be also possible that a promoter region binding negative regulator is eliminated, which would result in strong GUS staining. The same result could be observed, if a strong activation site was created by site-directed mutagenesis.

The GUS staining observed for the *pTT10\_3:GUS/pBI101-GTW* seems to be far weaker than for the same promoter fragment in *pBIB-Hyg* vector. Moreover, blue staining is not detectable at the flower/early silique developmental stages, compared to Fig. 2.7, B. to D., whereas it is very weak in older siliques, compared to 2.7, E. Staining was extremely weak, even without potassium ferricyanide and potassium ferrocyanide. No staining was detected in other plant organs. The negative control *pTT10\_2:GUS/pBI101-GTW* gave no detectable GUS staining, whereas this promoter fragment in *pBIB-Hyg* was characterized by remaining, faint blue staining in the funiculus and seed abscission zone. The GUS activity seems to be abolished in all the sub-dissection constructs (*3A*, *3B* and *3C*). For the site-directed mutagenesis, faint blue staining is observed for most of the M1 and M5, T1 plants tested, whereas it seems to be detectable in less than a half of the T1 plants tested for M2. Interestingly, GUS staining was also abolished in plants carrying the M4 construct. Only 5 primary transformants have been obtained for M3 construct and GUS activity was observed in one of them.

To sum up the present results of the *TT10* promoter sub-dissection and site-directed mutagenesis it has to be noted that the GUS staining for the *pTT10\_3:GUS/pBI101-GTW* - the positive control - is much weaker than the one observed when this fragment was studied in *pBIB-Hyg*. This could be simply due to the introduced GATEWAY® recombination site in the 5'UTR region of the promoter. If we consider the weak GUS activity of the *pTT10\_3:GUS/pBI101-GTW* as a new reference, we could point out, that the region between *pTT10\_3* and *pTT10\_3C* contains the motifs required for the gene expression. The lack of detectable GUS activity for most of the M2 plants could suggest that DPBF or MYB2 would be the *cis*-acting elements required for *TT10* gene expression. This could be further supported by the fact that the GUS activity is detectable for the constructs M1 (within 3-3C region) and M5 (downstream of 3C), which could be treated as a positive GUS staining reference. Surprisingly, no blue staining was detected for the plants carrying M4 construct what could indicate also involvement of DOF CARE. Nevertheless, all those results should be reproduced, to confirm this preliminary study.

T1#	X=	<i>pTT10_X:GUS/pBI-101-GTW</i>										
		3	3C	3B	3A	2	M1	M2	M3	M4	M5	
1		+	-	-				+	-			
2		+	-/+	-				+	-		+	
3		-	-	-				+	-			
4		+	-	-/+				-	-	n/a		
5		+	-	-				+	+			
6		+	-/+	-				-	n/a		+	
7		-	-	-				-				
8		+	-	-				n/a			n/a	
9		n/a	-	-				-				
10		-	-	-				-				+
11		+	-	-				-				
12		n/a	-	n/a				-		n/a		
GUS+/TotalT1:		7/10	2/12	1/11	0/12	0/12	3/3	4/11	1/5	0/3	3/3	
Status:		+	-	-	repeat							

Table 2.2: Patterns of GUS activity observed in the seed coat after sub-dissection and site-directed mutagenesis of *TT10* promoter.

Table is summarizing the GUS staining observed in 5-15DAF old siliques of 12 independent T1, hygromycin resistant plants. The plus (+) sign is for the blue staining observed in seed coat. The minus (-) sign means that no GUS activity was detected in any plant organ tested (rosette leaf, cauline leaf, stem, flower and siliques). The (+) or (-) in the dashed rectangles refers to the siliques analyzed in bulk; in this case (+) stand for blue staining in most of the siliques in the well; (-) when there was no detectable GUS activity; The experiment has been repeated twice for the constructs *3*, *3C*, and *3B* and in this case (-/+) means that GUS staining was observed in one of the repetitions. n/a - not consistent, ectopic GUS activity or plant died; for M3: plants not available. Plants were not genotyped and T2 seeds are not yet available.

#### 2.1.4 Promoter GUS activity in protoplasts

In addition to *in planta* study of the promoter activity, transient expression experiments were undertaken with protoplasts from *Arabidopsis* cell suspension in order to obtain more robust quantitative analyses (see 3.2.3.8). PEG mediated transfection resulted in detectable GUS activity for all the constructs tested (*1* - *5*). This approach could have accelerated the analysis, but the results obtained *in planta* demonstrated that *pTT10\_1* and *pTT10\_2* do not drive detectable GUS activity. It seems that protoplasting, transfection or the culture conditions used can affect *TT10* promoter activity (Takeda et al., 2002). The shortest promoter fragment, *pTT10\_1*, contains elements sufficient for GUS expression in transfection experiment in protoplasts. Therefore it was necessary to pursue the regulation studies *in planta*.

## 2.2 Finding transcription factors

To progress further in the understanding of the regulation of *TT10* gene expression, several approaches have been used to find candidate transcription factors which could be involved. The first strategy was to study promoter activity with reporter gene in candidate regulatory mutant backgrounds which are involved in flavonoid biosynthesis. The second approach was to select candidate transcription factors, based on co-expression data from microarrays. Several other mutants were considered on the basis of the putative *cis*-acting regulatory elements found in *TT10* promoter, and associated with the function of the gene. Quantitative RT-PCR was used to evaluate *TT10* transcript accumulation. Finally, yeast one-hybrid screen was tried to fish out some of the transcription factors binding to the promoter.

### 2.2.1 *TT10* promoter activity in candidate regulatory mutant backgrounds

One of the ways to find out how a gene is regulated is to study promoter activity with reporter gene in a regulatory mutant background. This approach proved to be successful for *BAN*, which promoter activity is significantly changed in *tt2*, *tt8* and *ttg1* mutants (Debeaujon et al., 2003). The modification of the GUS pattern could be either a direct or an indirect effect. Pourcel et al. (2005) already demonstrated that the activity of the *TT10* promoter is not modified in *tt2*, *tt8* and *ttg1*. In this work we wanted to complete the analysis by looking whether *TT10* could be regulated by the TT1 zinc finger protein, (Sagasser et al., 2002) and/or the TT16 MADS box protein, (Nesi et al., 2002) and/or the TTG2 WRKY transcription factor (Johnson et al., 2002). All those genes are involved in the regulation of flavonoid biosynthesis in seed.

To introduce *TT10* promoter GUS constructs in the *tt1*, *tt16* and *ttg2* background, crosses with the plants carrying *pTT10:GUS/pBIB-Hyg* construct (Pourcel et al., 2005) have been made. Plants have been selected for the antibiotic resistance and seed *transparent testa* phenotypes, and glabra character of the leaves for *ttg2*. The GUS staining for *pTT10:GUS/pBIB-Hyg* construct in *tt1*, *tt16* and *ttg2* backgrounds, seems to be much stronger than in the control wild-type plants. Higher concentrations of the potassium ferricyanide and potassium ferrocyanide were used to prevent GUS diffusion.

Analysis of the GUS patterns revealed that in *tt1* background blue staining is first detectable in the endothelium layer (Fig. 2.10 A.) and later on in the outer integument layer 1 (Fig. 2.10 D. and G.), which resembles wild-type promoter activity. The detection of GUS activity in the *tt16* mutant background, at the globular stage of embryo development (Fig. 2.10 B.), is restricted to the pigment strand, but weak staining is also observed in endothelium. However, this can be due to the strong staining at the pigment strand and GUS diffusion. At the later stages, the promoter activity is visible in both endothelium and oi2 seed coat layer (Fig. 2.10 E. and H.). *TT10* promoter activity in *ttg2* mutant background closely resembles that in *tt16*, with the strong staining at the micropyle and chalaza/pigment strand zones. Although the faint, patchy GUS activity was detectable in the endothelium (Fig. 2.10 C.). It seems that at the later stages, two-layer expression patterns is conserved (Fig. 2.10 F. and I.). The GUS activity in *tt1*, *tt16* and *ttg2* does seem not to be affected at the seed abscission zone. No ectopic promoter activity was detected in leaves nor in stem.

The results could indicate that *TT10* expression might be regulated by TT1, TT16 and TTG2 in the endothelium. Nevertheless, this could be indirect regulation, because the morphology of the endothelium is affected in *tt1* and *tt16* mutant seeds (Nesi et al., 2002; Sagasser et al., 2002; Pourcel, 2006). In the *ttg2* mutant, endothelium morphology seems not to be affected, but cells might not be finally differentiated. It has been shown that TTG2 is required for trichome development (Sagasser et al., 2002) and endosperm differentiation (Dilkes et al., 2008). Moreover, the results of the GUS staining are difficult to interpret due to the possible GUS diffusion. It would be interesting to check by *in-situ* hybridization, if *TT10* mRNA is

detectable in the endothelium of *tt1*, *tt16* and *ttg2* mutants.

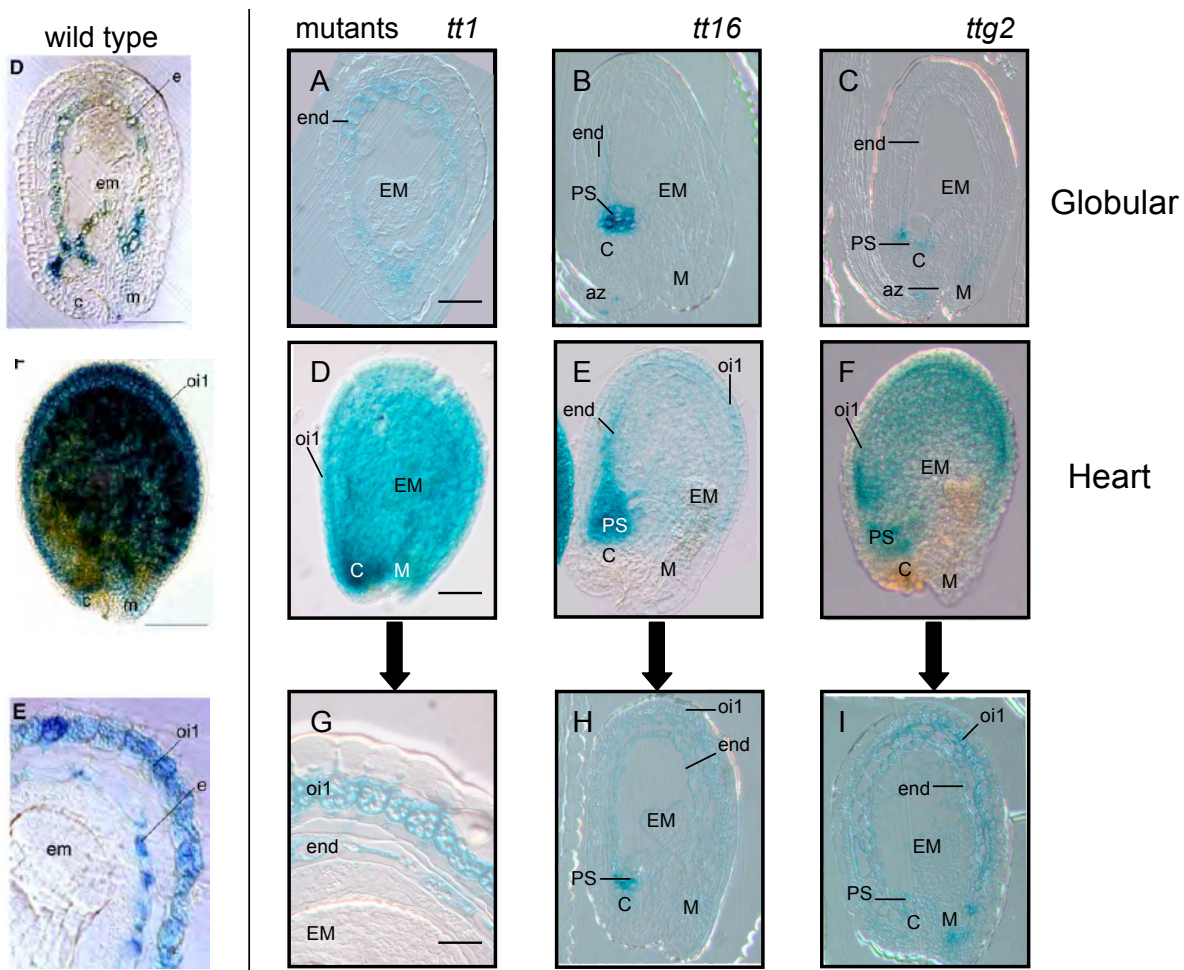


Figure 2.10: Pattern of *TT10* promoter activity in seeds of the candidate regulatory mutant backgrounds. Expression of the *pTT10:GUS/pBIB-Hyg* cassette in developing seeds of the candidate regulatory mutant backgrounds (*tt1* - left, *tt16* - middle and *ttg2* - right column). Promoter activity in the wild-type is depicted by pictures adapted from Pourcel et al. (2005).

(A) to (C) Seeds at the globular stage of the embryo development (3-4 DAF); (D) to (F), (H) and (I) Heart stage (6-7 DAF); (G) Bending cotyledon (8-9 DAF). GUS activity was observed with Nomarski optics on sections A to C and G to I and whole mounts D to F.

az, abscission zone; C, chalaza; end, endothelium; EM, embryo; M, micropyle; oi, outer integument; PS, pigment strand. Bars = 80 $\mu$ m A to C; 90 $\mu$ m D to F, H and I; 120 $\mu$ m G.

## 2.2.2 Confirmation of *TTG2* promoter activity in developing seeds

The seeds of *ttg2* mutant do not accumulate PAs and mucilage, but the endothelium structure seems not to be affected. Promoter activity has been reported to be detectable in the seed coat (Johnson et al., 2002). In this study, we wanted to test the hypothesis, that *TT10* expression could be regulated by *TTG2*. Indeed, *TTG2* encodes WRKY TF potentially binding W-box that have been found in the *TT10* promoter (see 2.1.2). For that reason, we wanted to check precisely if spatio-temporal activity of both promoters are similar in developing seeds. Figure 2.11 A. is showing GUS activity driven by the *TTG2* promoter, which is detectable in the endothelium and micropyle of the developing seeds at the early globular stage of embryo development. Staining seems to be very weak or absent in the chalaza - pigment strand zone. In addition,

at the early heart stage (Fig. 2.11 B. and C.), GUS activity is detected in the mucilage cell layer. Staining at the seed abscission zone was not detected.

Activity of *TTG2* and *TT10* promoters seem to start at similar developmental stages and they are both detectable in the endothelium. The *TT10* promoter activity seems to be more pronounced at the chalaza/pigment strand zone. From the heart stage of embryo development onwards, the activity of both promoters is different, with *TT10* being in outer integument 1, whereas *TTG2* restricted to the mucilage layer. It could be that *TTG2* is involved in the regulation of *TT10* in the endothelium. Another hypothesis could be that, they are both regulated by common, or distinct transcription factors resulting in the similar spatio-temporal activity. It would be interesting, to check by yeast-one hybrid and/or electrophoretic mobility shift assay if *TTG2* could bind to the *TT10* promoter. One more reason to verify the hypothesis of this interaction are W-boxes found in the *in silico* studies, which could be potential binding sites of WRKY transcription factors (see 2.1.2.2 and Fig. 2.6).

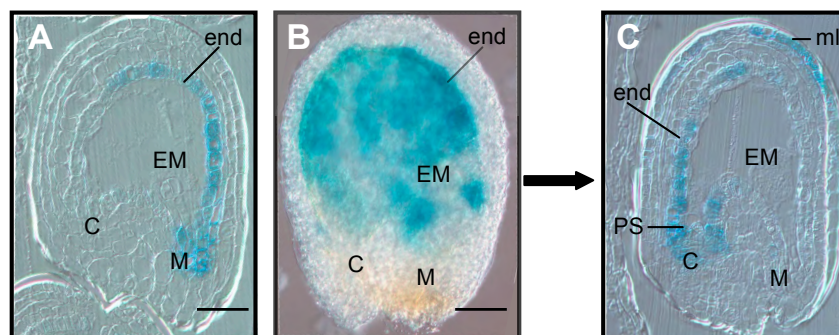


Figure 2.11: Pattern of *TTG2* promoter activity in wild-type seeds.

Expression of the *pTTG2::GUS/pBI101.1* cassette in developing seeds of wild-type plants.

(A) Seeds at the globular stage of embryo development (3-4 DAF); (B) and (C) heart stage (6-7 DAF). GUS activity was observed with Nomarski optics on sections A and C and whole mount B.

C, chalaza; end, endothelium; EM, embryo; M, micropyle; ml, mucilage layer; PS, pigment strand. Bars = 70µm A; 90µm B and C.

### 2.2.3 Analysis of public micro-array data - searching for co-expressed genes

The idea to look for the expression of potential regulators of the gene of interest is based on putative expression throughout time - developmental stage and space - tissue/cell. This is why we searched publicly available micro-array data to find regulatory genes co-expressed with *TT10*. Expression Angler program at BAR (<http://bar.utoronto.ca/>) was queried for the genes co-expressed with *TT10* in the AtGenExpress Seed Set. First 100 top hits, regardless of the r-value were searched for the transcription factors. The list of selected co-expressed genes is presented in Table 2.3. Details of the selection are given in the legend of the table.

#### 2.2.3.1 *TT10* expression in candidate regulatory mutant backgrounds obtained from co-expression analysis - seed set

In the section 2.2.1 it was of interest to check if *TT10* promoter activity was modified in previously identified regulatory mutant backgrounds. Here we decided to study the level of *TT10* mRNA, which could indicate, if certain genes could be involved in the regulation of *TT10* gene expression. The amount of information available for the co-expressed genes was rather scarce, therefore, we have checked, if their transcript was

present in developing flowers and/or seeds. Transcriptomic analysis (Fig. 2.12) showed, that At3g15500 is strongly expressed in developing flower and seeds, but also in the senescing leaf, whereas it is below the detection level in other tissues. At2g24430 seems to be seed specific, At5g66070 is relatively ubiquitous, but high amount of transcript were detected in developing pollen, whereas At5g62320/MYB99 expression seems to be restricted to early stages of flower development. At1g71450 and At5g07700/MYB76 were below the detection level. Homozygous mutant lines have been ordered from NASC, when available. All the plants have been genotyped to confirm the presence of T-DNA insertions and an homozygous MYB99 mutant line was selected from T3 plants.

	<b>AGI</b>	<b>r-value to bait, GeneChipID, Annotation</b>
	At5g48100	1.000 248735_at  TRANSPARENT TESTA 10, laccase
1	At5g41040	0.801 249289_at  transferase family protein
2	At5g64190	0.765 247324_at  similar to unknown protein [Arabidopsis thaliana]
3	At5g09480	0.735 245889_at  hydroxyproline-rich glycoprotein family protein
4	At3g53510	0.723 251944_at  ABC transporter family protein
5	At3g15500	0.721 258395_at  ANAC055_ATNAC3__ATNAC3 (ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 55); transcription factor
13	At2g40370	0.674 263825_at  LAC5__LAC5 (laccase 5); copper ion binding / oxidoreductase
41	At1g71450	0.650 259946_at  AP2 domain-containing transcription factor, putative
45	At2g24430	0.649 265685_at  ANAC038_ANAC039__ANAC038/ANAC039; transcription factor
48	At5g07700	0.647 250589_at  AtMYB76_MYB76__MYB76 (myb domain protein 76); DNA binding / transcription factor
71	At2g27070	0.630 250612_s_at  ARR13__ARR13 (ARABIDOPSIS RESPONSE REGULATOR 13); transcription factor/ two-component response regulator
77	At1g55580	0.628 265072_at  LAS_SCL18__LAS (LATERAL SUPPRESSOR); transcription factor
97	At5g62320	0.623 247451_at  AtMYB99_AtMYBCU15_MYB99__MYB99 (myb domain protein 99); DNA binding / transcription factor
98	At5g66070	0.623 247125_at  zinc finger (C3HC4-type RING finger) family protein

Table 2.3: List of selected genes co-expressed with *TT10* - Seed Set.

The genes co-expressed with *TT10* found with Expression Angler in the AtGenExpress Seed Set. Because of relatively low r-values, only the first five genes were listed, followed by the selection of putative transcription factors and also an other member of the laccase family. The genes with the position number in the grey background were the candidates which were studied and are described in this section.

Analysis of *TT10* expression and seed flavonoid composition have been realized for the following mutants:

- At3g15500 - gene coding for ATAF-like NAC-domain transcription factor, that does not contain sequences shared by CUC1, CUC2 and NAM (Takada et al., 2001; Ooka et al., 2003).
- At1g71450 - encodes a protein containing one Apetala 2 (AP2) domain and is a member of the DREB subfamily A-4 of ERF/AP2 transcription factor family (Nakano et al., 2006). It has been shown to be involved in salt stress, cold and ABA response (Ma et al., 2006; Wenqian et al., 2007).

- At2g24430 - gene coding for Arabidopsis NAC domain containing protein 038, Arabidopsis NAC domain containing protein 039 (Ooka et al., 2003). Induced by blue light in the *cryptochrome1* mutant (Folta et al., 2003).
- At5g07700 - encodes AtMYB76 transcription factor, which has been shown to be involved in regulation of aliphatic glucosinolate biosynthesis in *Arabidopsis* (Stracke et al., 2001; Gigolashvili et al., 2008).
- At5g66070 - encodes a zinc finger (3HC4-type RING finger) family protein, identical to RING-H2 finger protein ATL5P (Kosarev et al., 2002), which is a ubiquitin-ligase gene significantly up-regulated after chitoctase treatment (Libault et al., 2007).
- At5g62320 - encodes AtMYB99 transcription factor, expressed during pollen development, which has been shown to be involved in microsporogenesis (Stracke et al., 2001; Alves-Ferreira et al., 2007).

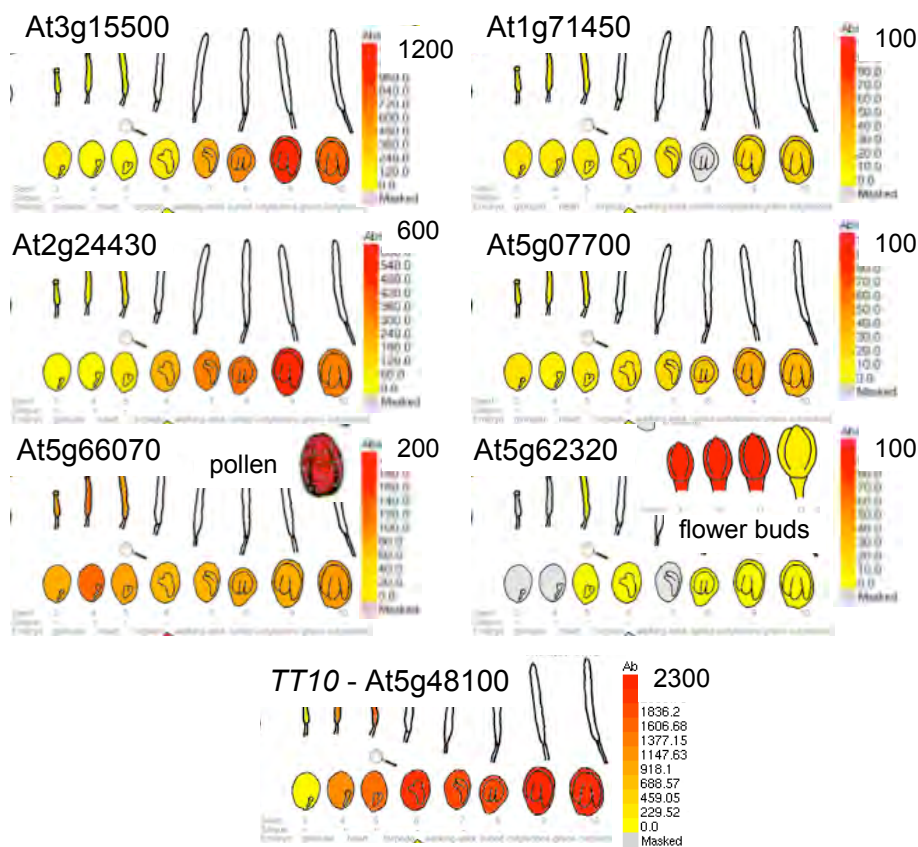


Figure 2.12: Expression levels of candidate regulators of *TT10* in developing seeds.

The expression profiles were visualized with eFP Browser. Various signal (expression potential) threshold values have been set to facilitate comparison of gene expression levels. Standard deviation filtering has been used to mask samples with the deviation greater than half of their expression value (grey). Samples where expression is relatively low or considered to be below the detection level are colored in yellow. Data are not available for separate siliques (white) from torpedo stage of embryo development onwards, where only dissected seeds were analyzed. Developing seed, silique and pollen where the gene expression potential is > 100 (detectable) are painted in red.

### 2.2.3.2 Results for regulatory mutants chosen from co-expression analysis

*TT10* expression in mutant backgrounds was not drastically down nor up-regulated. The largest difference was up to 25% decrease of *TT10* transcript accumulation in the *myb76* and *myb99* mutants (Fig. 2.13).



Those results should be confirmed with a second biological replicate before further interpretation. However, the MYB transcription factor family consists of almost 130 members (Stracke et al., 2001), therefore we could expect some functional redundancy that would explain limited impact of the *myb* mutants (Stracke et al., 2001). The other explanation for partial decrease of the mRNA level could be the involvement of two different mechanisms of regulation in endothelium and in the outer integument 1 seed coat layer. In such a case, the absence of *TT10* transcript in one layer would be masked by its expression in the other, what would be not possible to measure by RT-PCR.

Analysis of flavonoid composition (all the metabolomic analyses of flavonoid composition were done by Jean-Marc Routaboul) in the candidate mutants chosen on the basis of the co-expression data, revealed no significant differences in the amount of soluble PAs nor flavonols (Fig. 2.14). Surprisingly, in the At2g24430 mutant, we even observed a slight decrease of the soluble PAs fraction, which is in opposite to the expected *tt10* mutant phenotype.

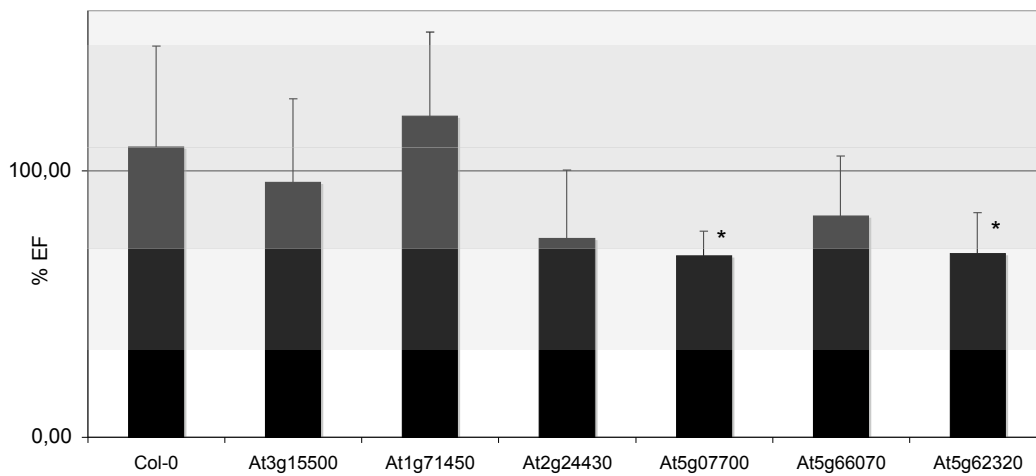


Figure 2.13: *TT10* transcript accumulation in candidate regulatory mutant background - co-expressed candidates.

*TT10* gene expression detected by quantitative RT-PCR, presented as a percentage of the expression of the reference gene *EF1*. Graph represents values for one biological repetition and averages  $\pm$  SE of three technical replicates. Dark and light grey boxes represent  $\pm$  1 SE and  $\pm$  2SE values respectively. Asterixes highlight samples where *TT10* expression was more than  $\pm$  1SE different from the control.

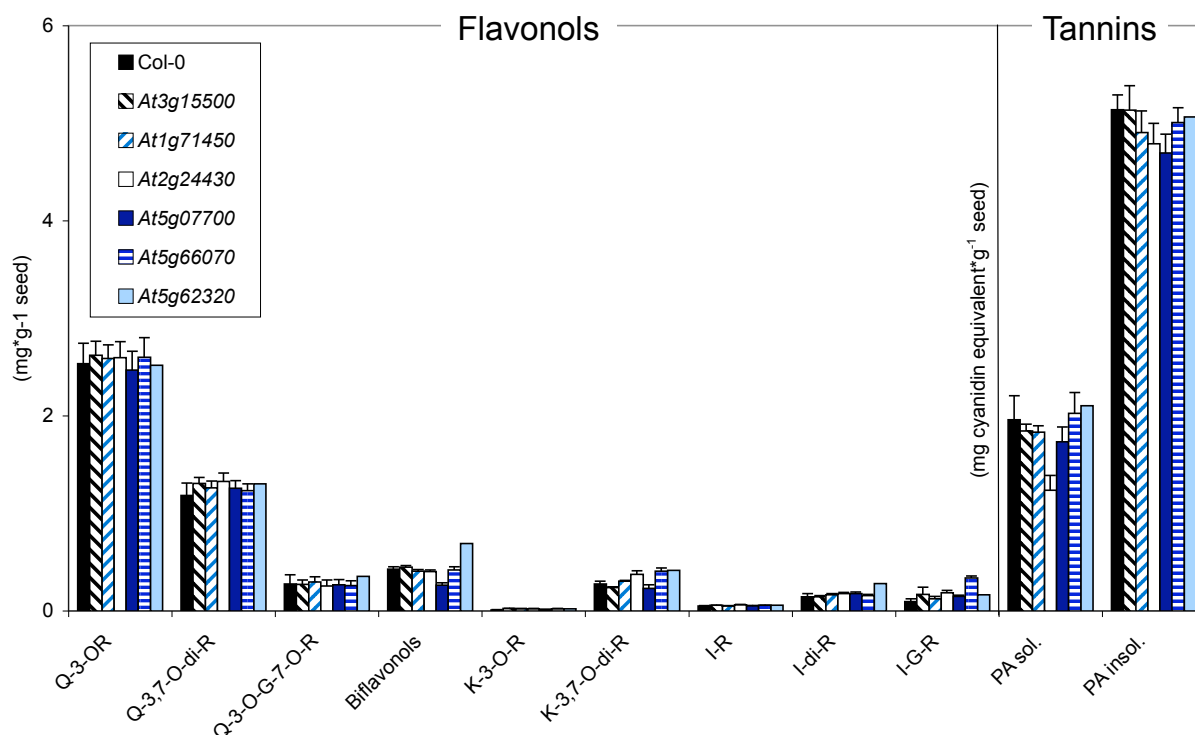


Figure 2.14: Flavonoid composition of mature seeds in the mutants for genes selected from co-expression analysis.

Analysis of flavonoid derivatives and tannins by LC-MS and after acid-catalyzed hydrolysis, respectively. Plants were grown in the greenhouse at the same time. Values represent averages  $\pm$  SE of three independent measurements. G, glucoside; I, isorhamnetin; K, kaempferol; Q, quercetin; R, rhamnoside.

### 2.2.3.3 Co-expression analysis - developmental set

Even if *TT10* is expressed mainly in seeds, limiting the search for the co-expressed genes to the AtGenExpress Seed Set could have resulted in biased results. Indeed this data set consists in the expression data for mature seeds samples subjected to various treatments, however *TT10* is expressed during seed development. This is reinforced by the fact that in this data set, *TT10* expression potential values are very low (max. 26.83), and only few samples could be considered after standard deviation filtering. For those reasons, it would have been more appropriate to search for co-expressed genes in the AtGenExpress Developmental Set. Expression Angler program at BAR was queried for genes co-expressed with *TT10* in the AtGenExpress Developmental Set. List of selected co-expressed genes is presented in Table 2.4. It would be interesting to investigate the potential role of transcription factors which were found to be co-expressed with *TT10*. A limitation of this approach is that, if the potential candidate gene regulating *TT10* expression is not seed specific, then it will not appear in the list of the co-expressed genes, because its expression profile will not fit with the one of *TT10*. One of the top priorities would be to check the role of AtMYB5, because in the phylogenetic studies, it is in the clade together with AtMYB11, AtMYB12, AtMYB75/PAP1, AtMYB90/PAP2, AtMYB111 and AtMYB123/TT2 (Stracke et al., 2001), which have been shown to be involved in the regulation of flavonoid biosynthetic genes (Lepiniec et al., 2006; Stracke et al., 2007). AtMYB5 encodes a transcription factor that acts as a negative regulator of trichome branching and plays a role in the correct formation of the seed coat, and possibly the formation of the underlying endosperm layers. Loss of function mutations lead to defects in seed coat mucilage, columella cells as well as trichome (Li et al., 2009) and decreased PA content (Gonzalez et al., 2008).

	AGI	r-value to bait, GeneChipID, Annotation
	At5g48100	1.000 248735_at  TRANSPARENT TESTA 10
1	At1g62000	0.985 264735_s_at  similar to unknown protein [Arabidopsis thaliana] (TAIR:AT1G62080.1)
2	At4g36700	0.983 246273_at  cupin family protein
3	At3g01570	0.974 259167_at  glycine-rich protein / oleosin
4	At5g40420	0.973 249353_at  OLEO2_PA23_OLEO2 (OLEOSIN 2)
5	At5g44120	0.973 249082_at  CRU1_CRA1 (CRUCIFERINA); nutrient reservoir
6	At4g00220	0.969 255701_at  JLO_LBD30 (JAGGED LATERAL ORGANS)
7	At1g47540	0.969 262431_at  trypsin inhibitor, putative
8	At5g38160	0.969 249547_at  protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
9	At3g25260	0.967 257841_at  proton-dependent oligopeptide transport (POT) family protein
10	At3g63040	0.963 251202_at  unknown protein
11	At5g03795	0.962 250893_at  oxidoreductase
12	At5g44360	0.961 249043_at  FAD-binding domain-containing protein
13	At4g28520	0.961 253767_at  CRU3_CRC_CRU3 (CRUCIFERIN 3)
14	At1g62290	0.960 264741_at  aspartyl protease family protein
15	At5g59170	0.960 247762_at  proline-rich family protein
16	At5g50260	0.957 248545_at  cysteine proteinase, putative
17	At3g27660	0.956 258240_at  OLEO4_OLEO4 (OLEOSIN4)
18	At3g49300	0.954 252253_at  proline-rich family protein
19	At5g62800	0.953 247421_at  seven in absentia (SINA) family protein
20	At4g25140	0.953 254095_at  OLEO1_OLEO1 (OLEOSIN1)
52	At3g26790	0.920 258258_at  FUSCA 3; DNA binding / transcription factor
59	At5g10140	0.916 250476_at  FLC_AGL25_FLF_(FLOWERING LOCUS C)
60	At3g20910	0.914 256806_at  CCAAT-binding transcription factor (CBF-B/NF-YA) family protein
64	At4g31060	0.909 253575_at  AP2 domain-containing transcription factor, putative
72	At5g17430	0.903 246426_at  DNA binding / transcription factor
74	At2g40370	0.901 263825_at  Laccase 5; copper ion binding / oxidoreductase
76	At2g27070	0.901 250612_s_at  ARABIDOPSIS RESPONSE REGULATOR 13; transcription factor/ two-component response regulator
80	At5g10510	0.894 250426_at  AIL6_AIL6 (AINTEGUMENTA-LIKE 6)
87	At3g13540	0.884 256985_at  AtM2_AtMYB5_ATMYB5__ATMYB5 (MYB DOMAIN PROTEIN 5); DNA binding / transcription factor
88	At5g57390	0.884 247904_at  AINTEGUMENTA-LIKE 5; DNA binding / transcription factor

Table 2.4: List of selected genes co-expressed with *TT10* - developmental set.

The genes co-expressed with *TT10* found with Expression Angler in the AtGenExpress Developmental Set. The first 20 genes are listed, followed by the selection of putative transcription factors position number in grey background. Other member of the laccase family have also been found (position 74).

## 2.2.4 Analysis of *TT10* expression in mutants involved in flavonoid biosynthesis

Although *TT10* is involved in one of the last known steps of flavonoid biosynthesis pathway, it is rather considered as a modification enzyme (Pourcel et al., 2005). Nevertheless, the regulation of *TT10* expression could be related to some of the known biosynthetic genes. *TT10* mRNA accumulation was analyzed in the mutants known to regulate flavonol and tannin biosynthetic genes. MYB11/PFG2, MYB12/PFG1 and MYB111/PFG3 have been shown to be involved in flavonol biosynthesis in vegetative tissues (Mehrtens et al., 2005; Stracke et al., 2007). Recently MYB12 has been shown to be necessary for quercetin-3-O-rhamnoside and biflavonol accumulation, but not tannins in *Arabidopsis* seeds (Jean-Marc Routaboul, unpublished results). We have also included the mutant and over-expressor of *AtMYB4*, which encodes a transcriptional

repressor of the C4H gene, involved in the biosynthesis of flavonoid precursors (Meissner et al., 1999; Jin et al., 2000). The *myb12* mutant accumulates almost 20% less *TT10* mRNA than the control wild-type plants and similar results were obtained for the triple mutant *myb11,12,111* (Fig. 2.15). However, the variation between the biological repetitions can reach up to 25%, therefore smaller differences should be considered as not significant before confirmation with second biological repetition. *TT10* expression is almost 30% reduced in over-expressor of AtMYB4, but not modified in the *myb4* mutant. This weak difference in the over-expressor, could be due to lower of the flavonoid precursor. *TT10* expression in *tt1*, *tt16* and *ttg2* mutants was already analyzed (Pourcel, 2006), but semiquantitative RT-PCR did not show significant differences. We confirmed those results with quantitative RT-PCR. As explained before (section 2.2.1), *TT10* expression could be regulated by those mutants in the endothelium, but we could still detect transcript present in the outer integument or the replum. The metabolomic analysis of those mutants does not correspond to the *tt10* mutant phenotype, and it is described elsewhere (Routaboul et al., 2006; Stracke et al., 2007).

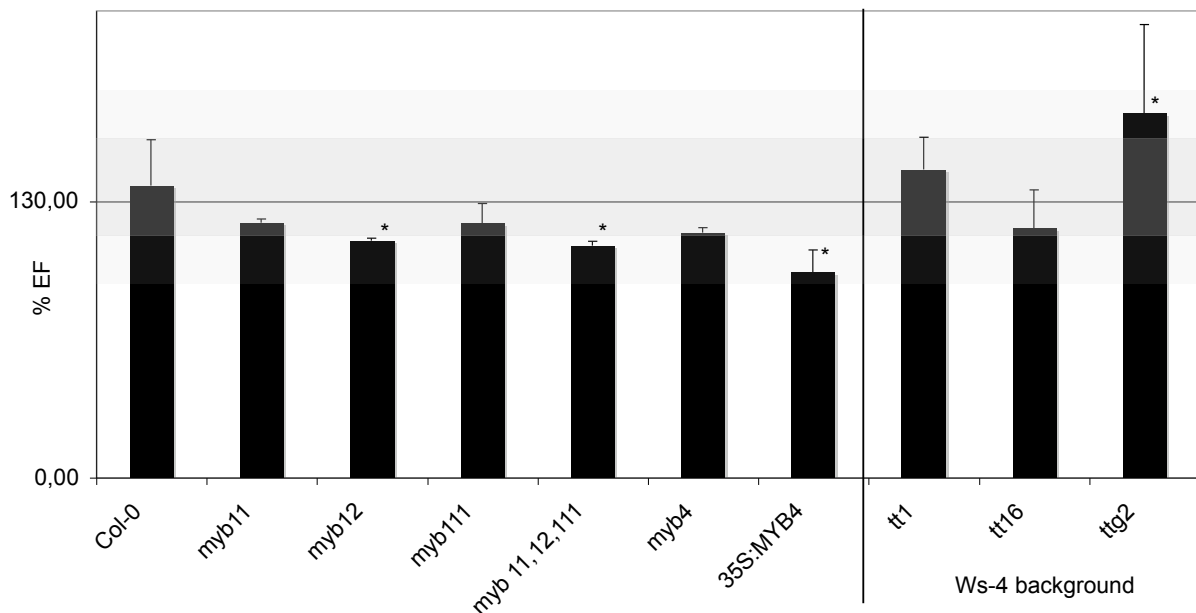


Figure 2.15: *TT10* transcript accumulation in candidate regulatory mutant backgrounds. *TT10* gene expression detected by quantitative RT-PCR is presented as a percentage of the expression of the reference gene *EF1*. Graph represents values for one biological repetition and averages  $\pm$  SE of three technical replicates. Asterisks highlight samples where *TT10* expression was more than  $\pm$  1SE different from the control.

### 2.2.5 Candidate genes related to stress response which could be involved in the regulation of *TT10* expression

The interest in the possibility of a stress-related regulation of the *TT10* emerged from the *in silico* analysis of the promoter sequence (section 2.1.2). Briefly, it was considered, that *TT10* expression during seed development could be related to the stress response and/or protective function of the flavonoids. Expression in response to biotic and abiotic stress could depend on the crosstalk between hormone signaling pathways regulated by abscisic acid, salicylic acid (SA), jasmonic acid (JA) and ethylene, as well as ROS signaling pathways (Fujita et al., 2006). Based on this assumption and putative *cis*-acting regulatory elements found in the sequence of the *TT10* promoter, we decided to study *TT10* expression in several mutants which

have been shown to be key components of the SA, JA and ethylene signaling networks. Expression of many pathogen related (PR) genes in response to salicylic acid can be directly regulated by WRKY70 (Li et al., 2004, 2006), or mediated by NPR1. In response to pathogen attack, many other WRKY transcription factors are activated by NPR1/TGA complex (Eulgem and Somssich, 2007). We included AtWRKY23 in our study, because plants overexpressing that gene accumulate more quercetin-3-O-rhamnoside, that could indicate involvement in flavonol metabolism and link their role with stress response (Godelieve Gheysen, personal communication). Recently AtWRKY23 has been shown to be involved in pathogenesis of cyst nematode *Heterodera shachtii* (Grunewald et al., 2008). Another reason to investigate involvement of SA are recent studies, demonstrating the importance of SA for seed quality, *e.g.* protein translation, priming of seed metabolism, synthesis of antioxidant enzymes and mobilization of seed storage proteins (Rajjou et al., 2006). Part of this study was carried out with plants carrying the bacterial gene *NahG*, encoding SA hydrolase. Those plants are deprived of an active form of salicylic acid (Rajjou et al., 2006) and for that reason they were included in the present study. To question the role of the stress response which could be also mediated by jasmonic acid, two mutants were used. The first one, called *jasmonate resistant 1 (jar1)* is impaired in JA perception (Staswick et al., 1992). Loss of function mutants are defective in a variety of responses to jasmonic acid. JAR1 is involved in pathogen defense, sensitivity to ozone, and wound responses (Lorenzo and Solano, 2005). The second one, coding for bHLH family transcription factor is AtMYC2 known as *jasmonate insensitive 1 (jin1)*. It has been shown to be a key component of the JA pathway and to act as a general integrator of different environmental stresses (Lorenzo et al., 2004; Lorenzo and Solano, 2005). We have also used the *constitutive triple response1 (ctr1)* mutant, which is a negative regulator of the ethylene signaling pathway, to check if *TT10* expression could be regulated by this hormone (Kieber et al., 1993; Guo and Ecker, 2004).

The amount of *TT10* transcript was not modified in those mutants, except for *npr1-1* and *npr1-5* where it was reduced (Fig. 2.16). Surprisingly even in the *ctr1* mutant plant, which had reduced growth phenotype and small siliques, *TT10* accumulated like in the control wild-type plants. We have reproducibly observed 50-70% decrease of *TT10* mRNA level in *npr1-1* and *npr1-5*. Surprisingly, another allele *npr1-2* is not showing any difference. However, it has to be noticed that *npr1-5* is in Nossen accession. Taking into account the natural variation demonstrated for *TT10* transcript accumulation (see 2.3.1) appropriate control for that sample is missing.

The analysis of flavonoid composition of mature seeds did not revealed any significant differences with the wild type for *wrky23*, *wrky70* and *ctr1* mutants (Fig. 2.17). Surprisingly, the *jar1* mutant accumulates less quercetin-3-O-rhamnoside and biflavonols and exhibits no significant difference in soluble PAs. It suggests that the expression of flavonol biosynthetic genes or any other process leading to their accumulation is affected in that mutant.

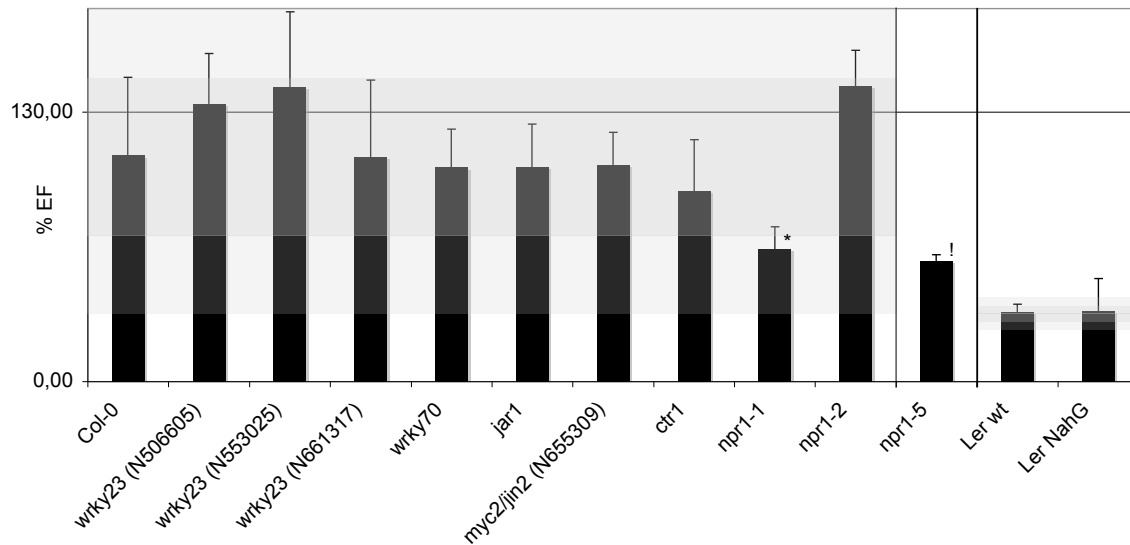


Figure 2.16: *TT10* transcript accumulation in candidate regulatory mutant backgrounds - stress related candidates.

*TT10* gene expression detected by quantitative RT-PCR is presented as a percentage of the expression of the reference gene *EF1*. Values represent averages  $\pm$  SE of three technical replicates. Asterixes highlight samples where *TT10* expression was more than  $\pm$  1SE different from the control. Experiments was repeated twice. (!) is pointing at the *npr1-5* mutant which is in Nossen background.

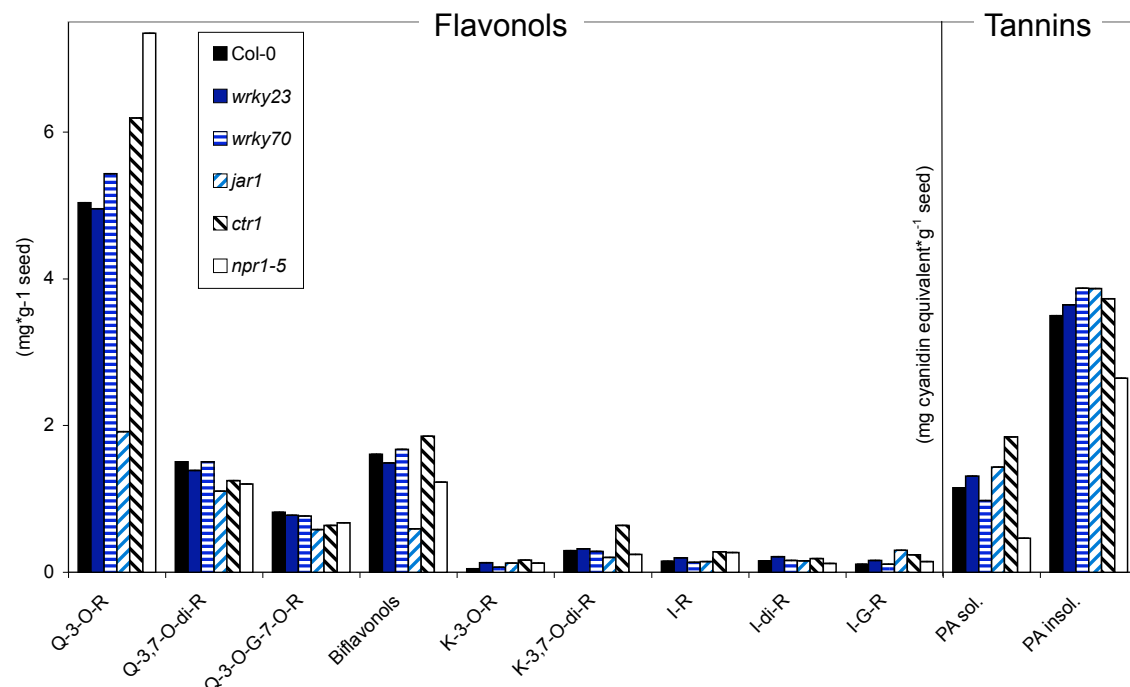


Figure 2.17: Flavonoid composition of mature seeds of the mutants - stress related candidates. Analysis of flavonol derivatives and tannins by LC-MS and after acid-catalyzed hydrolysis, respectively. Plants were grown in the greenhouse at the same time. Values represents averages  $\pm$  SE of three independent measurements. G, glucoside; I, isorhamnetin; K, kaempferol; Q, quercetin; R, rhamnoside. The *wrky23* allele is N661317.

## 2.2.6 Yeast one hybrid screening - fishing out transcription factors binding *TT10* promoter

In parallel to the above-described approaches, a transcription factor library created by the REGIA consortium (Paz-Ares and Consortium, 2002) was planned to be screened to find transcription factors binding to the *TT10* promoter. For yeast one-hybrid experiments, the *pTT10\_3* promoter fragment found in the dissection experiment to be required for GUS activity in the seed coat (see 2.1.3.1) was used. Briefly, the promoter fragment was PCR-amplified with primers including *SmaI* and *SacII* restriction sites (Appendix Tab. 5.5) and subcloned into *pCR®-Blunt II-Topo®*. The *SmaI-SacII* fragments were cloned into the MCS site of the *pHISi* vector in front of the *HIS3* gene. Integrity of the constructs was confirmed by restriction analysis and sequencing. In the next step, *Saccharomyces cerevisiae*, haploid MAT $\alpha$  yeast strain EGY48 was transformed. The procedure requires that this strain can form diploid form with the yeast strains of the REGIA library clones carrying the cDNA of the transcription factor (see 3.1.3.2 and 3.1.6). The diploid clones carrying the transcription factor binding to the sequence of interest - *pTT10\_3* promoter fragment, could express the *HIS3* gene and grow on the selective medium. However, in many cases, one of the native yeast transcription factors can activate this expression (autoactivation). For that reason, before library screening, it is necessary to check, if yeast can grow on the selective medium. For the *TT10* promoter fragment studied, all the yeast clones tested were growing on the selective medium SD-*ura-his* (see Fig. 2.18) Still, in the case of autoactivation, sometimes it is possible to screen the library, if the interaction is not too strong and 3-amino-1,2,4-triazole (3-AT) a competitive inhibitor of the *HIS3* gene product can be used to control leaky expression of *HIS3* gene. However, in the case of the *TT10* minimal promoter, even 5mM concentration of 3-AT was not sufficient. To overcome this problem, smaller promoter fragments could be used. Promoter region considered to be required for the seed coat specificity and other required for the remaining GUS activity in the funiculus were cloned and tested as described above. Those two fragments also resulted in the yeast clones growing on the selective medium, even in the presence of 5mM 3-AT. The *TT10* promoter fragments tested were not compatible with the yeast one-hybrid system and it was not possible to screen the REGIA library of transcription factors (Fig. 2.18).

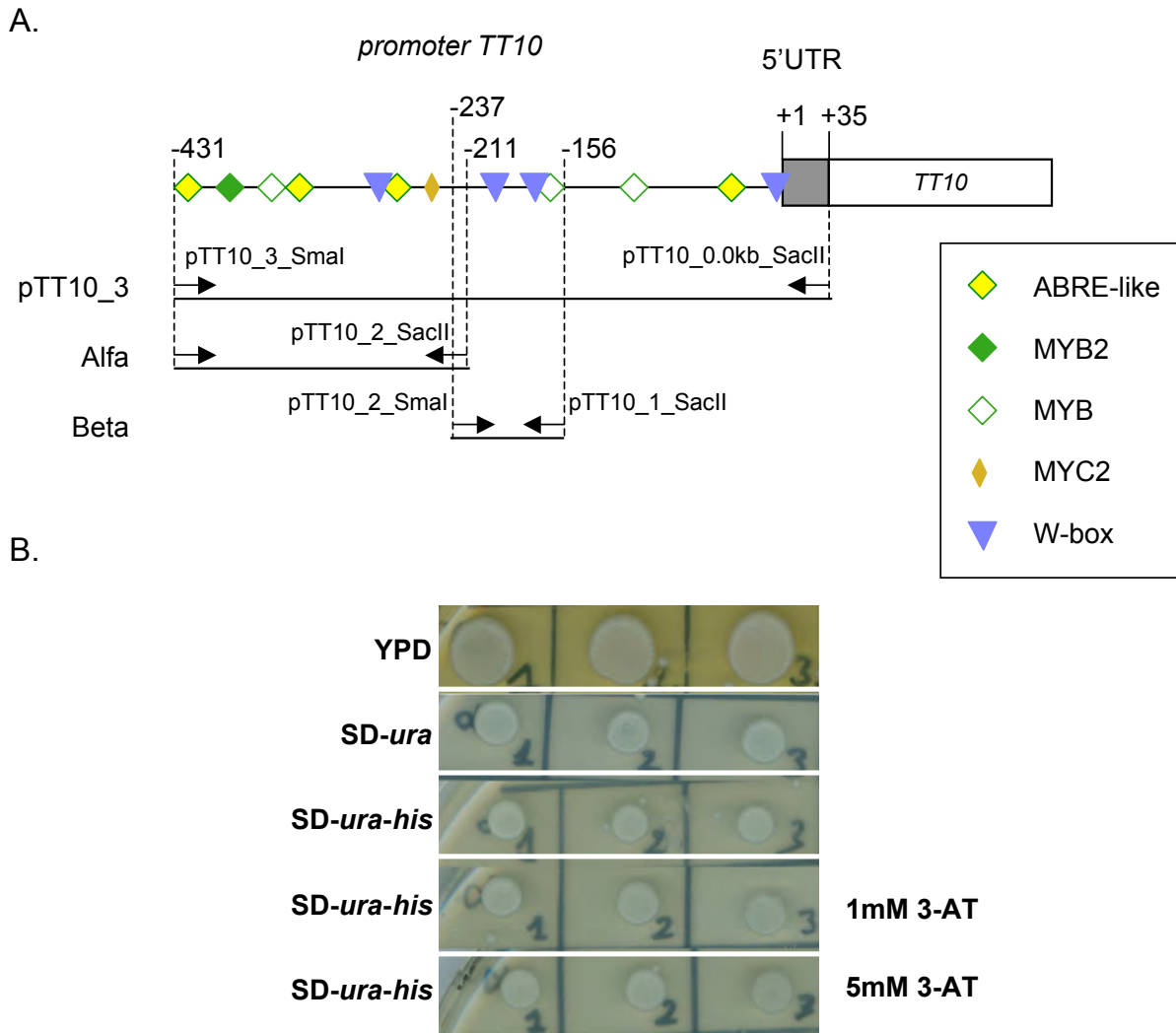


Figure 2.18: Summary of yeast one-hybrid experiment.

(A) Representation of the *TT10* promoter region. Selected putative *cis*-acting regulatory elements are presented with symbols. Alfa and Beta - promoter regions required for the seed coat and funiculus activity respectively. Primers used for cloning are represented with arrows.

(B) EGY48 yeast strain carrying *pTT10\_3/pHISi* vector integrated in the *ura3* genomic region. Growth on the full (YPD) and selective medium (SD-*ura*, SD-*ura-his*). 3-AT, 3-amino-1,2,4-triazole.



## 2.3 Insights from natural variation

Differences for *TT10* transcript accumulation among accessions have been demonstrated by Pourcel et al. (2005), showing low level of *TT10* mRNA in *Ler* in comparison with Col-0 and Ws-4. Recently, two populations of recombinant inbred lines (RILs), Cvi-0 x Col-0 and Bay-0 x Sha were studied for flavonoid composition in mature seeds. Consistent with the results of quantitative trait locus (QTL) mapping, a gene in the *TT10* interval was found to be responsible for the variation of insoluble proanthocyanidin, quercetin-3-O rhamnoside and biflavonol accumulation (Jean-Marc Routaboul, unpublished results). Three more accessions (Cvi-0, Bay-0 and Sha) were included in this study.

The objective of this part of the work, was to unravel the molecular basis of the observed differences in *TT10* transcript accumulation. Gene expression and therefore promoter regions are thought to be fundamental factors in adaptive evolution. It has been demonstrated that functional changes in promoters can arise from a few mutations and that the promoter regions are major determinant of the functional genetic variation (Bentsink et al., 2006; de Meaux et al., 2006). As a starting point we assumed that differences in *TT10* transcript accumulation would possibly be due to polymorphism in its promoter sequence. *TT10* promoter activity was studied as well as the link between transcript accumulation and flavonoid composition.

### 2.3.1 Differences in *TT10* transcript accumulation

The *TT10* transcript accumulation was studied in 6 *Arabidopsis* accessions using quantitative Real Time PCR (qRT-PCR). This technique was preferred for its high sensitivity to detect small differences in gene expression. The expression of flavonoid biosynthetic genes and flavonoid accumulation are known to be affected by environmental conditions (*e.g.* light). For that reason, plants were grown at the same time in soil in the greenhouse in order to collect appropriate material. Open flowers were marked and siliques which were developed from them were collected 7 days after flowering (DAF). This time point was chosen because of the the first maximum of *TT10* expression in siliques (see section 2.1.1 and 3.2.5).

The qRT-PCR analysis revealed that *TT10* expression level is within that of the reference gene *ELONGATION FACTOR 1aA4* in *Ler*, Cvi-0 and Sha accessions, whereas it was twice more in Col-0 and Ws-4 and almost twice more in Bay-0 (Fig. 2.19 A.). The differences observed cannot be explained by primer efficiency which is comparable in all accessions (Appendix, Table 3.7), nor by the differential expression of the reference gene *EF1*. Similar results were obtained when *TT10* expression was normalized with two other reference genes, *APT1* (adenine phosphoribosyltransferase) and *TUB4a* (tubuline beta-4 chain). Expression of the *APT1* and *TUB4A* in siliques is presented in Appendix, Figure 3.3. The whole experiment was repeated three times resulting in nine data points used for the Student *t*-Test statistical analysis (Kirkman, 1996). For Ws-4, Col-0 and Bay-0, seven points were considered, because two others were out of the two standard deviation difference range from the mean. The experiment was accompanied by comparison of the developmental stages of the seeds in the silique material used for the RNA extraction (Fig. 2.19 B.). The microscopic analysis revealed no striking differences in embryo development, which were in the range of an early to late heart stage in all accessions. Plants from which samples were collected for the qRT-PCR were further grown for analysis of flavonoid composition in mature seeds (see 2.3.3).

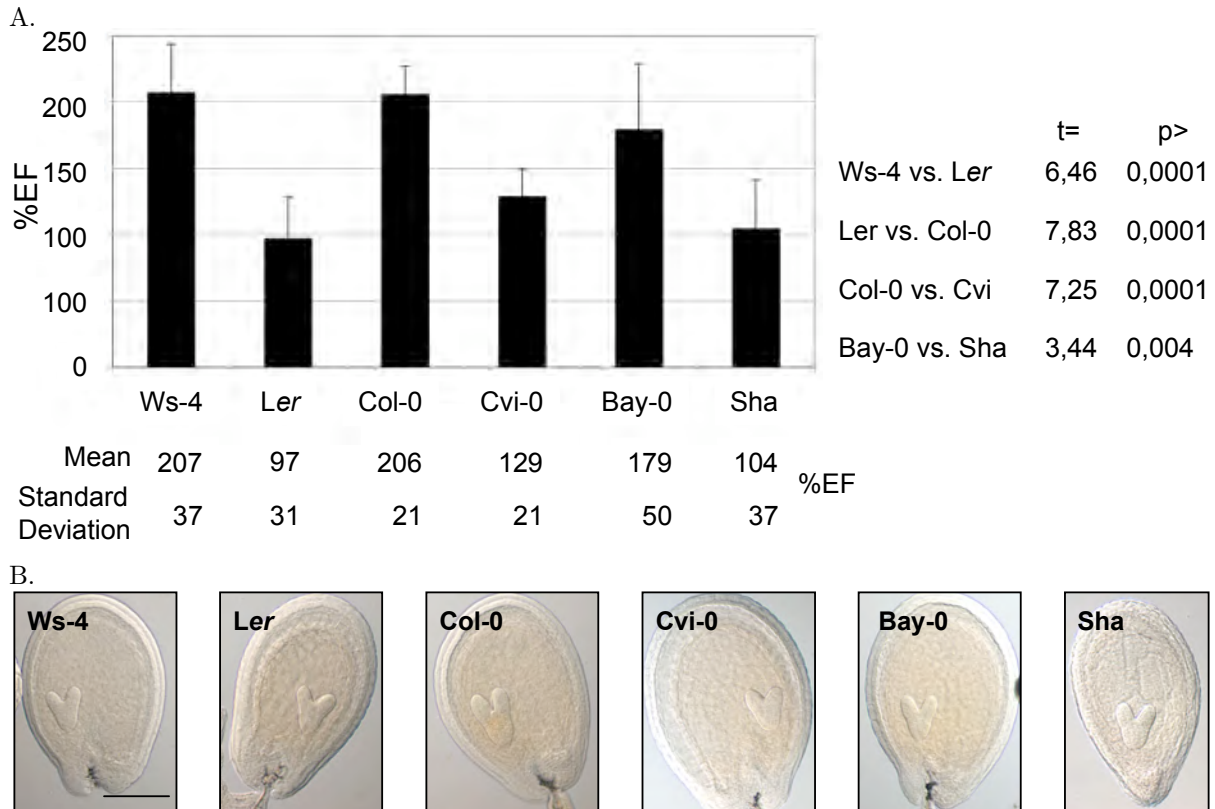


Figure 2.19: Assessment of *TT10* expression levels in various *Arabidopsis* accessions.

(A) Expression of *TT10* in various accessions, in siliques 7DAF measured by quantitative RT-PCR, presented as a percentage of the expression of the reference gene *EF1*. Mean and standard deviation values are given for each accession below the graph. The results of the Student *t*-Test are given for the accessions compared. Details of the analysis are in the text and section 3.2.5.

(B) Seeds were extracted from the siliques at 7 DAF and observed with Nomarski optics after mounting in chloralhydrate. The developmental stages are similar between accessions. Size bar = 90  $\mu$ m.

### 2.3.2 Analysis of promoter sequences

To try to explain the differences observed in *TT10* transcript accumulation, it was decided to compare the 2.0-kb region of the promoter which corresponds to the fragment studied with the GUS reporter gene. This region should contain all the *cis*-regulatory elements required not only for the tissue specificity, but also other elements, which could affect the level of gene expression. The primers used for the cloning of the 2.0-kb *TT10* promoter (Pourcel et al., 2005) did not result in PCR amplification on the genomic sequence of any accession except the positive control Col-0. Multiple attempts and optimizations did not resolve this problem. For details of various modifications of the PCR conditions, different polymerases and genomic DNA extractions see 3.2.4 and 3.2.4.5.

With the set of primers used for promoter dissection (2.1.3.1), it was possible to amplify the *pTT10\_3* - 465bp fragment, but not the 1.0-kb or 1.5-kb. PCR amplified fragments were cloned into *pCR®-Blunt II-Topo®* and sequenced. Two clones from two genomic DNA extractions for each of the accessions were sequenced to result in the consensus sequences aligned in Figure 2.20 and discussed in the text below. Two other approaches used to amplify 2.0-kb or longer *TT10* promoter region are described later (section 2.3.2.3).

### 2.3.2.1 Focus on the proximal part of the promoter

Analysis of the 465bp sequence of the *TT10* promoter of 6 *Arabidopsis* accessions: Ws-4, *Ler*, Col-0, Cvi-0, Bay-0 and Sha, revealed no major differences. Several common single nucleotide polymorphisms (SNPs) and one in-del were found. The analysis of the motifs affected by the SNPs is based on the *in silico* analysis and is referring to the Col-0 sequence. Each substitution is mentioned below, with the position in the promoter, the accession and eventually the consequences it could have.

Deletion in position -384bp (Ws-4, *Ler* and Bay-0) is located in the MYB2 *cis*-acting regulatory element, nevertheless this motif is still recognized *in silico* in this sequence. The T to C substitution in -330bp (*Ler*) position results in a loss of the I-box recognition. The next SNP, G to A substitution in position -262 results in loss of the MYC2 site which was found in Col-0, but not in the other accessions. Substitution C to T at -194 could change the context of the W-box recognition in Shahdara (Sha). Two base pair variation GG changed to TT -170/169 (Ws-4, *Ler*, Bay-0 and Cvi-0) results in the recognition of a new element called ANAERO1CONSENSUS (AAACAAA). This motif was found *in silico* in promoters of 13 genes induced in anaerobic conditions (Mohanty et al., 2005). Two next SNPs common for Ws-4, *Ler*, Bay-0 and Cvi-0, (T to C, -158; A to T, -146) are in the region of three *cis*-acting regulatory elements located nearby and result in the disappearance of the ACTCAT and MYB motifs, and modified context of the W-box. Substitution C to T in Ws-4 at position -100 gives a new element found in PLACE database and called LECPLEACS2 (TAAAATAT), which could be involved in ethylene response (Matarasso et al., 2005). The nearby T to A (-87; Ws-4, *Ler*, Bay-0 and Cvi-0) modification is changing the context of the MYB site, but the MYBCORE motif is still recognized. The G to A substitution (-34) abolished recognition of the SORLREP3AT motif in Bay-0 accession. One SNP was found in the 5'UTR region at position +8 (G to A) and it could change context of the W-box and GT1CONSENSUS motifs. The T to G substitution at -277 (Cvi-0) position does not seem to affect any known *cis*-acting regulatory element. All the SNPs described are indicated with the asterixes on the sequence alignment below.

```

-431bp
Col-0  ATTAACACAA AAACGTAACG AGTTTCGGAA TATTTCTGGT GTTGGTTTAT
Sha    ATTAACACAA AAACGTAACG AGTTTCGGAA TATTTCTGGT GTTGGTTTAT
Ws-4   ATTAACACAA AAACGTAACG AGTTTCGGAA TATTTCTGGT GTTGGTT-AT
Ler    ATTAACACAA AAACGTAACG AGTTTCGGAA TATTTCTGGT GTTGGTT-AT
Bay-0  ATTAACACAA AAACGTAACG AGTTTCGGAA TATTTCTGGT GTTGGTT-AT
Cvi-0  ATTAACACAA AAACGTAACG AGTTTCGGAA TATTTCTGGT GTTGGTTTAT
Consensus ATTAACACAA AAACGTAACG AGTTTCGGAA TATTTCTGGT GTTGGTTtAT
                                         Δ

-381bp
Col-0  ATAAAATAAA GGTAATGCTT GGCCTAACAA AACTCATAAG TGATTTTATT
Sha    ATAAAATAAA GGTAATGCTT GGCCTAACAA AACTCATAAG TGATTTTATT
Ws-4   ATAAAATAAA GGTAATGCTT GGCCTAACAA AACTCATAAG TGATTTTATT
Ler    ATAAAATAAA GGTAATGCTT GGCCTAACAA AACTCATAAG TGATTTTATT
Bay-0  ATAAAATAAA GGTAATGCTT GGCCTAACAA AACTCATAAG TGATTTTATT
Cvi-0  ATAAAATAAA GGTAATGCTT GGCCTAACAA AACTCATAAG TGATTTTATT
Consensus ATAAAATAAA GGTAATGCTT GGCCTAACAA AACTCATAAG TGATTTTATT

-331bp
Col-0  TTTATCTTTAG TAAAGACTTT GAACTAGTTG ATTGATGTCA CGTCTCTTAC
Sha    TTTATCTTTAG TAAAGACTTT GAACTAGTTG ATTGATGTCA CGTCTCTTAC
Ws-4   TTTATCTTTAG TAAAGACTTT GAACTAGTTG ATTGATGTCA CGTCTCTTAC
Ler    TCATCTTTAG  TAAAGACTTT GAACTAGTTG ATTGATGTCA CGTCTCTTAC
Bay-0  TTTATCTTTAG TAAAGACTTT GAACTAGTTG ATTGATGTCA CGTCTCTTAC
Cvi-0  TTTATCTTTAG TAAAGACTTT GAACTAGTTG ATTGATGTCA CGTCTCTTAC
Consensus TtATCTTTAG TAAAGACTTT GAACTAGTTG ATTGATGTCA CGTCTCTTAC
                                         *

```

-281bp				←MYC2→			
Col-0	TATATATGAT	ATAACATGTG	AAAATATAAT	CTAGAGTTAG	ATAAATACGA		
Sha	TATATATGAT	ATAACATGTA	AAAATATAAT	CTAGAGTTAG	ATAAATACGA		
Ws-4	TATATATGAT	ATAACATGTA	AAAATATAAT	CTAGAGTTAG	ATAAATACGA		
Ler	TATATATGAT	ATAACATGTA	AAAATATAAT	CTAGAGTTAG	ATAAATACGA		
Bay-0	TATATATGAT	ATAACATGTA	AAAATATAAT	CTAGAGTTAG	ATAAATACGA		
Cvi-0	TATAGATGAT	ATAACATGTA	AAAATATAAT	CTAGAGTTAG	ATAAATACGA		
Consensus	TATAtATGAT	ATAACATGTa	AAAATATAAT	CTAGAGTTAG	ATAAATACGA		
	*		*				
-231bp				W-box→			
Col-0	CTTTACTGTT	TTCGTAACT	TGAAAGTTTC	GTTTGACCAA	GAAATGTAGT		
Sha	CTTTACTGTT	TTCGTAACT	TGAAAGTTTC	GTTTGACTAA	GAAATGTAGT		
Ws-4	CTTTACTGTT	TTCGTAACT	TGAAAGTTTC	GTTTGACCAA	GAAATGTAGT		
Ler	CTTTACTGTT	TTCGTAACT	TGAAAGTTTC	GTTTGACCAA	GAAATGTAGT		
Bay-0	CTTTACTGTT	TTCGTAACT	TGAAAGTTTC	GTTTGACCAA	GAAATGTAGT		
Cvi-0	CTTTACTGTT	TTCGTAACT	TGAAAGTTTC	GTTTGACCAA	GAAATGTAGT		
Consensus	CTTTACTGTT	TTCGTAACT	TGAAAGTTTC	GTTTGACcAA	GAAATGTAGT		
				*			
-181bp			ACTCAT	←W-box	MYB→		
Col-0	TAAATATTTG	TGGGAATGTG	AACTCATTAG	TCAAAACGGA	CCAAGAAGAT		
Sha	TAAATATTTG	TGGGAATGTG	AACTCATTAG	TCAAAACGGA	CCAAGAAGAT		
Ws-4	TAAATATTTG	TTTGAATGTG	AACCCATTAG	TCAAATCGGA	CCAAGAAGAT		
Ler	TAAATATTTG	TTTGAATGTG	AACCCATTAG	TCAAATCGGA	CCAAGAAGAT		
Bay-0	TAAATATTTG	TTTGAATGTG	AACCCATTAG	TCAAATCGGA	CCAAGAAGAT		
Cvi-0	TAAATATTTG	TTTGAATGTG	AACCCATTAG	TCAAATCGGA	CCAAGAAGAT		
Consensus	TAAATATTTG	TttGAATGTG	AACcCATTAG	TCAAAtCGGA	CCAAGAAGAT		
		**	*	*			
-131bp					MYB		
Col-0	GATGCTTACC	GAAAACATAT	TTTGTA AAAAT	ACAAAATTCT	TATATCCGAA		
Sha	GATGCTTACC	GAAAACATAT	TTTGTA AAAAT	ACAAAATTCT	TATATCCGAA		
Ws-4	GATGCTTACC	GAAAACATAT	TTTGTA AAAAT	ATAAAAATTCT	TATAACCGAA		
Ler	GATGCTTACC	GAAAACATAT	TTTGTA AAAAT	ACAAAATTCT	TATAACCGAA		
Bay-0	GATGCTTACC	GAAAACATAT	TTTGTA AAAAT	ACAAAATTCT	TATAACCGAA		
Cvi-0	GATGCTTACC	GAAAACATAT	TTTGTA AAAAT	ACAAAATTCT	TATAACCGAA		
Consensus	GATGCTTACC	GAAAACATAT	TTTGTA AAAAT	AcAAAATTCT	TATAaCCGAA		
				*	*		
-81bp					SORLRREP3		
Col-0	TGTAAC TAAT	TTGAGAAAGA	CGGTCTCTAT	ACGTATATAA	AATATATGTA		
Sha	TGTAAC TAAT	TTGAGAAAGA	CGGTCTCTAT	ACGTATATAA	AATATATGTA		
Ws-4	TGTAAC TAAT	TTGAGAAAGA	CGGTCTCTAT	ACGTATATAA	AATATATGTA		
Ler	TGTAAC TAAT	TTGAGAAAGA	CGGTCTCTAT	ACGTATATAA	AATATATGTA		
Bay-0	TGTAAC TAAT	TTGAGAAAGA	CGGTCTCTAT	ACGTATATAA	AATATATATA		
Cvi-0	TGTAAC TAAT	TTGAGAAAGA	CGGTCTCTAT	ACGTATATAA	AATATATGTA		
Consensus	TGTAAC TAAT	TTGAGAAAGA	CGGTCTCTAT	ACGTATATAA	AATATATgTA		
					*		
-31bp	SORLRREP3			←W-box	GT1CONS	SENSUS	
Col-0	TATATGCCTC	TCTATTCTTT	GATGAACTTC	AGTCAGGGTA	ATTAATTTAC		
Sha	TATATGCCTC	TCTATTCTTT	GATGAACTTC	AGTCAGGATA	ATTAATTTAC		
Ws-4	TATATGCCTC	TCTATTCTTT	GATGAACTTC	AGTCAGGATA	ATTAATTTAC		
Ler	TATATGCCTC	TCTATTCTTT	GATGAACTTC	AGTCAGGATA	ATTAATTTAC		
Bay-0	TATATGCCTC	TCTATTCTTT	GATGAACTTC	AGTCAGGATA	ATTAATTTAC		
Cvi-0	TATATGCCTC	TCTATTCTTT	GATGAACTTC	AGTCAGGATA	ATTAATTTAC		
Consensus	TATATGCCTC	TCTATTCTTT	GATGAACTTC	AGTCAGGAtA	ATTAATTTAC		
				+1	*		
+20bp		+35					
Col-0	TAAAAC TCTT	CCAAA ATG					
Sha	TAAAAC TCTT	CCAAA ATG					
Ws-4	TAAAAC TCTT	CCAAA ATG					
Ler	TAAAAC TCTT	CCAAA ATG					
Bay-0	TAAAAC TCTT	CCAAA ATG					
Cvi-0	TAAAAC TCTT	CCAAA ATG					
Consensus	TAAAAC TCTT	CCAAA ATG					

Figure 2.20: Promoter sequence alignment and comparison of the putative CAREs present in various accessions.

Minimal promoter (*pTT10\_3* - 465bp) was sequenced from six *Arabidopsis* accessions. Asterixes (\*) are showing the single nucleotide polymorphisms (SNPs) and delta ( $\Delta$ ) is for the deletion. 5'UTR is underlined and transcription initiation site is marked +1. Putative *cis*-acting regulatory elements in the zones of the SNPs are presented, with arrows over the motif sequence, based on the Col-0 sequence analysis used as a reference. Differences due to the SNP's are discussed in the text. (For the detailed analysis of the putative CAREs, see Fig. 2.6 and section 2.1.2).

### 2.3.2.2 SNP vs. *TT10* transcript accumulation

It was mentioned at the beginning of this section, that functional changes in promoters can arise from a few mutations. It does not seem likely that at this point of the study, we could associate certain nucleotide variations observed among accessions with low or high level of *TT10* transcript accumulation.

### 2.3.2.3 Analysis of long promoter fragments

The elements required for high level of gene expression could be different from the elements required for the tissue specificity of the promoter. They are located upstream the 465bp analyzed here. The results of the MUG experiment with the dissected promoter fragments (see section 2.1.3.3) support the hypothesis that the motifs required for a strong gene activation are upstream from the *pTT10\_3* promoter fragment. Two approaches have been used to clone 2.0-kb sequence of *TT10* promoter region in various accessions.

### 2.3.2.4 BAC library screening

In a previous study (Pourcel et al., 2005), the *TT10* promoter from Columbia accession has been amplified on the MDN11 BAC template. Here, the failure to amplify the 2.0-kb promoter fragments, could be caused by the technical difficulty of the PCR on genomic DNA. Fortunately, the *Ler* and *Cvi* libraries were available from Wim Soppe's Group at the Plant Breeding and Genetics Department in Max Planck Institute of Cologne (MPIK). The BAC libraries were screened for the bacterial clone(s) carrying the *TT10* genomic regions at the MPIK.

Briefly, two probes have been prepared, one in 5' and one in 3' region of the *TT10* coding sequence. This approach allowed to ensure high specificity of screening, and simultaneously decreased the probability of cross hybridization with the sequence of the other member(s) of the laccase family. Screening resulted in 8 signals/bacterial clones for *Cvi* and 14 for *Ler*. Both BAC libraries have  $\pm 10$  time coverage of the *Arabidopsis* genome, therefore this number of hybridization signals could be an indication for good clones.

The PCR on bacterial colonies with the 3Fw/3Rev primers which are *TT10* specific, gave a product of an expected size for all of them. The BAC DNA was extracted from *Cvi* F.22/4 and *Ler* D.11/3 clones. Surprisingly, the attempts to amplify *TT10* promoter fragment on the BAC templates gave the same results as with the genomic, *pTT10\_3* being the longest fragment possible to amplify.

### 2.3.2.5 Going further in the promoter: amplification of longer fragments

Not being able to amplify bigger fragment than 465bp, could be due to the sequence polymorphism which was different from the one in the Col-0. Several primers have been designed upstream in the promoter region: -2.5-kb, -3.0-kb, -3.5-kb, -4.0-kb, -4.5-kb and -8.5-kb (CATMA primer in the 3'UTR of At5g48110 gene). The Col-0 genomic DNA was used as a template to validate primers and to set PCR conditions for AccuTaq(TM) LA DNA Polymerase (Sigma), which is suitable for long PCR amplifications up to 20-kb.

Finally, a single band product has been obtained for Ler with the pTT10 -4.5-kb primer and for Cvi with the At5g48110 CATMA F primer (Fig. 2.21). Both PCR products were smaller than the band obtained for the Col with the respective primer, suggesting insertion in the Col-0 sequence and/or deletions in *Ler* and Cvi. PCR products have been cloned into *pCR*(®)-*Blunt II-Topo*(®) for sequencing.

Sequencing of the borders of PCR fragment amplified in *Ler* accession revealed that the corresponding fragment contains *TT10* promoter region. However, before the time when all the technical difficulties were resolved, in June 2007 comparison of the 2.2-kb Columbia and 1.5-kb *Ler* sequence of the *TT10* promoter revealed more than 700bp deletion and several smaller in-dels differences between those accessions (Ishihara, 2007). The SNPs in the proximal 465bp were as described above.

Cloning of the PCR fragment amplified in Cvi accession was of very low efficiency, because of very big size. Sequencing of the PCR fragment in the only one clone obtained, revealed that the sequence did not correspond to the *TT10* promoter region. Identity of the the fragment amplified in Cvi background should be confirmed.

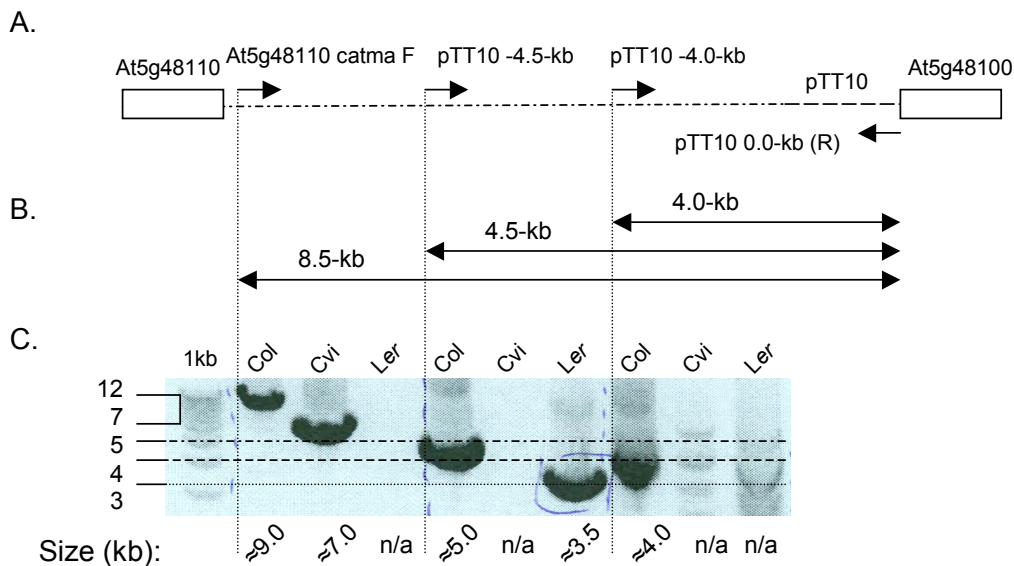


Figure 2.21: Polymorphism for length of *TT10* promoter amplification.

(A) Organization of the *TT10* genomic region in Col-0 accession. 8.5-kb intergenic region is depicted with the dashed line (not in scale), the genes with open rectangles and primers with arrows indicating their orientation. For sequence of the primers, see Appendix, Tab. 5.4).

(B) Arrows (not in scale) are representing a region which is amplified and expected PCR products size is given.

(C) Agarose Gel migration of the PCR reaction and product visualization with ethidium bromide (colors inverted). n/a - no PCR amplification or smear without clear major band. PCR templates: Col - BAC MDN11; Cvi - BAC F.22/4; Ler - BAC D.11/3. Similar results were obtained on the genomic DNA of respective accessions.

### 2.3.3 Flavonoid composition in matures seeds of various *Arabidopsis* accessions

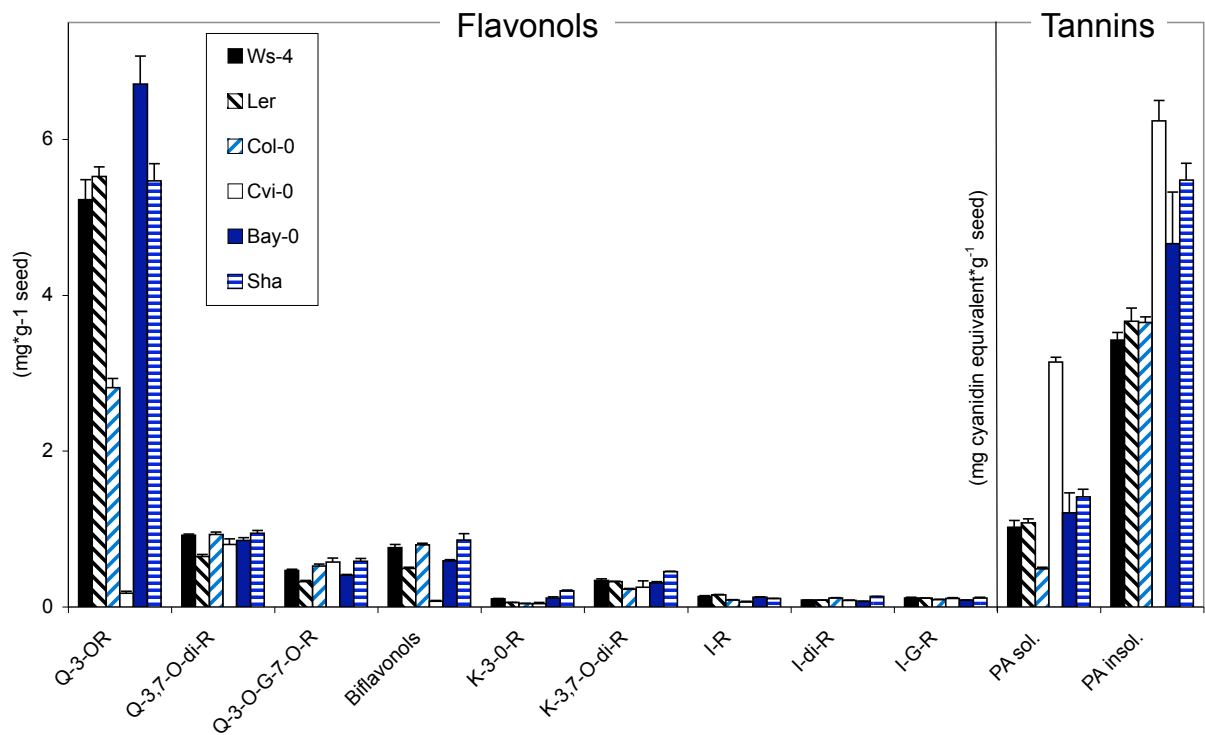
Previous studies uncovered many quantitative differences of flavonoid composition among accessions (Routaboul et al., 2006); Routaboul, unpublished results) which could be an adaptive trait of the plants living in different environments. *Transparent Testa 10* mutant (*tt10-2*) has been shown to accumulate 4.6-fold more soluble proanthocyanidins (soluble PAs) than the corresponding (Ws-4) wild-type plants. The analysis of

flavonols revealed that quercetin-3-O-rhamnoside (Q-3-OR) is 50% more, whereas biflavonols are 12-fold less abundant in *tt10-2* mature seed. Here we were trying to associate the amount of those compounds with *TT10* transcript accumulation (see section 2.3.1). Six *Arabidopsis* accessions, and seven *tt10* mutant alleles were grown at the same time in the greenhouse in the long day conditions. Flavonoids were extracted from freshly harvested seeds and measured, combining acid-catalyzed hydrolysis and liquid chromatography-mass spectrometry (LC-MS), as described elsewhere (Pourcel et al., 2005; Routaboul et al., 2006). The visual inspection of seed color, confirmed the pale-brown color of all the *tt10* alleles and pictures are presented in Appendix Fig. 5.2 (A.). Moreover, the Col-0 seeds seemed to be slightly darker than those of *Ler* and *Cvi-0*, whereas those of *Ws-4*, *Bay-0* and *Sha* slightly brighter. The amount of all the flavonoids measured is comparable among accessions tested. The exception was *Cvi-0*, which accumulated much more soluble PAs, whereas almost no Q-3-OR and biflavonols. In contrast the Col-0 accession accumulated less soluble PAs and Q-3-O-R than other accessions, but the amount of biflavonols was not affected (Fig. 2.22 (A)). All the *tt10* alleles accumulates more soluble PAs and Q-3-OR, but less biflavonols, than the corresponding wild-type plants, as expected.

### **Flavonoid composition vs. transcript accumulation**

Unfortunately we did not find correlation between the amount of the *TT10* mRNA (section 2.3.1 and Fig. 2.19) and flavonoid composition of the mature seeds (section 2.3.3 and Fig. 2.22).

A.



B.

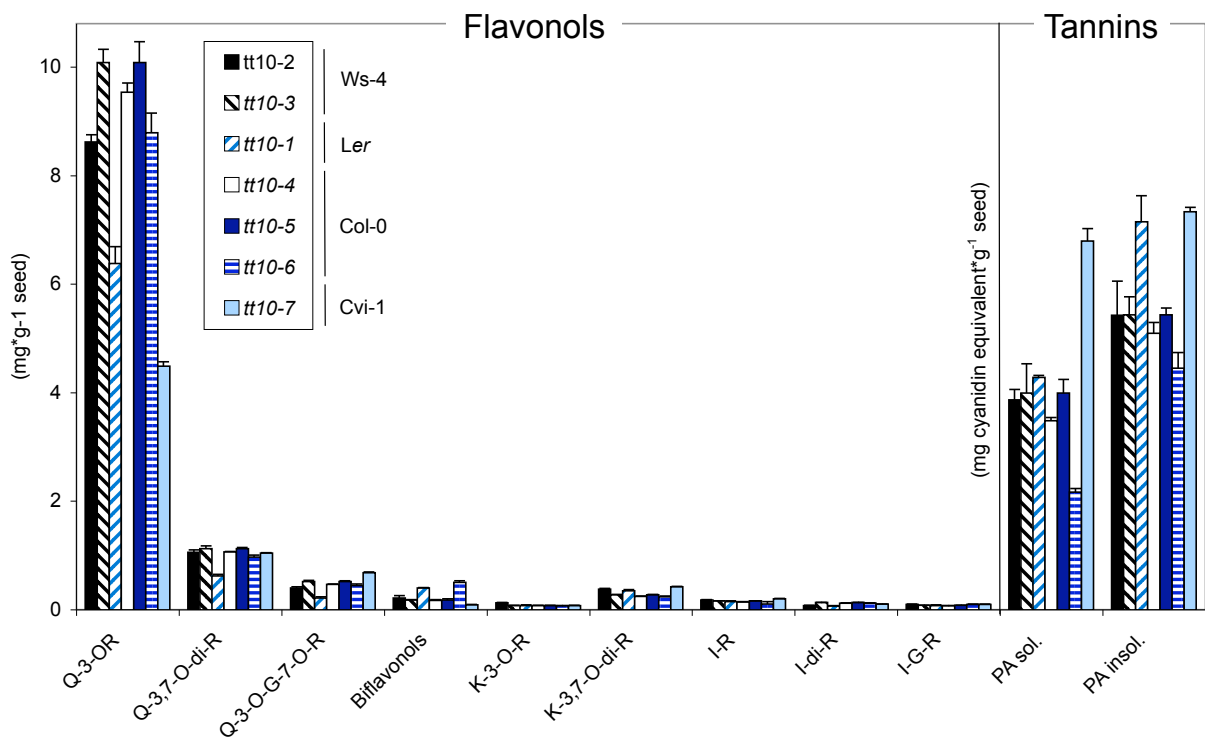


Figure 2.22: Flavonoid composition of mature seeds of various accessions and *tt10* alleles. (A) and (B) Flavonoid derivatives were analyzed by LC-MS and tannins were measured after acid-catalyzed hydrolysis. Values represent averages  $\pm$  SE of three independent measurements. G, glucoside; I, isorhamnetin; K, kaempferol; Q, quercetin; R, rhamnoside.



### 2.3.4 Functional complementation of *Ler* accession with Col-0 *TT10* genomic region

In this section, to assess if the level of *TT10* expression in *Ler* could be increased by the expression of Col-0 *TT10* gene, namely a 8.0-kb *TT10* genomic construct which was previously used to complement the *tt10* mutant (Pourcel et al., 2005). This construct contains the full *TT10* coding sequence and more than 2.0-kb promoter fragment. Analysis of the independent transformant lines shows that the *TT10* transcript amount can reach much higher levels than the one detected in *Ler* or Col-0 wild-type plants (Fig. 2.23). This could be explained by multiple T-DNA inserts. Analysis of segregation for hygromycin resistance showed that for most of the lines more than 90% of T2 plants were resistant to the antibiotic suggesting at least two loci, and only *Ler C'* was segregating 3:1 ratio (resistant : not resistant).

Flavonoid composition analysis revealed no decrease in the soluble PAs fraction in any of the transgenic lines (Fig. 2.24). Surprisingly, one of the lines (*Ler A*) contained even more soluble PAs than the control *Ler* wild-type plants. Four lines accumulated much less Q-3-O-R than *Ler* and Col-0, which could suggest that a 'complementation' occurs, but at the same time there was also a significant decrease in the biflavonol amount. The abundance of other flavonoids was not modified in the transformants lines. Transformant seeds at harvest seemed to be darker than those of the Col-0 wild-type plants (Appendix, Fig. 5.2).

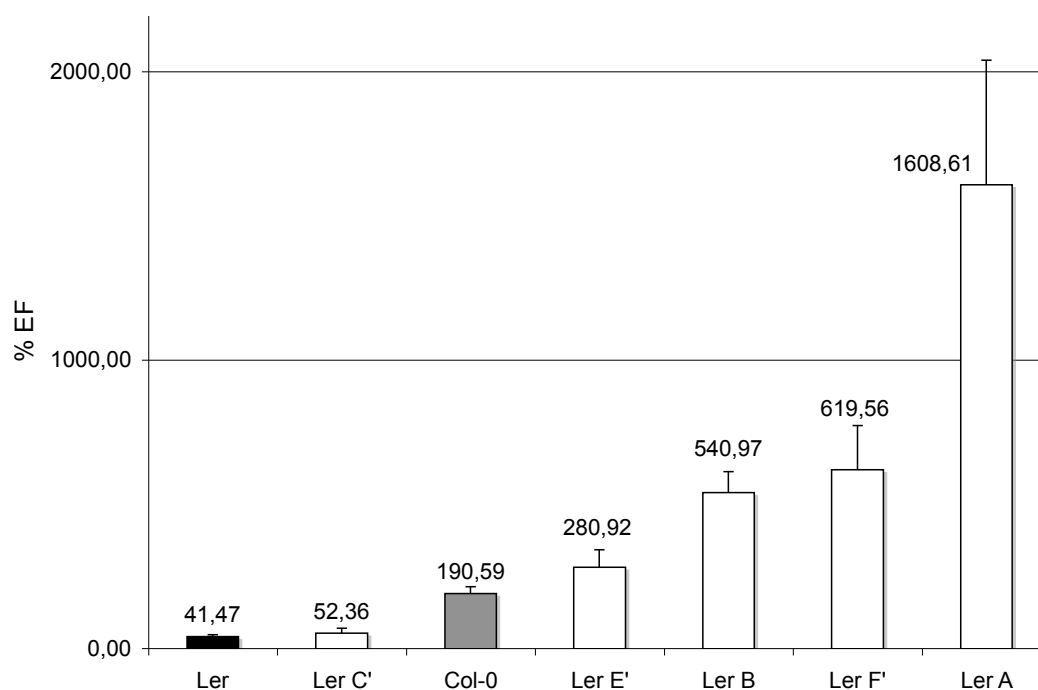


Figure 2.23: *TT10* transcript accumulation in *Ler* plants transformed with the Col-0 wild-type *TT10* genomic region.

*TT10* gene expression detected by quantitative RT-PCR, presented as a percentage of the expression of the reference gene *EF1*. The control wild-type plants are represented by the full bars: *Ler* - black, Col-0 - grey; *Ler* transformants - white/open bars. Values represent averages  $\pm$  SE of three technical replicates. Experiment was repeated twice with the same tendency of the transcript accumulation between the lines. RNA was extracted from siliques 7DAF of the T2 antibiotic resistant plants.

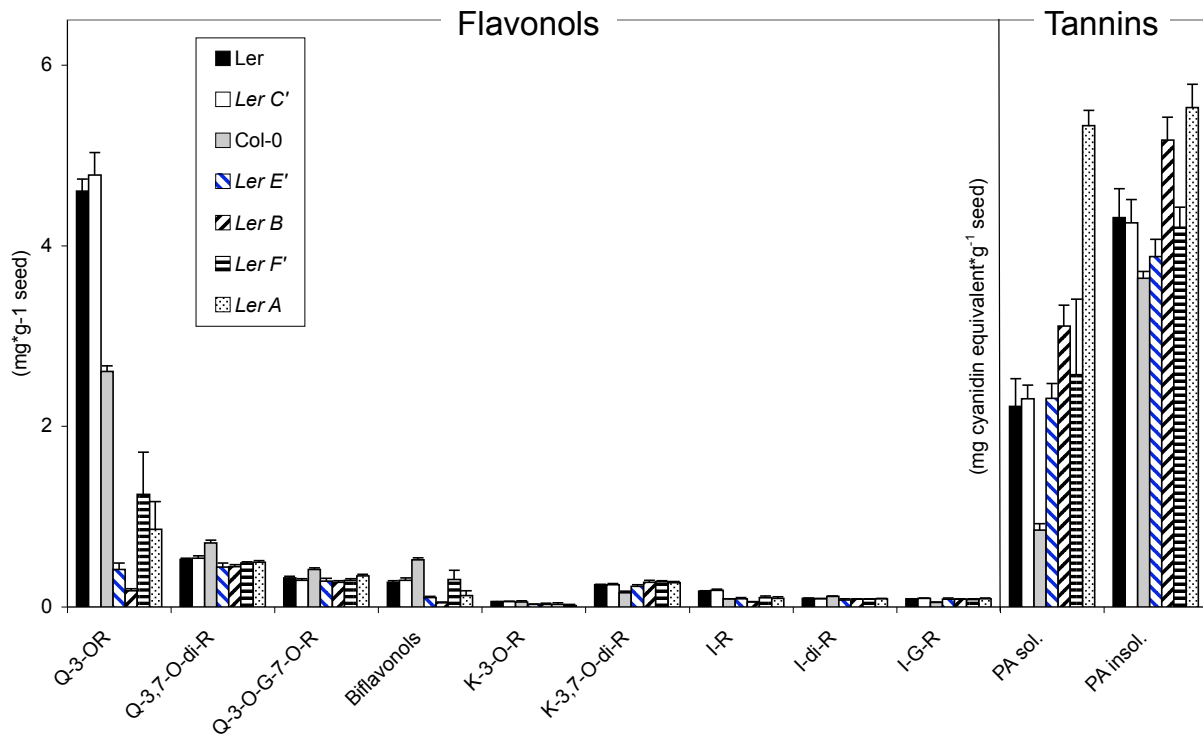


Figure 2.24: Flavonoid composition of mature seeds from *Ler* lines complemented with the Col-0 wild-type genomic region.

The order of the lines is set from left to right according to *TT10* transcript accumulation level. Analysis of flavonol derivatives by LC-MS and tannins after acid-catalyzed hydrolysis. Values represent averages  $\pm$  SE of three independent measurements. G, glucoside; I, isorhamnetin; K, kaempferol; Q, quercetin; R, rhamnoside.

### 2.3.5 Functional analysis of the Columbia promoter in other backgrounds

To investigate further if the expression of the *TT10* gene could be *cis*- or *trans*- regulated, we have studied activity of the *TT10* promoter fused to the GUS protein, in other accessions. Wild-type *Ler*, *Ws-4*, *Cvi-0* and *Bay-0* plants were transformed with *pTT10\_5* and *pTT10\_6* promoter fragments and the GUS activity was studied revealing no qualitative differences. Moreover, the dissected promoter fragments were also studied in *Ws-4* background, confirming previous quantitative results obtained in Col-0 background (see 2.1.3 and 2.1.3.1). We did not obtain any transformants for *Sha*.

## 2.4 Stress induction of the *TT10* promoter activity

Several aspects, namely *in silico* analysis (see 2.1.2), strong GUS staining in aborted seed (Pourcel et al., 2005) and the putative function of *TT10* (discussed in 2.1.2 and 2.2.5), motivated our interest in the hypothesized role of stress conditions regulating *TT10* gene expression. Pourcel et al. (2005) and Liang et al. (2006a) discussed the differences observed for *TT10* transcript accumulation in roots reported by McCaig et al. (2005), as being due to the stressing conditions during plant growth. Moreover, preliminary results with the *pTT10:GUS/pBIB-Hyg* in Ws-4, #21 plants were strongly suggesting that *TT10* could be ectopically induced in certain conditions. However, many conditions tested did not result in the reproducible GUS activity with *pTT10:GUS/pBIB-Hyg* in Ws-4, #21 plants as well as other transformant lines used. Conditions of the experiments and final observations are described below.

### 2.4.1 Abiotic stresses

Analysis of the publicly available transcriptomic data (Genevestigator, 07.02.2007), suggested, that *TT10* could be ectopically induced by various biotic and abiotic stresses. The highest upregulation of *TT10* expression was found for osmotic, salt, drought, heat, wounding and oxidative stresses, whereas highest downregulation was observed in sucrose, methyl jasmonate, ABA, senescence and hydrogen peroxide treatments. However, only the logarithmic differences presenting the fold induction between the samples and controls were taken in account. Absolute values and number of repetition were not considered.

Because most of the transcriptomic data were obtained with 7 and/or 14-15 days old seedlings, that developmental stage was also used in stress experiments aiming at finding ectopic GUS activity driven by the *TT10* promoter. Conditions of the experiment are described here because they are important factors which could affect ectopic GUS activity. Briefly, depending on the experiment plants were grown on either filter paper wetted with water or agar solidified B5 medium supplemented with 1% sucrose or liquid 0,5x B5 medium supplemented with 0,5% sucrose. After sowing, plates were stratified for 4 days and then grown in the phytotron at 16h light / 8h dark photoperiod, except for etiolated and constant light grown plants (Fig. 2.25). Stock solutions of hormones were prepared in ethanol or DMSO and diluted in water or culture medium. All the solutions for *in vitro* culture were filter-sterilized, except B5 medium which was autoclaved. Concentrations of hormones and other chemicals were based on the information found for the appropriate transcriptomics experiment, but also 10 to 100 times higher (Tab. 2.5).

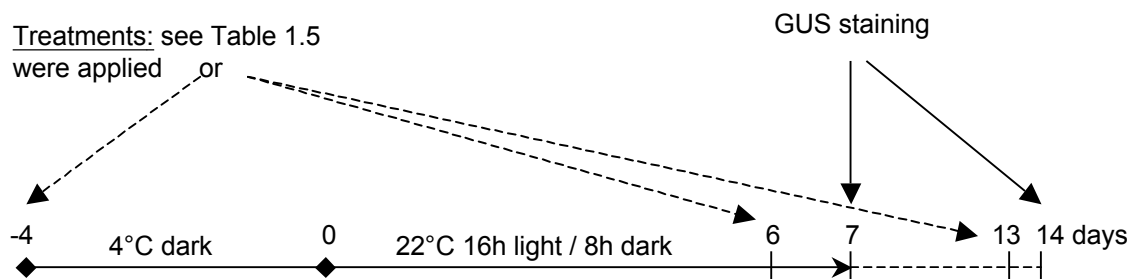


Figure 2.25: Stress experiments - experimental design.

For the experiments, 8 transgenic lines were used: three carrying 2.0-kb *TT10* promoter GUS construct in Col-0, four carrying 2.0-kb *TT10* promoter GUS construct in Ws-4 background (including #21) and one line carrying *Pro35Sdual:uidA* construct as a positive control of GUS staining. Apart from treatments listed in Table 2.5, plants were also grown in constant light and darkness. For all of the conditions tested, no

reproducible GUS activity was observed.

Treatment	Concentration / intensity
NaCl	50 - 150 - 300mM
mannitol	100 - 300 - 600mM
sucrose	100mM
(+)-catechin hydrate	50-500mM
(-)-epicatechin	50-500mM
quercetin rhamnoside hydrate	50-500mM
ABA	0,1 - 10 - 100 $\mu$ M
GA 4-7	10 - 50 $\mu$ M
methyl jasmonate	100 $\mu$ M - 1mM
jasmonic acid	100 $\mu$ M - 1mM
salicylic acid	100 $\mu$ M - 1mM
ethanol	1-5%
heat	30 - 40°C

Table 2.5: Concentration of hormones, salts and other treatments in stress experiments.

Seedlings were also submitted to progressive dehydration (Fig. 2.26), grown on the vertical plates or exposed to UV light under the flow hood and then GUS-stained within 4-24 hours after the treatment. None of those treatments resulted in clear inducible GUS pattern. However, it is possible, that *TT10* promoter could be ectopically induced because the blue staining was sometimes observed. Three patterns were observed. The first one considered plants grown in multiwell plates in liquid medium (Fig. 2.27 C.), where plantlets exhibited GUS activity in the root tip. It seems that it could be nutritionally related and depend on the density of sowing and developmental stage of the plantlets. The second pattern also considered the root tip, which is very often blue-stained in the seedlings grown on filter paper (Fig. 2.26). This staining also seems to be sowing-density dependent, however it could also be related to mechanical stress and root outgrowth into the filter. The third, often observed GUS activity, concerns plantlets grown on solid medium (Fig. 2.27 D.). The blue staining was sometimes observed in plantlets grown *in vitro* for 14 days or more. It seems to be associated with the vitrified tissues of the seedling. Surprisingly, those GUS patterns were observed only in Col-0 background. Sometimes, faint GUS activity was also observed in the upper part of the roots of plants grown in soil (Fig. 2.27 B.). Progressive dehydration (Fig. 2.26) was established to check if the limiting water availability could result in induction of the GUS activity driven by the *TT10* promoter. However, the preliminary results obtained in this system with the *pTT10:GUS/pBIB-Hyg* in Ws-4 (#21) did not result in any GUS induction in other transgenic lines and was later not reproducible with #21 either.

#### 2.4.2 Wounding and biotic stresses

Many of the transgenic lines analyzed, carrying GUS construct with various lengths of the *TT10* promoter sometimes result in blue staining in rosette leaves. This suggested, that *TT10* could be ectopically induced by unknown factors in that organ. Rosette leaves at various stages of development were pierced with a metal needle and GUS stained within 3-48 hours after. Although, blue staining was sometimes detectable, it was never associated with the wounded tissue (Fig. 2.27 A.). Moreover, GUS activity was never detected in wounded stem nor senescing leaves (data not shown). Because all the plants were grown in the greenhouse, this exceptional spotty GUS staining could be associated with bacteria which could grow on leaves and exhibit  $\beta$ -glucuronidase activity.

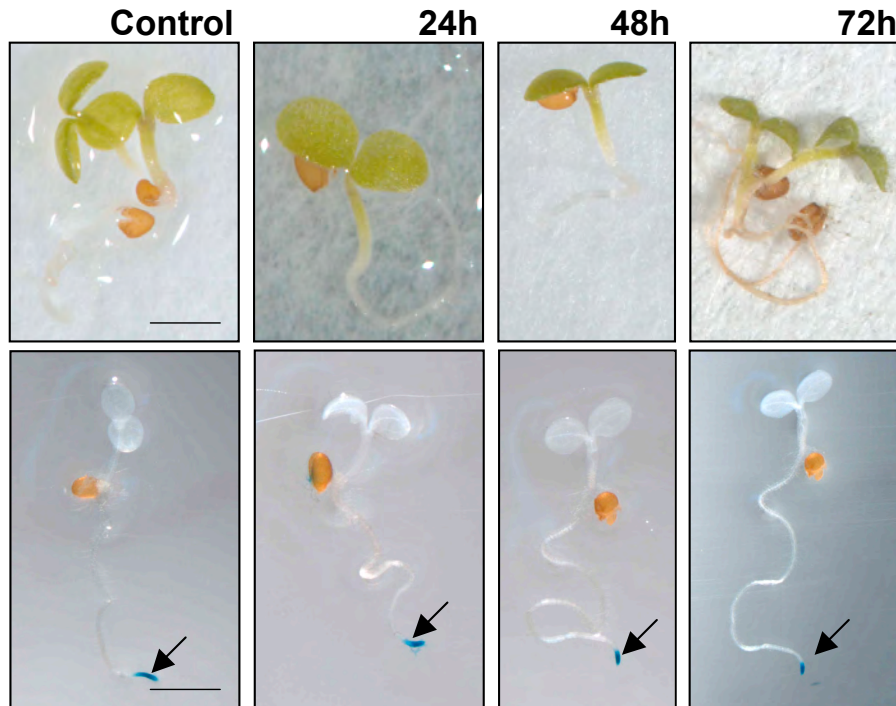


Figure 2.26: Plantlets subjected to progressive dehydration.

Ectopic expression of the *pTT10:GUS/pBIB-Hyg* cassette in wild-type Col-0 plants is shown. 6-day-old plantlets were subjected to progressive dehydration. Time indicates hours Petri dishes were outside the magenta box. Upper pictures demonstrate evaporation of water and drying filter paper. Bottom pictures, GUS staining of the corresponding plantlets. Arrow is pointing GUS staining in the root tip. Bar = 1mm.

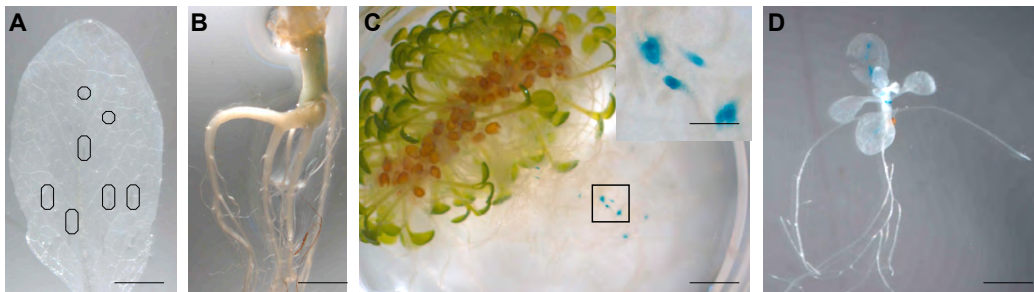


Figure 2.27: Ectopic GUS activity of *TT10* promoter.

Ectopic expression of the *pTT10:GUS/pBIB-Hyg* cassette in wild-type Col-0 plants is shown.

(A) Pierced wounded leaf. Circles highlight piercing in the leaf;

(B) Root of 5 week-old plant grown in soil;

(C) 7-day-old seedlings grown *in vitro* in 0,5x B5 liquid medium, 0,5% sucrose;

(D) 14-days-old seedlings grown on agar solidified B5 medium.

Bars = 300 $\mu$ m magnification in C, 2mm A and D, 3mm C, 7mm B.

## 2.5 Identification of laccases expressed in seeds

Several research groups are interested in the function of the laccase genes (Brown et al., 2005; McCaig et al., 2005; Cai et al., 2006; Abdel-Ghany and Pilon, 2008). However, apart from *LAC15/TT10* (Pourcel et al., 2005), little is known about their role in plant physiology. One of the secondary objectives of this thesis was to find out if other members of the laccase family, which in *Arabidopsis* consists of 17 members, were expressed in seeds and if they could be involved in flavonoid metabolism.

### 2.5.1 Expression data

The first aspect of this study was to identify the laccases which are expressed in *Arabidopsis* seeds. RNA was extracted from seeds of the Col-0 wild type plants dissected from siliques at 5 and 10 days after flowering (DAF). RT-PCR revealed that only *LAC15/TT10* is significantly detected in *Arabidopsis* seeds (Fig. 2.28 A.). However in seeds at 10 DAF, a weak signal is also detected for *LAC3*, *LAC5*, *LAC12*, in order of decreasing signal strength. It is important to mention that a strong band of *TT10* PCR product is detected with 32 cycles, whereas for other laccases even at 35 cycles only faint signal was detected (Fig. 2.28 A.). The situation is different when RNA is extracted from the whole siliques at 10 DAF. In this case, *LAC4* and *LAC5* seem to be expressed at comparable levels to *LAC15/TT10*. A strong signal is also detected for *LAC12*, *LAC11*, *LAC17*, *LAC3* in order of the signal strength.

We have also analyzed the expression of the laccase genes in the transcriptomic data available (Fig. 2.28 B.). The sample from the transcriptomic analysis (AtGenExpress: Developmental series), named Seed Stage 4 with/whole silique, which corresponds to the stage of 5 DAF. Relatively strong signal in that sample, observed for *LAC4* and *LAC17* could come from replum and/or valve. Transcriptomic sample Seeds Stage 8 without Siliques, should correspond to the extracted seeds 10DAF (Fig. 2.28 A. and B.). Comparison of those two samples clearly confirms that *LAC15/TT10* is the most expressed laccase in seeds. However, relatively strong signal was also detected for *LAC5*. All the other laccases seem to be not significantly expressed. It is important to note that strong signal for *LAC4*, *LAC17* and *LAC2* in the transcriptomic analysis, Seed Stage 4 with/whole silique, is also detected by RT-PCR in siliques 3-8 DAF (data not shown). The expression remains strong for *LAC4*, was hardly detectable for *LAC17* and below the detection level for *LAC2*, until 18 DAF. Later stages were not analysed (data not shown).

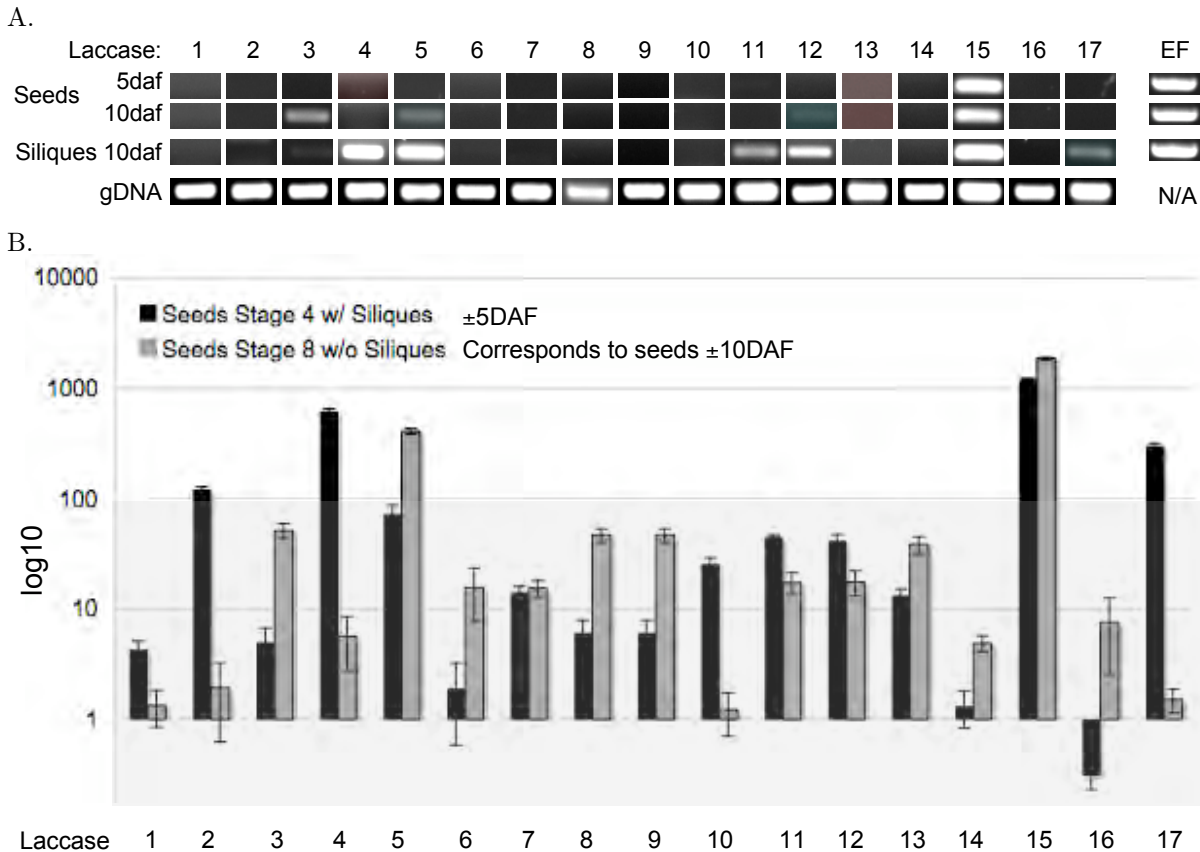


Figure 2.28: Comparison of the expression pattern of the laccase multigene family in seed and silique. (A) Expression of the laccase family genes in seeds and siliques of the wild-type Col-0 plants. cDNA diluted x100, PCR at 35 cycles, 1/5 of the reaction volume for EF was deposited. Picture was taken at the limit of signal saturation for EF, therefore the band intensities can be compared to estimate the amount of transcript. The list of primers is given in Appendix, Tab. 5.9. (B) Laccase multigene family expression based on publicly available transcriptomic data. Bars represent gene expression level in Seeds Stage 4 with Siliques (Black) and isolated seeds, Seeds Stage 8 without Siliques (Grey). All the bars covered by the shaded box should be considered as not significant (below detection level). Data are presented in log<sub>10</sub> scale to highlight fold differences in the expression level of the laccase family gene members. Data used for the graph were retrieved *via* eFP browser and correspond to AtGenExpress: Developmental series (siliques and seeds), hybridization slides: ATGE\_77\_D, ATGE\_77\_E, ATGE\_77\_F, ATGE\_82\_A, ATGE\_82\_B and ATGE\_82\_C.

## 2.5.2 Flavonoid composition of mature seeds of the laccase mutants

*LAC15/TT10* has been shown to be involved in flavonoid metabolism and is required for the brown color of the wild-type *Arabidopsis* seeds (Pourcel et al., 2005). Although it is the only laccase mutant easily distinguishable for its *tt* phenotype (Pourcel et al., 2005; Cai et al., 2006), we decided to precisely analyze several laccase mutants for their flavonoid composition. Metabolomic analysis of the composition of the mature seeds confirmed the flavonol and PA phenotype of *tt10* mutant, whereas no significant differences in flavonoid composition were observed for *lac3*, *lac5* and *lac12* mutants (Fig. 2.29). It seems that the laccases tested are not involved in flavonoid oxidation, or their role is minor and could be masked by the presence of *TT10*. It would be interesting to analyze flavonoid composition of the laccase double mutant with *tt10*.

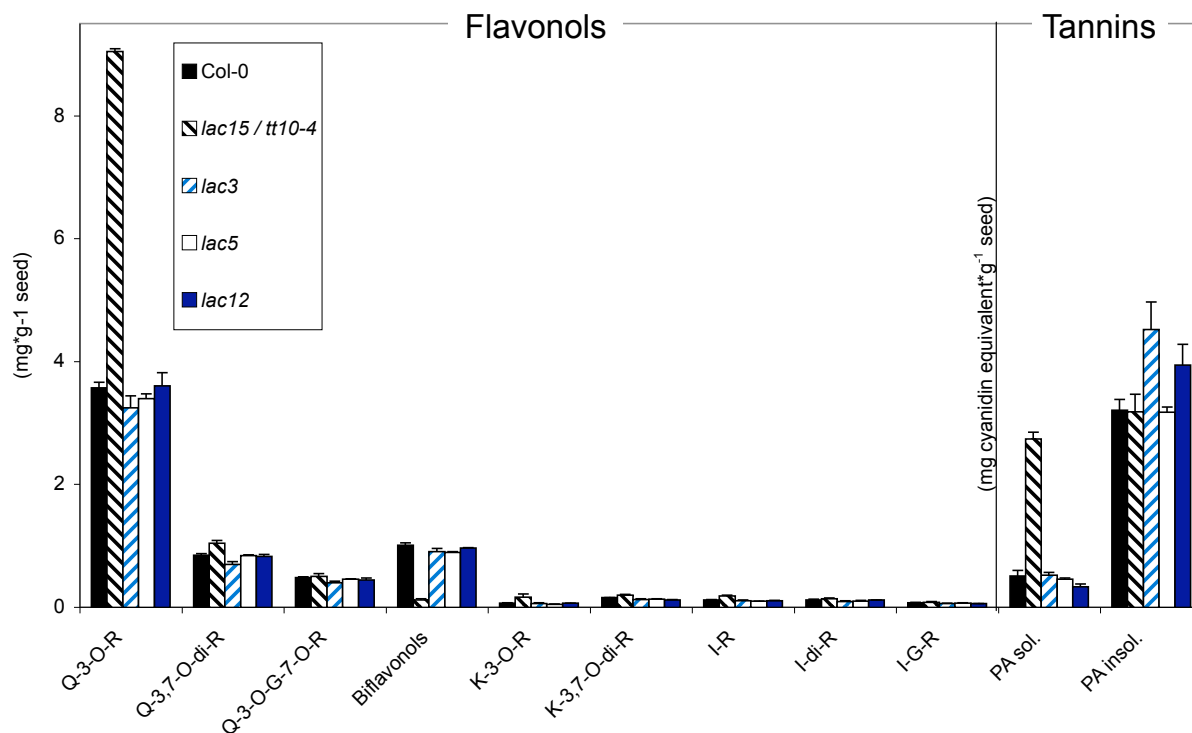


Figure 2.29: Flavonoid composition of mature seeds of the laccase mutants.

Analysis of flavonol derivatives by LC-MS and tannins after acid-catalyzed hydrolysis. Values represent averages  $\pm$  SE of three independent measurements. G, glucoside; I, isorhamnetin; K, kaempferol; Q, quercetin; R, rhamnoside.

### 2.5.3 Promoter GUS activities

Several laccase promoters were cloned in collaboration with Serge Berthet (Secondary Cell Wall and Lignification Group, INRA-IJPB, Versailles) and help from Christian Dubos. Here we present the results for those which were found to be expressed in seeds: *LAC3*, *LAC5* and *LAC12* to validate the expression data, study tissue-specific expression and investigate promoter activity, after fusion with the *uidA* reporter gene. The genomic sequences of the laccase promoters were retrieved from TAIR and used to design primers. It was considered that 2.0-kb sequence downstream from the translation initiation site of the gene should contain all the elements required for the tissue specificity of gene expression. However, in case of *LAC3* and *LAC12*, the neighboring gene was located too close and only the sequence up to the 3'UTR of that genes were considered. The respective promoter fragments (*LAC3* 1.44-kb, *LAC5* 1.9-kb, *LAC12* 1.65-kb) have been PCR-amplified with the primers containing *attB1* and *attB2* GATEWAY® recombination sites (see Appendix Tab. 5.10) and introduced into the *pDONR207* vector by BP recombination. Then promoter fragments were transferred into binary vector *pBI101GUS-GTW* by LR recombination (Baudry et al., 2006; Dubos et al., 2008). The integrity of all the constructs have been checked by sequencing and plants have been transformed. Presented characterization of each of the promoter patterns is based on the analysis of the GUS activity in 8 independent transformant lines. All the plants have been genotyped for the presence of corresponding constructs. The same promoter activity pattern was observed in T2 plants. In this study we wanted to have a full overview of the GUS activities driven by the laccase promoters, therefore all the plant organs are presented. However, the focus was given on the seed GUS activity.



### 2.5.3.1 LAC3

The *LAC3* transcript was detected only in the seedling root (McCaig et al., 2005) and in the developing seed (this study). The 1.44-kb *LAC3* promoter resulted in detectable GUS activity mainly in secondary roots (Fig. 2.30 H.) and seedling root (Fig. 2.30 I.). Localized blue staining was also detected in the sepal and petal abscission zones of the receptacle (Fig. 2.30 F. and G.). No GUS activity was detected in various stages of the flower and seed development, even when no potassium ferricyanide and potassium ferrocyanide were used.

### 2.5.3.2 LAC5

The *LAC5* transcript seems to be abundant in the whole plant (McCaig et al., 2005). However it was hardly detectable in developing seeds and only at the late stages of silique development (see 2.5.1). The GUS activity driven by 1.9-kb *LAC5* promoter seems to be associated with aging of the various plant organs (Fig. 2.31). The blue staining is not visible in young flower buds, but the progression of GUS activity is observed from the open flower stage onward, first in the sepals and later petals (Fig. 2.31 A. and B.). Strong blue staining of the stigma and filament, was observed at the open flower and young silique stage, when stamens were still attached (Fig. 2.31 B. and C.). It seems that staining intensity increases with the age of the silique (Fig. 2.31 D.), however strong GUS activity is also detected in abnormal siliques (Fig. 2.31 E. and F.) and aborted seeds (Fig. 2.31 H.). Surprisingly, some blue staining was also detectable in the funiculus of the neighboring seeds the development of which seems to be normal (Fig. 2.31 G.). Stronger staining of the funiculus was observed in old siliques (14-20 DAF). It is important to note that no GUS staining was detected in developing seeds. Faint blue signal was observed in the longitudinal sections of the stem (Fig. 2.31 L.). No GUS activity was detected in rosette and cauline leaves (Fig. 2.31 M. and N.), however in older leaves blue staining seems to colocalize with the dying/senescing tissue (Fig. 2.31 O. and P.). Strong GUS activity was observed in the hypocotyl and root of the seedling (Fig. 2.31 K.) grown *in vitro*. Sometimes faint blue staining was observed in the secondary roots of the plants grown in the soil (Fig. 2.31 Q.).

### 2.5.3.3 LAC12

The *LAC12* transcript was detected only in the stem (McCaig et al., 2005) and at low level in developing siliques and seeds (see 2.5.1). The 1.65-kb promoter of *LAC12* resulted in strong GUS activity in stem and root (Fig. 2.32 E. and J.). No GUS activity was detected in upper part of the stem and inflorescence (Fig. 2.32 A.) and most of the leaves (Fig. 2.32 B. and E.). Surprisingly, in all independent transformant lines, some of the leaves had detectable GUS activity in the main vessel (pointed by arrow at Fig. 2.32 C.) or at the border of the leaf (Fig. 2.32 D.). In seedlings blue staining was easily visible at the hypocotyl and in patchy pattern in roots (Fig. 2.32 M.). The GUS staining in the siliques (Fig. 2.32 F. to I.) is first detectable at the abscission zone (Fig. 2.32 F., H. and magnified in I.). Some staining was also observed at the upper part of the silique, in the replum (Fig. 2.32 G.). GUS activity was also detected at the micropyle of the seed (Fig. 2.32 L.). It seems that some GUS activity was also present in the seed coat, however it was extremely weak (Fig. 2.32 K.).

### 2.5.3.4 Summary

The reporter gene activity driven by the laccase promoters resembles their transcript accumulation. Even though we have detected some transcript of *LAC3*, *LAC5* and *LAC12* in developing seeds (section 2.5.1), only promoter of *LAC12* gave detectable GUS activity in seeds. It is possible, that during seed excision, the

sample was contaminated by some cells coming from replum or valve, what resulted in weak signal detected by RT-PCR.

The laccase promoter activity is compared to that observed and discussed previously for *LAC15/TT10* (Tab. 2.6). It seems that *TT10* is the only laccase significantly present in *Arabidopsis* seeds, which would explain that a mutant in that gene results in a clear phenotype. It would be interesting to check if any other laccase driven by the *TT10* promoter could oxidize flavonoids and complement *transparent testa 10* phenotype. It seems that some of the function of laccases could be redundant or at least partially overlapping, however, probably due to the tissue specificity only *LAC15/TT10* can oxidize flavonoids in *Arabidopsis* and ensure brown color of the wild-type seeds.

	<b>Plant Organ</b>	<i>LAC3</i>	<i>LAC5</i>	<i>LAC12</i>	<i>LAC15</i>
Seed	Integuments			-	
	Micropyle			+	
	Chalaza	-	- / + stress	-	+
	Funiculus			-	
	Abscission zone			-	
Silique - Flower	Flower bud	-	-	-	+
	Stigma/style - Transmitting tissue	-	late stages / senescence +	especially abscission zone +	+
	Abscission zone	+			+
	Receptacle	-			-
	Peduncle	-			-
	Stamen	-	+ filament	-	-
	Pollen		-		-
Other organs	Cauline/Rosette leaf, trichomes	-	old leaves +	+	-
	Stem	-	+		
	Root	secondary roots	+		
Seedling	Cotyledons	-	-	-	
	Hypocotyl	+	+	+	-
	Root				

Table 2.6: Comparison of laccase promoter activities.

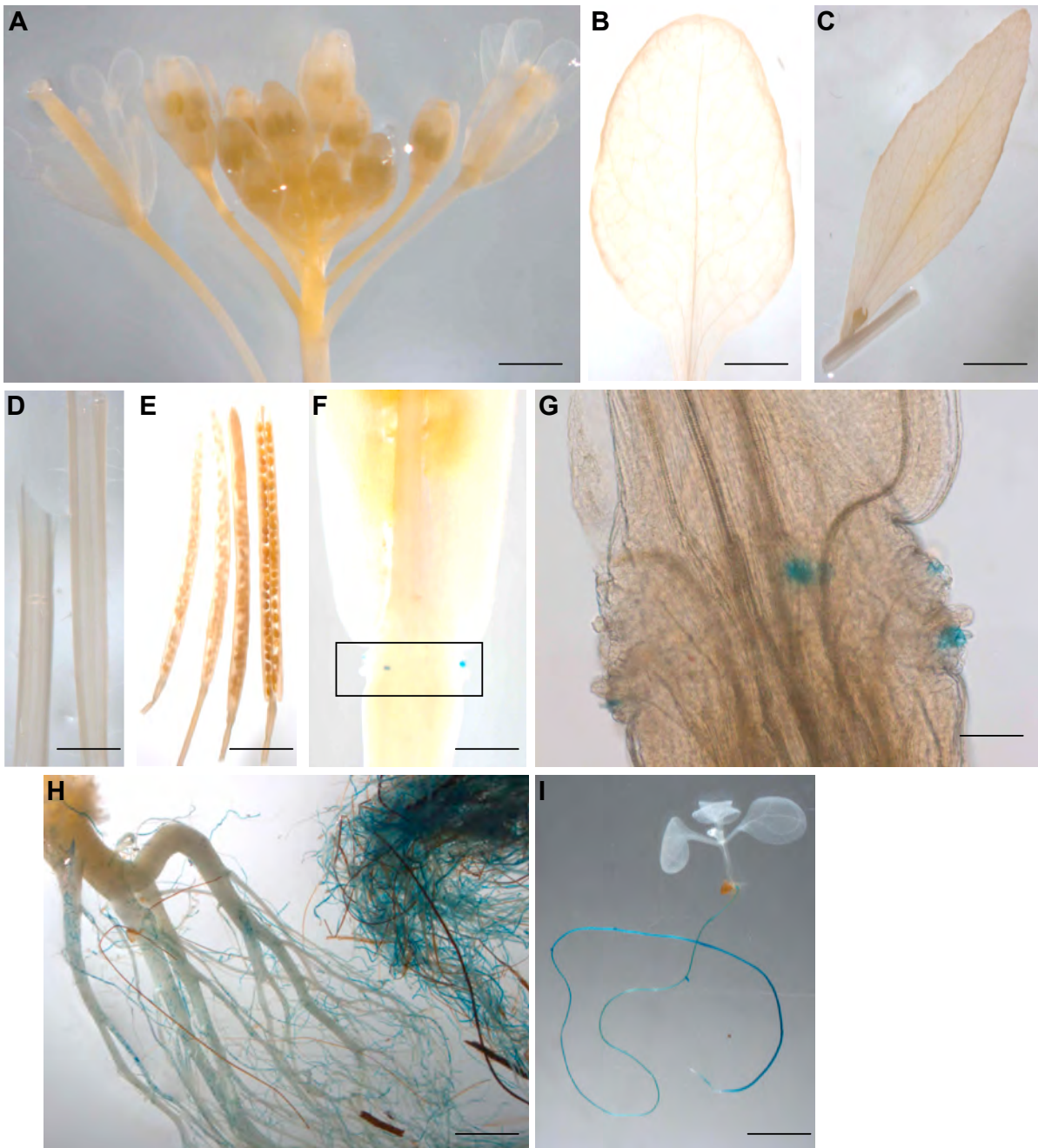


Figure 2.30: Pattern of *LAC3* promoter activity in wild-type plants.

Expression of the *pLAC3:GUS/pBI101-GTW* cassette in wild-type Col-0 plants. (A) Inflorescence of 5-weeks-old plant; (B) rosette leaf; (C) cauline leaf; (D) stem; (E) various stages of developing siliques (from left:  $\pm 6$ , 10, 14, 20 DAF); (F) bottom part of the silique  $\pm 14$  DAF; (G) abscission zone of the silique,  $\pm 14$  DAF; (H) root; (I) 10 days old seedling.

GUS activity was observed with binocular microscope on whole mounts except G where whole mount samples in chloralhydrate, were observed with microscope, using Nomarski optics.

Potassium ferricyanide and potassium ferrocyanide were not used, except E to H, 1mM. Bars = 100 $\mu$ m G, 280 $\mu$ m F, 500 $\mu$ m D, 2mm A and I, 3mm B, C and E, 4mm H. DAF - days after flowering

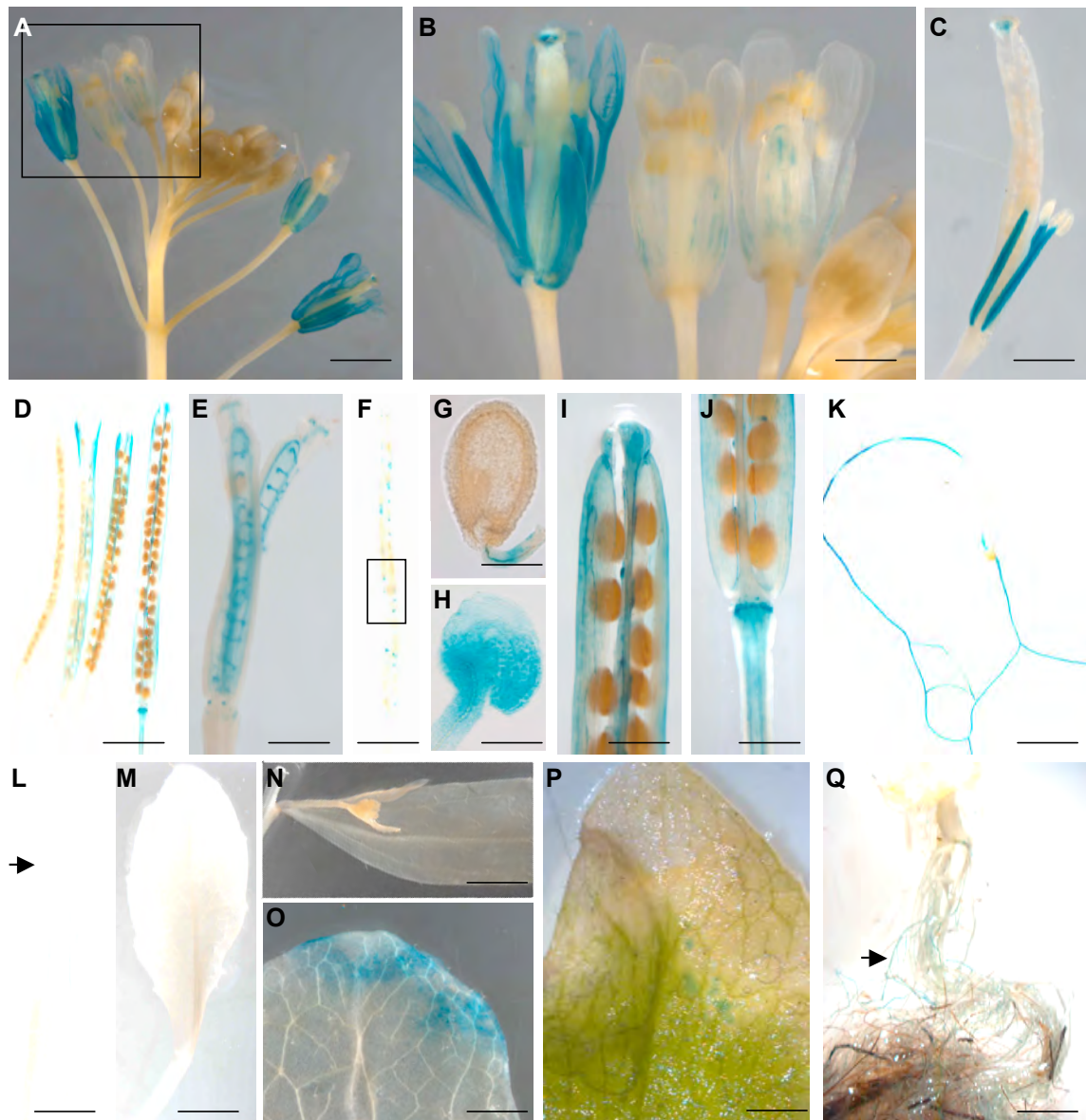


Figure 2.31: Pattern of *LAC5* promoter activity in wild-type plants.

Expression of the *pLAC5:GUS/pBI101-GTW* cassette in wild-type Col-0 plants. (A) Inflorescence of 5 weeks old plant; (B) developing flowers and young silique, senescing sepals and petals; (C) young silique,  $\pm 2-3$ DAF; (D) various stages of developing siliques (from left:  $\pm 6, 10, 14, 20$  DAF); (E) and (F) abnormal siliques,  $\pm 8$  DAF; (G) seed  $\pm 8$  DAF; (H) aborted seed; (I) top and (J) bottom part of a  $\pm 20$  DAF silique; (K) 10 days old seedling; (L) stem; (M) rosette leaf; (N) cauline leaf; (O) and (P) old/senescing cauline leaf; (Q) root. Arrows are pointing weak GUS staining.

GUS activity was observed with binocular microscope on whole mounts, except, (G) and (H) where, whole mount samples in chloralhydrate, were observed with microscope, using Nomarski optics.

Different concentrations of potassium ferricyanide and potassium ferrocyanide were used to prevent GUS diffusion: L and O 1mM, A to J 2mM, K, M, N, P and Q 0mM. Bars = 60 $\mu$ m H, 80 $\mu$ m G, 300 $\mu$ m F, 400 $\mu$ m I and J, 500 $\mu$ m E, 680 $\mu$ m B, 800 $\mu$ m L, 820 $\mu$ m C, 1mm O and P, 2mm A and K, 3mm D, M and N, 7mm Q. DAF - days after flowering

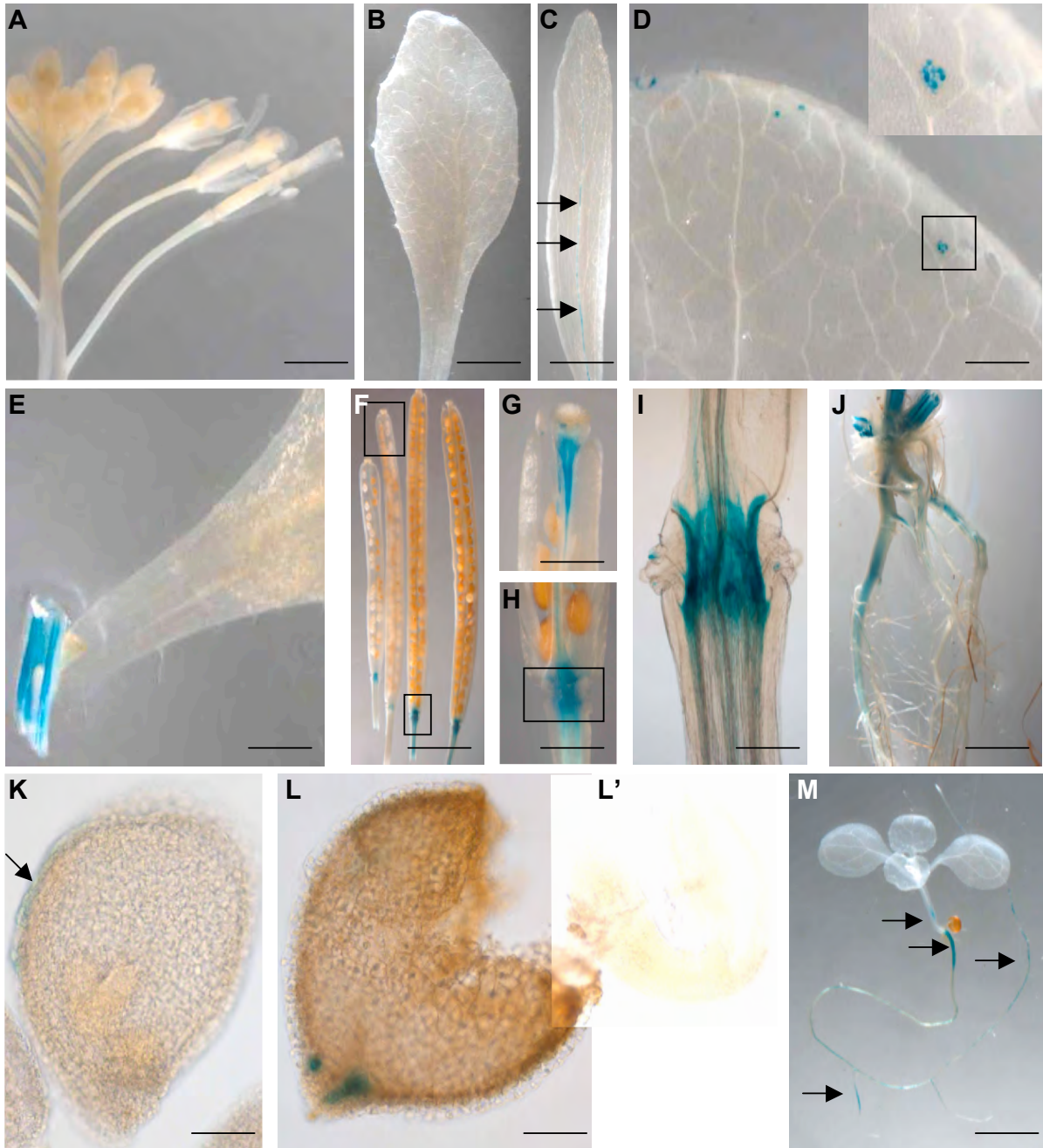


Figure 2.32: Pattern of *LAC12* promoter activity in wild-type plants.

Expression of the *pLAC12:GUS/pBI101-GTW* cassette in wild-type Col-0 plants. (A) Inflorescence of 5 weeks old plant; (B) rosette and (C) cauline leaf; (D) old rosette leaf; (E) part of stem and cauline leaf; (F) various stages of developing siliques (from left:  $\pm 6$ , 10, 14, 20 DAF); (G) top and (H) bottom part of a  $\pm 10$  DAF and  $\pm 14$  DAF siliques respectively; (I) abscission zone of the silique,  $\pm 14$  DAF; (J) root; (K) seed,  $\pm 7$  DAF; (L) seed coat and (L') dissected embryo,  $\pm 14$  DAF, superposition of 2 pictures; (M) 10 days old seedling. Arrows are pointing week GUS staining.

GUS activity was observed with binocular microscope on whole mounts, except I, K, L, and L' where, whole mount samples in chloralhydrate, were observed with microscope, using Nomarski optics.

Different concentrations of potassium ferricyanide and potassium ferrocyanide were used to prevent GUS diffusion: C, D, K and L 1mM, E to J 2mM, A, B, and M 0mM. Bars = 40 $\mu$ m L, 50 $\mu$ m K, 300 $\mu$ m I, 500 $\mu$ m D, 700 $\mu$ m G and H, 1mm A and E, 2mm M, 3mm B and F, 7mm C and J. DAF - days after flowering

## 2.6 Cell suspension as a tool to study flavonoid biosynthesis

Cell suspension maintained in the Seed Biology Laboratory was originally established from *Arabidopsis* Columbia accession, and is a gift from Geneviève Ephritikhine from Membrane Transport and Signalisation Group, Institut des Sciences du Végétal, CNRS, Gif-sur-Yvette. Manipulation and characterization of the cell suspension was one of the secondary objectives of this thesis. Several preliminary results are discussed in this section. The ultimate goal would be to use cell suspension as a simplified system to study flavonoid biosynthesis and metabolism.

**Flavonoid pathway genes in cell suspension** The first step to study flavonoid biosynthesis in the simplified system was to check which, if any, genes of the pathway are expressed in cell suspension. To prepare material for RNA extraction, cell suspension was transferred to an Eppendorf and centrifuged to remove supernatant and frozen in liquid nitrogen. However, the material prepared that way was not possible to grind due to the remaining liquid which after freezing consolidated the sample. Therefore cell lysis was further done by sonication to destroy the cells and release the nucleic acids for further extraction according to the manual of the Total RNA extraction kit (Sigma). Surprisingly, the RNA yield was comparable for sonicated and liquid nitrogen frozen samples (Fig. 2.33 A.). It seems that freezing and normal condition lysis are sufficient to extract good quality RNA from the cell suspension used in our laboratory. RT-PCR revealed that several genes of the flavonoid biosynthetic pathway were expressed and *TT10* is among them (Fig. 2.33 B.). However we did not detect chalcone synthase, which is the first enzyme of the flavonoid pathway. It would be interesting to check if any flavonoids are accumulating in the cell suspension in standard culture conditions.

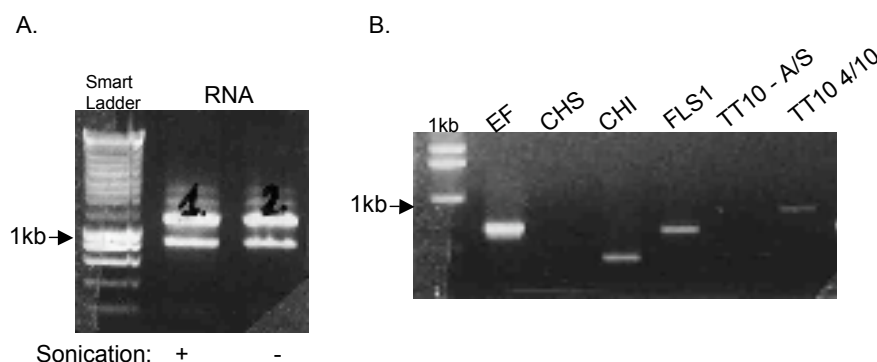


Figure 2.33: Total RNA and flavonoid pathway genes expressed in the cell suspension.

(A) Total RNA extracted from the cells grown in the suspension. Comparison of the extraction method efficiency. (B) RT-PCR detection of the genes involved in the flavonoid biosynthesis expressed in the cell suspension. EF - reference gene *Elongation Factor 1 $\alpha$ A4*, CHS - *Chalcone Synthase*, CHI - *Chalcone Isomerase*, FLS1 - *Flavonol Synthase 1*, TT10 A/S - *Transparent Testa 10* - ATG - STOP primers.

To further investigate which genes of the flavonoid pathway were expressed, sequences of the primers used to prepare the probes in the CATMA arrays, were retrieved from the CATMA GST database. Those primers guarantee gene specificity and result in a 200 - 500bp PCR product. The list of the primers, for which specificity was validated on the genomic DNA is listed in Appendix, Table 5.14.

This project was discontinued, due to the focus on the analysis of the developmental and stress regulation of *TT10* gene expression.



# Chapter 3

## Materials and Methods

### 3.1 Materials

#### 3.1.1 Plant Material

*Arabidopsis thaliana* L. Heynh. (thale cress) is a small flowering plant that is used as a model organism to study plant biology it is also a model for molecular and genetic analyses of the flavonoid biosynthesis pathway (Winkel-Shirley, 2001b; Routaboul et al., 2006). *Arabidopsis* is a member of the *Brassicaceae* (mustard) family, which includes cultivated species (*e.g.* cabbage, cauliflower, rapeseed, radish). Small size, rapid life cycle, sequenced genome, many available protocols and other resources are making *Arabidopsis* a suitable tool to study gene expression. More details about *Arabidopsis* can be found at TAIR - <http://www.arabidopsis.org/portals/education/aboutarabidopsis.jsp>.

##### 3.1.1.1 *Arabidopsis thaliana* - wild type accessions

On the course of evolution *Arabidopsis thaliana* was dispersed among many habitats, where isolated populations have accumulated mutations in their genome. The mutations which are the result of the adaptation to the environment are now used to study the molecular basis of gene expression. The wild type accessions used in our experiments are listed in the Table 3.1.

##### 3.1.1.2 *Arabidopsis thaliana* - mutants

To facilitate gene discovery and to speed up progress in acquiring knowledge about plant biology, a collection of mutants has been created. In my studies, I used the collection of mutants previously studied in the laboratory, as well as new candidate mutants which were acquired at Nottingham Arabidopsis Stock Centre (NASC) and at the Versailles Arabidopsis thaliana Resource Center. *Transparent testa* mutant 47.1/Cvi-1 has been kindly provided by Dr. Wim Soppe MPI Cologne, Germany. After allelism test at the Seed Biology laboratory, the line was designated *tt10-7*. The *myb4* mutant was kindly provided by Dr. Cathie Martin and all *pfg* mutants by Dr. Bernd Weisshaar. The laccase mutant lines *lac3*, *lac5* and *lac12* were provided by Dr. Lise Jouanin. Details of the mutant lines are listed in Tables 3.2 and 3.3.



Accession	Location	Country	Habitat / Altitude (m)	NASC ID
Ws-4	Wassilewskija / Dnjepr River	Belarus (USSR)	sandy ryefield / 100-200	N5390
Ler-0	Gorzow Wielkopolski (Landsberg)	Poland	n/a / 1-100	NW20
Col-0	Gorzow Wielkopolski (Landsberg) / Columbia	Poland / USA	greenhouse / 1-100	N1092
Cvi-0	Cape Verdi Islands	Cape Verde Islands	rocky wall with moss / 1200	N902
Bay-0	Bayreuth	Germany	fallow land / 300-400	N954
Sha	Shakdara River (Pamir)	Tadjikistan	mountains / 3400	N929

Table 3.1: Wild type accessions of *Arabidopsis thaliana* used in this work.

Several information about the wild type accessions of *Arabidopsis thaliana* are presented: i) abbreviated name of the accession, ii) place of origin and general habitat information, iii) NASC ID number. Ws-4 is the parental line for Versailles T-DNA tagged lines. Information about location and habitat for this line are assumed to be like for Ws-0 (N1602). Landsberg *erecta*, Ler-0 (NW20) originates from original Landsberg ecotype - La-0 (N1298) after mutagenesis. Col-0 is coming from the name of university where these *Arabidopsis* seeds were taken from, to make a stock and store accessions. It originates from the Landsberg nonirradiated population. Sha, Shahdara, called also Shakdara. m - meters  
For details about the lines see: NASC at <http://arabidopsis.info/>.

<i>tt10</i> mutant alleles	Background	Mutagen	Collection	NASC ID	Reference
<i>tt10-1</i>	Ler-0	EMS	M.Koornneef, CS110	N110	M.Koornneef
<i>tt10-2</i>	Ws-4	T-DNA	INRA Versailles, CPI13	n/a	Pourcel <i>et al.</i> 2005
<i>tt10-3</i>	Ws-4		INRA Versailles, CQK31	n/a	
<i>tt10-4</i>	Col-0		SALK_002972	N502972	
<i>tt10-5</i>	Col-0		SALK_128292	N128292	
<i>tt10-6</i>	Col-0		SALK_114753	N114753	
<i>tt10-7</i>	Cvi-1	EMS	MPIK Cologne W.Soppe, 47.1	n/a	W.Soppe, unpublished

Table 3.2: Alleles of the *tt10* mutant studied in this work.

Table is summarizing information about *tt10* mutant alleles. Pictures of the mutant and corresponding wild type seeds are presented in Appendix, section 5.3. SALK institute: <http://signal.salk.edu/>.

Table 3.3: *Arabidopsis thaliana* mutant lines used in this study.

Table presents information about the mutant lines used in the study. For the clarity of the presentation, mutants presented in the table are divided in sub-groups, according to the experiments where they were used and how they are described in the Results. First column is giving the name of the mutated gene, second AGI reference number. Exact identification of the line (allele) is possible with the NASC ID. Transcription factor family, or short description of the protein is also given. The background of each mutant is indicated. Mutagen agent and when appropriate T-DNA insertion details are given. Literature refers to the first description of the mutant allele or when important features of the mutation were described. Mutant collections of the SALK and Versailles institutes were described by Alonso *et al.* (2003) and Bechtold *et al.* (1993) respectively. N.B. Mutant lines from the Versailles collection are in Ws-4 background, formerly Ws-2 as stated in publications cited.

Table 3.3 Part 1.

Mutant	AGI	Protein affected	Background	Mutagen	Collection / Allele	NASC ID	Comments	
<i>tt</i> - regulatory mutants	<i>tt1</i>	At1g34790	Ws-4	T-DNA	Versailles - DXL6	n/a	(Debeaujon and Lepiniec, unpublished; Sagasser <i>et al.</i> , 2002)	
	<i>tt16</i>	At5g23260	Ws-4	T-DNA	Versailles - DXT32; <i>tt16-1</i>	n/a	(Nesi <i>et al.</i> 2002); Kanamycine resistant; Insertion: intron	
	<i>ttg2</i>	At2g37260	Ws-4	T-DNA (untagged)	Versailles - CTA18; <i>ttg2-2</i>	n/a	(Johnson <i>et al.</i> , 2002, Debeaujon <i>et al.</i> , 2003)	
	<i>genes in the flavonoid pathway</i>	<i>pf1/myb12</i>	At2g47460	Col-0	En-1 Transposon	<i>myb12-1f</i>	n/a	(Mehrtens <i>et al.</i> 05); deletion; loss of function
		<i>pf2/myb11</i>	At3g2610	Col-0	T-DNA	SALK_077068	N577068	
		<i>pf3/myb111</i>	At5g49330	Col-0	T-DNA	GABI-Kat GK291D01	n/a	(Stracke <i>et al.</i> 2007); Insertion: intron ( <i>pf3</i> ), exon ( <i>pf3</i> )
		<i>myb11,12,111</i>	- / -	Col-0	En-1 / T-DNA	triple mutant	n/a	
		<i>myb12</i>	At2g47460	Col-0	T-DNA	SALK_046675C	N665933	Insertion: exon
	<i>myb4</i>	At4g38620	Col-0	dSpm Transposon	Sainsbury Lab (Meissner <i>et al.</i> 1999)	n/a	(Jim <i>et al.</i> , 2000); Insertion: exon	
	<i>Regulators of the early-biosynthetic</i>	<i>wrky23</i>	At2g47260	Col-0	T-DNA	SALK_003943C	N661317	(Grunewald <i>et al.</i> , 2008); Insertion: promoter
			At5g56400	Col-0	T-DNA	SALK_053025	N553025	
			At5g03730	Col-0	T-DNA	SALK_006605	N506605	
		<i>wrky70</i>	At5g56400	Col-0	T-DNA	SALK_025198C	N665593	(Li <i>et al.</i> , 2004); Insertion: exon
<i>ctr1</i>		At5g03730	Col-0	diepoxybutane	<i>ctr1-1</i>	N8057	(Kieber <i>et al.</i> 1993)	
<i>jar1</i>		At2g46370	Col-0	EMS	<i>jar1-1</i>	N8072	(Staswick <i>et al.</i> 1992)	
<i>myc2/jin1</i>		At1g32640	Col-0	T-DNA	SALK_017005C	N655309	(Berger <i>et al.</i> , 1996); Insertion: exon, loss-of-function	
			Col-0	EMS	SALK_061267C	N656547	Insertion: exon	
			Col-0	EMS	<i>npr1-1</i>	N3726	(Cao <i>et al.</i> , 1994)	
<i>Mutants in genes involved in (a)biotic stress</i>		<i>npr1</i>	At1g64280	Nossen	EMS	<i>npr1-2</i>	N3801	(Zipfel <i>et al.</i> , 2004); Nossen, carries a T-DNA containing the kanamycin resistance gene and the <i>PR-1a:tms2</i> reporter gene
			Nossen	T-DNA	<i>npr1-5</i>	N3724		

Table 3.3 Part 2.

Mutant	AGI	Protein affected	Background	Mutagen	Collection / Allele	NASC ID	Comments
<i>nac3</i>	At3g15500	ATAF-like NAC-domain TF	Col-0	T-DNA	SALK_014331 (D)/(R)	N514331	(Takada <i>et al.</i> , 2001); Insertion: exon
-	At1g71450	DREB subfamily A-4 of ERF/AP2 TF	Col-0	T-DNA	SALK_136922C	N667434	(Ma <i>et al.</i> , 2006); Insertion: promoter
<i>nac038/nac039</i>	At2g24430	NAC domain TF	Col-0	T-DNA	SALK_025040C	N660776	(Folta <i>et al.</i> , 2003); Insertion: intron
<i>myb76</i>	At5g07700	MYB domain protein	Col-0	T-DNA	SALK_055242C	N662521	(Stracke <i>et al.</i> , 2001); Insertion: intron
-	At5g66070	Zinc finger TF	Col-0	T-DNA	SALK_148182C	N658053	(Libault <i>et al.</i> , 2007); Insertion: exon
<i>myb99</i>	At5g62320	MYB domain protein	Col-0	T-DNA	SALK_003193 (F)	N503193	(Alves-Ferreira <i>et al.</i> , 2007); Insertion: exon
<i>lac3</i>	At2g30210	Laccase gene	Col-0	T-DNA	SALK_31901	n/a	(Cai <i>et al.</i> , 2006); Insertion: intron; loss of-function
<i>lac5</i>			Ws-4	T-DNA	Versailles - EAQ31	n/a	n/a
<i>lac5</i>	At2g40370		Col-0	T-DNA	SALK_63466	n/a	(Cai <i>et al.</i> , 2006); Insertion: exon; loss-of-function
<i>lac12</i>	At2g05390		Col-0	T-DNA	SALK_47455	n/a	Insertion: exon

### 3.1.1.3 *Arabidopsis thaliana* - transgenic plants

Transgenic lines used in this study are mentioned below.

- i) *Pro35Sdual:uidA* used as positive control in the GUS assays (Debeaujon et al., 2003),
- ii) *pTTG2:GUS* (line 5) to check TTG2 promoter activity in the testa during seed development (Johnson et al. (2002); provided by Dr. David R. Smyth),
- iii) *pTT10:GUS/pBIB-Hyg* #3, #16, #21, #22; all in Ws-4 background (Pourcel et al., 2005), used as positive control for the *TT10* promoter activity studies and to prepare crosses (#21, see Table 3.4D.) with the candidate regulatory mutants. *pTT10:GUS/pBIB-Hyg* is also referred to as *pTT10\_6* meaning 2.0-kb *TT10* promoter.
- iv) *p35Sdual:MYB4* to check *TT10* expression in *MYB4* overexpressing plants (Jin et al. (2000); provided by Dr. Cathy Martin),
- v) *NahG (Ler)* to analyze the involvement of SA in *TT10* expression (Rajjou et al. (2006); gift of Dr. Xinnian Dong).

Transgenic lines prepared in this study are listed in the tables below:

A.

Name of the construct	Ws-4	Col-0			Ler-0	Cvi-0	Bay-0
	<i>pBIB-Hyg</i>	<i>pBIB-Hyg</i>	<i>pCambia 1300</i>	<i>pBI101 GTW</i>	<i>pBIB-Hyg</i>	<i>pBIB-Hyg</i>	<i>pBIB-Hyg</i>
pTT10_1:GUS	24	12	n/a	n/a	n/a	n/a	n/a
pTT10_2:GUS	10	24	n/a	12	n/a	n/a	n/a
pTT10_3:GUS	14	24	24	12	n/a	n/a	n/a
pTT10_4:GUS	28	24	n/a	n/a	n/a	n/a	n/a
pTT10_5:GUS	16	24	n/a	n/a	16	16	8
pTT10_6:GUS	(*)	24	n/a	n/a	8	n/a	10

B.

Name of the construct	Col-0	
	<i>pCambia 1300</i>	<i>pBI101 GTW</i>
pTT10_3A:GUS	24	12
pTT10_3B:GUS	24	12
pTT10_3C:GUS	24	12
pTT10_3_M1:GUS	24	12
pTT10_3_M2:GUS	24	12
pTT10_3_M3:GUS	24	5
pTT10_3_M4:GUS	24	9
pTT10_3_M5:GUS	24	12

C.

Name of the construct	Ler-0
	<i>pBIB-Hyg</i>
8kb TT10 (Pourcel et al. 2005)	5

D.

Name of the construct	Crosses (female x male)	
	pTT10_6 in <i>tt1</i> (Ws-4)	#21 x <i>tt1</i>
pTT10_6 in <i>tt16</i> (Ws-4)	#21 x <i>tt16</i>	<i>tt16</i> x #21
pTT10_6 in <i>ttg2</i> (Ws-4)	#21 x <i>ttg2</i>	<i>ttg2</i> x #21

Table 3.4: Transgenic lines generated during this work.

Tables (A) and (B) are summarizing the number of independent transformants analyzed depending on vector and accession. Each line has been genotyped for the presence of the corresponding construct and a *uidA* gene. T2 plant segregation for antibiotic resistance was analyzed. In bold is given the name of the construct referring to its length (see Results, section 2.1.3.1). (A) *TT10* promoter dissection, (\*) - (Pourcel et al., 2005), (B) *TT10* promoter 'sub'-dissection and site-directed mutagenesis, (C) *Ler-0* plants transformed with 8-kb genomic (Col-0) region of the *TT10* gene. T2 plants were antibiotic resistant, (D) Plants with transgene *pTT10:GUS/pBiB-Hyg* #21 introduced by cross (see above 3.1.1.3 iii)) into *tt1*, *tt16* and *ttg2* mutants (see Table 3.3). Selection was based on antibiotic resistance and *transparent testa* and *glabra* (*ttg2*) phenotypes.

### 3.1.1.4 *Arabidopsis thaliana* - cell suspension

The cell suspension culture of *Arabidopsis thaliana* maintained at the Seed Biology Laboratory was a gift from Dr. Geneviève Ephritikhine at Membrane Transport and Signalisation Group, Institut des Sciences du Végétal, CNRS Gif-Sur-Yvette. It was originally obtained from seedlings of Columbia accession by Axelos et al. (1992). The cell suspension culture is composed by small, near-uniform clumps of cells and presents green color.

## 3.1.2 Bacterial Strains

### 3.1.2.1 *Escherichia coli*, strain DH10B

Electro-competent strain, used to multiply all the vectors and for routine cloning.

Genotype: F- *mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*)  $\Phi$ 80*lacZ* $\Delta$ M15  $\Delta$ *lacX74* *deoR* *recA1* *endA1* *araD139*  $\Delta$ (*ara*, *leu*)7697 *galU* *galK*  $\lambda$  *rpsL* (StrR) (Invitrogen).

### 3.1.2.2 *Escherichia coli*, One shot® TOP 10

Electro-competent strain, used for low-efficiency cloning and to recover vectors created by PCR in the site-directed mutagenesis experiment.

Genotype: F- *mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*)  $\Phi$ 80*lacZ* $\Delta$ M15  $\Delta$ *lacX74* *recA1* *araD139*  $\Delta$ (*ara*, *leu*)7697 *galU* *galK* *rpsL* (StrR) *endA1* *nupG* (Invitrogen; users manual no: 28-0126).

### 3.1.2.3 *Escherichia coli*, DB3.1

Electro-competent strain used to multiply GATEWAY™ vectors. Strain is resistant to *CcdB*.

Genotype: F- *gyrA462* *endA1* D(*sr1-recA*) *mcrB* *mrr* *hsdS20*(rb-, mb-) *supE44* *ara-14* *galK2* *lacY1* *proA2* *rpsL20*(Smr) *xyl-5* L- *leu* *mtl-1* (Invitrogen).

### 3.1.2.4 *Agrobacterium tumefaciens*, strain 58C1pMP90

The strain has been described by Koncz et al. (1984) and is used to transform *Arabidopsis thaliana* by the floral dip technique (see 3.2.3.5). The 58C1 strain has a gene for rifampicin resistance in its genome and gentamycin resistance gene on the helper plasmid Ti (pMP90).

## 3.1.3 Yeast strain

### 3.1.3.1 *Saccharomyces cerevisiae*, strain EGY48

Strain used for yeast one hybrid screen of the REGIA bank of transcription factors (see 3.1.6).

Genotype: *MAT $\alpha$* , *his3*, *trp1*, *ura3*, LexAop(x6)-LEU2

### 3.1.3.2 Yeast strains in the REGIA transcription factor library

Detail of the yeast used to create the REGIA transcription factor library are not known (Paz-Ares and Consortium, 2002). However the clones carrying the CDSs of the transcription factors are haploid: *MAT $\alpha$* . The library screening is based on their ability to mate with the *MAT $\alpha$*  (e.g. EGY48) strain to form diploid forms.

### 3.1.4 Plasmids and expression vectors

Table 3.5 presents plasmids used and created during this PhD.

Vector name	Origin / provided by	Used for	Selection / resistance
<i>pCR®-Blunt II-Topo®</i>	Invitrogen	blunt end PCR products cloning	kanamycin
<i>pGEM-T</i>	Promega	T-end PCR products cloning	B-galactosidase / ampicillin
<i>pBS-GUS</i>	Debeaujon et al., 2003		ampicillin
<i>pTT10:GUS/pBS</i>	Pourcel et al., 2005	template for <i>TT10</i> promoter fragments amplification and site-directed mutagenesis	ampicillin
<i>pTT10:GUS/pBIB-HYG</i> <i>8-kb TT10/pBIB-HYG</i>		plant transformation	kanamycin; hygromycin in plants
<i>pBIB-HYG</i>	Becker, 1990	<i>SmaI/KpnI</i> sub-cloning from <i>pBS-GUS</i> vector and plant transformation	kanamycin; hygromycin in plants
<i>pCAMBIA 1300</i>	www.cambia.org	<i>KpnI/PstI</i> subcloning from <i>pBS-GUS</i> vector and plant transformation	kanamycin; hygromycin in plants
<i>pBI101GUS GTW</i>	Invitrogen (Baudry et al. 2006)	<i>attR1/attR2</i> , Recombination (LR) from <i>pDONR207</i> to prepare transcriptional fusions with <i>GUS</i> reporter gene	chloramphenicol / kanamycin; kanamycin in plants
<i>pDONR207</i>	Invitrogen	<i>attP1/attP2</i> - BP Recombination - PCR fragments entry vector	chloramphenicol / gentamycin
<i>pHISi</i>	Clontech	<i>SmaI/SacII</i> , DNA / promoter fragments analysis in one hybrid experiments	Ampicilline in <i>E.coli</i> ; Insertion in <i>ura3</i> locus of the yeast, SD <i>his</i> -selection in the EGY48 strain
<i>pTT10_X:GUS/pTopo®</i>	this work	cloning of PCR fragments of <i>TT10</i> promoter	kanamycin
<i>pTT10_X,Y:GUS/pBS</i>		<i>XhoI/NcoI</i> promoter fragments cloning in front of <i>GUS</i> reporter gene; transient expression in protoplasts	ampicillin
<i>pTT10_X:GUS/pBIB-HYG</i>		plant transformation	kanamycin; hygromycin in plants
<i>pTT10_Y:GUS/pCAMBIA1300</i>		plant transformation	kanamycin; hygromycin in plants
<i>pTT10_Z/pDONR207</i>		sub-cloning	chloramphenicol / gentamycin
<i>pTT10_Z:GUS/pBI101GTW</i>		plant transformation	chloramphenicol / kanamycin; kanamycin in plants
<i>pTT10_3(Accession)/pTOPO®</i>		cloning of PCR fragments of <i>TT10</i> promoter from accessions; sequencing	kanamycin
<i>pTT10_3(Accession):GUS/pBS</i>		<i>XhoI/NcoI</i> promoter fragments cloning in front of <i>GUS</i> reporter gene; transient expression in protoplasts	ampicillin
<i>pTT10 4.5kb (Ler)/pTOPO®</i>		cloning of PCR fragments of <i>TT10</i> promoter from accessions; sequencing	kanamycin
<i>pLAC3, 4, 5 and 12/pDONR207</i>		sub-cloning	chloramphenicol / gentamycin
<i>pLAC3, 4, 5 and 12/pBI101GTW</i>		plant transformation	chloramphenicol / kanamycin; kanamycin in plants

Table 3.5: List of vectors used in this study.

Table summarizes plasmids used for molecular cloning, sequencing, transient expression and *Arabidopsis* transformation. Promoter - GUS constructs cloned during this work are listed as a pool, where X in the name of the construct means promoter fragments from 1 to 5 (for details see Results, section 2.1.3.1), Y is for site-directed mutagenesis constructs and control of pTT10\_3, and Z = Y + negative control of pTT10\_2. *TT10* promoter cloned from different accessions are referring to 465bp = *pTT10\_3* promoter amplified on genomic DNA of Ws-4, *Ler*-0, Col-0, Cvi-0, Bay-0 and Sha. Promoters of the LAC3, LAC4, LAC5 and LAC12 were amplified on the Col-0 genomic DNA and are listed together. GATEWAY™ vectors can be multiplied only in *Escherichia coli*, DB3.1 strain.

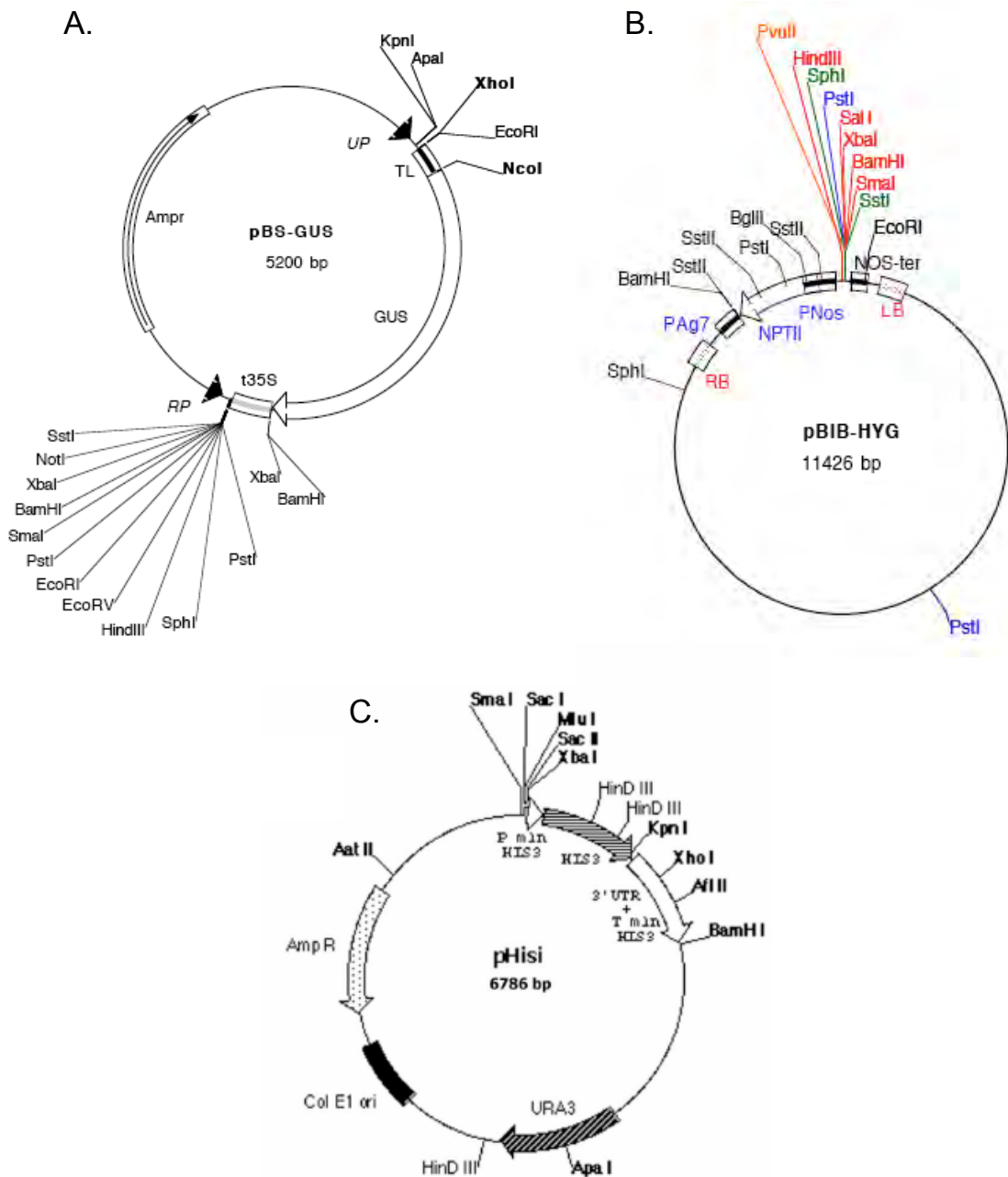


Figure 3.1: Maps of the vectors used for cloning.

(A) *pBS/GUS* vector used to prepare transductional fusions of the promoter and GUS reporter gene, subcloning and transient expression. (B) Binary vector *pBIB-Hyg* used for transfer of the constructs in the T-DNA cassette into plants. (C) *pHISi* vector used for the transcriptional fusion of the nucleotide sequence and *HIS3* reporter gene.

### 3.1.5 Cvi and *Ler* BAC library at MPI, Cologne

BAC genomic library of two *Arabidopsis* accessions, *Ler* and *Cvi* was available in Wim Soppe lab (Max Planck Institute for Plant Breeding Research, Cologne, Germany). Briefly, libraries contain  $\approx 10\times$  coverage of the *Arabidopsis* genome. The *Cvi* genomic library 109 is in pIndigoBAC-5 vector (Chloramphenicol

resistance), cloned in HindIII restriction site with the average insert size 142kb. In the Ler 114 BiBAC2 (Kanamycin selection) library fragments were cloned in BamHI manner, with 162kb average size of the insert.

### 3.1.6 Transcription factor library

To identify transcription factors binding to the promoter in a yeast one hybrid experiment, a transcription factor library created by REGIA (REGulatory Gene Initiative in Arabidopsis) consortium was used. Library consists of more than 2000 clones of a normalized full length open reading frame transcription factors in a yeast (Paz-Ares and Consortium, 2002).

## 3.2 Methods

### 3.2.1 Handling of bacteria

#### 3.2.1.1 Culture

Bacteria are grown in LB medium (Luria-Bertani:  $5g \cdot l^{-1}$  yeast extract,  $10g \cdot l^{-1}$  bacto-tryptone,  $10g \cdot l^{-1}$  NaCl) at 37°C for *E. coli* and 28°C for *A. tumefaciens*. If needed, the medium is supplemented with antibiotics (purchased from Duchefa, Haarlem, Netherlands), as shown in Table 3.6. The medium can be liquid or solidified with 1.6 % Bacto agar.

Antibiotics	Stock (mg/ml)	Solvent	Final concentration (mg/l)		
			Bacteria	Plants	
				T1 selection	Routine
Ampicillin	100	water	100	n/a	n/a
Chloramphenicol	12,5	ethanol 50°	12,5	n/a	n/a
Gentamycin	50	ethanol	50	n/a	n/a
Hygromycin	50	water	50	50	10
Kanamycin	100	water	100	100	10
Rifampicin	50	DMSO	50	n/a	n/a

Table 3.6: Antibiotics

#### 3.2.1.2 Preparation and transformation of electro-competent bacteria

A bacterial culture initiated in 5 ml LB medium is incubated overnight at 37°C, under agitation at 220 rpm. One ml of this pre-culture is used to inoculate a 500 ml LB culture which is grown until reaching the exponential phase of growth ( $OD_{600nm} = 0.4 - 0.6$ ). Bacteria are centrifuged 10 min at 6000 rpm and the pellet resuspended in 500 ml 10% sterile glycerol previously cooled at 4°C. Four successive centrifugations are realized with a progressive reduction of glycerol volume used for resuspension: 250, 125, 5 and 1 ml. Bacteria are aliquoted as 10  $\mu$ l fractions and stored at -80°C.

### 3.2.2 Handling of yeast

#### 3.2.2.1 Culture

For routine culture, yeasts (*Saccharomyces cerevisiae*) are grown in YPD medium ( $10g \cdot l^{-1}$  yeast extract,  $20g \cdot l^{-1}$  bacto-peptone;  $10g \cdot l^{-1}$   $\alpha$ -glucose; Clontech, Mountain View, USA) at 28°C. The medium can be



liquid, or solidified with 2,4 % Bacto agar. For transformant selection, culture is realized in minimal SD (synthetic defined) medium (Clontech) without uracil and/or histidine.

### 3.2.2.2 Preparation and transformation of yeast

The LIAC/SS carrier DNA/PEG method (Gietz and Woods, 2002) was applied (see The TRAF0 page at <http://home.cc.umanitoba.ca/~gietz/method.html> for details). Briefly, 5 ml YPD culture is inoculated and grown o/n with shaking at 28°C. This culture is used to inoculate 50 ml YPD medium which is grown until reaching a cell density of  $2 \cdot 10^7 \text{ cells} \cdot \text{ml}^{-1}$  (quantity for 10 transformations). After harvest and rinsing in water, the cells are resuspended in 500  $\mu\text{l}$  100 mM lithium acetate (LiAc) at  $2 \cdot 10^9 \text{ cells} \cdot \text{ml}^{-1}$ . The transformation mix consists ină: 240  $\mu\text{l}$  PEG (50% w/v), 36  $\mu\text{l}$  1.0M LiAc, 50  $\mu\text{l}$  SS-DNA as carrier ( $2.0 \text{ mg} \cdot \text{ml}^{-1}$ ),  $x \mu\text{l}$  plasmid DNA (0.1-10  $\mu\text{g}$ ),  $34 - x \mu\text{l}$  sterile bidistilled water). After aliquoting in 50- $\mu\text{l}$  fractions, cells are pelleted. 355  $\mu\text{l}$  TRAF0 mix and 5  $\mu\text{l}$  plasmid DNA are added on the pellet and vortexed. The mix is incubated at 28°C for 30 min before being heat-shocked at 42°C for 30 min. After pelleting and resuspension in water, 2-200  $\mu\text{l}$  of the mix is plated on selection medium.

## 3.2.3 Handling of *Arabidopsis thaliana*

### 3.2.3.1 Seed sterilization

The sterilization solution is prepared by dissolving a tablet containing 1.5 g active chloride (Inovchloreă; Inov'Chem, Brest, France) in 50 ml distilled water and 5 ml of this solution is diluted in 45 ml 96° ethanol (final solution exhibiting 1° active chloride). Seeds are incubated during 10 min in the sterilization solution, rinsed 2 times with 96° ethanol and finally dried overnight under a flow hood.

### 3.2.3.2 Plants growth conditions *in vitro*

Seeds are sown on 'Arabidopsis' medium, containingă 3·1g.l<sup>-1</sup> Gamborg's B5 medium (macro- and microelements, vitaminsă; Duchefa, Netherlands), 1% (w/v) sucrose, 1% (w/v) agar and 5ml·l<sup>-1</sup> pH indicator BCP (bromocresol purple) at 0.16 %, everything being autoclaved 20 min at 120°C. Seeds are stratified during 3 days at 4°C in the dark to break dormancy and synchronize germination and afterwards transferred in a growth cabinet (16h light at 20°C / 8h dark at 18°C; 60% relative humidity; 200 mE·m<sup>-2</sup>·s<sup>-1</sup> light intensity at the Petri plate level) to get plantlets. For the analysis of antibiotic resistance, seeds are germinated on 'Arabidopsis' medium supplemented with relevant antibiotics, as shown in Table 3.6. Around 12 days after transfer of germination plates to light, resistant versus sensitive plantlets can be numbered.

### 3.2.3.3 Plants growth conditions in greenhouse

Plants are grown in soil (Tref Substrates) and regularly watered with a nutritive solution (Fertil). A photoperiod of 13h with a minimal light intensity of 100 mE·m<sup>-2</sup>·s<sup>-1</sup> is ensured by additional lightning. Temperature is around 22°C during the day and 18°C during the night. Hygrometry is not controlled.

### 3.2.3.4 Crosses

A flower bud from the mother plant is opened by inserting the tip of one pair of forceps between petals and sepals. All immature anthers are removed with the other pair of forceps. An anther is taken from an open, mature flower from the father plant and tapped on the stigma until covering it with pollen grains. The cross is marked with a colored thread and documented (mother, father, date, color code).

### 3.2.3.5 Plant transformation by the floral dip method

Plant transformation was performed as described by Clough and Bent (1998). *Arabidopsis* adult plants were dipped into a solution containing *Agrobacterium*, 5% sucrose, and 50 ml · l<sup>-1</sup> of the surfactant Silwet (Witco, Geneva, Switzerland).

### 3.2.3.6 Maintenance of cell suspensions

Cells have been maintained in JPL medium (Jouanneau and Péaud-Lenoël, 1967) under continuous light at 22 ° C and subcultured at one-week intervals with 10% culture in 100 ml fresh medium.

### 3.2.3.7 Genotyping

Genotyping of the transformant and the mutant lines was applied whenever it was necessary to screen plants for the presence of a desired insert. PCR was done with a gene-specific primer and a vector specific primer.

### 3.2.3.8 Transient expression of promoter:GUS constructs in *Arabidopsis* protoplasts

Protoplasts were isolated from *Arabidopsis* cell suspensions, and GUS constructs were transiently expressed in protoplasts by PEG-mediated transformation as described in Marmagne et al. (2004). Briefly, *Arabidopsis* suspension cells were digested in Gamborg's B5 medium supplemented with 0.17 M glucose, 0.17 M mannitol, 1% cellulase, and 0.2% macerozyme. The protoplasts were purified by floatation in Gamborg's B5 medium supplemented with 0.28 M sucrose.

For transformation,  $0.2 \cdot 10^6$  protoplasts were mixed with 5 µg of plasmid DNA in a solution containing PEG 6.000 25%, 0.45 M mannitol, 0.1 M Ca(NO<sub>3</sub>), pH 9, and incubated in the dark for 20 min. Then the PEG was washed twice with 0.275 M Ca(NO<sub>3</sub>) and the protoplasts were re-suspended in Gamborg's B5 medium supplemented with 0.17 glucose, 0.17 mannitol, and maintained in this medium until microscopic observation.

## 3.2.4 DNA methods

### 3.2.4.1 Isolation of plasmid DNA from *E. coli*

Plasmid DNA was purified with the Qiagen Plasmid Maxi/Midi kit (Qiagen, Hilden, Germany) or GenElute<sup>TM</sup> Plasmid Miniprep kit (Sigma, USA), following the manufacturer's recommendations.

### 3.2.4.2 DNA fragments purification

DNA fragment after restriction were purified with GeneElute<sup>TM</sup> Agarose Spin Columns (Sigma), or JET-SORB (Genomed) kits according to the manual's instructions.

### 3.2.4.3 Isolation of genomic DNA from plants

The protocol uses the detergent properties of CTAB (cetyltrimethylammonium bromide) leading to cell lysis and to protein and polysaccharide denaturation. Around 3 g rosette leaves are ground in liquid nitrogen and transferred in a Falcon tube containing 11 ml extraction buffer (2% CTAB w/v, 1.4M NaCl, 20 mM EDTA, 0.2% β-mercaptoethanol v/v, 100 mM Tris-HCl pH 8.0) and incubated 30 min at 60°C with shaking. Organic residues (lipids, chlorophylls, ...) are extracted with 11 ml chloroform/isoamylalcohol (24:1 v/v) and eliminated by centrifugation (10 min, 10.000g). Nucleic acids from the supernatant are precipitated with 1 vol. isopropanol after o/n incubation at 4°C and centrifugation (20 min, 400g). The pellet is

resuspended in 500  $\mu\text{l}$  water and RNAs are eliminated by incubating at 37°C during 30 min with 30  $\mu\text{l}$  RNaseA (Sigma) at 10  $\text{mg} \cdot \text{ml}^{-1}$ . After deproteinisation by centrifugation (10 min, 10.000g) in 1 vol. phenol/chloroform/isoamylalcohol (25:24:1 v/v/v) and then 1 vol. chloroform/isoamylalcohol (24:1 v/v), DNA is precipitated by 1 vol. isopropanol and 0.1 vol. sodium acetate, then rinsed with 70% ethanol and resuspended in 100  $\mu\text{l}$  water.

#### 3.2.4.4 Rapid DNA extraction for genotyping

This protocol was used for DNA extraction for plant genotyping. Briefly, young leaf 5-7 mm in diameter was ground in liquid nitrogen in an Eppendorf tube. 60  $\mu\text{l}$  250 mM NaOH was added and sample was boiled for one minute. To neutralize, 60  $\mu\text{l}$  250 mM HCl and 30  $\mu\text{l}$  Tris-HCl 0.5 N, pH 8.0-Triton 0.25% v/v are added. Sample is boiled for one minute and centrifugated (2 min, 12000g) to pellet lysis remainings. 1  $\mu\text{l}$  of the supernatant is used in standard PCR reaction.

#### 3.2.4.5 Polymerase chain reaction (PCR)

PCR (Mullis et al., 1986) was employed to amplify DNA fragments for cloning, for genotyping of transgenic *Arabidopsis* plants, and for screening of transformed bacterial colonies, as well as performing site-directed mutagenesis. General cycling conditions were: 94°C to 98°C for 2 minutes, 35 cycles of 94°C for 30 seconds, 55°C to 65°C for 30 seconds and 72°C (1kb/minute), and a final extension step of 72°C for 1 minute.

#### PCR for cloning

When a fragment was going to be cloned, a proof-reading DNA polymerase was used. A normal reaction of 50 $\mu\text{l}$  contained 1X enzyme buffer, 0.2mM of each dNTP, 10 pmol of forward primer, 10 pmol of reverse primer, 1 unit of DNA polymerase, DNA template and water. Cycling conditions for Phusion High-Fidelity DNA polymerase (Finnzymes, Finland) were: 98°C for 30 seconds of initial denaturation followed by, 35 cycles of 98°C for 7 seconds, 55°C for 15 seconds and 72°C for up to 3 minutes (1kb/20 seconds), and a final extension step of 72°C for 1 minute. AccuTaq<sup>TM</sup> LA DNA polymerase (Sigma, USA) was used for amplification of long DNA fragments, with cycling conditions: 98°C for 1 minute of initial denaturation, followed by 35 cycles of 94°C for 15 seconds, 55°C for 20 seconds and 68°C for up to 20 minutes (1kb/minute), and a final extension step of 68°C for 10 minutes. Reaction buffer composition was optimized for magnesium concentration (2-4 mM) and supplemented with DMSO (1-4% final concentration) according to the manufacturer manual.

#### PCR on bacterial colonies

Colony PCR was applied whenever it was necessary to screen bacterial colonies for the presence of a desired insert. Colony PCR was done with a gene-specific primer and a vector specific primer. A typical 20  $\mu\text{l}$  reaction consisted of 1X Taq DNA polymerase buffer, 0.2mM of each dNTP, 10 pmol of forward primer, 10 pmol of reverse primer, 0.25 unit of Taq DNA polymerase (New England Biolabs) and water. The PCR mix was distributed into reaction tubes but no template DNA was added. Instead a sterilized toothpick was used to touch a bacterial colony on a plate and then the colony was mixed with the PCR mix in each tube. Cycling conditions were almost the same as the general cycling conditions, the only change was in the first step, that was established as 95°C for 5 minutes.

#### 3.2.4.6 PCR fragments purification

PCR fragment purification was done with GeneElute<sup>TM</sup> Agarose Spin Columns (Sigma), or MinElute PCR purification Kit (Qiagen) kits according to the manual's instructions.

### 3.2.4.7 Sequencing

Sequences were realized at the Seed Biology Laboratory, on an Applied Biosystems Abi Prism 310 Genetic Analyzer sequencer using the BigDye-terminator v3.1 chemistry.

### 3.2.4.8 Site-directed mutagenesis

The QuikChange® Site-Directed Mutagenesis Kit from Stratagene was used according to manufacturer instructions (from 30 ng of the original plasmid). Primers were design using the primerX software available at the <http://www.bioinformatics.org/primerx/>.

## 3.2.5 RNA methods

### 3.2.5.1 RNA isolation and cDNA synthesis

Frozen tissues were ground in liquid nitrogen and total RNA was extracted with the RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions, completed with a on-column DNase treatment (Qiagen DNase 'Rnase-free') to avoid DNA contamination. The only change to the protocol was addition of polyvinylpyrrolidone (Sigma) for tissue grounding. The RNA extracts were treated with 30 U of RNase-free DNase I (Qiagen) and were eluted with 40 µl of diethyl pyrocarbonate-treated water. For RT-PCR studies, DNA-free RNA was converted into first-strand cDNA using the SuperScript pre-amplification system for first-strand cDNA synthesis (Gibco BRL) with oligo(dT)22.

N.B. For the purpose of reliable results, high quality RNA is obligatory. RNA integrity was checked on agarose gel and ratio of the 28S/18S has been verified, and only samples with ratio above 1.8:1 were used. Quantification of the RNA was made with NanoDrop (Thermo scientific, Fisher).

To ensure reliable comparison of gene expression between *Arabidopsis* accessions and mutants open flowers were marked and siliques at the same developmental stage were collected (Fig. 3.2). For all comparison studies plants were grown in the greenhouse at the same time. Samples, were collected from three plants, except At5g62320/MYB99 where only one homozygous plant was available. RNA was extracted from three siliques from three different plants, collected and immediately frozen in liquid nitrogen. In each experiment the same amount of RNA was used for cDNA synthesis. Three technical replicates of qRT-PCR were realized for each cDNA synthesis.

### 3.2.5.2 RT-PCR

The cDNA samples were diluted 10- and 100-fold, and 1 µl of dilution was amplified in a 20-µl standard PCR mixture. First expression of the control gene *EF1* was checked and that depending on the gene expression level, PCR reactions were carried out on different cDNA dilutions.

### 3.2.5.3 Quantitative real-time RT-PCR (qRT-PCR)

qRT-PCR was carried out as described in Baud et al. (2004), the reaction was performed on the LightCycler Instrument (Roche, Meylan, France) with the LightCycler-FastStar DNA Master SYBR Green I kit for PCR (Roche) according to the manufacturer's protocol. Each reaction was performed with 5 µl of 1:10 to 1:100 dilution of the first cDNA strands in a total volume of 20 ml. The reactions were incubated at 95 °C for 8 min to activate the hot start recombinant Taq DNA polymerase, followed by 45 cycles of 10 s at 95 °C, 6 s at 55 °C, and 20 s at 72 °C. The primer set for the *TT10* gene, namely 3Fw-TT10 and 3Rev-TT10 (see Appendix, Tab. 5.11) was previously designed by Lucille Pourcel (unpublished results). The specificity of

the PCR amplification was checked with a heat dissociation protocol (from 65 °C to 95 °C) following the final cycle of the PCR. PCR products were then purified and sequenced (Pourcel, unpublished). The efficiency of the primer set was calculated by performing real time PCR on several dilutions of first strands (see section 3.2.5.5). The results obtained for the different tissues analyzed were standardized to the constitutive *EF1* gene expression level (Liboz et al., 1990). We confirmed that the level of *EF1* expression was comparable between accessions studied by analysis of expression of two other reference genes (see 3.2.5.4).

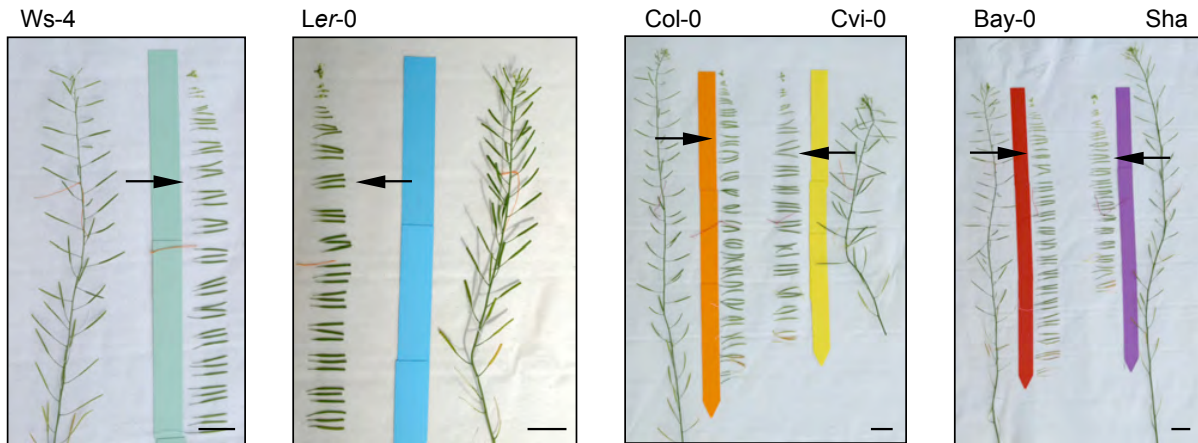


Figure 3.2: Flower marking and silique sample collection.

Sample preparation for *TT10* expression comparison between accessions. *Arabidopsis* accessions used in the study were developing different number of siliques per day. Picture is showing different height of stem and siliques developed in 23 days counted from first flower appearance. Arrows are showing siliques 7 days after flowering.

Bar = 2cm

#### 3.2.5.4 Normalization controls for qRT-PCR

In qRT-PCR, expression of the gene of interest is presented as an amount of the control gene, which amount is thought to be equal among the samples (e.g. tissues; accessions in this study). Here we demonstrate, that the expression of two other reference genes is comparable between the samples. We can assume that together with *Elongation Factor 1aA4*, (At1g07920, At5g60390) they are reliable control genes and can be used for normalization of gene expression in qRT-PCR experiment.

#### 3.2.5.5 Primer efficiency

Quantitative RT-PCR results could be artefactual due to the polymorphism in the primer binding sites and/or secondary structures of the cDNA. By showing that the primer efficiency is comparable between the samples, we assumed that the amplification was not affected by those factors in *Arabidopsis* accessions analyzed.

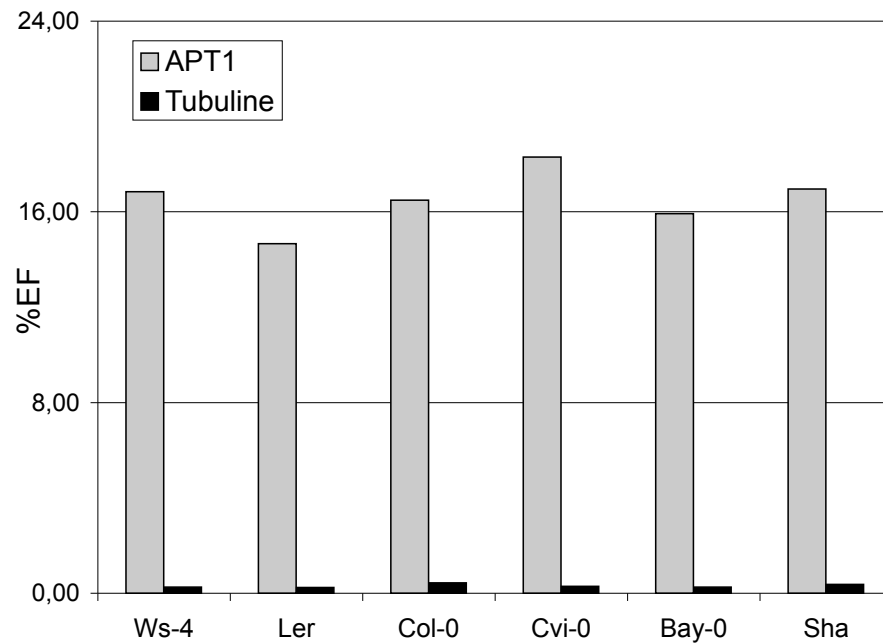


Figure 3.3: Expression of *APT1* and *TUBULINE* in siliques 7DAF.

The expression of two reference genes, *APT1* (adenine phosphoribosyltransferase, At1g27450) and *TUBULINE* (tubuline beta-4 chain, At5g44340) have been expressed as a percentage of the *Elongation Factor 1aA4* gene expression.

cDNA Dilution	Sample / number of cycles						H <sub>2</sub> O
	Ws-4	Ler	Col-0	Cvi-0	Bay-0	Sha	
x10	18,45	18,93	17,53	19,48	19,08	17,92	n/a
x100	22,08	22,89	21	23,06	22,69	21,44	
x1000	25,4	26,01	24,34	26,43	25,96	24,92	
x10000	28,61	29,03	27,84	29,55	29,04	28,23	
X100000	<b>30,7</b>	<b>30,87</b>	<b>30,43</b>	<b>31,04</b>	<b>30,69</b>	<b>30,56</b>	<b>31,68</b>
Slope	3,38	3,341	3,428	3,359	3,315	3,441	n/a
Error	0,0538	0,126	0,0166	0,0577	0,0662	0,0282	
<b>Primer efficiency</b>	<b>97-98%</b>	<b>99%</b>	<b>99%</b>	<b>98-99%</b>	<b>100%</b>	<b>95-96%</b>	

Table 3.7: Efficiency of the 3Fw/3Rw set of primers.

The efficiency of the 3Fw/3Rw set of primers is comparable and close to 100% in all the *Arabidopsis* accessions tested. The efficiency was calculated by the curve fitting for the pair of primers used to study *TT10* gene expression. The highest cDNA dilution sample (x100000) was not included.

### 3.2.6 Histochemical methods for GUS detection and quantification

#### 3.2.6.1 Histochemical detection of GUS activity

The histochemical analysis of GUS activity was performed as described by Debeaujon et al. (2003). Tissues were prefixed in 90% (v/v) ice-cold acetone:water for 30 min at 4°C. After three rinses in 0.1 M phosphate buffer, pH 7.2 (PB), they were transferred in a PB solution containing 2 mM 5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucuronide (Duchefa, Haarlem, The Netherlands), 0.1% (v/v) Triton X-100:water, 0.5 to 10 mM each of potassium ferrocyanide and potassium ferricyanide, depending on the construct, and 10 mM Na-EDTA. Vacuum was applied for 1 h before incubating for 12 h at 37°C in the dark. Chlorophyll was removed by room temperature incubation in 70% (v/v) ethanol:water. Afterward, stained tissues were either cleared with a chloral hydrate solution (chloral hydrate:distilled water:glycerol (8:2:1,w/v/v) as whole mounts or embedded in resin after fixation. Namely, developing seeds for resin embedding were fixed in 2% (v/v) glutaraldehyde and 1% (v/v) paraformaldehyde in 0.1 M phosphate buffer, pH 7.2. After incubation at 4°C for 24 h, samples were washed in phosphate buffer, dehydrated using a series of graded ethanol solutions, and embedded in Technovit 7100 resin (Heraeus Kulzer, Wehrheim, Germany) according to the manufacturer's instructions. Sections (4-7  $\mu$ m) were made on a rotary Jung RM 2055 microtome (Leica Microsystems, Heidelberg, Germany) equipped with metallic blades (Heraeus Kulzer). Observations were performed with an Axioplan 2 microscope (Zeiss, Jena, Germany) equipped with bright-field optics. Microscopic observations were performed with Nomarski differential interference contrast optics.

#### 3.2.6.2 Quantification of GUS activity (MUG assay)

The fluorometric assay was performed as described by Jefferson et al. (1987) on extracts made from transgenic and control plants. The reaction product of the assay, namely 4-methylumbelliferone, was measured kinetically, in duplicate, using a fluorometer (Fluoroskan II; Labsystems, Helsinki, Finland). Protein concentration was estimated for each extract using the method of Bradford (1976).

Results of measured GUS activity are expressed in pmol/min/mg protein and are representation of experiment which was repeated twice with the same tendency. For each construct 2 or 3 independent transformant lines were studied. Two samples for each line were analyzed separately (sample consist of three siliques of the same age collected from 3 plants). Standard deviation bars are representation of the differences for all the samples of the certain construct.

### 3.2.7 Analysis of seed flavonoids

The results of metabolomic analysis of flavonoid composition presented, was done by Jean-Marc Routaboul.

The analysis of flavonoid composition from mature seeds was modified from Routaboul et al. (2006) as follows. Fifteen mg of seeds were ground for 1 min at maximum speed using a FastPrep-24 sample preparation system (MP Biomedical) in 2ml-tube containing 1 1/4 inch ceramic sphere and 1 ml acetonitrile / water (75:25; v/v). For some experiments, a solvent designed to maximize procyanidins extraction was used (MeOH / Acetone / water / TFA, 32:40:28:0.05; v/v/v/v). Samples were sonicated for 20 min at 4°C and centrifugated at 13000rpm for 10 min. Following centrifugation, the pellet was extracted further with 1 ml acetonitrile / water (75:25; v/v) overnight at 4°C. The two extracts were pooled. An aliquote of the final extract was filtered on a 0.45 $\mu$ m syringe filter before being analyzed by mass spectrometry. The pellet was preserved for insoluble PA analysis.

Analyses of soluble and insoluble PAs were carried out according to Porter et al. (1986), with minor changes, on 500  $\mu$ l aliquots of the final extracts and on the entire remaining pellets, respectively. Three ml

of butanol-HCl reagent (butanol / concentrated HCl, 95:5, v/v) and 0.1 ml of the ferric reagent (2% ferric ammonium sulfate in 2N HCl) were added. Soluble anthocyanins were then measured at 550 nm. The tubes were put in a boiling water bath adjusted at 98°C for 60 min. After cooling the tube on ice, PA-related absorbance was recorded at 550 nm. The amount of PAs or anthocyanins are expressed in  $mg \cdot g^{-1}$  cyanidin equivalent using a calibration curve.

LC-MS analysis was conducted using a 'Quattro LC' with an ESI 'Z-spray' interface (MicroMass Co, Manchester, UK), MassLynx software, an Alliance 2695 RP-HPLC system (Waters, USA), and a Waters 2487 UV detector set at 280 nm. An Uptisphere C18 column (150x2 mm, 5  $\mu$ m, Interchrom) was used with a mix comprising solvent A acetonitrile / water, (95:5, v/v, 0.2% acetic acid) and solvent B acetonitrile / water (5:95, v/v, 0.2% acetic acid) with a gradient profile (starting with 10:90, A/B, v/v, for 5 min; linear gradient up to 70:30, A/B over 35 min; a washing step 100:0, A/B for 15 min and final equilibration at 10:90, A/B for 20 min) at a flow rate of 0.2 ml/min. Relative quantification was based on the area of major MS signals. Flavonol contents were expressed relative to quercetin-3-O-rhamnoside and rutin (Extrasynthese, France) for mono glycosylated and di-glycosylated flavonols, respectively. ESI source parameters were optimized using these standards.

### 3.2.8 Routine techniques

Routine techniques, such as DNA agarose gel, DNA precipitation, DNA ligation, DNA cleavage with restriction endonucleases and DNA concentration measurement were done according to Sambrook and Russel (2001).

### 3.2.9 Internet resources and services

#### 3.2.9.1 The Arabidopsis Information Resource - TAIR

<http://www.arabidopsis.org/>

TT10 at TAIR

#### 3.2.9.2 The Arabidopsis Thaliana Integrated Database - ATIDB

<http://www.atidb.org/>

TT10 at ATIDB

#### 3.2.9.3 Integrative database around plant genomes - FLAGdb++ v.3.9

<http://urgv.evry.inra.fr/FLAGdb> (Samson et al., 2004)

#### 3.2.9.4 The Complete Arabidopsis Transcriptome Micro Array database - CATMA Database

<http://www.catma.org/>

CATMA Database (Crowe et al., 2003)

#### 3.2.9.5 AtcisDB at AGRIS

<http://arabidopsis.med.ohio-state.edu/AtcisDB/>

AtcisDB is the Arabidopsis cis-regulatory element database at Arabidopsis Gene Regulatory Information Server. It can be searched for *cis*-acting regulatory elements with AtcisDB Search Engine (Molina and Grotewold, 2005; Palaniswamy et al., 2006).



### 3.2.9.6 The Genevestigator - expression database and meta-analysis system

<https://www.genevestigator.com/gv/index.jsp>

The database has been first queried in January 2006, then systematically verified. (Last updated: 7.04.2009). (Zimmermann et al., 2005; Hruz et al., 2008).

### 3.2.9.7 The Bio-Array Resource for Arabidopsis Functional Genomics - BAR

<http://bar.utoronto.ca/> (Toufighi et al., 2005)

#### Arabidopsis eFP Browser

<http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi> eFP Browser is a visualization tool which pictures gene expression for easier interpretation. More over, presented data are normalized, signal threshold can be set up and standard deviation filtering can be used to keep only statistically significant results. Signal values lower than 100 (expression potential) should be considered as background (Bassel, personal communication; Dubos, personal communication). (Schmid et al., 2005; Kilian et al., 2007; Winter et al., 2007).

#### Expression Angler

[http://bar.utoronto.ca/ntools/cgi-bin/ntools\\_expression\\_angler.cgi](http://bar.utoronto.ca/ntools/cgi-bin/ntools_expression_angler.cgi) Expression angler is program calculating the correlation coefficients for expression for all gene expression vectors compared to the gene of interest. (Toufighi et al., 2005).

### 3.2.9.8 GeneCAT - Gene Co-expression Analysis Toolbox

<http://genecat.mpg.de/cgi-bin/Ainitiator.py>

GeneCAT is composed of 5 tools (e.g ExpressionProfiling, ExpressionTree, Co-expression analysis) for gene expression analysis (Cite Mutwil)

### 3.2.9.9 DIURNAL

<http://diurnal.cgrb.oregonstate.edu/>

The DIURNAL search tool provides easy access and visualization of circadian/diurnal gene expression data. (Mockler et al., 2007).

Data from ATH1 chip were normalized with gcRMA method (Wu et al., 2004).

### 3.2.9.10 PLACE and Signal Scan - A Database of Plant *Cis*-acting Regulatory DNA Elements

<http://www.dna.affrc.go.jp/PLACE/signalscan.html>

SIGNAL Scan (Prestridge, 1991) is a computer program that scans DNA sequences for known transcriptional elements based on the PLACE database (Higo et al., 1999). Last update of PLACE database: 08.01.2007.

### 3.2.9.11 Plant CARE

<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>

Plant Cis-Acting Regulatory Element database and a portal to tools for in silico analysis of promoter sequences (Lescot et al., 2002).

**3.2.9.12 Plant PAN - Plant Promoter Analysis Navigator**

[http://plantpan.mbc.nctu.edu.tw/gene\\_group/index.php](http://plantpan.mbc.nctu.edu.tw/gene_group/index.php)

Plant PAN is a plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements (Chang et al., 2008).

**3.2.9.13 ATHENA**

<http://www.bioinformatics2.wsu.edu/Athena>

Athena is a web-based application with large set of data visualization, mining, and analysis tools related to the control of gene expression (O'Connor et al., 2005).

**3.2.9.14 Basic Local Alignment Search Tool - BLAST**

<http://blast.ncbi.nlm.nih.gov/Blast.cgi> (Altschul et al., 1990)

**3.2.9.15 Multiple Sequence Alignment**

<http://bioinfo.genotoul.fr/multalin/multalin.html> (Corpet, 1988)

**3.2.9.16 PrimerX - Primer design for the site directed mutagenesis**

PrimerX is a program for automated design of mutagenic primers for site-directed mutagenesis. <http://www.bioinformatics.org/primerx/>



## Chapter 4

# Discussion, Conclusions and Perspectives

### 4.1 Regulation of *TT10* gene expression

The modalities of gene expression and regulation of gene expression are related to the gene function. House-keeping genes are typically constitutively expressed at relatively constant level across many or all known conditions. On the other hand, inducible genes are the ones, which expression is either responsive to the cell cycle, developmental stage or environmental change. Those regulation systems would ensure, that the functional protein is produced where and when it is needed. In case of *TRANSPARENT TESTA 10* gene, it seems that the expression is tightly regulated during plant development and is mainly restricted to developing seeds and siliques. It has been shown that *TT10* is expressed in two layers of the seed coat, where it colocalizes with, and is involved in metabolism of flavonols and tannins (Pourcel et al., 2005; Pourcel, 2006). It has also been reported that the *tt10* mutant exhibits a delay in root elongation and seeds accumulated nearly 30% less extractable lignin than the wild type (Liang et al., 2006a). In the present study we were mainly interested in how *TT10* expression is regulated during plant development, however previously raised concept of stress induced regulation (Pourcel et al., 2005; Pourcel, 2006; Liang et al., 2006a) was also investigated.

#### 4.1.1 Transcript accumulation

Previous studies reported *TT10* transcript in root, leaf, stem and flowers (McCaig et al., 2005; Cai et al., 2006; Liang et al., 2006a; Abdel-Ghany and Pilon, 2008), however in the present study we were able to detect *TT10* mRNA only in flower, developing siliques and seeds, what is consistent with the results of Pourcel et al. (2005). Surprisingly, we have found that during silique development expression was bi-phasic with the maximum of transcript accumulation at 7 and 15 DAF. This feature could be explained by gene expression, first in the endothelium layer, and second peak could be due to the expression in the outer integument 1 (*oi1*). However, promoter activity studies revealed, that GUS staining in *oi1* was already detectable at 8 DAF (Pourcel et al. (2005), and this study). Most probably, the first peak of expression could reflect *TT10* expression in seeds, whereas the second peak could come from the expression in the valve or replum, because RNA was extracted from the whole siliques (without receptacle and peduncule). Another possible explanation for the intermediate decrease of transcript amount, would be the massive transcriptional reprogramming of gene expression, which is observed during the transition stage of the seed development (Borisjuk et al., 2004). Moreover, Ruuska et al. (2002) have shown several distinct expression patterns of

gene expression during *Arabidopsis* seed filling. It would be interesting to compare *TT10* expression profiles in dissected seeds and other components of the silique to find out if the observed bi-phasic expression is due to the different time of expression in those organs.

#### 4.1.2 Developmental vs. stress regulated expression

Quantitative RT-PCR used in this study did not reveal any *TT10* transcript accumulation in other tissue than siliques and seeds. Previously observed expression in other organs (Pourcel et al., 2005; Liang et al., 2006a) may be due to some 'stress' or different plant culture conditions. McCaig et al. (2005) could detect *TT10* mRNA in leaves and seedling, using 40 cycle PCR amplification. However, this system set for high sensitivity of transcript detection, could result in artifacts. In comparison, the 35-cycle PCR used here to study the expression profiles of the laccase family genes in seeds, revealed that *TT10* is the most expressed laccase in developing seeds (see section 2.5.1), whereas a 40-cycle PCR resulted in detectable signals for all the samples (data not shown). According to Liang et al. (2006a), the weak expression of *AtLAC15* in roots observed by Cai and Wu (unpublished data), was confirmed by *Arabidopsis* microarray data ([www.arabidopsis.org](http://www.arabidopsis.org), no precise data set was cited). Present analysis of the transcriptomic data (see 2.1.1.2), found *TT10* expression only in developing siliques, seeds and pollen. However, this could be due to very strict cut-off values. Curiously, *TT10* signature tag in the massively parallel signature sequencing (MPSS, <http://mpss.udel.edu/at/>, Meyers et al., 2004), was found in only one out of two 21 day root samples, and moreover not in 24-48 hours post-fertilization silique sample (McCaig et al., 2005). On the basis of these data, the presence of *TT10* in roots and leaves is questionable.

Reporter gene activity driven by the *TT10* promoter gave further support for the expression data. GUS activity, as previously reported by Pourcel et al. (2005), was mainly detected in the seeds, replum and transmitting tissue. Faint staining was also observed in the abscission zone of the petals and sepals, but not in the petals and sepals themselves (see Fig. 2.7). Similar GUS pattern was also reported by Liang et al. (2006a), regarding silique staining. However GUS activity was not reported in young siliques as well as precise information about the staining in the seed coat was missing.

To investigate potential stress-induced regulation of *TT10* gene expression, several experiments have been undertaken (see section 2.4). First analysis of publicly available transcriptomic data, as well as comments reported in the discussion by Pourcel et al. (2005), suggested that some stress conditions could modify *TT10* expression. Nevertheless, when strict cut-off and standard deviation filtering was used, *TT10* signal was below the detection level (discussed in Results 2.1.1.2 and 2.4). Moreover, in similar conditions as the ones described for the microarrays, transgenic seedlings carrying promoter fused to the reporter gene did not result in any detectable GUS staining. Surprisingly, in some cases blue staining was observed in the root tip and vitrified plantlets grown *in vitro*, but not in a reproducible manner. It would be possible that *TT10* could be expressed in the root tip, but in our experiments, proper combination of the developmental stage/treatment were not clearly identified to name the stress which could be involved. A possible induction could be due to oxidative or osmotic stress, because most of the experiments were carried out in the liquid medium and seem to be density of sowing dependent. The experiments carried out on the filter paper also resulted in blue staining of the root tip, which could be due to mechanical stress in this case.

Until now, *TT10* has been shown to be involved in flavonoid metabolism in seeds (Pourcel et al., 2005). Therefore it could also be involved in flavonol metabolism in roots, or other defense or detoxification-related process. Recently Liang et al. (2006b) have shown that *ZmLAC1* expression was induced in roots under salt stress conditions in seedlings grown in hydroponics. Moreover, several *Arabidopsis* laccase gene

transcript increased under PEG solution, which caused inhibition of root growth (Liang et al., 2006b). In our conditions we were able to detect induction of GUS activity driven by *TT10* promoter neither in salt stress nor in seedlings grown with mannitol. Both PEG and mannitol treatments are commonly used to mimic osmotic stress conditions (Bouchabke-Coussa et al., 2008). It is possible that GUS activity is very weak and/or that when whole seedlings were floating in the culture solution, detectable gene induction due to stress conditions does not occur. Moreover stress response could be affected by sucrose which was always present in the growth media. It is possible that in *Arabidopsis*, other laccases but not *TT10* are induced in roots in response to salt stress and/or osmotic stress.

The function of laccases is associated with defense against pathogen attack and wound healing (Mayer and Staples, 2002). However we were not able to detect GUS activity in wounded organs (leaves, stems). Previously, strong GUS activity was observed in aborted seeds (Pourcel et al., 2005). Trying to find a link between this phenomenon with stress response, we have emasculated *Arabidopsis* flowers which developed in seedless siliques in which strong GUS staining was observed in degenerated ovules (data not shown). It seems that activation of *TT10* expression does not require either pollination or fertilization, which is surprising, because unfertilized ovules do not synthesize any PAs (Debeaujon et al., 2003). Most of our experiments suggest that *TT10* expression is mainly under developmental regulation. Those results are in agreement with some previous studies that had shown that polyphenol oxidase (PPO) genes are mostly expressed in young developing tissues, with expression declining as development continues. Anyway, examples of strictly developmentally regulated as well as stress-induced PPO gene expression are well documented (Goldman et al., 1998; Constabel et al., 2000 and references therein). We could still consider that level of *TT10* expression during seed and silique development could be stress-modulated. Nevertheless, it would be good to accompany studies of GUS induction by measurement of RNA accumulation, because elements required for stress response could be located outside the promoter, e.g. in the first intron (Rombauts et al., 2003; Rose et al., 2008).

It is important to add that some experiments aiming at comparing *TT10* expression between various accessions were realized on silique samples collected from plants grown in the a growth chamber with controlled environmental conditions. In those samples, detected amounts of *TT10* mRNA were comparable between accessions (data not shown). We could consider that the growth conditions in the culture chamber were more stable than those in the greenhouse, where temperature, light intensity and humidity variations could influence *TT10* expression. It would be interesting to compare *TT10* transcript accumulation in plants grown in various light intensities or photoperiods. We could hypothesize that response of different accessions is not the same and that those differences could be a result of plant adaptation to their habitats. Last but not least, it is important to note that all dissected promoter fragments studied in protoplasts resulted in detectable GUS activity. It may be only coincidence, but it could also indicate, that *TT10* promoter contains *cis*-acting regulatory elements which are involved in stress response caused by protoplasting/wounding (Takeda et al., 2002). Interestingly, laccase activity was required for the cell-wall reconstruction in regenerating tobacco protoplasts (de Marco and Roubelakis-Angelakis, 1997).

### 4.1.3 Regulatory components

A promoter is a functionally defined 5' part of the gene required for its transcription. It can be divided in the core part, which is required for assembly of the RNA polymerase II complex at the right position for directing basal level of transcription and the distal parts containing elements regulating spatio-temporal expression. Regulatory elements can also be located downstream, for instance in the first intron of the

gene itself. It is not precisely defined and is species-dependent how far upstream or downstream regulatory elements and enhancers are located. Moreover, the three dimensional organization of DNA and chromatin structure constitutes an important part of gene regulation (reviewed in Rombauts et al., 2003).

The *TT10* promoter studied by Pourcel et al. (2005) corresponds to a 2.0-kb region upstream from the translation initiation site, including 5' UTR. Interestingly, recently published results demonstrated that a 1.5-kb *TT10* promoter fragment without 5' UTR was able to drive GUS expression in a tissue specific manner (Liang et al., 2006a). Moreover Liang et al. (2006a) considered 161bp upstream from the translation initiation site as 5'UTR, not the 34bp which were previously defined by RACE PCR (Pourcel et al., 2005). During this PhD, we wanted to precisely define a 'minimal' promoter containing all the elements sufficient for the spatio-temporal transcription of the gene in the seed coat. We were also interested in identifying transcription factors which could bind specific DNA sequences in the 'minimal' promoter. Promoter dissection and all the other results giving insight into the regulation of *TT10* expression are discussed below.

#### 4.1.3.1 Promoter and *Cis*-acting regulatory elements

Thanks to the 5' promoter dissection, a 194bp (between *pTT10\_3* and *pTT10\_2*) region required for seed coat specific expression has been identified. The results of the promoter sub-dissection could further narrow down that region to 63bp (between *pTT10\_3* and *pTT10\_3C*; see section 2.1.3.5). Moreover, site-directed mutagenesis suggests that the putative *cis*-acting regulatory element recognized as MYB2 binding site located in that region could be functional. This hypothesis is supported by the fact that several flavonoid biosynthetic genes are regulated by MYB genes (Debeaujon et al., 2003; Lepiniec et al., 2006) and another member of the laccase family has been recently demonstrated to be regulated by MYB58 and MYB63 transcription factors (Zhou et al., 2009). Results of site-directed mutagenesis suggested also that a DOF binding site for plant specific zinc finger transcription factors could be required for *TT10* expression. Those are 4 bp motifs, associated with carbon metabolism genes and frequently found in the genome (Yanagisawa, 2000). In the *TT10* promoter there are 5 DOF repeats within a 150 bp region, which seems to be more than average distribution of that motif in the *Arabidopsis* genome. In site-directed mutagenesis experiment, one of the substitutions was made in the DOF motif overlapping with I-box motif (see M4 on Fig. 2.6), which is a conserved sequence in promoters of light regulated genes (Terzaghi and Cashmore, 2003). It would be therefore interesting to further investigate the importance of both the DOF and I-box binding motifs.

Realizing the promoter dissection experiment we have observed that the detectable GUS activity in seeds was abolished with 271bp (*pTT10\_2*) promoter fragment, but faint staining was still detectable in the funiculus. Debeaujon et al. (2003) suggested that micropyle and chalaza zone could be particularly important to protect seed from pathogen invasion and to ensure moderate desiccation of the seed during maturation. Funiculus and seed abscission zone are in direct contact with chalaza, therefore it could be possible that to ensure proper protection of that zone, expression of the *TT10* could be regulated differently in this area. The role of *TT10* in this region as well as in the replum would rather be involvement in lignification of the dehiscence zones of the seed and silique. Moreover, this could be achieved in concert with other laccases and contribute to the pod shatter and seed dispersal (Liljegren et al., 2000).

It seems that element(s) required for high level of gene expression are localized upstream from the region required for tissue specificity. Existence of enhancer element(s), stabilizing the complex involved in the spatio-temporal regulation of gene expression would not be surprising (Rombauts et al., 2003). An alternative explanation would be that *TT10* expression is regulated by two different mechanisms in the endothelium and outer integument layer 1. Analysis of GUS activity was mainly done by observation of

the whole mount samples, therefore absence of GUS activity in endothelium would be masked by strong staining in the outer integument. Preliminary results of the GUS activity with the seed sections did not clearly answer this question (data not shown).

Liang et al. (2006a) showed GUS activity driven by the *TT10* promoter without the 5'UTR. It suggests that the 5'UTR region of the promoter is not required for the GUS activity and that weak GUS activity observed for the constructs with *pBI101-GTW* vector, are rather vector- related. Similar drawback, with short promoter sequences was also observed for *TT2* and *TT16* (Christian Dubos, personal communication).

One more interesting aspect of the *TT10* promoter could be discussed on the basis of the transient activity of all the promoter fragments studied in protoplasts. Even the shortest 171bp promoter fragment contained all the elements required for the GUS expression in this system. A possible explanation would be that protoplasting or the culture conditions are creating stressful conditions sufficient for ectopic GUS activity (de Marco and Roubelakis-Angelakis, 1997; Takeda et al., 2002). On the other hand, cells grown in the cell suspension are not differentiated and we could imagine that they miss negative regulator such as *MYB4* or *MYBL2* (Jin et al., 2000; Dubos et al., 2008) which are normally present in the plant.

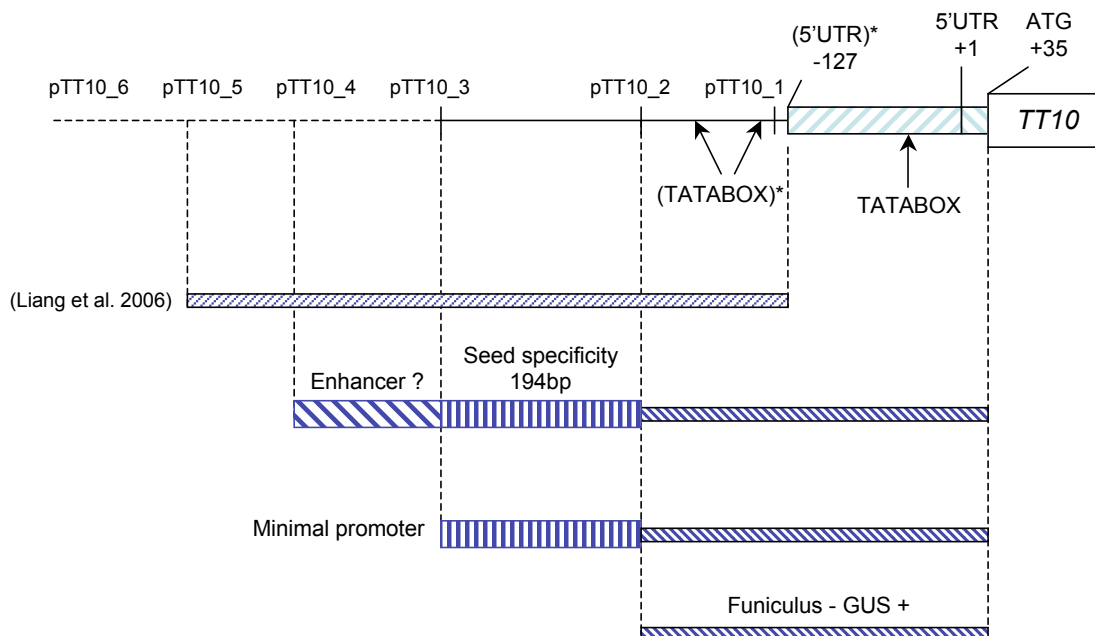


Figure 4.1: Organization of the regulatory elements in the *TT10* Promoter.

Figure is summarizing promoter regions found to be required for seed, funiculus and strong expression of the *TT10* gene. Asterixes (\*) are highlighting the alternative 5'UTR and TATABOXES considered by Liang et al. (2006a).

It would be interesting to determine if the putative MYB2 binding site is functional and if an unidentified complex MYB/bHLH could regulate expression of *TT10*. One more indication for possible involvement of such a complex is *in silico* identified putative MYC2 binding motif around 150bp from MYB2 site. Abe et al. (2003) had found the combination of that motifs in many ABA responsive genes. Of interest is also the fact that the expression of several pathogen-related genes is upregulated in plants overexpressing MYC2/MYB2 (Abe et al., 2003). Taking into account that *TT10* is a laccase involved in flavonoid biosynthesis, it is possible that one of the MYB family transcription factors is involved in its regulation. The reasons are multiple. Analysis of the phylogenetic relationship of the MYB TF involved in flavonoid biosynthesis shows that they form cluster (Stracke et al., 2001). Moreover, the genes regulating the early steps of the phenylpropanoid



pathway (*MYB75/PAP1*, *MYB90/PAP2*) or flavonols (*MYB11*, *MYB12*, *MYB111*) are even more closely related and are forming clades in this cluster. We could expect that one of the MYBs in this cluster could be involved in regulation of *TT10* expression. Interestingly, by co-expression analysis we have found *MYB5* TF, which is also in this clade. Moreover, two other MYB TF, *MYB58* and *MYB63*, have been recently shown to regulate expression of *LAC4* (Zhou et al., 2009). Those MYBs cluster together, but not with the above-mentioned TF involved in the regulation of phenylpropanoid genes (Stracke et al., 2001). Based on all those data, we can make hypothesis that *TT10* could be regulated by one of the MYB transcription factors. However, if it would be one more related to the MYBs involved in regulation of the flavonoid biosynthetic genes or one(s) regulating *LAC4* is difficult to predict.

One more important point needs to be added to the study of the involvement of the MYB transcription factors in the regulation of the *TT10* expression. Two mutants chosen from the co-expression studies, *myb76* and *myb99* showed decrease in *TT10* transcript accumulation. It is possible that this slight decrease is not sufficient to observe some impact on flavonoids, because analysis of the flavonoid composition of mature seeds from those plants did not reveal any differences compared with wild-type. It might be that regulation of *TT10* expression is ensured by several transcription factors with redundant function and therefore studies should consider multiple mutants in these regulatory genes. It would be interesting to check if *TT10* expression could be further down-regulated in double/triple mutants of the MYBs being in the same clades with *myb76* and *myb99* (Stracke et al., 2001). It could be also interesting to see, if the GUS pattern driven by the *TT10* promoter is modified in those mutants. Regarding the putative MYB binding sites found by *in silico* analysis, it could be also interesting to check if they could bind to the *TT10* promoter.

#### 4.1.3.2 Candidate transcription factors for developmental regulation

*In silico* analysis and promoter dissection studies were aiming at localizing regions in the upstream sequence of the *TT10* gene which could contain *cis*-acting regulatory elements and together with the TATA box would constitute the promoter of the gene. In parallel, we have also looked for transcription factors which could be involved in the regulation of *TT10* expression. Some results were already obtained by Pourcel et al. (2005), showing that *TT10* is not regulated by the complex regulating *BAN* expression (Debeaujon et al., 2003; Baudry et al., 2004, 2006).

Results obtained during this work, are consistent with the preliminary data indicating that *TT1*, *TT16* and *TTG2* could regulate *TT10* (Pourcel, 2006). In this case, the introduction of the promoter:GUS cassette into the regulatory mutant backgrounds was done by crossing, not by transformation (Pourcel, 2006), therefore any problem due to the influence of the insert position on promoter activity has been minimized. Moreover, the observation of GUS activity on seed sections revealed delayed *TT10* promoter activity in *tt1* and *tt16* mutants. This could be explained by defects in endothelium development in the seeds of those plants (Sagasser et al., 2002; Nesi et al., 2002). A similar pattern of promoter activity was observed in the endothelium of the *ttg2* mutant, which seems to be normally developed. However, the cells could be not properly differentiated (Johnson et al., 2002). The *TTG2* gene has been also shown to be involved in mucilage accumulation and its promoter activity was reported to be active in the whole seed coat (Johnson et al., 2002). We have shown that it is only active in endothelium and mucilage layers. We could consider that *TT10* expression is regulated by those genes in the endothelium. However, the observed modification of the GUS pattern could be an indirect effect. Promoter activity in the outer integument 1 layer could be regulated differentially from the one in endothelium, because the GUS activity in that layer seems to be not affected in *tt1*, *tt16* and *ttg2* mutants. It is possible that the regulation of gene expression in the seed

coat is very complex, e.g. regulation of the BAN gene in the endothelium seems to be different from that in micropyle and chalaza zones (Debeaujon et al., 2003). Taking in account that *TT10* is also involved in flavonol metabolism (Pourcel et al., 2005), we could expect that its expression in the outer integument 1, is similar to the one of genes of that branch of the flavonoid pathway. Expression of *TT10* was only slightly reduced in *myb11*, *myb12*, *myb111* which are mutants in genes known to be required for flavonol biosynthesis in vegetative parts (Stracke et al., 2007) and seeds (Routaboul, unpublished results).

Current model for the regulation of PA biosynthesis in the seed coat involves *TT2*, *TT8* and *TTG1* which are required for the *BAN* expression for example. Moreover, *TTG2* seems to act downstream from *TTG1*, whereas *TT16* upstream from *TT2*, and *TT1* seems to be necessary for *BAN* expression in a few cells at the endothelium base. Also fertilization is required for PA accumulation in general (Debeaujon et al., 2003). Taking in account that *TT10* expression is not modified in *tt2*, *tt8*, and *ttg1* mutants and only slightly modified in *tt1*, *tt16* and *ttg2*, we should rather expect distinct regulators.

Interestingly, *Arabidopsis* lacks polyphenol oxidases of the catechol oxidase (CO) type (Pourcel et al., 2005). Therefore we may hypothesize that laccase such as *TT10* has been recruited evolutionary to compensate for the lack of CO. This is relevant with the fact that both *TT10* and COs have flavonoids as substrates. *TT10* is not directly involved in the biosynthesis of flavonoids, but only in the oxidative polymerization of the final products (Pourcel et al., 2005), therefore regulation of its expression could evolve as a special function of one of the laccase family genes in lack of the catechol oxidases in *Arabidopsis* (Constabel et al., 2000; Thipyapong et al., 2004; Pourcel et al., 2005). It would be interesting to find out if the *TT10* function could be replaced by other laccases as well as a catechol oxidases from other species.

#### 4.1.3.3 Candidate transcription factors for stress regulation

Previously discussed potential stress regulation of *TT10* led us to study of transcript accumulation in mutants in genes involved in pathogen and stress response (see section 2.2.5). We did not observe any significant differences in the amount of *TT10* mRNA in any of the mutants considered, except *npr1-1*. Surprisingly, transcript accumulation in *npr1-2* allele was comparable to the wild type. However both alleles studied come from the EMS mutagenesis and in both of them protein is detected (Cao et al., 1997; Somssich, 2003). Moreover, it has been shown that *npr1-1* mutant was non-functional with respect to systemic acquired resistance, but in an NPR1-GFP overexpressor line carrying *npr1-1* mutation, the *npr1-1* protein was part of the detectable oligomeric complex. The *npr1-2* mutation resulted in a cysteine to tyrosine conversion and is thought to result in a non-functional protein (summarized in Somssich, 2003). NPR1 function is well studied and it is thought to play an important role in the cross-talk between salicylic and jasmonic acid-dependent defense pathways (Somssich, 2003; Dong, 2004; Pieterse and Loon, 2004). It would be possible that the mutation in the *npr1-1* line resulted in inactivation of the protein domain required for recruitment of factors involved in *TT10* expression. We have found that *TT10* expression was also decreased in *npr1-5* allele. However recently we realized that this mutant contains a T-DNA insertion carrying kanamycin resistance gene and the *PR-1a:tms2* reporter gene and is in Nossen background. Taking into account the natural variation occurring for *TT10* gene expression, we were not able to conclude on that result. It would be necessary to check other alleles of the *npr1* mutant. It might be also interesting to include mutants in TGA proteins which belong to bZIP family of transcription factors. These TFs has been shown to interact with NPR1 (reviewed in Pieterse and Loon (2004) and Dong, 2004).

We have found several putative W-boxes in the *TT10* promoter sequence, which are binding sites for WRKY TFs. This family of TF is involved in regulation of plant defense processes (Pandey and Somssich, 2009). Moreover, elicitor-induced expression of transcription factors and metabolic reprogramming of

secondary metabolism have been demonstrated in *Medicago truncatula* cell suspension, in which WRKY transcription factors expression were one of the most affected by yeast elicitor and methyl jasmonate treatment. Several laccases have been shown to be induced by that treatment (Naoumkina et al., 2008). Interestingly a laccase has recently been proposed as a candidate gene to explain a QTL for resistance to nematode in soybean (Iqbal et al., 2008). Other recent report demonstrated that knocking down expression of *WRKY23* resulted in lower infection of the cyst nematode *Heterodera schachtii* (Grunewald et al., 2008). Moreover plants overexpressing *WRKY23* were accumulating more quercetin-3-O-rhamnoside (Godelieve Gheysen, personal communication). Following the indications that laccase could be involved in nematode resistance, and that *WRKY23* was shown to mediate infection rate of the cyst nematode we were interested in possible relation with *TT10*. Analysis of the transcript accumulation revealed, that *TT10* expression was not modified in three alleles of the *wrky23* studied. We could assume that *WRKY23* is not involved in regulation of *TT10* expression. However all the lines studied had T-DNA insertion in the promoter region of the gene and they are not knockouts (Grunewald et al., 2008). It is possible that even low expression was sufficient to sustain *TT10* transcription at the wild-type level.

Studies of *TT10* expression in other mutants affected in plant defense genes did not reveal any differences in the transcript accumulation. However, all our experiments concerned not stressed plant. We could expect that in those conditions plants would have some basic level of pathogen related genes, which expression is only induced in response to environmental challenges. Induction of the defense genes would be part of plant plasticity and adaptation strategy, therefore it would be interesting to test how expression of *TT10* and flavonoid biosynthetic genes is maintained in those mutants under various stresses (e.g. in response to pathogen attack, dehydration, different light intensities).

Last, during our quest to find regulator of *TT10* expression, we have observed that the *jar1* mutant which is affected in jasmonic acid perception (Lorenzo et al., 2004), was accumulating less quercetin-3-O-rhamnoside and biflavonols, which could suggest some involvement in the biosynthesis of those flavonols. Interestingly, *jar1* is located in the interval found in the QTL mapping for the flavonoid composition in mature seeds (Jean-Marc Routaboul, unpublished results).

#### 4.1.3.4 What have we learned from natural variation?

Study of natural variation is a novel approach to discover new functions especially those involved in adaptation to a specific habitat. Phenotypic differences are often due to allelic variations at several loci and the contribution of each locus to the phenotype can be quite small (Weigel and Nordborg, 2005). It has been previously shown that single nucleotide polymorphism between gene alleles could result in differential gene expression or protein activity (de Meaux et al., 2005; Bentsink et al., 2006; Loudet et al., 2007). Here we wanted to find out if level of *TT10* expression, which varies between accession, could be attributed to polymorphism in *cis*-acting regulatory elements in the promoter. We have found several SNPs in *Arabidopsis* accessions studied, but we were not able to link them with the level of transcript accumulation, nor flavonoid composition of matures seeds.

Comparison of long promoter sequences between Col-0 (2.2-kb) and *Ler* (1.5-kb) revealed several indels and more than 700bp deletion in *Ler* accession (Ishihara, 2007). Interestingly we have found that ATCOPIA42 transposon identified in *TT10* promoter by FLAGdb++ server (Samson et al., 2004), may correspond to this large deletion (see Appendix, Fig. 5.7). It seems that, in *Ler*, the transposon was excised or inserted in Col-0. We were not able to identify the remaining LTR sequences. It is important to note that the transposon is located in the region which was found to be required for high *GUS* expression. Interestingly the transposon region in the *TT10* promoter is methylated (see section 4.1.3.7, Chromatine structure), suggesting that it may be inactive. Indeed in plants, silent transposable elements are methylated

(Rabinowicz et al., 2003). It is difficult to predict which effect transposons can have on gene expression. Both situations were reported, one with which transposon promoters can activate neighboring gene expression and act as an enhancer, and the other with which they can also act as insulators, and result in a decrease or complete block of gene expression (Dorsett, 1999; Valenzuela and Kamakaka, 2006). It would be interesting to check what is the effect of the transposon present in *TT10* promoter on its expression.

The study of the *TT10* Col-0 promoter in other accessions backgrounds (see section 2.3.5) revealed that tissue specific activity was not altered. Moreover, fragments studied in the promoter dissection experiment (see section 2.1.3.1) pointed at the same region being required for the seed coat activity of the promoter in Col-0 and Ws-4. Interestingly also the upstream element required for high GUS activity in Col-0 background is also required for strong GUS induction in Ws-4. Those results suggest that the differences in the *trans*-regulation (i.e. due to transcription factor) are not likely because *cis*-acting regulatory elements in the promoter sequence from Col-0 accessions are recognized by transcription factors in other accessions and result in the same tissue specific promoter activity.

#### 4.1.3.5 Transcription factor library screening

In case of yeast one- and two- hybrid screening auto-activation is a technical problem encountered with some promoters and proteins. Very often it can be overcome by adding 3-AT, the competitive inhibitor of the HIS3 product. However in small scale experiments, the concentration could be raised up to 25-50mM (Baudry et al., 2004; Xiao, 2006). For screening of the REGIA transcription factor library it was not advised to use more than 5mM 3-AT because of the the sensitivity of the yeast and lack of detection of weak interactions. We have tried to use smaller fragments of the *TT10* promoter, but this strategy did not eliminate the auto-activation. It would be good to consider even smaller promoter fragments e.g. 63bp (region between *pTT10\_3* and *pTT10\_3C*) promoter fragment containing the MYB2 binding site or even isolated motifs identified in the site-directed mutagenesis experiment. Another solution could be use of another MAT $\alpha$  yeast strain which could enable to study longer promoter fragments (Xiao, 2006).

Yeast one-hybrid screen, could give an indication of the class/family of the transcription factors which could be involved in regulation of the target gene expression. However later on their functionality should be confirmed *in planta*. In some cases, interaction found in yeast, would not be possible *in planta*, because of differential spatio-temporal presence of the transcription factor and expression of the gene considered in the experiment.

#### 4.1.3.6 Co-expressed genes

We have analyzed *TT10* expression and flavonoid composition of mature seeds of several mutants chosen on the basis of co-expression analysis. Apart from *myb76* and *myb99*, all the other mutants analyzed accumulated wild-type amount of *TT10* mRNA and flavonoids. However other group of genes co-expressed with *TT10* during plant development have been proposed and should be more relevant, because they are co-expressed during seed development (see section 2.2.3.3). Interestingly, among those new candidates was MYB5, which may be the most promising candidate to study. The relevance of the other genes should be considered based on the literature studies.

It is important to keep in mind that co-expressed genes are not necessarily involved in regulation of *TT10* expression. It is possible that *TT10* and co-expressed genes are regulated by the same factor. To test this hypothesis it would be required to compare their promoter sequences for presence of the common *cis*-acting regulatory elements. On the other hand, the list of the co-expressed genes miss those which are expressed in several tissues.

#### 4.1.3.7 Epigenetics?

Recently small RNAs (smRNAs) have been recognized as one of the players for regulation of gene expression in plants. Depending on the formation process, smRNAs can be divided in several classes as recently reviewed by Voinnet (2009). smRNAs regulate gene expression in a process called post-transcriptional gene silencing (PTGS), RNA silencing or RNA interference (RNAi). Many plant responses related to adaptation to changing environment have been shown to be regulated by smRNAs (Voinnet, 2008). Interestingly, several laccases have been shown to be predicted targets of *miR397* (LAC2, LAC4 and LAC17), *miR408* (LAC3, LAC12 and LAC13) and *miR857* (LAC7) (Jones-Rhoades et al., 2006), which are induced in seedling grown in hydroponics conditions with copper-depleted medium (Abdel-Ghany and Pilon, 2008). Those *miRNAs* are thought to be involved in the regulation of expression of non essential copper proteins (Abdel-Ghany and Pilon, 2008). There are no predicted *miRs* for *LAC15*. Surprisingly Abdel-Ghany and Pilon (2008) reported that *LAC15* was ubiquitously expressed and its transcript level was moderately affected by copper availability in the growth medium. Moreover, it was observed that *LAC15* expression was 5-100-fold higher (depending on the plant organ tested) in plants grown with high copper concentration compared to the no copper condition (Abdel-Ghany and Pilon, 2008). It is possible, that in the plants grown in the hydroponic conditions with various concentration of copper available, *TT10* expression could be induced. However, the only organs in which *TT10* expression was assessed were leaf, stem, root and flower. In all cases expression was very low in range of 100-10000-fold lower than the reference gene actin (Abdel-Ghany and Pilon, 2008). We could question the significance of those results because of very low level of *TT10* expression in those experiments. Moreover no miRNA was predicted to target *TT10*. On the other hand it could explain that in our conditions, plants grown in soil had limited availability of copper and therefore we were unable to detect *TT10* transcript in tissues other than siliques.

Promoters are often regarded as linear stretches of DNA, whereas regulation of gene expression is enabled by interaction of all the components in the three dimensional structure. DNA is packed into chromatin and often access to the DNA template for transcription is limited. Moreover DNA methylation has been shown to repress transcription initiation (Rombauts et al., 2003). We checked publicly available data (<http://epigenomics.mcdb.ucla.edu/>, UCSC Genome Bioinformatics) for the DNA and histone methylation of the *TT10* gene and promoter. These data revealed two methylated regions in the promoter and histone H3K27m3 methylation at the beginning of the fifth exon and on the border between the fifth exon and fifth intron. The experiments were carried out with whole seedlings (10-14 days) or aerial part of 5-week-old plants (Zhang et al., 2006, 2007). For the map of methylated regions see Appendix, Fig. 5.6.

Interestingly, recent work has shown that a high level of  $\approx 24$  nt small interfering RNA is present in *Ler* accession compared to Col-0. These siRNAs were directing DNA methylation and heterochromatinization at the *hAT* transposon element adjacent to the promoter of FLOWERING LOCUS C gene in *Ler*, whereas the same element was not affected in Col-0 (Zhai, 2008). Methylation could be one of the mechanisms to switch off *TT10* expression in plant organs other than siliques and seeds or affect the level of its expression in various accessions.

## 4.2 *TT10* and flavonoid metabolism in seeds

### 4.2.1 *TT10* and flavonoid metabolism

Pourcel et al. (2005) have demonstrated that *TT10* is involved in oxidative polymerization of flavonols and tannins. During this work we analyzed the correlation between *TT10* expression levels and the amount of soluble tannins as well as accumulation of quercetin-3-O-rhamnoside (Q-3-O-R) and biflavonols. For that reason *Ler* accession has been transformed with Col-0 genomic sequence to increase the amount of *TT10*.

We have achieved not only to bring the *TT10* transcript level to the one observed in Col-0, but we also observed until 40-fold higher expression in one line. Normally the soluble fraction of PAs is regarded as not or slightly oxidized and not cell-wall nor protein-bound, whereas after oxidation they are regarded as insoluble/non extractable. By overexpressing *TT10*, we were expecting to increase oxidation and to extract less soluble PAs from the 'complemented' lines. Visual inspection of the seeds revealed that they were darker than the *Ler* wild-type control, but surprisingly none of the lines accumulated less soluble PAs. Interestingly, four transgenic lines accumulated less Q-3-O-R, but expected increase in biflavonol amount was not observed. The results obtained are puzzling and difficult to interpret. Moreover increase of the transcript accumulation may not always be correlated with the amount of active protein, therefore further comments are only hypothetical.

The most important fact was that in none of the lines studied a decrease of soluble PA fraction was observed. We could hypothesize that the wild-type *TT10* expression is sufficient to metabolize tannins available in *Ler* and for some unknown reason oxidation can not go beyond that limit. Indeed, *TT10* has been shown to trigger oxidative polymerization of tannins and flavonols, but subsequent reactions on which color and probably the extractability of the oxidized phenols depends are still not well studied (Nicolas et al., 1993; Pourcel et al., 2005). It is highly probable that not only oxidation but also subsequent steps leading to formation of the insoluble fractions of PAs are driven enzymatically. However, this process could also be spontaneous due to the highly chemically reactive nature of intermediate quinones.

Because of the natural variation observed for *TT10* expression we were also interested in the effect of mutation in *TT10* gene in other *Arabidopsis* accessions. Those studies clearly confirm the involvement of *TT10* in oxidation of tannins and flavonols, because all the alleles of *tt10* mutant accumulated more soluble tannins and quercetin-3-O-rhamnoside (Q-3-O-R) but less biflavonols than respective wild-type plants.

Interestingly *Cvi*, normally accumulates much more tannins than other accessions, but almost no Q-3-O-R and biflavonols. As expected the *tt10-7* mutant allele (in *Cvi* background) accumulates more soluble PAs and Q-3-O-R like other *tt10* mutant alleles. Regarding low level of biflavonols in the *Cvi*, one would expect that Q-3-O-R are not synthesized. However, detection of Q-3-O-R in the *tt10-7* mutant suggests that other enzyme or other unknown feature would be required to act in concert with *TT10* to synthesize biflavonols. It is also possible that in *Cvi*, biflavonols are further metabolized what would explain that they are detectable in very low amount in this accession.

It is clear that *TT10* is involved in tannin and flavonol oxidation, however it seems that some differences between accessions could have arisen in respect to the amount and classes of the flavonoids accumulated (Pourcel et al., 2005; Routaboul et al., 2006 Jean-Marc Routaboul, unpublished results). There are still many aspects of flavonoid metabolism to be discovered to fully understand their accumulation and oxidation.

### 4.2.2 Other functions?

Liang et al. (2006a) reported the involvement of *TT10* in lignin metabolism in seeds and root elongation. It is possible that *TT10* is involved in the lignification not only in the seeds, but also in the replum where its promoter activity was observed. It is also possible that in certain conditions *TT10* could be involved in

flavonol metabolism in roots. However, until now those functions have not been demonstrated.

### 4.3 Regulation in the multi-gene family

It seems that laccases have acquired their specific functions very early during evolution and they are involved in resistance to various stresses (Mayer and Staples, 2002; Thipyapong et al., 2004). One of the functions related to plant defense would be oxidation of flavonoids (Pourcel et al., 2005). Interestingly in *Arabidopsis* genome no typical catechol oxidases were found (Pourcel et al., 2005, 2007). Laccases have a broad spectrum of substrates and they can oxidize ortho- and para-diphenols. Therefore, they could be involved in various aspects of phenylpropanoid metabolism (Mayer and Staples, 2002; Pourcel et al., 2007). Pourcel et al. (2005) have proposed that in absence of catechol oxidase, laccases may have evolved to fulfill this lack (i.e. catalyzing o-diphenol oxidation).

Interestingly, in this study we showed that *TT10* is the only member of the family expressed in seeds. This fact answers the question why a single mutant in this laccase gene results in a clear phenotype. It seems that *TT10* is a laccase specialized in flavonoid oxidation by co-localizing with the substrates in plant tissues.

Strong GUS staining driven by *TT10* promoter in aborted seeds was previously associated with senescence/cell death (Pourcel et al., 2005). Interestingly, during the characterization of expression profiles and GUS activity driven by promoters of other laccases, we have observed that the *LAC5* promoter resulted in strong GUS activity in degenerated ovules and aborted seeds. It seems that in such a case of developmental defect, *TT10* together with *LAC5* could be involved in protection of a fragile zone caused by cell death. Analysis of the transcriptomic data suggests that *LAC5* is highly induced during senescence (data not shown) and profile of GUS activity in aborted seeds observed for *TT10* and *LAC5* highly resembles that of *SENESCENCE ASSOCIATED GENE 12* promoter (data not shown).

It is possible, that in certain conditions several laccases act in concert to ensure sufficient protection of plant. Other example of overlapping expression of several laccases would be silique replum in which in addition to *TT10*, GUS activity driven by the promoters of *LAC5*, *LAC12* (this study), *LAC4* (data not shown) and *LAC2* (Serge Berthet, unpublished results) were also detected. This function of the laccase in various organs could overlap and therefore studies of single mutants would not result in identification of significant phenotypes, with *TT10* being an exception.

### 4.4 Final remarks

We should acknowledge the involvement of the *TT10* in plant protection against various stresses/environmental challenges. It would rather be a preformed defense by preparation of the oxidized flavonoid shield - seed coat - to protect the embryo, than direct response to pathogen attack. Regulation of the *TT10* expression and importance of flavonoid oxidation still needs to be further investigated. Moreover studies of other members of laccase family will bring interesting information about evolutionary specialization of gene function by tissue specific gene expression.

# Chapter 5

## Appendices

### 5.1 *In silico* analysis of the *TT10* promoter

*TT10* (Col-0) promoter sequence used in this study:

```
cacactgattttgcttggaatgcgccaagcaaacatTTTTcacttattccttattaataataaattttttacaagaattct
attagctcggcaattacttgacatattctacccgaatagtcaacctctactaaaaacaatcaaagcctaattccttcat
ccgtttcttatcacacacctgggaagagaaagagaattacaacaccactTTTTaagcttttctaggcttcgaaagcta
tacacgtatggcaatattctagagtgtttcggcggctcaatgcctatttctagaaaattatactattctccaagaatcg
atgagcgggatctttccttgaacgtcaaacctatctcttttctgttgggattttccttgatttccaagcctatcactttc
cttgtttgagatccttcttgatatttaagtctatctctttccttatttgttgagaagaatattaccaaatcttttcaac
cgactttcgtccttgatcaccacgctaagtcaagcgtcttggcacgtcacatcaaccactccagctccatcatgtc
agcttacgctccatgtcagcaattcagacgtttgatgtagatttgatgcagcttctctgatatccaagagtttctaag
tccaccatctcgttgcctgatatccatcaactgttattctgatatctaacagctcctgatatccatcagctcgttgtt
gtcagaatctgcatcttcacaccacaccataatccattgctagtgaatttttgagtaggcttgatcattgtccttctc
ccttgggtccttgagaggattcatcaccatcacataacatataactcccctgtcaaacacagcaaatgattgcttact
aagacactgttctgcacagtctcatttgcctgaacctgaatgaggaggtagctaggataattaccaaggctggaga
tcactatttattttttatcttccatttctgctataatctatagaagatgacatgaagaggttgacaaacaaaaaaaa
aacaatatcatgaagagaagccattgggctaaaatcaaaaaaaaaaaaaaaaaatttgcaccaaacaaatcaaatgt
ttataggggaacatttctcgggttattaaagttttcaaattagaaaaagaaaaaaaaagaagattccttgaagttca
ccagtttaagtttctcgtaatgcattagaaatcgacatctcaccagtcaaaaatcttgacaaatattcgcaaattagt
tttaaaaacattcttccagtaaaatgtaaacattaaaaaagaacatctctttcgtttcttttattgttagcgaacat
ctctttcgtaaaaaagaaaattgttttcaggaatcttaaaattctatagcttaatacgtacgttacgtacatgtagctt
tggcgtggaatgtagctccccgacatctaattagatgataactcgttttcatttaattcagtgaacaatgtaactatg
cacgtaagaactaaattaacacaaaaacgtaacgagtttcggaatatttctgggtgttggtttatataaaataaagtaa
tgcttggcctaacaaaactcataagtgttttattttatcttagtaagactttgaactagttgattgtagtcacgtct
cttactatataatgatataacatgtgaaaatataatctagagttagataaaatcgcactttactgttttcgttaactgaaa
gtttcgtttgaccaagaaatgtagttaaatatttggggaatgtgaactcattagtcaaaacggaccaagaagatgatgc
ttaccgaaaacatattttgtaaaatacaaaattcttatatccgaatgtaactaatttgagaagacgggtctctatagta
tataaaatataatgatataatgcctctctattcttggatgaactcagtcagggttaataatttactaaaactcttccaaa atg
```

#### 5.1.1 Putative *cis*-acting regulatory DNA elements found in PLACE database using the Signal Scan

The PLACE database (Higo et al., 1999) was queried with 2000bp *TT10* promoter sequence. The putative *cis*-regulatory elements retrieved with the Signal Scan (Prestridge, 1991) program are listed in the Table 5.1. The CARE location is given in the order of occurrence on the promoter sequence (from -1964-kb to +36bp).



Factor/Name	Location (Strand)	Sequence			
			AGCBOXNPGLB	271 (-)	AGCCGCC
CACTFTPPCA1	3 (+)	YACT	GCCCORE	271 (-)	GCCGCC
ARR1AT	6 (+)	NGATT	CAATBOX1	278 (+)	CAAT
-300ELEMENT	35 (-)	TGHAAARK	POLLEN1LELAT52	287 (-)	AGAAA
GT1CONSENSUS	36 (-)	GRWAAW	POLLEN1LELAT52	292 (+)	AGAAA
GT1GMSCAM4	36 (-)	GAAAAA	GT1CONSENSUS	293 (+)	GRWAAW
GTGANTG10	40 (-)	GTGA	TATABOX2	297 (-)	TATAAAT
CACTFTPPCA1	41 (+)	YACT	CACTFTPPCA1	302 (+)	YACT
-10PEHVPSBD	45 (+)	TATTCT	SP8BFIBSP8BIB	302 (+)	TACTATT
TATABOX3	51 (+)	TATTAAT	-10PEHVPSBD	305 (+)	TATTCT
TATABOX3	52 (-)	TATTAAT	ARR1AT	316 (-)	NGATT
POLASIG3	55 (+)	AATAAT	BS1EGCCR	324 (+)	AGCGGG
POLASIG1	58 (+)	AATAAA	NODCON1GM	330 (-)	AAAGAT
BOXIINTPATPB	77 (-)	ATAGAA	OSE1ROOTNODE	330 (-)	AAAGAT
CPBCSPOR	80 (+)	TATTAG	DOFCOREZM	332 (-)	AAAG
CAATBOX1	91 (+)	CAAT	POLLEN1LELAT52	333 (-)	AGAAA
CACTFTPPCA1	95 (+)	YACT	ACGTATERD1	341 (+)	ACGT
WBOXATNPR1	98 (+)	TTGAC	HEXMOTIFTAH3H4	341 (+)	ACGTCA
WRKY71OS	99 (+)	TGAC	ACGTATERD1	341 (-)	ACGT
BIHD1OS	99 (-)	TGCA	TGACGTMAMY	341 (-)	TGACGT
ROOTMOTIFTAPOX1	103 (+)	ATATT	ASF1MOTIFCAMV	342 (-)	TGACG
-10PEHVPSBD	104 (+)	TATTCT	WBOXATNPR1	343 (-)	TTGAC
-10PEHVPSBD	113 (-)	TATTCT	WRKY71OS	343 (-)	TGAC
WBOXHVIS01	118 (-)	TGACT	CIACADIANLELHC	345 (+)	CAANNNNATC
WBOXNTERF3	118 (-)	TGACY	GATABOX	351 (-)	GATA
WBOXATNPR1	119 (-)	TTGAC	NODCON2GM	354 (+)	CTCTT
WRKY71OS	119 (-)	TGAC	OSE2ROOTNODE	354 (+)	CTCTT
CACTFTPPCA1	128 (+)	YACT	DOFCOREZM	356 (-)	AAAG
SEF4MOTIFGM7S	131 (-)	RTTTTTR	POLLEN1LELAT52	358 (-)	AGAAA
ANAERO1CONSENSUS	134 (+)	AAACAAA	RAV1AAT	363 (-)	CAACA
ARR1AT	139 (-)	NGATT	ARR1AT	368 (+)	NGATT
DOFCOREZM	143 (+)	AAAG	ECCRCAH1	369 (+)	GANTTNC
DOFCOREZM	154 (-)	AAAG	GT1CONSENSUS	370 (-)	GRWAAW
MYBCOREATCYCB1	161 (-)	AACGG	GT1CONSENSUS	371 (-)	GRWAAW
POLLEN1LELAT52	164 (-)	AGAAA	ARR1AT	378 (+)	NGATT
IBOXCORENT	166 (-)	GATAAGR	ECCRCAH1	379 (+)	GANTTNC
IBOX	167 (-)	GATAAG	GT1CONSENSUS	380 (-)	GRWAAW
IBOXCORE	168 (-)	GATAA	CIACADIANLELHC	385 (+)	CAANNNNATC
GATABOX	169 (-)	GATA	GATABOX	391 (-)	GATA
GTGANTG10	171 (-)	GTGA	GTGANTG10	393 (-)	GTGA
DPBFCOREDCC3	175 (+)	ACACNNG	CACTFTPPCA1	394 (+)	YACT
EBOXBNNAPA	176 (+)	CANNTG	DOFCOREZM	396 (-)	AAAG
MYCCONSUSAT	176 (+)	CANNTG	GATABOX	420 (+)	GATA
RAV1BAT	176 (+)	CACCTG	ROOTMOTIFTAPOX1	421 (+)	ATATT
EBOXBNNAPA	176 (-)	CANNTG	TATABOXOSPAL	422 (+)	TATTTAA
MYCCONSUSAT	176 (-)	CANNTG	GATABOX	432 (-)	GATA
NODCON2GM	184 (-)	CTCTT	NODCON2GM	435 (+)	CTCTT
OSE2ROOTNODE	184 (-)	CTCTT	OSE2ROOTNODE	435 (+)	CTCTT
POLLEN1LELAT52	187 (+)	AGAAA	DOFCOREZM	437 (-)	AAAG
DOFCOREZM	189 (+)	AAAG	TATABOX5	443 (+)	TTATTT
NODCON2GM	190 (-)	CTCTT	RAV1AAT	448 (-)	CAACA
OSE2ROOTNODE	190 (-)	CTCTT	-10PEHVPSBD	456 (-)	TATTCT
ECCRCAH1	194 (+)	GANTTNC	P1BS	457 (+)	GNATATNC
2SSEEDPROTBANAPA	200 (+)	CAAACAC	P1BS	457 (-)	GNATATNC
CANBNNAPA	200 (+)	CNAACAC	ROOTMOTIFTAPOX1	458 (-)	ATATT
PROXBNNAPA	200 (+)	CAAACACC	ROOTMOTIFTAPOX1	459 (+)	ATATT
CACTFTPPCA1	207 (+)	YACT	GTGANTG10	463 (-)	GTGA
DOFCOREZM	209 (-)	AAAG	ARR1AT	469 (-)	NGATT
DOFCOREZM	218 (-)	AAAG	NODCON1GM	470 (-)	AAAGAT
GT1CONSENSUS	219 (-)	GRWAAW	OSE1ROOTNODE	470 (-)	AAAGAT
GT1GMSCAM4	219 (-)	GAAAAA	DOFCOREZM	472 (-)	AAAG
POLLEN1LELAT52	221 (-)	AGAAA	MYBCORE	477 (-)	CNGTTR
DOFCOREZM	234 (+)	AAAG	CBFHV	479 (+)	RYCGAC
ZDNAFORMINGATCAB1	242 (-)	ATACGTGT	DRE2COREZMRAB17	479 (+)	ACCGAC
ABRELATERD1	243 (-)	ACGTG	DRECRTCOREAT	479 (+)	RCCGAC
ACGTATERD1	244 (+)	ACGT	LTRECOREATCOR15	480 (+)	CCGAC
ACGTATERD1	244 (-)	ACGT	ECCRCAH1	482 (+)	GANTTNC
CAATBOX1	252 (+)	CAAT	DOFCOREZM	484 (-)	AAAG
ROOTMOTIFTAPOX1	253 (-)	ATATT	DOFCOREZM	491 (-)	AAAG
ROOTMOTIFTAPOX1	254 (+)	ATATT	GTGANTG10	497 (-)	GTGA
-10PEHVPSBD	255 (+)	TATTCT	WBOXHVIS01	508 (-)	TGACT
CACTFTPPCA1	263 (-)	YACT	WBOXNTERF3	508 (-)	CTGACY
LTRE1HVBLT49	267 (-)	CCGAAA	WBOXNTERF3	508 (-)	TGACY

WRKY71OS	509 (-) TGAC	CACTFTPPCA1	877 (+) YACT
DOFCOREZM	515 (+) AAAG	CACTFTPPCA1	885 (+) YACT
RHERPATEXPA7	526 (+) KCACGW	SURECOREATSULTR11	899 (-) GAGAC
ABRELATERD1	527 (-) ACGTG	EBOXBNNAPA	903 (+) CANNTG
ACGTATERD1	528 (+) ACGT	MYCCONSENSUSAT	903 (+) CANNTG
HEXMOTIFTAH3H4	528 (+) ACGTCA	EBOXBNNAPA	903 (-) CANNTG
ACGTATERD1	528 (-) ACGT	MYCCONSENSUSAT	903 (-) CANNTG
TGACGTVMMY	528 (-) TGACGT	CCAATBOX1	909 (+) CCAAT
ASF1MOTIFCAMV	529 (-) TGACG	CAATBOX1	910 (+) CAAT
WRKY71OS	530 (-) TGAC	CACTFTPPCA1	931 (-) YACT
GTGANTG10	531 (-) GTGA	MY BST1	939 (+) GGATA
CACTFTPPCA1	541 (+) YACT	SREATMSD	939 (-) TTATCC
BIHD1OS	557 (+) TGTCA	GATABOX	940 (+) GATA
WRKY71OS	558 (-) TGAC	GT1CONSENSUS	940 (+) GRWAAW
CGCGBOXAT	566 (+) VCGCGB	IBOXCORE	940 (+) GATAA
CGCGBOXAT	566 (-) VCGCGB	GT1CONSENSUS	944 (-) GRWAAW
BIHD1OS	575 (+) TGTCA	GTGANTG10	961 (-) GTGA
WRKY71OS	576 (-) TGAC	CACTFTPPCA1	962 (+) YACT
EECCRCAH1	580 (-) GANTTNC	POLASIG1	967 (-) AATAAA
CAATBOX1	581 (+) CAAT	MARTBOX	968 (+) TTWTWTTWTT
ACGTATERD1	589 (+) ACGT	TATABOX5	968 (+) TTATTT
ACGTATERD1	589 (-) ACGT	GT1CONSENSUS	975 (-) GRWAAW
ARR1AT	600 (+) NGATT	IBOXCORE	976 (-) GATAA
GATABOX	619 (+) GATA	GATABOX	977 (-) GATA
TATCCAOSAMY	621 (+) TATCCA	ARR1AT	995 (-) NGATT
GATABOX	621 (-) GATA	BOXIINTPATPB	1000 (+) ATAGAA
MYBST1	621 (-) GGATA	WRKY71OS	1008 (+) TGAC
NODCON2GM	626 (-) CTCTT	BIHD1OS	1008 (-) TGTCA
OSE2ROOTNODULE	626 (-) CTCTT	NODCON2GM	1015 (-) CTCTT
EECCRCAH1	628 (+) GANTTNC	OSE2ROOTNODULE	1015 (-) CTCTT
GATABOX	660 (+) GATA	WBOXATNPR1	1022 (+) TTGAC
AMYBOX2	662 (+) TATCCAT	WRKY71OS	1023 (+) TGAC
TATCCAOSAMY	662 (+) TATCCA	BIHD1OS	1023 (-) TGTCA
TATCCAYMOTIFOSRAMY3D	662 (+) TATCCAY	ANAERO1CONSENSUS	1027 (+) AAACAAA
GATABOX	662 (-) GATA	MARTBOX	1031 (-) TTWTWTTWTT
MYBST1	662 (-) GGATA	MARTBOX	1032 (-) TTWTWTTWTT
GAREAT	672 (-) TAACAAR	MARTBOX	1033 (-) TTWTWTTWTT
-10PEHVPSBD	677 (+) TATTCT	CAATBOX1	1043 (+) CAAT
GATABOX	683 (+) GATA	ROOTMOTIFTAPOX1	1044 (-) ATATT
GATABOX	685 (-) GATA	GATABOX	1046 (-) GATA
MYBCORE	689 (-) CNGTTR	NODCON2GM	1053 (-) CTCTT
GATABOX	700 (+) GATA	OSE2ROOTNODULE	1053 (-) CTCTT
AMYBOX2	702 (+) TATCCAT	CAATBOX1	1063 (-) CAAT
TATCCAOSAMY	702 (+) TATCCA	CCAATBOX1	1063 (-) CCAAT
TATCCAYMOTIFOSRAMY3D	702 (+) TATCCAY	I IATCYTC	1065 (+) TGGGCY
GATABOX	702 (-) GATA	RBCSCONSENSUS	1073 (+) AATCCAA
MYBST1	702 (-) GGATA	ARR1AT	1073 (-) NGATT
RAV1AAT	717 (-) CAACA	MARTBOX	1078 (-) TTWTWTTWTT
BIHD1OS	720 (+) TGTCA	MARTBOX	1079 (-) TTWTWTTWTT
WRKY71OS	721 (-) TGAC	MARTBOX	1080 (-) TTWTWTTWTT
ARR1AT	726 (-) NGATT	MARTBOX	1081 (-) TTWTWTTWTT
GTGANTG10	737 (-) GTGA	MARTBOX	1082 (-) TTWTWTTWTT
ARR1AT	752 (-) NGATT	MARTBOX	1083 (-) TTWTWTTWTT
CAATBOX1	757 (-) CAAT	MARTBOX	1084 (-) TTWTWTTWTT
CACTFTPPCA1	763 (-) YACT	MARTBOX	1085 (-) TTWTWTTWTT
GTGANTG10	764 (+) GTGA	ANAERO1CONSENSUS	1104 (+) AAACAAA
SEF4MOTIFGM7S	769 (+) RTTTTTR	CAATBOX1	1115 (+) CAAT
CACTFTPPCA1	776 (-) YACT	MYBCORE	1141 (+) CNGTTR
CAATBOX1	789 (-) CAAT	TAAAGSTKST1	1148 (+) TAAAG
ARR1AT	817 (+) NGATT	NTBBF1ARROLB	1148 (-) ACTTTA
ANAERO3CONSENSUS	821 (+) TCATCAC	DOFCOREZM	1149 (+) AAAG
GTGANTG10	824 (-) GTGA	GT1CONSENSUS	1153 (-) GRWAAW
GTGANTG10	831 (-) GTGA	POLLEN1LELAT52	1164 (+) AGAAA
MYB1AT	836 (+) WAACCA	GT1CONSENSUS	1165 (+) GRWAAW
CACTFTPPCA1	844 (+) YACT	GT1GMSCAM4	1165 (+) GAAAAA
BIHD1OS	852 (+) TGTCA	DOFCOREZM	1168 (+) AAAG
WBOXATNPR1	853 (-) TTGAC	POLLEN1LELAT52	1170 (+) AGAAA
WRKY71OS	853 (-) TGAC	GT1CONSENSUS	1171 (+) GRWAAW
2SSEEDPROTBANAPA	855 (+) CAAACAC	GT1GMSCAM4	1171 (+) GAAAAA
CANBNNAPA	855 (+) CNAACAC	MARTBOX	1172 (-) TTWTWTTWTT
CAATBOX1	867 (-) CAAT	DOFCOREZM	1179 (+) AAAG
ARR1AT	869 (+) NGATT	ARR1AT	1184 (+) NGATT
CAATBOX1	871 (-) CAAT	DOFCOREZM	1189 (-) AAAG

GTGANTG10	1198 (-)	GTGA	ACGTATERD1	1548 (+)	ACGT
POLLEN1LELAT52	1211 (-)	AGAAA	ACGTATERD1	1548 (-)	ACGT
POLLEN1LELAT52	1227 (+)	AGAAA	ECCRCRH1	1555 (+)	CANTTNC
ARR1AT	1230 (-)	NGATT	LTRE1HVBLT49	1558 (-)	CCGAAA
CBFHV	1231 (+)	RYCGAC	ROOTMOTIFTAPOX1	1564 (-)	ATATT
GTGANTG10	1240 (-)	GTGA	ROOTMOTIFTAPOX1	1565 (+)	ATATT
WBOXPCWRKY1	1245 (-)	TTTGACY	POLLEN1LELAT52	1568 (-)	AGAAA
WBOXHVIS01	1245 (-)	TGACT	DPBFCOREDCDC3	1571 (-)	ACACNNG
WBOXNTERF3	1245 (-)	TGACY	CANBNNAPA	1574 (-)	CNAACAC
WBOXATNPR1	1246 (-)	TTGAC	RAV1AAT	1575 (-)	CAACA
WRKY71OS	1246 (-)	TGAC	REALPHALGLHCB21	1577 (-)	AACCAA
SEF4MOTIFGM7S	1248 (-)	RTTTTTR	MYB1AT	1578 (-)	WAACCA
CCA1ATLHCB1	1249 (+)	AAMAATCT	TATAPVTRNALEU	1581 (+)	TTTATATA
ARR1AT	1252 (-)	NGATT	TATABOX4	1582 (-)	TATATAA
ROOTMOTIFTAPOX1	1263 (-)	ATATT	TATABOX4	1583 (+)	TATATAA
ROOTMOTIFTAPOX1	1264 (+)	ATATT	TATAPVTRNALEU	1583 (-)	TTTATATA
SEF4MOTIFGM7S	1283 (-)	RTTTTTR	TATABOX5	1589 (-)	TTATTT
DOFCOREZM	1293 (-)	AAAG	POLASIG1	1590 (+)	AATAAA
CACTFTPPCA1	1298 (-)	YACT	TAAAGSTKST1	1592 (+)	TAAAG
CAATBOX1	1311 (+)	CAAT	DOFCOREZM	1593 (+)	AAAG
POLASIG2	1312 (+)	AATTAATA	GT1CONSENSUS	1596 (+)	GRWAAW
DOFCOREZM	1319 (+)	AAAG	AMYBOX1	1610 (+)	TAACARA
POLLEN1LELAT52	1321 (+)	AGAAA	GAREAT	1610 (+)	TAACAAR
NODCON2GM	1329 (+)	CTCTT	MYBGAHV	1610 (+)	TAACAAA
OSE2ROOTNODULE	1329 (+)	CTCTT	PREATPRODH	1617 (+)	ACTCAT
DOFCOREZM	1331 (-)	AAAG	CACTFTPPCA1	1624 (-)	YACT
POLLEN1LELAT52	1337 (-)	AGAAA	GTGANTG10	1625 (+)	GTGA
DOFCOREZM	1340 (-)	AAAG	ARR1AT	1626 (+)	NGATT
POLASIG1	1342 (-)	AATAAA	POLASIG1	1630 (-)	AATAAA
CAATBOX1	1345 (-)	CAAT	TATABOX5	1631 (+)	TTATTT
NODCON2GM	1361 (+)	CTCTT	GT1CONSENSUS	1635 (-)	GRWAAW
OSE2ROOTNODULE	1361 (+)	CTCTT	IBOXCORE	1636 (-)	GATAA
DOFCOREZM	1363 (-)	AAAG	GATABOX	1637 (-)	GATA
DOFCOREZM	1374 (+)	AAAG	NODCON1GM	1638 (-)	AAAGAT
POLLEN1LELAT52	1376 (+)	AGAAA	OSE1ROOTNODULE	1638 (-)	AAAGAT
GT1CONSENSUS	1377 (+)	GRWAAW	DOFCOREZM	1640 (-)	AAAG
CAATBOX1	1381 (-)	CAAT	TAAAGSTKST1	1640 (-)	TAAAG
ARR1AT	1393 (-)	NGATT	CACTFTPPCA1	1644 (-)	YACT
BOXIINTPATPB	1403 (-)	ATAGAA	TAAAGSTKST1	1646 (+)	TAAAG
GARE2OSREP1	1408 (-)	TAACGTA	DOFCOREZM	1647 (+)	AAAG
ACGTATERD1	1409 (+)	ACGT	TBOXATGAPB	1651 (+)	ACTTTG
ACGTATERD1	1409 (-)	ACGT	DOFCOREZM	1652 (-)	AAAG
ACGTABOX	1416 (+)	TACGTA	ARR1AT	1664 (+)	NGATT
ACGTABOX	1416 (-)	TACGTA	CAATBOX1	1666 (-)	CAAT
ACGTATERD1	1417 (+)	ACGT	BIHD1OS	1671 (+)	TGTCA
ACGTATERD1	1417 (-)	ACGT	WRKY71OS	1672 (-)	TGAC
CURECORECR	1419 (+)	GTAC	RHERPATEXPA7	1673 (+)	KCACGW
CURECORECR	1419 (-)	GTAC	GTGANTG10	1673 (-)	GTGA
GARE2OSREP1	1420 (-)	TAACGTA	ABRELATERD1	1674 (-)	ACGTG
ACGTATERD1	1421 (+)	ACGT	ACGTATERD1	1675 (+)	ACGT
ACGTATERD1	1421 (-)	ACGT	ACGTATERD1	1675 (-)	ACGT
ACGTABOX	1425 (+)	TACGTA	SURECOREATSULTR11	1677 (-)	GAGAC
ACGTABOX	1425 (-)	TACGTA	NODCON2GM	1679 (+)	CTCTT
ACGTATERD1	1426 (+)	ACGT	OSE2ROOTNODULE	1679 (+)	CTCTT
ACGTATERD1	1426 (-)	ACGT	CACTFTPPCA1	1683 (+)	YACT
CURECORECR	1428 (+)	GTAC	GATABOX	1693 (+)	GATA
CURECORECR	1428 (-)	GTAC	EBOXBNNAPA	1700 (+)	CANNTG
DOFCOREZM	1438 (-)	AAAG	MYCATERD1	1700 (+)	CATGTG
SV40COREENHAN	1445 (+)	GTGGWWHG	MYCCONSUSUSAT	1700 (+)	CANNTG
S1FBOXSORPS1L21	1450 (+)	ATGGTA	EBOXBNNAPA	1700 (-)	CANNTG
LTRECOREATCOR15	1461 (+)	CCGAC	MYCATRD22	1700 (-)	CACATG
GATABOX	1478 (+)	GATA	MYCCONSUSUSAT	1700 (-)	CANNTG
IBOXCORE	1478 (+)	GATAA	GTGANTG10	1703 (+)	GTGA
POLASIG2	1493 (-)	AATTAATA	GT1CONSENSUS	1705 (+)	GRWAAW
CACTFTPPCA1	1501 (-)	YACT	ROOTMOTIFTAPOX1	1708 (-)	ATATT
GTGANTG10	1502 (+)	GTGA	ARR1AT	1713 (-)	NGATT
CAATBOX1	1507 (+)	CAAT	GATABOX	1725 (+)	GATA
RHERPATEXPA7	1520 (+)	KCACGW	GT1CONSENSUS	1725 (+)	GRWAAW
ABRELATERD1	1521 (-)	ACGTG	IBOXCORE	1725 (+)	GATAA
ACGTATERD1	1522 (+)	ACGT	NTBBF1ARROLB	1735 (+)	ACTTTA
ACGTATERD1	1522 (-)	ACGT	DOFCOREZM	1736 (-)	AAAG
POLLEN1LELAT52	1527 (+)	AGAAA	TAAAGSTKST1	1736 (-)	TAAAG
SEF4MOTIFGM7S	1543 (-)	RTTTTTR	CACTFTPPCA1	1739 (+)	YACT

DOFCOREZM	1758 (+)	AAAG	REBETALGLHCB21	1878 (-)	CGGATA
WBOXPCWRKY1	1767 (+)	TTTGACY	POLLEN1LELAT52	1899 (+)	AGAAA
ELRECOREPCR1	1768 (+)	TTGACC	DOFCOREZM	1901 (+)	AAAG
WBOXATNPR1	1768 (+)	TTGAC	SURECOREATSULTR1	1908 (-)	GAGAC
WBOXNTERF3	1769 (+)	TGACY	ACGTABOX	1915 (+)	TACGTA
WRKY71OS	1769 (+)	TGAC	ACGTABOX	1915 (-)	TACGTA
POLLEN1LELAT52	1775 (+)	AGAAA	ACGTATERD1	1916 (+)	ACGT
TATABOXOSPAL	1785 (-)	TATTTAA	ACGTATERD1	1916 (-)	ACGT
ROOTMOTIFTAPOX1	1788 (-)	ATATT	TATABOX4	1919 (+)	TATATAA
ROOTMOTIFTAPOX	1789 (+)	ATATT	TATAPVTRNALEU	1919 (-)	TTTATATA
GTGANTG10	1803 (+)	GTGA	LECPLEACS2	1923 (+)	TAAAATAT
PREATPRODH	1807 (+)	ACTCAT	ROOTMOTIFTAPOX	1926 (-)	ATATT
WBOXPCWRKY1	1814 (-)	TTTGACY	SORLREP3AT	1932 (+)	TGTATATAT
WBOXHVISO1	1814 (-)	TGACT	-10PEHVPSBD	1948 (+)	TATTCT
WBOXNTERF3	1814 (-)	TGACY	DOFCOREZM	1952 (-)	AAAG
WBOXATNPR1	1815 (-)	TTGAC	WBOXHVISO1	1966 (-)	TGACT
WRKY71OS	1815 (-)	TGAC	WBOXNCHN48	1966 (-)	CTGACY
MYBCOREATCYCB1	1820 (+)	AACGG	WBOXNTERF3	1966 (-)	TGACY
TCA1MOTIF	1828 (-)	TCATCTTCTT	WRKY71OS	1967 (-)	TGAC
LTRE1HVBLT49	1844 (+)	CCGAAA	GT1CONSENSUS	1972 (+)	GRWAAW
ROOTMOTIFTAPOX1	1852 (+)	ATATT	CACFTPPCA1	1983 (+)	YACT
-300ELEMENT	1858 (+)	TGHAAARK	NODCON2GM	1991 (+)	CTCTT
GATABOX	1878 (-)	GATA	OSE2ROOTNODULE	1991 (+)	CTCTT
MYBST1	1878 (-)	GGATA			

Table 5.1: Putative *cis*-acting regulatory elements found in *TT10* promoter with Signal Scan.

### 5.1.2 CAREs found in *AtcisDB*

The list of putative *cis*-acting regulatory elements had been obtained by querying *AtcisDB Search Engine* (Molina and Grotewold, 2005) for locus ID At5g48100 (up dated 16.04.2009). The image of the genomic region of the At5g48100 locus is a visualization of the data presented in the table (adapted from *AtcisDB*).

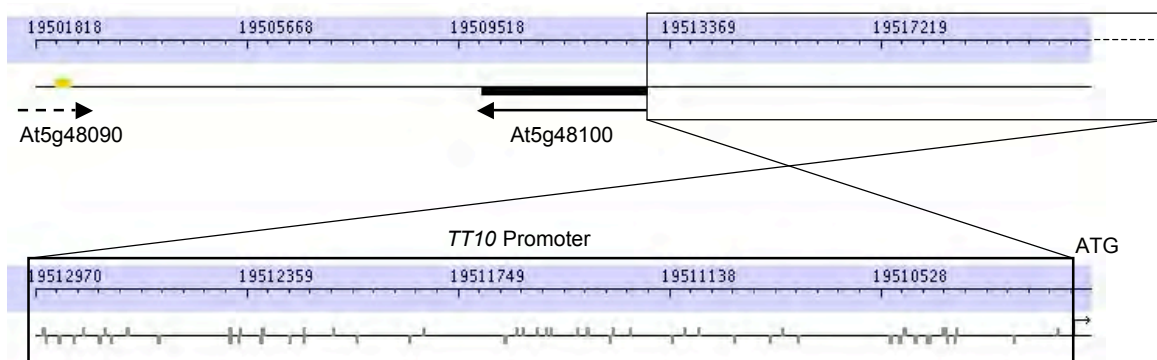


Figure 5.1: Visualization of *cis*-acting regulatory elements found with *AtcisDB* search engine. Blue lines represent chromosome 5 and At5g48100/*TT10* gene - black box. *TT10* is transcribed from negative strand and At5g48090 from positive strand - yellow box. Bottom part of the figure is focusing on the promoter sequence analyzed. Grey boxes are representing putative *cis*-acting regulatory elements identified and listed in the Table 5.2.

Table 1.2

Binding Site (BS) Name	BS Genome Start	BS Genome End	Binding Site Sequence	Binding Site Family/TF
AtMYC2 BS in RD22	19510310	19510315	cacatg	BHLH
AtMYC2 BS in RD22	19512829	19512834	cacatg	BHLH
Bellringer/replumless/pennywise BS1 IN AG	19510139	19510146	aaattaa	Homeobox
Bellringer/replumless/pennywise BS1 IN AG	19510495	19510502	aaattaa	Homeobox
Bellringer/replumless/pennywise BS2 IN AG	19512403	19512410	aaattagt	Homeobox
Bellringer/replumless/pennywise BS3 IN AG	19512402	19512409	actaattt	Homeobox
ATB2/AtbZIP53/AtbZIP44/GBF5 BS in ProDH	19512747	19512752	actcat	bZIP
ATB2/AtbZIP53/AtbZIP44/GBF5 BS in ProDH	19512937	19512942	actcat	bZIP
W-box promoter motif	19512898	19512903	ttgacc	WRKY
W-box promoter motif	19512943	19512948	ttgact	WRKY
W-box promoter motif	19512374	19512379	ttgact	WRKY
W-box promoter motif	19511247	19511252	ttgact	WRKY
ARF1 binding site motif	19510305	19510310	tgctc	ARF
CCA1 binding site motif	19512379	19512386	aaaaatct	MYB-related
DPBF1&2 binding site motif	19510309	19510315	acacatg	bZIP
DPBF1&2 binding site motif	19511305	19511311	acacctg	bZIP
DPBF1&2 binding site motif	19512700	19512706	acaccag	bZIP
HSEs binding site motif	19512311	19512320	agaagattct	HSF
HSEs binding site motif	19512310	19512319	agaatctct	HSF
MYB binding site promoter	19510455	19510462	caccaacc	MYB
MYB2 binding site motif	19512610	19512618	taactcgtt	MYB
MYB4 binding site motif	19512231	19512237	accaaac	MYB
MYB4 binding site motif	19510455	19510461	accaaacc	MYB
RAV1-A binding site motif	19511128	19511132	caaca	ABI3VP1
RAV1-A binding site motif	19512704	19512708	caaca	ABI3VP1
RAV1-A binding site motif	19511846	19511850	caaca	ABI3VP1
RAV1-A binding site motif	19511577	19511581	caaca	ABI3VP1
RAV1-A binding site motif	19511492	19511496	caaca	ABI3VP1
RAV1-A binding site motif	19510013	19510017	caaca	ABI3VP1
RAV1-B binding site motif	19511306	19511311	cacctg	ABI3VP1
LFY consensus binding site motif	19510843	19510848	ccaatg	LFY
LFY consensus binding site motif	19511885	19511890	ccattg	LFY
LFY consensus binding site motif	19512039	19512044	ccaatg	LFY
LFY consensus binding site motif	19512191	19512196	ccattg	LFY
LFY consensus binding site motif	19512191	19512196	ccaatg	LFY
LFY consensus binding site motif	19511051	19511056	ccattg	LFY
ABRE-like binding site motif	19511370	19511377	tacgtgta	...
ARF binding site motif	19510305	19510310	tgctc	...
BoxII promoter motif	19510389	19510394	ggtaa	...
BoxII promoter motif	19510425	19510430	ggtaa	...
DRE-like promoter motif	19511608	19511616	aaccgactt	...
EveningElement promoter motif	19510467	19510475	aaaatatct	...
GATA promoter motif [LRE]	19510332	19510337	tgataa	...
GATA promoter motif [LRE]	19510762	19510767	agataa	...
GATA promoter motif [LRE]	19512607	19512612	tgataa	...

GATA promoter motif [LRE]	19512854	19512859	agataa	...
GATA promoter motif [LRE]	19512765	19512770	agataa	...
GATA promoter motif [LRE]	19512105	19512110	agataa	...
GATA promoter motif [LRE]	19511560	19511565	agatag	...
GATA promoter motif [LRE]	19511519	19511524	tgatag	...
GATA promoter motif [LRE]	19511479	19511484	agatag	...
GATA promoter motif [LRE]	19511297	19511302	tgataa	...
GATA promoter motif [LRE]	19511090	19511095	agataa	...
GATA promoter motif [LRE]	19510347	19510352	agatag	...
GATA promoter motif [LRE]	19510338	19510343	agatag	...
GCC-box promoter motif	19511400	19511405	gccgcc	...
Ibox promoter motif	19510763	19510768	gataag	...
Ibox promoter motif	19511296	19511301	gataag	...
L1-box promoter motif	19510946	19510953	taaatgca	...
L1-box promoter motif	19510808	19510815	taaatgta	...
T-box promoter motif	19512781	19512786	actttg	...
Z-box promoter motif	19511371	19511378	atactgtg	...
SORLIP2	19510383	19510387	gggcc	...

Table 5.2: Putative *cis*-acting regulatory elements found in *TT10* promoter with *AtcisDB*.

## 5.2 Primers

All the primers were purchased from SIGMA. They were desalted, except primers used for the site-directed mutagenesis which were PAGE-purified. Primers were delivered lyophilized and dissolved in water after arrival to obtain 100 $\mu$ M stock solution.

Name	5'-3' sequence	Orientation on promoter	Restriction site
pTT10 6	gta <b>CTCGAG</b> cacactgattttgcttggaaatg	F	<b>XhoI</b>
pTT10 5	gta <b>CTCGAG</b> tccatcatgtcagcttac	F	<b>XhoI</b>
pTT10 4	gta <b>CTCGAG</b> tagaagatgacatgaagagg	F	<b>XhoI</b>
pTT10 3	gta <b>CTCGAG</b> attaacacaaaaacgtaacgag	F	<b>XhoI</b>
pTT10 3C	gta <b>CTCGAG</b> aaataaaggtaatgcttggcc	F	<b>XhoI</b>
pTT10 3B	gta <b>CTCGAG</b> aactagttgattgatgtcacg	F	<b>XhoI</b>
pTT10 3A	gta <b>CTCGAG</b> tactatatatgatataacatgtg	F	<b>XhoI</b>
pTT10 2	gta <b>CTCGAG</b> atacgaactttactgttttcg	F	<b>XhoI</b>
pTT10 1	gta <b>CTCGAG</b> aagaagatgatgcttaccg	F	<b>XhoI</b>
pTT10:ATG:GUS	acaggacgta <b>CCATGG</b> tggaagag	R	<b>NcoI</b>

Table 5.3: Primers used for promoter dissection to clone fragments in *pBS-GUS* vector.

Name	5'-3' sequence	Orientation on promoter
At5g48110 catma F	tacttccaagctagtgaatgaagag	F
pTT10 -4.5kb	cagcatgtattttcctcatgag	F
pTT10 -4.0kb	taaaccgggtgattttccgcc	F
pTT10 -3.5kb	ttaaaccatttaacaactcaacc	F
pTT10 -3.0kb	ttacaaaatgaaacgatgcatg	F
pTT10 -2.5kb	acaagcgtgaagaagactgg	F
pTT10 0.0kb	ataaattgaagaaggaatgtgac	R

Table 5.4: Primers used for amplification of long promoter fragments in *Ler* and *Cvi* accessions.

Name	5'-3' sequence	Orientation on promoter	Restriction site
pTT10 3 SmaI	gta <b>CCCGGG</b> attaacacaaaaacgtaacgag	F	<b>SmaI</b>
pTT10 2 SmaI	gta <b>CCCGGG</b> atacgactttactgttttcg	F	<b>SmaI</b>
pTT10 2 SacII	gta <b>CCGCGG</b> agttaacgaaaacagtaaagtcg	R	<b>SacII</b>
pTT10 1 SacII	gta <b>CCGCGG</b> acaaaatatgttttcggttaagc	R	<b>SacII</b>
pTT10 0.0 kb SacII	gta <b>CCGCGG</b> aagagtttttagtaaattaattacc	R	<b>SacII</b>

Table 5.5: Primers used to amplify promoter fragments for transcription factor library screening with *pHISi* vector.

Name of the construct and primers used	5'-3' sequence
M1 F	CTAAATTAACACAAAATTTTAACGAGTTTCGGAATA
M1 R	TATTCGAAACTCGTTAAAATTTGTGTTAATTTAG
M2 F	GTTTCGGAATATTTCTGCTATCCATTTATATAAAAATAAAGG
M2 R	CCTTTATTTTATATAAATGGATAGCAGAAATATTCGAAAC
M3 F	GCTTGGCCTAACATTTTTTCATAAGTGATTT
M3 R	AAATCACTTATGAAAAATGTTAGGCCAAGC
M4 F	GTGATTTTATTTATTTTTTAGTAAAATTTTGAAGTAGTTGATTGATGTC
M4 R	GACATCAATCAACTAGTTCAAATTTTTACTAAAAATAAAATAAAATCAC
M5 F	GAACTAGTTGATTGATGATAAACTCTTACTATATATG
M5 R	CATATATAGTAAGAGTTTTATCATCAATCAACTAGTTC

Table 5.6: Primers used for site-directed mutagenesis on *pTT10\_3* promoter fragment in *pBS-GUS* vector.

Name	5'-3' sequence	Orientation on promoter	Probe for:	Probe size
MDN-ATG0	gcttaccgaaaacatattttgtaa	F	5' region	820bp
MDN-Rev 800 TT10	caaaattatggggacttcatgg	R		
3Fw-TT10	gtataaaaatacaatgaccgtgcctagg	F	3' region	990bp
MDN-rv-0-long	gttatgacacaagattaatacattgtttatc	R		

Table 5.7: Primers used for probes preparation for BAC library screening.

Name	5'-3' sequence	Orientation on promoter	GATEWAY recombination site
pTT10 3	attaacacaaaaacgtaacgag	F	<i>attB1</i>
pTT10 3C	aaataaaggtaatgcttgcc	F	<i>attB1</i>
pTT10 3B	aactagttgattgatgacg	F	<i>attB1</i>
pTT10 3A	tactatataatgataacatgtg	F	<i>attB1</i>
pTT10 2	atacgactttactgttttcg	F	<i>attB1</i>
LAC15 R1	tttggaagagtttttagtaaattaattacc	R	<i>attB2</i>
<i>attB1</i> :	ggggacaagttgtacaaaaaacgaggct		
<i>attB2</i> :	ggggaccactttgtacaagaaagctgggt		

Table 5.8: Primers used for promoter dissection - GATEWAY® cloning.

Name	5'-3' sequence	Comments	AGI
LAC1 5GST	TTCCGTAAACACACGGCG	CATMA GST	At1g18140
LAC1 3GST	AGAATCTCTTCCCTGCGTAACC		
LAC2 5GST	TGTGTGGCTGATGCATTG	CATMA GST	At2g29130
LAC2 3GST	CATTCCAAGAAAGGAAAAAGG		
LAC3 5GST	GTGGTCAAATCGCTCCTGTT	CATMA GST	At2g30210
LAC3 3GST	TTGTGAAGATTCCTGGGGTG		
LAC4 5GST	GGTTTGGTTCATGCATTGTCACTT	CATMA GST	At2g38080
LAC4 3GST	CCACGCATTAATCACCCTC		
LAC5 5GST	TCAATACAAATCCGCCCC	CATMA GST	At2g40370
LAC5 3GST	GTCGATGGTTACGAAGAGA		
LAC6 5GST	ATGGATGGGCTCAGAAGT	CATMA GST	At2g46570
LAC6 3GST	GATTTTCTGTTGTGAGTGTG		
LAC7 5GST	AGCAATCACTCTTTCTGTG	CATMA GST	At3g09220
LAC7 3GST	AAGCGCATGGTTCTGTAA		
LAC8 F	ATGGGTTGTCTCAGATTTATCGCTAATA	McCaig et al. 2005	At5g01040
LAC8 R	TGCTCTCGTACAAATCAGTTCAAGCAT		
LAC9 F	ATGGGTTGTCTCAGATTTATCGCTAATA	McCaig et al. 2005	At5g01050
LAC9 R	GTGTAGAGAAATGGAAGCATGGTTCGTAAA		
LAC10 5GST	ATTAAGCCGGGTCACAGCTAC	CATMA GST	At5g01190
LAC10 3GST	GTATGATGACTTCTTCTATGTGG		
LAC11 5GST	TGAGATGGGAACACTACTCTGAT	CATMA GST	At5g03260
LAC11 3GST	CATTTCTTATGCAATTTCCGGG		
LAC12 5GST	CCAACCCCAAAACGCTATTAC	CATMA GST	At5g05390
LAC12 3GST	CCTTGGCACC GGCTTTTAG		
LAC13 5GST	CACAGTAGATGGCTGCGA	CATMA GST	At5g07130
LAC13 3GST	TTGGCTGGAGCATCTGTAG		
LAC14 5GST	CCGGATTAAACCGATTCTACCCT	CATMA GST	At5g09360
LAC14 3GST	GTCGAAATGGAAGATTCTCAC		
LAC15 5GST	CATCGTTTTTCAGGATATTACCCAAC	CATMA GST	At5g48100
LAC15 3GST	GAGTTTTGAGGACACATGCGAAG		
LAC16 5GST	TTTACCAGTGAGAGCAGGAA	CATMA GST	At5g58910
LAC16 3GST	GTGTGACCGATATAGTGGAG		
LAC17 5GST	GTCTCTTCTTTCTTACAAGCTTCGC	CATMA GST	At5g60020
LAC17 3GST	TGGGTTTACATAGATCTACACTTGC		

Table 5.9: Primers used to study laccase family genes expression.

CATMA GST primers are coming from the CATMA project: <http://www.catma.org/database/>

Name	5'-3' sequence	Orientalion on promoter	GATEWAY recombination site	Promoter Size
LAC3 F1	aagtcaaacttcacgatttacg	F	<i>attB1</i>	1.44 kb
LAC3 R1	tgtttgcttcaactgtgagaag	R	<i>attB2</i>	
LAC4 F2	gtttgatccagttttggttcg	F	<i>attB1</i>	1.95 kb
LAC4 R1	ctcctctctatctttctcttc	R	<i>attB2</i>	
LAC5 F2	caaattcagatgggaggtgg	F	<i>attB1</i>	1.9 kb
LAC5 R1	tgatgtgatttctgtaagcttc	R	<i>attB2</i>	
LAC12 F1	ggaatggttgattactattaagc	F	<i>attB1</i>	1.65 kb
LAC12 R1	taagtgtgtttaagctttgag	R	<i>attB2</i>	

*attB1*: ggggacaagtttgtaaaaaagcaggct  
*attB2*: gggaccacttttgtaaaaaagcaggct

Table 5.10: Primers used for laccase promoter cloning.



Name	5'-3' sequence	T <sub>m</sub> (°C) of the product
3Fw-TT10	gtataaaaaatacaatgaccgtgcctagg	82,7
3Rev-TT10	cagaatctgctgatttggctctct	
qEF F	ctggaggttttgaggctggat	82,4
qEF R	ccaagggtgaaagcaagaaga	
APT1 F	cctattgcgttggctattg	85,2
APT1 R	tcttcaactcactactcgttc	
TUB4a F	tttctcagtgtttccttctcc	83,2
TUB4a R	cctcattgtccaaaaccatac	

Table 5.11: Primers used in quantitative RT-PCR.

Name	5'-3' sequence	Name of the mutant line	
		NASC	SALK
At1g71450 LP	caaaaatgcataattccaaagc	N667434	SALK_136922C
At1g71450 RP	aggggaagttgagacgagcttc		
At2g24430 LP	gtcgcaagtcaccttgtctc	N660776	SALK_025040C
At2g24430 RP	tccgttcgataagctcatgtc		
At3g15500 LP	taaacgatgagcगतagcgag	N514331	SALK_014331 (D)/(R)
At3g15500 RP	aaaggaacccaaaccaattgg		
At5g07700 LP	ccttgatgtagagctcgatcg	N662521	SALK_055242C
At5g07700 RP	attgtcacgagtgcatgacac		
At5g62320 LP1	ggattaagaaaggggtccatgg	N503193	SALK_003193 (F)
At5g62320 RP1	tcaggtcagggacattcaaag		
At5g66070 LP	catgacgacatcactgtggtc	N658053	SALK_148182C
At5g66070 RP	ctaagtgccgtccagagtcag		
Sig LB1	cggaaccaccatcaaacag	nd	nd

Table 5.12: Primers used for genotyping of SALK T-DNA mutants.

Primers used for genotyping were designed with T-DNA Primer Design available at <http://signal.salk.edu/tdnaprimers.2.html>.

Name	5'-3' sequence	Used for:
EF1 Alpha A4 UP	ATGCCCCAGGACATCGTGATTTTCAT	expression control for RT-PCR
EF1 Alpha A4 LOW	TTGGCGGCACCCTTAGCTGGATCA	
nos3A	CATCGCAAGACCGGCAACAGG	constructs verification and sequencing in pBIB-Hyg vector
pBIB Hyg	CCATGTAAGCCCACTGCAAGCTACC	
UP	GTA AACGACGGCCAGT	Used for pTOPO and pBS construct verification and sequencing
RP	GGAAACAGCTATGACCATG	
GUS1	CTGATCAATTCCACAGTTTTCGCG	Used for GUS constructs verification and sequencing

Table 5.13: Other primers used in the study.

Name	5'-3' sequence	AGI	CATMA GST	PROBE / PCR product size
TT1 GST 5	TGGAGTCACCACCCTATAACGAGA	At1g34790	CATMA1a33100	158
TT1 GST 3	CAAGTTTATGCGATCAATAAGCGG			
TT2 GST 5	TTTCGATAGGATCCATTCGGAGTTT	At5g35550	CATMA5a30760	207
TT2 GST 3	CAACAAGTGAAGTCTCGGAGC			
TT3 GST 5	GAACGAGGCGCATTACTC	At5g42800	CATMA5a38610	191
TT3 GST 3	TTGAAGGTACGTTATATTCGGG			
TT4 GST 5	CTGCCGCTCAGACCATCCT	At5g13930	CATMA5a12150	400
TT4 GST 3	CCGAAACCAAACAAGACACC			
TT5 GST 5	GCGTTTTTCGAAAGATGATAG	At3g55120	CATMA3a48130	221
TT5 GST 3	CAGTTCCTTTGGCTAGTTT			
TT6 GST 5	TTTTGAGCAATGGGAGGT	At3g51240	CATMA3a44230	271
TT6 GST 3	TCGACAGGCTTGTCACCTC			
TT7 GST 5	GGTACATCCTAAGCCAAGGT	At5g07990	CATMA5a07240	173
TT7 GST 3	ATGACTTGATTTGGCGTTACAC			
TT8 GST 5	TTTGTGCTTGACATTCCTCAG	At4g09820	CATMA4c42186	220
TT8 GST 3	ATCCTTCTCTAGAGTTGGCATCA			
TT12 GST 5	TGAGCTCCACAGAGACATAC	At3g59030	CATMA3a52070	195
TT12 GST 3	AGACCACTATAGAGGCTCCAG			
TT16/ABS GST 5	AGCAATTCGACCATGAGA	At5g23260	CATMA5a20750	163
TT16/ABS GST 3	TTACGCTCACGGACTTTG			
TT18 GST 5	TATGCAAACGATCAAGCCA	At4g22880	CATMA4a24645	160
TT18 GST 3	TGTAATCACTTGGTGTCTTAGG			
TT19 GST 5	AGAGCACCGAGCCATCGT	At5g17220	CATMA5a15500	300
TT19 GST 3	CCCCGAGCCTTAACCATCT			
TTG1 GST 5	TGGATAATTCAGCTCCAGATTCGTT	At5g24520	CATMA5a22120	152
TTG1 GST 3	GTTGTTGTAATCTTCGAGGAAGC			
TTG2 GST 5	GGTCAAGTTTCAGAGATTGTGT	At2g37260	CATMA2a35540	361
TTG2 GST 3	TTCTTTTGCTTCTGCTTGG			
BAN GST 5	ACATGATCAAGCCGGCGATAC	At1g61720	CATMA1a50805	196
BAN GST 3	CAGTTAAAAGGCTTCTCCTCTGTG			
AHA10 GST 5	CGGCTACATTAATTGCTGTGT	At1g17260	CATMA1a16290	181
AHA10 GST 3	TTACGGTCCAATACGAGATTCC			
FLS1 GST 5	AGGATTACAGTTACCGCA	At5g08640	CATMA5a07985	203
FLS1 GST 3	ACATGTCAACATCACACACA			
F30GT 2 GST 5	GGAGCAATGGATGAATGCC	At2g18560	CATMA2c47270	167
F30GT 2 GST 3	GTTTTGCCCATTCAAAGAGA			
Q30MT GST 5	GACCAAAAATCCTGAAGC	At5g54160	CATMA5a50073	218
Q30MT GST 3	AGCTTTCATGAGAACCCTTG			
F30GT 5 GST 5	CCCAATATATCCTCCGACCT	At5g65550	CATMA5a60880	344
F30GT 5 GST 3	GCAGTCTTTCGAGGGTCA			
78D 1 GST 5	CTCCATGAAAGGAAGCTCTTTA	At1g30530	CATMA1a28550	212
78D 1 GST 3	CAGGTACAATGAAAATCAACAACCC			
78D 2 GST 5	TCTTCTCTTTAGCCCTCTCCT	At5g17050	CATMA5c64278	499
78D 2 GST 3	CATTTATCTCCGTCGCCATA			
TDS 2 GST 5	ACCAAGTAACAGCATGGA	At3g17980	CATMA3a17450	239
TDS 2 GST 3	TAGTTGTTTCGTGTAAATGCC			
TDS 6 GST 5	TCGGGTTTGAGACGGAAGG	At5g05270	CATMA5a04460	174
TDS 6 GST 3	TAAGGGGCCATGGTTCTTAGGTT			
PAP 1 GST 5	CACAGTTAACAACGACTGCAAC	At1g56650	CATMA1a47745	356
PAP 1 GST 3	AGGTGAAAGCAAACCTATACAC			
PAP 2 GST 5	AAGCTGATGCGATTGTTC	At1g66390	CATMA1c72010	200
PAP 2 GST 3	CTGAAATCTTTACTTTCATACAG			
PAL 1 GST 5	AGATTAACGGGGCACACAAG	At2g37040	CATMA2a35330	242
PAL 1 GST 3	AGTTGAGATCGCAGCCACTT			
PAL 2 GST 5	CAACAACACTAACATTTGCTCTTC	At3g53260	CATMA3c57749	236
PAL 2 GST 3	TGCCAAAGTCTTCGTAGTAACC			
PAL 4 GST 5	TAAATCAGCAGTGAGTCAGGTGG	At3g10340	CATMA3a09340	242
PAL 4 GST 3	GGAAAACCGATGTCCGGG			
C4H GST 5	GACCAGAGAGGTTCTTTGA	At2g30490	CATMA2a28746	400
C4H GST 3	GCAATCGTAGAACGAACCA			
4CL3 GST 5	GAGATCGGTGCGTTATTGGA	At1g65060	CATMA1a54346	204
4CL3 GST 3	AAGGATGGCTTGAGGGAGAC			
MYB4 GST 5	AAGGGCATGGAAGTCAACAAC	At4g38620	CATMA4a40092	200
MYB4 GST 3	CTCCAAGCTTCGAAAGCCC			
MYB12 GST 5	CACAACCTCATTAGGAAGCC	At2g47460	CATMA2a45915	290
MYB12 GST 3	CAATCATCTCCATAGTAGTCAC			
MYBL2 GST 5	ATGTGGTGAACAATTTTGCC	At1g71030	CATMA1a60325	200
MYBL2 GST 3	CTCATCGGAATAGAAGAAGC			

Table 5.14: CATMA GST primers used to study expression of flavonoid biosynthesis genes.

## 5.3 Seeds colors

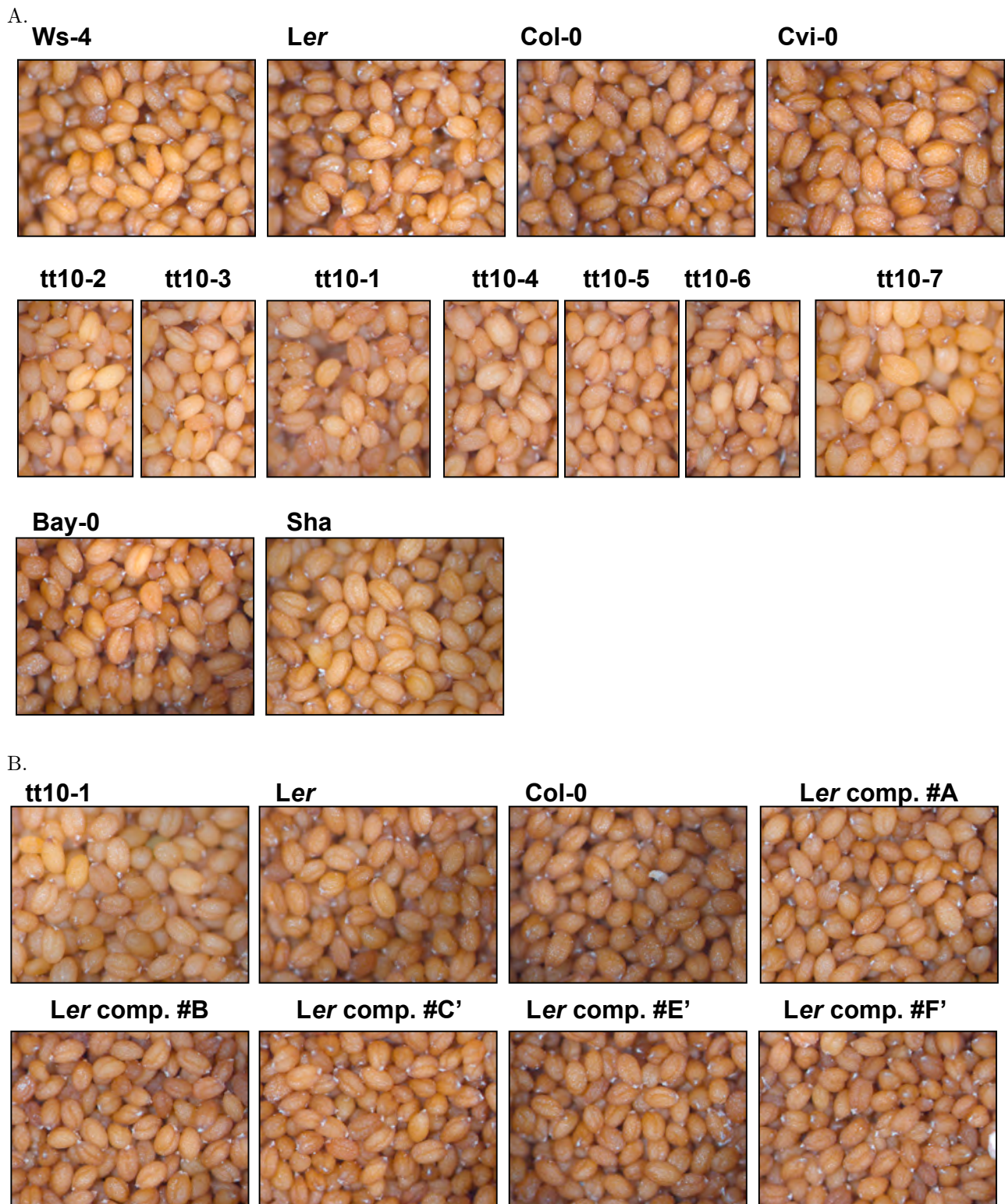


Figure 5.2: Seed colors of various *Arabidopsis* accessions and *tt10* mutant alleles, (A). Seed color of the plants used in the Ler 'complementation' experiment, (B).

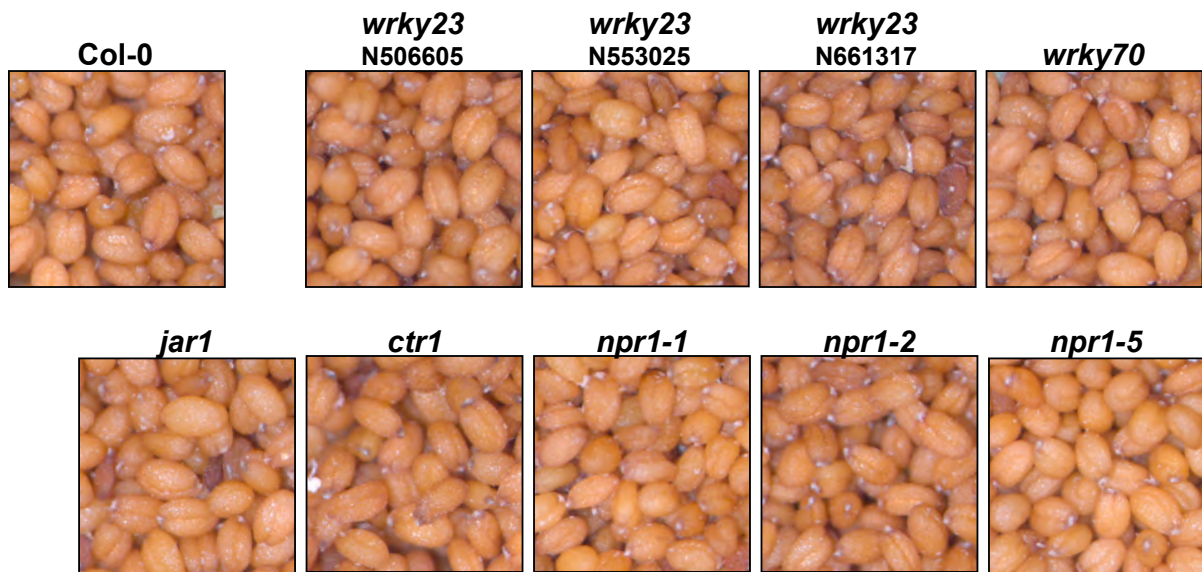


Figure 5.3: Seed color of various mutants used in the study - Part 1.

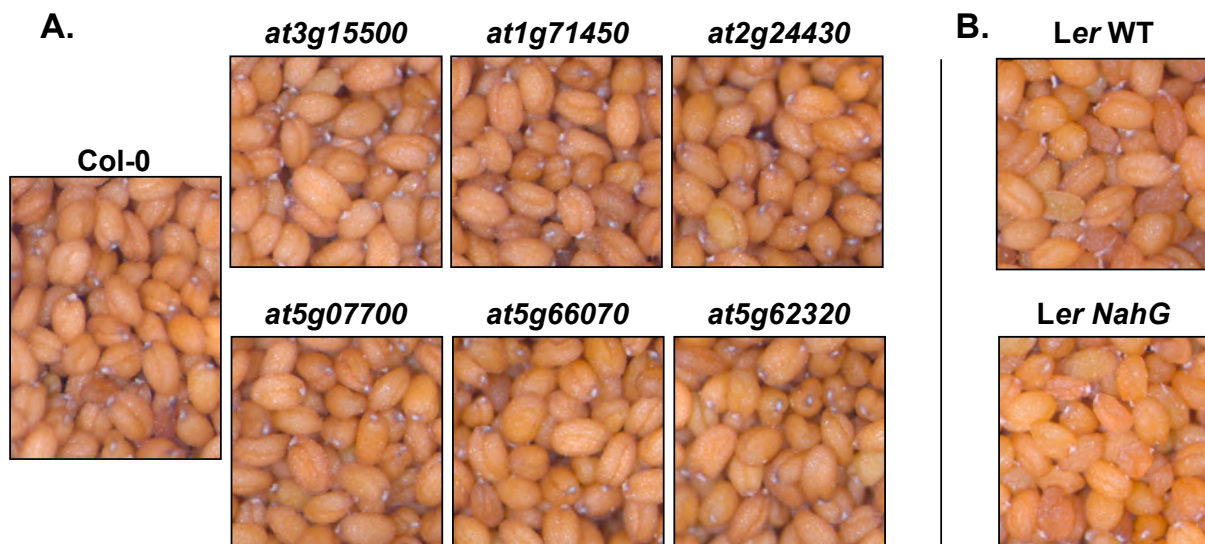
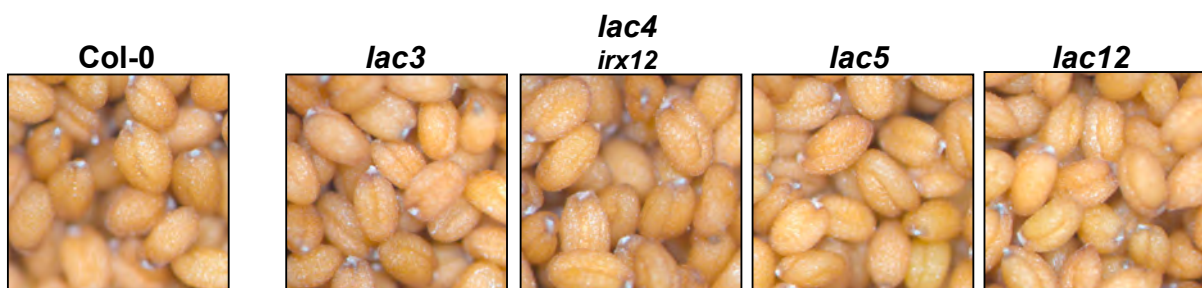
Figure 5.4: Seed color of various mutants used in the study - Part 2. (A) Mutants chosen from co-expression studies. (B) Mutant carrying *NahG* gene and *Ler* wild-type control.

Figure 5.5: Seed color of mutants in laccase gene.

## 5.4 Map of the methylation in the *TT10* gene

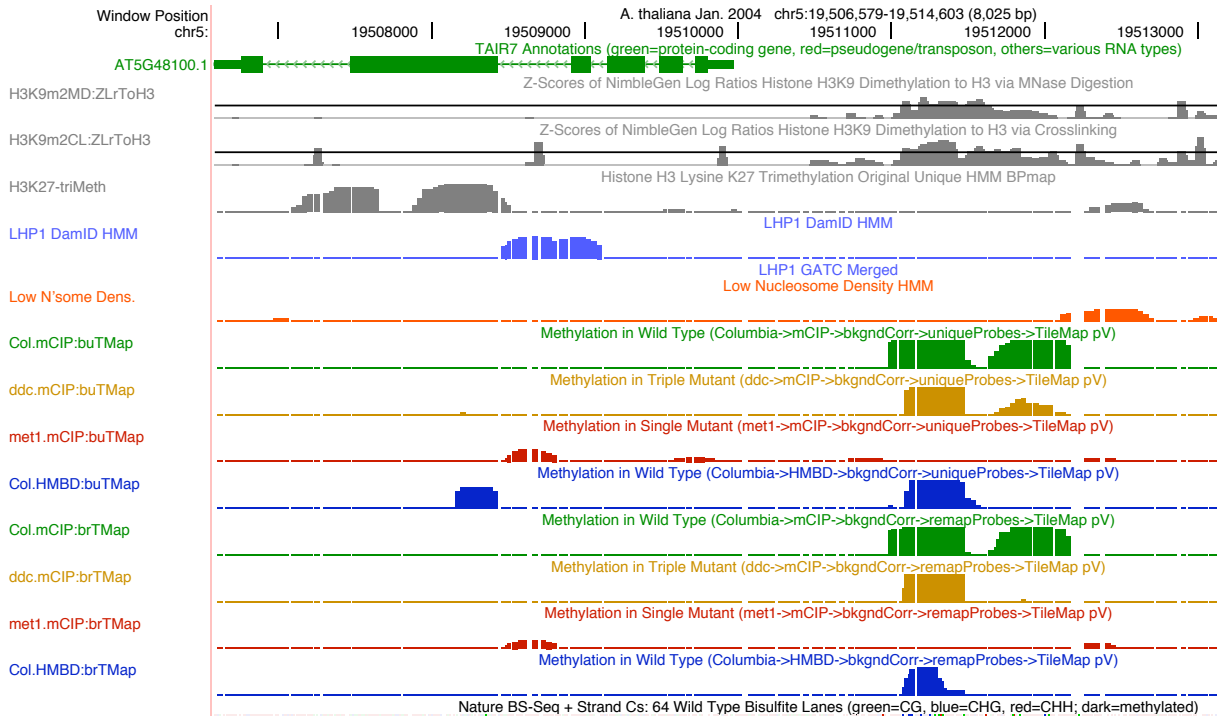


Figure 5.6: Methylation map of the *TT10* gene region.

Methylation map of the DNA in seedlings have been pictured by the Genome Browser at <http://epigenomics.mcdb.ucla.edu/> (UCSC Genome Bioinformatics).

### 5.5 Transposon in *TT10* promoter

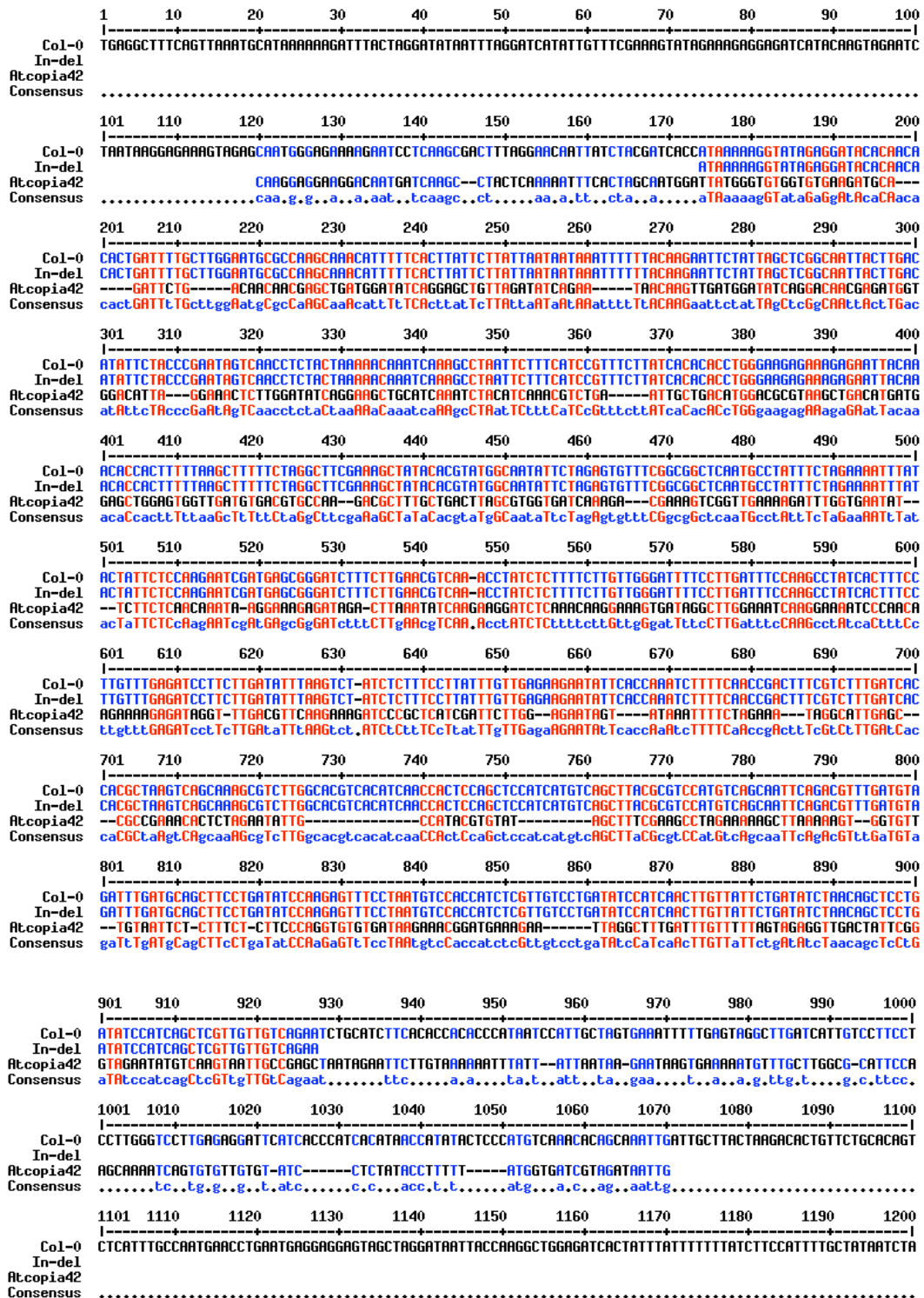


Figure 5.7: Alignment of transposon sequence.

Alignment of the Col-0 wild-type intergenic region -2.2-kb to -1.0-kb of *TT10* promoter. In-del is a promoter fragment identified by Hirofumi Ishihara which is not present in *Ler*. Atcopia42 is a transposon sequence retrieved from FLAGdb++ found in *TT10* promoter region.



# Bibliography

- Abdel-Ghany SE, Pilon M (2008) MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in Arabidopsis. *The Journal of Biological Chemistry* 283: 15932–15945, URL <http://www.ncbi.nlm.nih.gov/pubmed/18408011>, PMID: 18408011.
- Abe H, Urao T, Ito T, Seki M, Shinozaki K, et al. (2003) Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *The Plant Cell* 15: 63–78, URL <http://www.ncbi.nlm.nih.gov/pubmed/12509522>, PMID: 12509522.
- Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, et al. (1997) Role of arabidopsis MYC and MYB homologs in drought- and abscisic acid-regulated gene expression. *The Plant Cell* 9: 1859–68, URL <http://www.ncbi.nlm.nih.gov/pubmed/9368419>, PMID: 9368419.
- Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, et al. (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science (New York, NY)* 301: 653–7, URL <http://www.ncbi.nlm.nih.gov/pubmed/12893945>, PMID: 12893945.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410, URL <http://www.ncbi.nlm.nih.gov/pubmed/2231712>, PMID: 2231712.
- Alves-Ferreira M, Wellmer F, Banhara A, Kumar V, Riechmann JL, et al. (2007) Global expression profiling applied to the analysis of Arabidopsis stamen development. *Plant Physiology* 145: 747–762, URL <http://www.ncbi.nlm.nih.gov/pubmed/17905860>, PMID: 17905860.
- Aron PM, Kennedy JA (2008) Flavan-3-ols: nature, occurrence and biological activity. *Molecular Nutrition & Food Research* 52: 79–104, URL <http://www.ncbi.nlm.nih.gov/pubmed/18081206>, PMID: 18081206.
- Axelos M, Couric C, Mazzolini L, Bardet C, Lescure B (1992) A protocol for transient gene expression in Arabidopsis thaliana protoplasts isolated from cell suspension cultures. *Plant Physiology and Biochemistry* pp. 123–128.
- Baud S, Vaultier M, Rochat C (2004) Structure and expression profile of the sucrose synthase multigene family in Arabidopsis. *Journal of Experimental Botany* 55: 397–409, URL <http://www.ncbi.nlm.nih.gov/pubmed/14739263>, PMID: 14739263.
- Baudry A, Caboche M, Lepiniec L (2006) TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in Arabidopsis thaliana. *The Plant Journal: For Cell and Molecular Biology* 46: 768–779, URL <http://www.ncbi.nlm.nih.gov/pubmed/16709193>, PMID: 16709193.



- Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, et al. (2004) TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *The Plant Journal: For Cell and Molecular Biology* 39: 366–380, URL <http://www.ncbi.nlm.nih.gov/pubmed/15255866>, PMID: 15255866.
- Bechtold N, Ellis J, Pelletier G (1993) In planta *Agrobacterium*-mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *CR Acad Sci Série III-Sciences de la Vie* 316: 1194–99.
- Bentsink L, Jowett J, Hanhart CJ, Koornneef M (2006) Cloning of DOG1, a quantitative trait locus controlling seed dormancy in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* 103: 17042–17047, URL <http://www.ncbi.nlm.nih.gov/pubmed/17065317>, PMID: 17065317.
- Berger S, Bell E, Mullet JE (1996) Two Methyl Jasmonate-Insensitive Mutants Show Altered Expression of AtVsp in Response to Methyl Jasmonate and Wounding. *Plant Physiology* 111: 525–531, URL <http://www.ncbi.nlm.nih.gov/pubmed/12226307>, PMID: 12226307.
- Borevitz JO, Xia Y, Blount J, Dixon RA, Lamb C (2000) Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *The Plant Cell* 12: 2383–2394, URL <http://www.ncbi.nlm.nih.gov/pubmed/11148285>, PMID: 11148285.
- Borisjuk L, Rolletschek H, Radchuk R, Weschke W, Wobus U, et al. (2004) Seed development and differentiation: a role for metabolic regulation. *Plant Biology (Stuttgart, Germany)* 6: 375–386, URL <http://www.ncbi.nlm.nih.gov/pubmed/15248120>, PMID: 15248120.
- Bouchabke-Coussa O, Quashie M, Seoane-Redondo J, Fortabat M, Gery C, et al. (2008) ESKIMO1 is a key gene involved in water economy as well as cold acclimation and salt tolerance. *BMC Plant Biology* 8: 125, URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2630945>, PMC2630945.
- Bradford M (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein dye-binding. *Anal Biochem* p. 248–254.
- Brown DM, Zeef LA, Ellis J, Goodacre R, Turner SR (2005) Identification of Novel Genes in *Arabidopsis* Involved in Secondary Cell Wall Formation Using Expression Profiling and Reverse Genetics. *The Plant Cell* 17: 2281–2295, URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1182489>, PMC1182489.
- Buchel AS, Brederode FT, Bol JF, Linthorst HJ (1999) Mutation of GT-1 binding sites in the Pr-1A promoter influences the level of inducible gene expression in vivo. *Plant Molecular Biology* 40: 387–96, URL <http://www.ncbi.nlm.nih.gov/pubmed/10437823>, PMID: 10437823.
- Busk PK, Pagès M (1998) Regulation of abscisic acid-induced transcription. *Plant Molecular Biology* 37: 425–35, URL <http://www.ncbi.nlm.nih.gov/pubmed/9617810>, PMID: 9617810.
- Cai X, Davis EJ, Ballif J, Liang M, Bushman E, et al. (2006) Mutant identification and characterization of the laccase gene family in *Arabidopsis*. *Journal of Experimental Botany* 57: 2563–9, URL <http://www.ncbi.nlm.nih.gov/pubmed/16804053>, PMID: 16804053.
- Cao H, Bowling SA, Gordon AS, Dong X (1994) Characterization of an *Arabidopsis* Mutant That Is Nonresponsive to Inducers of Systemic Acquired Resistance. *The Plant Cell* 6: 1583–1592, URL <http://www.ncbi.nlm.nih.gov/pubmed/12244227>, PMID: 12244227.

- Cao H, Glazebrook J, Clarke JD, Volko S, Dong X (1997) The Arabidopsis NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. *Cell* 88: 57–63, URL <http://www.ncbi.nlm.nih.gov/pubmed/9019406>, PMID: 9019406.
- Cao X, Costa LM, Biderre-Petit C, Kbhaya B, Dey N, et al. (2007) Abscisic acid and stress signals induce *Viviparous1* expression in seed and vegetative tissues of maize. *Plant Physiology* 143: 720–731, URL <http://www.ncbi.nlm.nih.gov/pubmed/17208960>, PMID: 17208960.
- Caselle M, Cunto FD, Provero P (2002) Correlating overrepresented upstream motifs to gene expression: a computational approach to regulatory element discovery in eukaryotes. *BMC Bioinformatics* 3: 7, URL <http://www.ncbi.nlm.nih.gov/pubmed/11876822>, PMID: 11876822.
- Chang W, Lee T, Huang H, Huang H, Pan R (2008) PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. *BMC Genomics* 9: 561, URL <http://www.ncbi.nlm.nih.gov/pubmed/19036138>, PMID: 19036138.
- Chen C, Chen Z (2002) Potentiation of developmentally regulated plant defense response by *AtWRKY18*, a pathogen-induced Arabidopsis transcription factor. *Plant Physiology* 129: 706–16, URL <http://www.ncbi.nlm.nih.gov/pubmed/12068113>, PMID: 12068113.
- Chinnusamy V, Gong Z, Zhu J (2008) Abscisic acid-mediated epigenetic processes in plant development and stress responses. *Journal of Integrative Plant Biology* 50: 1187–95, URL <http://www.ncbi.nlm.nih.gov/pubmed/19017106>, PMID: 19017106.
- Chinnusamy V, Ohta M, Kanrar S, Lee B, Hong X, et al. (2003) *ICE1*: a regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis. *Genes & Development* 17: 1043–54, URL <http://www.ncbi.nlm.nih.gov/pubmed/12672693>, PMID: 12672693.
- Chung K, Igari K, Uchida N, Tasaka M (2008) New perspectives on plant defense responses through modulation of developmental pathways. *Molecules and Cells* 26: 107–12, URL <http://www.ncbi.nlm.nih.gov/pubmed/18562800>, PMID: 18562800.
- Ciolkowski I, Wanke D, Birkenbihl RP, Somssich IE (2008) Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant Molecular Biology* 68: 81–92, URL <http://www.ncbi.nlm.nih.gov/pubmed/18523729>, PMID: 18523729.
- Claus H (2004) Laccases: structure, reactions, distribution. *Micron* (Oxford, England: 1993) 35: 93–96, URL <http://www.ncbi.nlm.nih.gov/pubmed/15036303>, PMID: 15036303.
- Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal: For Cell and Molecular Biology* 16: 735–43, URL <http://www.ncbi.nlm.nih.gov/pubmed/10069079>, PMID: 10069079.
- Constabel CP, Yip L, Patton JJ, Christopher ME (2000) Polyphenol oxidase from hybrid poplar. Cloning and expression in response to wounding and herbivory. *Plant Physiology* 124: 285–295, URL <http://www.ncbi.nlm.nih.gov/pubmed/10982443>, PMID: 10982443.
- Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research* 16: 10881–90, URL <http://www.ncbi.nlm.nih.gov/pubmed/2849754>, PMID: 2849754.

- Crowe ML, Serizet C, Thareau V, Aubourg S, Rouzé P, et al. (2003) CATMA: a complete Arabidopsis GST database. *Nucleic Acids Research* 31: 156–8, URL <http://www.ncbi.nlm.nih.gov/pubmed/12519971>, PMID: 12519971.
- de Marco A, Roubelakis-Angelakis KA (1997) Laccase activity could contribute to cell-wall reconstitution in regenerating protoplasts. *Phytochemistry* 46: 421 – 425, URL <http://www.sciencedirect.com/science/article/B6TH7-3SCB6WX-2/2/220a1f5d4c900e2501f352a3ce849531>.
- de Meaux J, Goebel U, Pop A, Mitchell-Olds T (2005) Allele-specific assay reveals functional variation in the chalcone synthase promoter of Arabidopsis thaliana that is compatible with neutral evolution. *The Plant Cell* 17: 676–690, URL <http://www.ncbi.nlm.nih.gov/pubmed/15705952>, PMID: 15705952.
- de Meaux J, Pop A, Mitchell-Olds T (2006) Cis-regulatory evolution of chalcone-synthase expression in the genus Arabidopsis. *Genetics* 174: 2181–2202, URL <http://www.ncbi.nlm.nih.gov/pubmed/17028316>, PMID: 17028316.
- Debeaujon I, Léon-Kloosterziel KM, Koornneef M (2000) Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. *Plant Physiology* 122: 403–14, URL <http://www.ncbi.nlm.nih.gov/pubmed/10677433>, PMID: 10677433.
- Debeaujon I, Lepiniec L, Pourcel L, Routaboul J (2007) Seed coat development and dormancy. In: *Seed development, dormancy and germination*. Eds. K. Bradford and H. Nonogaki, Blackwell Publishing, 26–49 pp.
- Debeaujon I, Nesi N, Perez P, Devic M, Grandjean O, et al. (2003) Proanthocyanidin-accumulating cells in Arabidopsis testa: regulation of differentiation and role in seed development. *The Plant Cell* 15: 2514–31, URL <http://www.ncbi.nlm.nih.gov/pubmed/14555692>, PMID: 14555692.
- Devic M, Guillemainot J, Debeaujon I, Bechtold N, Bensaude E, et al. (1999) The BANYULS gene encodes a DFR-like protein and is a marker of early seed coat development. *The Plant Journal: For Cell and Molecular Biology* 19: 387–98, URL <http://www.ncbi.nlm.nih.gov/pubmed/10504561>, PMID: 10504561.
- Dilkes BP, Spielman M, Weizbauer R, Watson B, Burkart-Waco D, et al. (2008) The maternally expressed WRKY transcription factor TTG2 controls lethality in interploidy crosses of Arabidopsis. *PLoS Biology* 6: 2707–2720, URL <http://www.ncbi.nlm.nih.gov/pubmed/19071961>, PMID: 19071961.
- Dixon RA, Achnine L, Kota P, Liu C, Reddy MSS, et al. (2002) The phenylpropanoid pathway and plant defence: a genomics perspective. *Molecular Plant Pathology* 3: 371–390, URL <http://www.ingentaconnect.com/content/bsc/mpp/2002/00000003/00000005/art00008>.
- Dixon RA, Paiva NL (1995) Stress-Induced Phenylpropanoid Metabolism. *The Plant Cell* 7: 1085–1097, URL <http://www.ncbi.nlm.nih.gov/pubmed/12242399>, PMID: 12242399.
- Dixon RA, Xie D, Sharma SB (2005) Proanthocyanidins—a final frontier in flavonoid research? *The New Phytologist* 165: 9–28, URL <http://www.ncbi.nlm.nih.gov/pubmed/15720617>, PMID: 15720617.
- Dong X (2004) NPR1, all things considered. *Current Opinion in Plant Biology* 7: 547–552, URL <http://www.ncbi.nlm.nih.gov/pubmed/15337097>, PMID: 15337097.
- Dorsett D (1999) Distant liaisons: long-range enhancer-promoter interactions in Drosophila. *Current Opinion in Genetics & Development* 9: 505–514, URL <http://www.ncbi.nlm.nih.gov/pubmed/10508687>, PMID: 10508687.

- Dubos C, Gourrierc JL, Baudry A, Huet G, Lanet E, et al. (2008) MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *The Plant Journal: For Cell and Molecular Biology* 55: 940–53, URL <http://www.ncbi.nlm.nih.gov/pubmed/18532978>, PMID: 18532978.
- Earley KW, Haag JR, Pontes O, Opper K, Juehne T, et al. (2006) Gateway-compatible vectors for plant functional genomics and proteomics. *The Plant Journal* 45: 616–629, URL <http://dx.doi.org/10.1111/j.1365-313X.2005.02617.x>.
- Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. *Trends in Plant Science* 5: 199–206, URL <http://www.ncbi.nlm.nih.gov/pubmed/10785665>, PMID: 10785665.
- Eulgem T, Somssich IE (2007) Networks of WRKY transcription factors in defense signaling. *Current Opinion in Plant Biology* 10: 366–71, URL <http://www.ncbi.nlm.nih.gov/pubmed/17644023>, PMID: 17644023.
- Finkelstein RR, Lynch TJ (2000) The *Arabidopsis* abscisic acid response gene *ABI5* encodes a basic leucine zipper transcription factor. *The Plant Cell* 12: 599–609, URL <http://www.ncbi.nlm.nih.gov/pubmed/10760247>, PMID: 10760247.
- Folta KM, Pontin MA, Karlin-Neumann G, Bottini R, Spalding EP (2003) Genomic and physiological studies of early cryptochrome 1 action demonstrate roles for auxin and gibberellin in the control of hypocotyl growth by blue light. *The Plant Journal: For Cell and Molecular Biology* 36: 203–214, URL <http://www.ncbi.nlm.nih.gov/pubmed/14535885>, PMID: 14535885.
- Foster R, Izawa T, Chua NH (1994) Plant bZIP proteins gather at ACGT elements. *The FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 8: 192–200, URL <http://www.ncbi.nlm.nih.gov/pubmed/8119490>, PMID: 8119490.
- Fujita M, Fujita Y, Noutoshi Y, Takahashi F, Narusaka Y, et al. (2006) Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Current Opinion in Plant Biology* 9: 436–42, URL <http://www.ncbi.nlm.nih.gov/pubmed/16759898>, PMID: 16759898.
- Gietz RD, Woods RA (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods in Enzymology* 350: 87–96, URL <http://www.ncbi.nlm.nih.gov/pubmed/12073338>, PMID: 12073338.
- Gigolashvili T, Engqvist M, Yatusovich R, Müller C, Flügge U (2008) HAG2/MYB76 and HAG3/MYB29 exert a specific and coordinated control on the regulation of aliphatic glucosinolate biosynthesis in *Arabidopsis thaliana*. *The New Phytologist* 177: 627–642, URL <http://www.ncbi.nlm.nih.gov/pubmed/18042203>, PMID: 18042203.
- Goldman MH, Seurinck J, Marins M, Goldman GH, Mariani C (1998) A tobacco flower-specific gene encodes a polyphenol oxidase. *Plant Molecular Biology* 36: 479–485, URL <http://www.ncbi.nlm.nih.gov/pubmed/9484488>, PMID: 9484488.
- Goldsbrough AP, Albrecht H, Stratford R (1993) Salicylic acid-inducible binding of a tobacco nuclear protein to a 10 bp sequence which is highly conserved amongst stress-inducible genes. *The Plant Journal: For Cell and Molecular Biology* 3: 563–71, URL <http://www.ncbi.nlm.nih.gov/pubmed/8220463>, PMID: 8220463.

- Gonzalez A, Zhao M, Leavitt JM, Lloyd AM (2008) Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *The Plant Journal: For Cell and Molecular Biology* 53: 814–827, URL <http://www.ncbi.nlm.nih.gov/pubmed/18036197>, PMID: 18036197.
- Gourrierc JL, Li YF, Zhou DX (1999) Transcriptional activation by Arabidopsis GT-1 may be through interaction with TFIIA-TBP-TATA complex. *The Plant Journal: For Cell and Molecular Biology* 18: 663–8, URL <http://www.ncbi.nlm.nih.gov/pubmed/10417717>, PMID: 10417717.
- Grunewald W, Karimi M, Wiczorek K, de Cappelle EV, Wischnitzki E, et al. (2008) A role for AtWRKY23 in feeding site establishment of plant-parasitic nematodes. *Plant Physiology* 148: 358–368, URL <http://www.ncbi.nlm.nih.gov/pubmed/18599655>, PMID: 18599655.
- Gubler F, Kalla R, Roberts JK, Jacobsen JV (1995) Gibberellin-regulated expression of a myb gene in barley aleurone cells: evidence for Myb transactivation of a high-pI alpha-amylase gene promoter. *The Plant Cell* 7: 1879–91, URL <http://www.ncbi.nlm.nih.gov/pubmed/8535141>, PMID: 8535141.
- Guo H, Ecker JR (2004) The ethylene signaling pathway: new insights. *Current Opinion in Plant Biology* 7: 40–49, URL <http://www.ncbi.nlm.nih.gov/pubmed/14732440>, PMID: 14732440.
- Hartmann U, Sagasser M, Mehrtens F, Stracke R, Weisshaar B (2005) Differential combinatorial interactions of cis-acting elements recognized by R2R3-MYB, BZIP, and BHLH factors control light-responsive and tissue-specific activation of phenylpropanoid biosynthesis genes. *Plant Molecular Biology* 57: 155–171, URL <http://dx.doi.org/10.1007/s11103-004-6910-0>.
- Haughn G, Chaudhury A (2005) Genetic analysis of seed coat development in Arabidopsis. *Trends in Plant Science* 10: 472–477, URL <http://www.ncbi.nlm.nih.gov/pubmed/16153880>, PMID: 16153880.
- Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research* 27: 297–300, URL <http://www.ncbi.nlm.nih.gov/pubmed/9847208>, PMID: 9847208.
- Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, et al. (2008) Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes. *Advances in Bioinformatics* 2008: 5 pages, URL <http://www.hindawi.com/GetArticle.aspx?doi=10.1155/2008/420747>.
- Huang N, Sutliff TD, Litts JC, Rodriguez RL (1990) Classification and characterization of the rice alpha-amylase multigene family. *Plant Molecular Biology* 14: 655–68, URL <http://www.ncbi.nlm.nih.gov/pubmed/2102847>, PMID: 2102847.
- Hudson ME, Quail PH (2003) Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. *Plant Physiology* 133: 1605–16, URL <http://www.ncbi.nlm.nih.gov/pubmed/14681527>, PMID: 14681527.
- Iqbal M, Ahsan R, Afzal A, Jamai A, Meksem K, et al. (2008) Multigeneic QTL: The Laccase Encoded within the Soybean Rfs2/rhg1 Locus Inferred to Underlie Part of the Dual Resistance to Cyst Nematode and Sudden Death Syndrome. *Curr Issues Mol Biol* p. i11–19, URL <http://www.horizonpress.com/cimb/abstracts/v11/i11.html>.
- Ishihara H (2007) Analysis of Two Single Trait Loci Affecting Flavonol Glycoside Accumulations in Arabidopsis thaliana Natural Variations. Doctor of natural sciences, Bielefeld University, URL [http://bieson.ub.uni-bielefeld.de/volltexte/2007/1183/pdf/Hirofumi\\_Ishihara2007.pdf](http://bieson.ub.uni-bielefeld.de/volltexte/2007/1183/pdf/Hirofumi_Ishihara2007.pdf).

- Izawa T, Foster R, Nakajima M, Shimamoto K, Chua NH (1994) The rice bZIP transcriptional activator RITA-1 is highly expressed during seed development. *The Plant Cell* 6: 1277–87, URL <http://www.ncbi.nlm.nih.gov/pubmed/7919992>, PMID: 7919992.
- Jefferson RA, Kavanagh TA, Bevan MW (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO Journal* 6: 3901–3907, URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=553867&rendertype=abstract>, PMC553867.
- Jin H, Cominelli E, Bailey P, Parr A, Mehrrens F, et al. (2000) Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *EMBO J* 19: 6150–6161.
- Johnson CS, Kolevski B, Smyth DR (2002) TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of Arabidopsis, encodes a WRKY transcription factor. *The Plant Cell* 14: 1359–75, URL <http://www.ncbi.nlm.nih.gov/pubmed/12084832>, PMID: 12084832.
- Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. *Annual Review of Plant Biology* 57: 19–53, URL <http://www.ncbi.nlm.nih.gov/pubmed/16669754>, PMID: 16669754.
- Jouanneau J, Péaud-Lenoël C (1967) Croissance et synthèse des protéines de suspensions cellulaires de tabac sensibles à la kinétine. *Physiol Plant* pp. 834–850.
- Kieber JJ, Rothenberg M, Roman G, Feldmann KA, Ecker JR (1993) CTR1, a negative regulator of the ethylene response pathway in Arabidopsis, encodes a member of the raf family of protein kinases. *Cell* 72: 427–441, URL <http://www.ncbi.nlm.nih.gov/pubmed/8431946>, PMID: 8431946.
- Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, et al. (2007) The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *The Plant Journal: For Cell and Molecular Biology* 50: 347–63, URL <http://www.ncbi.nlm.nih.gov/pubmed/17376166>, PMID: 17376166.
- Kim SY, Chung HJ, Thomas TL (1997) Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the Dc3 promoter using a modified yeast one-hybrid system. *The Plant Journal: For Cell and Molecular Biology* 11: 1237–51, URL <http://www.ncbi.nlm.nih.gov/pubmed/9225465>, PMID: 9225465.
- Kirkman T (1996) Statistics to use. URL <http://www.physics.csbsju.edu/stats/>.
- Koornneef M (1990) Mutations affecting the testa colour in Arabidopsis. *Arabidopsis Information Service* pp. 1–4.
- Kosarev P, Mayer KFX, Hardtke CS (2002) Evaluation and classification of RING-finger domains encoded by the Arabidopsis genome. *Genome Biology* 3: RESEARCH0016, URL <http://www.ncbi.nlm.nih.gov/pubmed/11983057>, PMID: 11983057.
- Lepiniec L, Debeaujon I, Routaboul J, Baudry A, Pourcel L, et al. (2006) Genetics and biochemistry of seed flavonoids. *Annual Review of Plant Biology* 57: 405–30, URL <http://www.ncbi.nlm.nih.gov/pubmed/16669768>, PMID: 16669768.
- Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, et al. (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucl Acids Res* 30: 325–327, URL <http://nar.oxfordjournals.org/cgi/content/abstract/30/1/325>.

- Li J, Brader G, Kariola T, Palva ET (2006) WRKY70 modulates the selection of signaling pathways in plant defense. *The Plant Journal: For Cell and Molecular Biology* 46: 477–491, URL <http://www.ncbi.nlm.nih.gov/pubmed/16623907>, PMID: 16623907.
- Li J, Brader G, Palva ET (2004) The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. *The Plant Cell* 16: 319–331, URL <http://www.ncbi.nlm.nih.gov/pubmed/14742872>, PMID: 14742872.
- Li SF, Milliken ON, Pham H, Seyit R, Napoli R, et al. (2009) The Arabidopsis MYB5 transcription factor regulates mucilage synthesis, seed coat development, and trichome morphogenesis. *The Plant Cell* 21: 72–89, URL <http://www.ncbi.nlm.nih.gov/pubmed/19136646>, PMID: 19136646.
- Liang M, Davis E, Gardner D, Cai X, Wu Y (2006a) Involvement of AtLAC15 in lignin synthesis in seeds and in root elongation of Arabidopsis. *Planta* 224: 1185–1196, URL <http://dx.doi.org/10.1007/s00425-006-0300-6>.
- Liang M, Haraldsen V, Cai X, Wu Y (2006b) Expression of a putative laccase gene, ZmLAC1, in maize primary roots under stress. *Plant, Cell & Environment* 29: 746–753, URL <http://www.ncbi.nlm.nih.gov/pubmed/17087459>, PMID: 17087459.
- Libault M, Wan J, Czechowski T, Udvardi M, Stacey G (2007) Identification of 118 Arabidopsis transcription factor and 30 ubiquitin-ligase genes responding to chitin, a plant-defense elicitor. *Molecular Plant-Microbe Interactions: MPMI* 20: 900–911, URL <http://www.ncbi.nlm.nih.gov/pubmed/17722694>, PMID: 17722694.
- Liboz T, Bardet C, Thai ALV, Axelos M, Lescure B (1990) The four members of the gene family encoding the Arabidopsis thaliana translation elongation factor EF-1 alpha are actively transcribed. *Plant Molecular Biology* 14: 107–110, URL <http://www.ncbi.nlm.nih.gov/pubmed/2101309>, PMID: 2101309.
- Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, et al. (2000) SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. *Nature* 404: 766–770, URL <http://dx.doi.org/10.1038/35008089>.
- Lorenzo O, Chico JM, Sánchez-Serrano JJ, Solano R (2004) JASMONATE-INSENSITIVE1 encodes a MYC transcription factor essential to discriminate between different jasmonate-regulated defense responses in Arabidopsis. *The Plant Cell* 16: 1938–1950, URL <http://www.ncbi.nlm.nih.gov/pubmed/15208388>, PMID: 15208388.
- Lorenzo O, Solano R (2005) Molecular players regulating the jasmonate signalling network. *Current Opinion in Plant Biology* 8: 532–540, URL <http://www.ncbi.nlm.nih.gov/pubmed/16039901>, PMID: 16039901.
- Loudet O, Saliba-Colombani V, Camilleri C, Calenge F, Gaudon V, et al. (2007) Natural variation for sulfate content in Arabidopsis thaliana is highly controlled by APR2. *Nature Genetics* 39: 896–900, URL <http://www.ncbi.nlm.nih.gov/pubmed/17589509>, PMID: 17589509.
- Ma S, Gong Q, Bohnert HJ (2006) Dissecting salt stress pathways. *Journal of Experimental Botany* 57: 1097–1107, URL <http://www.ncbi.nlm.nih.gov/pubmed/16510518>, PMID: 16510518.
- Marles MAS, Ray H, Gruber MY (2003) New perspectives on proanthocyanidin biochemistry and molecular regulation. *Phytochemistry* 64: 367–383, URL <http://www.ncbi.nlm.nih.gov/pubmed/12943753>, PMID: 12943753.

- Marmagne A, Rouet M, Ferro M, Rolland N, Alcon C, et al. (2004) Identification of new intrinsic proteins in *Arabidopsis* plasma membrane proteome. *Molecular & Cellular Proteomics*: MCP 3: 675–691, URL <http://www.ncbi.nlm.nih.gov/pubmed/15060130>, PMID: 15060130.
- Marusek CM, Trobaugh NM, Flurkey WH, Inlow JK (2006) Comparative analysis of polyphenol oxidase from plant and fungal species. *Journal of Inorganic Biochemistry* 100: 108–123, URL <http://www.ncbi.nlm.nih.gov/pubmed/16332393>, PMID: 16332393.
- Matarasso N, Schuster S, Avni A (2005) A novel plant cysteine protease has a dual function as a regulator of 1-aminocyclopropane-1-carboxylic Acid synthase gene expression. *The Plant Cell* 17: 1205–1216, URL <http://www.ncbi.nlm.nih.gov/pubmed/15749766>, PMID: 15749766.
- Mayer AM (2006) Polyphenol oxidases in plants and fungi: going places? A review. *Phytochemistry* 67: 2318–2331, URL <http://www.ncbi.nlm.nih.gov/pubmed/16973188>, PMID: 16973188.
- Mayer AM, Staples RC (2002) Laccase: new functions for an old enzyme. *Phytochemistry* 60: 551–65, URL <http://www.ncbi.nlm.nih.gov/pubmed/12126701>, PMID: 12126701.
- McCaig BC, Meagher RB, Dean JFD (2005) Gene structure and molecular analysis of the laccase-like multicopper oxidase (LMCO) gene family in *Arabidopsis thaliana*. *Planta* 221: 619–36, URL <http://www.ncbi.nlm.nih.gov/pubmed/15940465>, PMID: 15940465.
- Mehrtens F, Kranz H, Bednarek P, Weisshaar B (2005) The *Arabidopsis* transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiology* 138: 1083–1096, URL <http://www.ncbi.nlm.nih.gov/pubmed/15923334>, PMID: 15923334.
- Meissner RC, Jin H, Cominelli E, Denekamp M, Fuertes A, et al. (1999) Function search in a large transcription factor gene family in *Arabidopsis*: assessing the potential of reverse genetics to identify insertional mutations in R2R3 MYB genes. *The Plant Cell* 11: 1827–1840, URL <http://www.ncbi.nlm.nih.gov/pubmed/10521515>, PMID: 10521515.
- Meyers BC, Lee DK, Vu TH, Tej SS, Edberg SB, et al. (2004) *Arabidopsis* MPSS. An online resource for quantitative expression analysis. *Plant Physiology* 135: 801–13, URL <http://www.ncbi.nlm.nih.gov/pubmed/15173564>, PMID: 15173564.
- Mockler TC, Michael TP, Priest HD, Shen R, Sullivan CM, et al. (2007) The DIURNAL project: DIURNAL and circadian expression profiling, model-based pattern matching, and promoter analysis. *Cold Spring Harbor Symposia on Quantitative Biology* 72: 353–63, URL <http://www.ncbi.nlm.nih.gov/pubmed/18419293>, PMID: 18419293.
- Mohanty B, Krishnan SPT, Swarup S, Bajic VB (2005) Detection and preliminary analysis of motifs in promoters of anaerobically induced genes of different plant species. *Annals of Botany* 96: 669–81, URL <http://www.ncbi.nlm.nih.gov/pubmed/16027132>, PMID: 16027132.
- Molina C, Grotewold E (2005) Genome wide analysis of *Arabidopsis* core promoters. *BMC Genomics* 6: 25, URL <http://www.biomedcentral.com/1471-2164/6/25>.
- Mullis K, Faloona F, Scharf S, Saiki R, Horn G, et al. (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology* 51 Pt 1: 263–273, URL <http://www.ncbi.nlm.nih.gov/pubmed/3472723>, PMID: 3472723.



- Mutwil M, Obro J, Willats WGT, Persson S (2008) GeneCAT—novel webtools that combine BLAST and co-expression analyses. *Nucleic Acids Research* 36: W320–6, URL <http://www.ncbi.nlm.nih.gov/pubmed/18480120>, PMID: 18480120.
- Nakano T, Suzuki K, Fujimura T, Shinshi H (2006) Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiology* 140: 411–432, URL <http://www.ncbi.nlm.nih.gov/pubmed/16407444>, PMID: 16407444.
- Nakashima K, Fujita Y, Katsura K, Maruyama K, Narusaka Y, et al. (2006) Transcriptional regulation of ABI3- and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of Arabidopsis. *Plant Molecular Biology* 60: 51–68, URL <http://www.ncbi.nlm.nih.gov/pubmed/16463099>, PMID: 16463099.
- Naoumkina MA, He X, Dixon RA (2008) Elicitor-induced transcription factors for metabolic reprogramming of secondary metabolism in *Medicago truncatula*. *BMC Plant Biology* 8: 132, URL <http://www.ncbi.nlm.nih.gov/pubmed/19102779>, PMID: 19102779.
- Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M, et al. (2000) The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of DFR and BAN genes in Arabidopsis siliques. *The Plant Cell* 12: 1863–78, URL <http://www.ncbi.nlm.nih.gov/pubmed/11041882>, PMID: 11041882.
- Nesi N, Debeaujon I, Jond C, Stewart AJ, Jenkins GI, et al. (2002) The TRANSPARENT TESTA16 locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. *The Plant Cell* 14: 2463–79, URL <http://www.ncbi.nlm.nih.gov/pubmed/12368498>, PMID: 12368498.
- Nesi N, Jond C, Debeaujon I, Caboche M, Lepiniec L (2001) The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *The Plant Cell* 13: 2099–114, URL <http://www.ncbi.nlm.nih.gov/pubmed/11549766>, PMID: 11549766.
- Nicolas J, Cheynier V, Fleuriet A, Rouet-Mayer M (1993) Polyphenols and enzymatic browning. *Polyphenolic Phenomena*, Scalbert, A. (ed.)- Paris (France): INRA: pp. 165–175.
- O'Connor TR, Dyreson C, Wyrick JJ (2005) Athena: a resource for rapid visualization and systematic analysis of Arabidopsis promoter sequences. *Bioinformatics* 21: 4411–4413, URL <http://bioinformatics.oxfordjournals.org/cgi/content/abstract/21/24/4411>.
- Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, et al. (2003) Gibberellin biosynthesis and response during Arabidopsis seed germination. *The Plant Cell* 15: 1591–604, URL <http://www.ncbi.nlm.nih.gov/pubmed/12837949>, PMID: 12837949.
- Ooka H, Satoh K, Doi K, Nagata T, Otomo Y, et al. (2003) Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes* 10: 239–247, URL <http://www.ncbi.nlm.nih.gov/pubmed/15029955>, PMID: 15029955.
- Palaniswamy SK, James S, Sun H, Lamb RS, Davuluri RV, et al. (2006) AGRIS and AtRegNet. A Platform to Link cis-Regulatory Elements and Transcription Factors into Regulatory Networks. *Plant Physiol* 140: 818–829, URL <http://www.plantphysiol.org/cgi/content/abstract/140/3/818>.

- Pandey SP, Somssich IE (2009) The role of WRKY transcription factors in plant immunity. *Plant Physiology* URL <http://www.ncbi.nlm.nih.gov/pubmed/19420325>, PMID: 19420325.
- Pauwels L, Inzé D, Goossens A (2009) Jasmonate-inducible gene: What does it mean? *Trends in Plant Science* 14: 87–91, URL <http://www.ncbi.nlm.nih.gov/pubmed/19162528>, PMID: 19162528.
- Pauwels L, Morreel K, Witte ED, Lammertyn F, Montagu MV, et al. (2008) Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured *Arabidopsis* cells. *Proceedings of the National Academy of Sciences of the United States of America* 105: 1380–1385, URL <http://www.ncbi.nlm.nih.gov/pubmed/18216250>, PMID: 18216250.
- Paz-Ares J, Consortium TR (2002) REGIA, An EU Project on Functional Genomics of Transcription Factors From *Arabidopsis thaliana*. *Comparative and Functional Genomics* 3: 102–8, URL <http://www.ncbi.nlm.nih.gov/pubmed/18628849>, PMID: 18628849.
- Pieterse CMJ, Loon LCV (2004) NPR1: the spider in the web of induced resistance signaling pathways. *Current Opinion in Plant Biology* 7: 456–464, URL <http://www.ncbi.nlm.nih.gov/pubmed/15231270>, PMID: 15231270.
- Planchais S, Perennes C, Glab N, Mironov V, Inzé D, et al. (2002) Characterization of cis-acting element involved in cell cycle phase-independent activation of *At*Arath;CycB1;1 transcription and identification of putative regulatory proteins. *Plant Molecular Biology* 50: 111–27, URL <http://www.ncbi.nlm.nih.gov/pubmed/12139003>, PMID: 12139003.
- Porter L, Hrstich L, Chan B (1986) The conversion of procyanidins and prodelphinidins to cyanidins and delphinidins. *Phytochemistry* p. 223–230.
- Pourcel L (2006) Identification du gene TT10 codant une laccase. Caracterisation de sa fonction dans l'oxydation des flavonoides de la graine chez *Arabidopsis thaliana*. These, Universite de Paris-Sud UFR Scientifique d'Orsay .
- Pourcel L, Routaboul J, Cheynier V, Lepiniec L, Debeaujon I (2007) Flavonoid oxidation in plants: from biochemical properties to physiological functions. *Trends in Plant Science* 12: 29–36, URL <http://www.ncbi.nlm.nih.gov/pubmed/17161643>, PMID: 17161643.
- Pourcel L, Routaboul J, Kerhoas L, Caboche M, Lepiniec L, et al. (2005) TRANSPARENT TESTA10 encodes a laccase-like enzyme involved in oxidative polymerization of flavonoids in *Arabidopsis* seed coat. *The Plant Cell* 17: 2966–80, URL <http://www.ncbi.nlm.nih.gov/pubmed/16243908>, PMID: 16243908.
- Prestridge DS (1991) SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. *Computer Applications in the Biosciences: CABIOS* 7: 203–6, URL <http://www.ncbi.nlm.nih.gov/pubmed/2059845>, PMID: 2059845.
- Rabinowicz PD, Palmer LE, May BP, Hemann MT, Lowe SW, et al. (2003) Genes and transposons are differentially methylated in plants, but not in mammals. *Genome Research* 13: 2658–2664, URL <http://www.ncbi.nlm.nih.gov/pubmed/14656970>, PMID: 14656970.
- Rajjou L, Belghazi M, Huguet R, Robin C, Moreau A, et al. (2006) Proteomic investigation of the effect of salicylic acid on *Arabidopsis* seed germination and establishment of early defense mechanisms. *Plant Physiology* 141: 910–23, URL <http://www.ncbi.nlm.nih.gov/pubmed/16679420>, PMID: 16679420.

- Rajjou L, Debeaujon I (2008) Seed longevity: survival and maintenance of high germination ability of dry seeds. *Comptes Rendus Biologies* 331: 796–805, URL <http://www.ncbi.nlm.nih.gov/pubmed/18926494>, PMID: 18926494.
- Rajjou L, Lovigny Y, Groot SPC, Belghazi M, Job C, et al. (2008) Proteome-wide characterization of seed aging in *Arabidopsis*: a comparison between artificial and natural aging protocols. *Plant Physiology* 148: 620–41, URL <http://www.ncbi.nlm.nih.gov/pubmed/18599647>, PMID: 18599647.
- Riva S (2006) Laccases: blue enzymes for green chemistry. *Trends in Biotechnology* 24: 219–226, URL <http://www.ncbi.nlm.nih.gov/pubmed/16574262>, PMID: 16574262.
- Rombauts S, Florquin K, Lescot M, Marchal K, Rouzé P, et al. (2003) Computational Approaches to Identify Promoters and cis-Regulatory Elements in Plant Genomes. *Plant Physiology* 132: 1162–1176, URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=167057>, PMC167057.
- Rose AB, Elfersi T, Parra G, Korf I (2008) Promoter-proximal introns in *Arabidopsis thaliana* are enriched in dispersed signals that elevate gene expression. *The Plant Cell* 20: 543–551, URL <http://www.ncbi.nlm.nih.gov/pubmed/18319396>, PMID: 18319396.
- Routaboul J, Kerhoas L, Debeaujon I, Pourcel L, Caboche M, et al. (2006) Flavonoid diversity and biosynthesis in seed of *Arabidopsis thaliana*. *Planta* 224: 96–107, URL <http://www.ncbi.nlm.nih.gov/pubmed/16395586>, PMID: 16395586.
- Ruuska SA, Girke T, Benning C, Ohlrogge JB (2002) Contrapuntal networks of gene expression during *Arabidopsis* seed filling. *The Plant Cell* 14: 1191–1206, URL <http://www.ncbi.nlm.nih.gov/pubmed/12084821>, PMID: 12084821.
- Sagasser M, Lu G, Hahlbrock K, Weisshaar B (2002) A. *thaliana* TRANSPARENT TESTA 1 is involved in seed coat development and defines the WIP subfamily of plant zinc finger proteins. *Genes & Development* 16: 138–49, URL <http://www.ncbi.nlm.nih.gov/pubmed/11782451>, PMID: 11782451.
- Sakai H, Aoyama T, Oka A (2000) *Arabidopsis* ARR1 and ARR2 response regulators operate as transcriptional activators. *The Plant Journal: For Cell and Molecular Biology* 24: 703–11, URL <http://www.ncbi.nlm.nih.gov/pubmed/11135105>, PMID: 11135105.
- Sambrook J, Russel D (2001) *Molecular cloning: a laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press .
- Samson F, Brunaud V, Duchene S, Oliveira YD, Caboche M, et al. (2004) FLAGdb++: a database for the functional analysis of the *Arabidopsis* genome. *Nucl Acids Res* 32: D347–350, URL [http://nar.oxfordjournals.org/cgi/content/abstract/32/suppl\\_1/D347](http://nar.oxfordjournals.org/cgi/content/abstract/32/suppl_1/D347).
- Santner A, Calderon-Villalobos LIA, Estelle M (2009) Plant hormones are versatile chemical regulators of plant growth. *Nature Chemical Biology* 5: 301–7, URL <http://www.ncbi.nlm.nih.gov/pubmed/19377456>, PMID: 19377456.
- Satoh R, Fujita Y, Nakashima K, Shinozaki K, Yamaguchi-Shinozaki K (2004) A novel subgroup of bZIP proteins functions as transcriptional activators in hypoosmolarity-responsive expression of the ProDH gene in *Arabidopsis*. *Plant & Cell Physiology* 45: 309–17, URL <http://www.ncbi.nlm.nih.gov/pubmed/15047879>, PMID: 15047879.

- Satoh R, Nakashima K, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2002) ACTCAT, a novel cis-acting element for proline- and hypoosmolarity-responsive expression of the ProDH gene encoding proline dehydrogenase in *Arabidopsis*. *Plant Physiology* 130: 709–19, URL <http://www.ncbi.nlm.nih.gov/pubmed/12376638>, PMID: 12376638.
- Scalbert A, Johnson IT, Saltmarsh M (2005) Polyphenols: antioxidants and beyond. *The American Journal of Clinical Nutrition* 81: 215S–217S, URL <http://www.ncbi.nlm.nih.gov/pubmed/15640483>, PMID: 15640483.
- Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, et al. (2005) A gene expression map of *Arabidopsis thaliana* development. *Nature Genetics* 37: 501–506, URL <http://www.ncbi.nlm.nih.gov/pubmed/15806101>, PMID: 15806101.
- Selmi C, Mao TK, Keen CL, Schmitz HH, Gershwin ME (2006) The anti-inflammatory properties of cocoa flavanols. *Journal of Cardiovascular Pharmacology* 47 Suppl 2: S163–171; discussion S172–176, URL <http://www.ncbi.nlm.nih.gov/pubmed/16794453>, PMID: 16794453.
- Shirley BW, Kubasek WL, Storz G, Bruggemann E, Koornneef M, et al. (1995) Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. *The Plant Journal: For Cell and Molecular Biology* 8: 659–671, URL <http://www.ncbi.nlm.nih.gov/pubmed/8528278>, PMID: 8528278.
- Simpson SD, Nakashima K, Narusaka Y, Seki M, Shinozaki K, et al. (2003) Two different novel cis-acting elements of *erd1*, a *clpA* homologous *Arabidopsis* gene function in induction by dehydration stress and dark-induced senescence. *The Plant Journal: For Cell and Molecular Biology* 33: 259–70, URL <http://www.ncbi.nlm.nih.gov/pubmed/12535340>, PMID: 12535340.
- Solano R, Nieto C, Avila J, nas LC, Diaz I, et al. (1995) Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB.Ph3) from *Petunia hybrida*. *The EMBO Journal* 14: 1773–84, URL <http://www.ncbi.nlm.nih.gov/pubmed/7737128>, PMID: 7737128.
- Somssich IE (2003) Closing another gap in the plant SAR puzzle. *Cell* 113: 815–816, URL <http://www.ncbi.nlm.nih.gov/pubmed/12837238>, PMID: 12837238.
- Stalberg K, Ellerstöm M, Ezcurra I, Ablov S, Rask L (1996) Disruption of an overlapping E-box/ABRE motif abolished high transcription of the *napA* storage-protein promoter in transgenic *Brassica napus* seeds. *Planta* 199: 515–9, URL <http://www.ncbi.nlm.nih.gov/pubmed/8818291>, PMID: 8818291.
- Staswick PE, Su W, Howell SH (1992) Methyl jasmonate inhibition of root growth and induction of a leaf protein are decreased in an *Arabidopsis thaliana* mutant. *Proceedings of the National Academy of Sciences of the United States of America* 89: 6837–6840, URL <http://www.ncbi.nlm.nih.gov/pubmed/11607311>, PMID: 11607311.
- Stracke R, Ishihara H, Huep G, Barsch A, Mehrstens F, et al. (2007) Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *The Plant Journal* 50: 660–677, URL <http://dx.doi.org/10.1111/j.1365-3113.2007.03078.x>.
- Stracke R, Werber M, Weisshaar B (2001) The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* 4: 447–456, URL <http://www.ncbi.nlm.nih.gov/pubmed/11597504>, PMID: 11597504.

- Takada S, Hibara K, Ishida T, Tasaka M (2001) The CUP-SHAPED COTYLEDON1 gene of Arabidopsis regulates shoot apical meristem formation. *Development* (Cambridge, England) 128: 1127–35, URL <http://www.ncbi.nlm.nih.gov/pubmed/11245578>, PMID: 11245578.
- Takeda J, Ito Y, Maeda K, Ozeki Y (2002) Assignment of UVB-responsive cis-element and protoplastization- (dilution-) and elicitor-responsive ones in the promoter region of a carrot phenylalanine ammonia-lyase gene (gDcPAL1). *Photochemistry and Photobiology* 76: 232–8, URL <http://www.ncbi.nlm.nih.gov/pubmed/12194222>, PMID: 12194222.
- Taylor LP, Grotewold E (2005) Flavonoids as developmental regulators. *Current Opinion in Plant Biology* 8: 317–323, URL <http://www.ncbi.nlm.nih.gov/pubmed/15860429>, PMID: 15860429.
- Terzaghi WB, Cashmore AR (2003) Light-Regulated Transcription. URL <http://arjournals.annualreviews.org/doi/abs/10.1146/annurev.pp.46.060195.002305>.
- Thain SC, Murtas G, Lynn JR, McGrath RB, Millar AJ (2002) The circadian clock that controls gene expression in Arabidopsis is tissue specific. *Plant Physiology* 130: 102–110, URL <http://www.ncbi.nlm.nih.gov/pubmed/12226490>, PMID: 12226490.
- Thipyapong P, Hunt MD, Steffens JC (2004) Antisense downregulation of polyphenol oxidase results in enhanced disease susceptibility. *Planta* 220: 105–117, URL <http://www.ncbi.nlm.nih.gov/pubmed/15300439>, PMID: 15300439.
- Thipyapong P, Steffens JC (1997) Tomato Polyphenol Oxidase (Differential Response of the Polyphenol Oxidase F Promoter to Injuries and Wound Signals). *Plant Physiology* 115: 409–418, URL <http://www.ncbi.nlm.nih.gov/pubmed/12223816>, PMID: 12223816.
- Thum KE, Kim M, Morishige DT, Eibl C, Koop HU, et al. (2001) Analysis of barley chloroplast psbD light-responsive promoter elements in transplastomic tobacco. *Plant Molecular Biology* 47: 353–66, URL <http://www.ncbi.nlm.nih.gov/pubmed/11587507>, PMID: 11587507.
- Toufighi K, Brady S, Austin R, Ly E, Provart N (2005) The Botany Array Resource: e-Northern, Expression Angling, and promoter analyses. *The Plant Journal* 43: 163, 153, URL <http://dx.doi.org/10.1111/j.1365-313X.2005.02437.x>.
- Ulker B, Somssich IE (2004) WRKY transcription factors: from DNA binding towards biological function. *Current Opinion in Plant Biology* 7: 491–498, URL <http://www.ncbi.nlm.nih.gov/pubmed/15337090>, PMID: 15337090.
- Valenzuela L, Kamakaka RT (2006) Chromatin insulators. *Annual Review of Genetics* 40: 107–138, URL <http://www.ncbi.nlm.nih.gov/pubmed/16953792>, PMID: 16953792.
- Voinnet O (2008) Post-transcriptional RNA silencing in plant-microbe interactions: a touch of robustness and versatility. *Current Opinion in Plant Biology* 11: 464–470, URL <http://www.ncbi.nlm.nih.gov/pubmed/18583181>, PMID: 18583181.
- Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. *Cell* 136: 669–687, URL <http://www.ncbi.nlm.nih.gov/pubmed/19239888>, PMID: 19239888.

- Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, et al. (1999) The TRANS-PARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. *The Plant Cell* 11: 1337–50, URL <http://www.ncbi.nlm.nih.gov/pubmed/10402433>, PMID: 10402433.
- Weigel D, Nordborg M (2005) Natural Variation in Arabidopsis. How Do We Find the Causal Genes? *Plant Physiol* 138: 567–568, URL <http://www.plantphysiol.org>.
- Weltmeier F, Ehlert A, Mayer CS, Dietrich K, Wang X, et al. (2006) Combinatorial control of Arabidopsis proline dehydrogenase transcription by specific heterodimerisation of bZIP transcription factors. *The EMBO Journal* 25: 3133–43, URL <http://www.ncbi.nlm.nih.gov/pubmed/16810321>, PMID: 16810321.
- Wenqian M, Pei H, Li J, Huimin T, Yuxian Z (2007) Characterization of three Arabidopsis AP2/EREBP family transcription factors that were involved in ABA sensitivity, freeze and salt tolerance. 18TH INTERNATIONAL CONFERENCE ON ARABIDOPSIS RESEARCH URL <http://www.arabidopsis.org/servlets/TairObject?type=publication&id=501722039>.
- Williams RJ, Spencer JPE, Rice-Evans C (2004) Flavonoids: antioxidants or signalling molecules? *Free Radical Biology & Medicine* 36: 838–849, URL <http://www.ncbi.nlm.nih.gov/pubmed/15019969>, PMID: 15019969.
- Winkel-Shirley B (2001a) Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiology* 126: 485–93, URL <http://www.ncbi.nlm.nih.gov/pubmed/11402179>, PMID: 11402179.
- Winkel-Shirley B (2001b) It takes a garden. How work on diverse plant species has contributed to an understanding of flavonoid metabolism. *Plant Physiology* 127: 1399–404, URL <http://www.ncbi.nlm.nih.gov/pubmed/11743081>, PMID: 11743081.
- Winkel-Shirley B (2002) Biosynthesis of flavonoids and effects of stress. *Current Opinion in Plant Biology* 5: 218–23, URL <http://www.ncbi.nlm.nih.gov/pubmed/11960739>, PMID: 11960739.
- Winkel-Shirley B (2006) The biosynthesis of flavonoids. In *The Science of Flavonoids*. Ed. grotewold, e., ed., Springer, 71–95 pp.
- Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, et al. (2007) An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* 2: e718, URL <http://www.ncbi.nlm.nih.gov/pubmed/17684564>, PMID: 17684564.
- Wu Z, Irizarry RA, Gentleman R, Martinez-Murillo F, Spencer F (2004) A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *Journal of the American Statistical Association* 99: 909–917, URL <http://pubs.amstat.org/doi/abs/10.1198/016214504000000683>.
- Xiao W (2006) Yeast protocols, vol. Vol. 313 of *Methods in Molecular Biology*. Edition: 2 ed., Humana Press, 392 pp.
- Xie Z, Zhang Z, Zou X, Huang J, Ruas P, et al. (2005) Annotations and functional analyses of the rice WRKY gene superfamily reveal positive and negative regulators of abscisic acid signaling in aleurone cells. *Plant Physiology* 137: 176–89, URL <http://www.ncbi.nlm.nih.gov/pubmed/15618416>, PMID: 15618416.

- Yanagisawa S (2000) Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. *The Plant Journal: For Cell and Molecular Biology* 21: 281–8, URL <http://www.ncbi.nlm.nih.gov/pubmed/10758479>, PMID: 10758479.
- Yu D, Chen C, Chen Z (2001) Evidence for an important role of WRKY DNA binding proteins in the regulation of NPR1 gene expression. *The Plant Cell* 13: 1527–40, URL <http://www.ncbi.nlm.nih.gov/pubmed/11449049>, PMID: 11449049.
- Zhai JAL (2008) Small RNA-Directed Epigenetic Natural Variation in *Arabidopsis thaliana*. *PLoS Genet* 4: e1000056, URL <http://dx.plos.org/10.1371/journal.pgen.1000056>.
- Zhang F, Gonzalez A, Zhao M, Payne CT, Lloyd A (2003) A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. *Development (Cambridge, England)* 130: 4859–4869, URL <http://www.ncbi.nlm.nih.gov/pubmed/12917293>, PMID: 12917293.
- Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, et al. (2007) Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. *PLoS Biology* 5: e129, URL <http://www.ncbi.nlm.nih.gov/pubmed/17439305>, PMID: 17439305.
- Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, et al. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* 126: 1189–1201, URL <http://www.ncbi.nlm.nih.gov/pubmed/16949657>, PMID: 16949657.
- Zhang Z, Xie Z, Zou X, Casaretto J, Ho TD, et al. (2004) A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. *Plant Physiology* 134: 1500–13, URL <http://www.ncbi.nlm.nih.gov/pubmed/15047897>, PMID: 15047897.
- Zhou J, Lee C, Zhong R, Ye Z (2009) MYB58 and MYB63 Are Transcriptional Activators of the Lignin Biosynthetic Pathway during Secondary Cell Wall Formation in Arabidopsis. *Plant Cell* 21: 248–266, URL <http://www.plantcell.org/cgi/content/abstract/21/1/248>.
- Zimmermann P, Hennig L, Gruissem W (2005) Gene-expression analysis and network discovery using Genevestigator. *Trends in Plant Science* 10: 407–9, URL <http://www.ncbi.nlm.nih.gov/pubmed/16081312>, PMID: 16081312.
- Zipfel C, Robatzek S, Navarro L, Oakeley EJ, Jones JDG, et al. (2004) Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature* 428: 764–7, URL <http://www.ncbi.nlm.nih.gov/pubmed/15085136>, PMID: 15085136.