Ecole Doctorale Environnements-Santé-STIC-E2S
UMR 1347 Agroecology INRA/ Université de Bourgogne/AgroSup Dijon

THÈSE

Pour obtenir le grade de

Docteur de l'Université de Bourgogne

Discipline: Sciences Vie

Spécialité : Physiologie moléculaire des plantes et biotechnologie

Par

Rana Muhammad ATIF

Le 17 Décembre 2012

Dissecting the factors controlling seed development in the model legume *Medicago truncatula*

Directeur de thèse
Sergio J OCHATT

Co-encadrant de thèse
Richard D THOMPSON

N. LEBORGNE-CASTEL	Professeur, Université de Bourgogne à Dijon	**Présidente**
B. SCHOEFS	Professeur, Université du Maine à Le Mans	**Rapporteur**
P. DRUART	Directeur de Recherche, CRA, Belgique	**Rapporteur**
S. OCHATT	Chargé de Recherche, INRA, Dijon	**Directeur de thése**
R. THOMPSON	Directeur de Recherche, INRA, Dijon	**Co-encadrant (Invité)**

Dedication

This thesis is dedicated to

My parents

To whom I owe the gift of my life,

Who are heaven for me in this world,

Whose prayers are always with me,

Who provided me everything, whenever I needed.

& My Murshid Molana Ilyas Attar Qadri (مدظلہ)

Acknowledgements

I gratefully acknowledge the support of all the people whose invaluable contributions helped me throughout my PhD research, and without the support of whom I would not have completed this Doctoral thesis work.

First of all, I should like to owe my sincere and profound gratitude to my honourable supervisor, **Dr. Sergio J. Ochatt** for the trust that he put into me, and for providing me an opportunity to join his lab for my Doctoral studies. It was a unique experience to work with him and his practical demonstrations in the lab are also basics of my work. Without appointments, his doors were always open for discussions and providing guidance. I also wish to express my gratitude to my respected co-supervisor, **Dr. Richard D. Thompson** for useful research suggestions throughout PhD research work. Despite his tight schedules, he always provided me substantial guidance and valuable suggestions. I thank both of them for their patience, advices, supervision, encouragement and constructive criticism that inspired me for hard work during my research and in the preparation of this thesis. Their undying efforts and experience paved my way through the difficult and tiresome path of research. Again thanks to both of you Sir.

I owe my gratitude to Dr. Christophe Salon, Director of Unit for encouraging me during my PhD research work. I would also like to thank Dr. M. Darmency, Dr. C. LeSignor, Dr. K. Gallardo and Mr. G. Aubert for helpful discussions at different steps of my research work. I also appreciate the important contributions of Mrs. M. Sanchez, Mr. L. Jacas and Mrs. C. Conreux for their significant technical assistance in experiments. Discussions with my colleagues Ms. M. Noguero, Mr. Adnan Niazi and especially Mr. Farrukh Azeem have been invaluable.

I am particularly grateful from the core of my heart to the Higher Education Commission (HEC) of Pakistan for the award of PhD scholarship and hence providing me an opportunity to conduct my doctoral studies in France. I also thank to SFERE (Paris France) for their help in the management of my scholarship and stay in France.

I would like to thank the reporters of my thesis, Dr. B. Schoefs and Dr. P. Druart for accepting and evaluating my PhD research work. I would also like to thank Dr. N. Leborgne-Castel for accepting to preside the jury of my thesis evaluation. Prof. B. Schoefs and Dr. R.

Sangwan are gratefully acknowledged for helpful discussions and advice given during the Committee Thesis Meetings.

My sincere friends have always inspired me of achieving high ideals in life with their moral support and encouraging attitudes. Especial thanks to Dr. Munawar, Dr. Iqrar, Dr. Usman, Dr. Raza, Dr. Hamid, Dr. Farhan, Dr. Shoaib and Faisal Nawaz for the motivation and guidelines in the critical phase of thesis. I would not forget the moral support of Sajjad bhai and Farasat bhai during the writing up of my thesis. I am grateful to Dr. Anees, Dr. Ashfaque, Dr. Amjad, Dr. Sabir, Dr. Ahmad, Abid, Ahsan, Shamshir and Houssem for their help and support during my stay in Dijon. Cookings of Dr. Amjad and Dr. Hamid would be remembered. I would not forget the time especially the get together events I spent with you guys. Again, I would like to thank all other friends in France and in Pakistan with whom I had a nice time.

I am grateful to all the Arakeen-e-Shura and zimadaraans of Dawateislami who inspired and helped me for higher studies. Especial thanks to Mohsin bhai, Rehan bhai, Rizwan bhais, Tahir bhai, Ahsan bhai, Syed Tahir bapu, Sajid bhais, Asghar bhai, Khushbakht bhai, Shamas bhai and other Attaries (that are not mentioned) for their support during my stay in France. I would also like to thank all the Attaries in Pakistan and foreign who helped and prayed for my success.

Last but not the least comes the most precious and important part of my acknowledgement. It is about my family members. Although I was far away from family their prayers were always with me. Thanks for their love, support, encouragement and prayers for me over the years throughout my studies. Thanks to my affectionate, admiring, friendly and loving wife, brothers and sister for their unreserved love, affection and prayers which enabled me to acquire this goal. Sweet memories of my grandfather (May his soul rest in peace with blessings of Allah), who passed away last year when I was in France are always with me in the rest of my PhD.

I found no words to express my gratitude and profound admiration to my loving parents who were always a source of inspiration and encouragement for me. I am forever indebted to my mother for her utmost care, endless patience and encouragement when it was most required. Her hand always raised in prayers for me. No one could be more important, precious, loving and kind in my life to me than my sweet mother. I cannot say anything except I love you and missed you Ammi Jaan.

Résumé

Les légumineuses sont une source riche pour l'alimentation humaine comme celle du bétail mais elles sont aussi nécessaires à une agriculture durable. Cependant, les fractions majeures des protéines de réserve dans la graine sont pauvres en acides aminés soufrés et peuvent être accompagné de facteurs antinutritionnels, ce qui affecte leur valeur nutritive. Dans ce cadre, *Medicago truncatula* est une espèce modèle pour l'étude du développement de la graine des légumineuses, et en particulier concernant la phase d'accumulation des protéines de réserve. Vu la complexité des graines de légumineuses, une connaissance approfondie de leur morphogenèse ainsi que la caractérisation des mécanismes sous-jacents au développement de l'embryon et au remplissage de la graine sont essentielles. Une étude de mutagenèse a permis d'identifier le facteur de transcription DOF1147 (DNA-binding with One Finger) appartenant à la famille Zn-finger, qui s'exprime dans l'albumen pendant la transition entre les phases d'embryogenèse et de remplissage de la graine. Lors de mon travail de thèse, il a été possible de générer plusieurs constructions pour l'analyse de l'expression de *DOF1147* ainsi que de la protéine DOF1147. Un protocole efficace pour la transformation génétique stable de *M. truncatula* a été établi et des études de localisation subcellulaire ont montré que DOF1147 est une protéine nucléaire. Un arbre phylogénétique a révélé différents groupes de facteurs de transcription DOF avec des domaines conservés dans leur séquence protéique. L'analyse du promoteur *in silico* chez plusieurs gènes-cible potentiels de DOF1147 a identifié les éléments *cis*-régulateurs de divers facteurs de transcription ainsi que des éléments répondant aux auxines (AuxREs), ce qui suggère un rôle possible de l'auxine pendant le développement de la graine. Une étude *in vitro* du développement de la graine avec divers régimes hormonaux, a montré l'effet positif de l'auxine sur la cinétique du développement de la graine, que ce soit en terme de gain de masse ou de taille, plus fort avec l'ANA que l'AIB. Grâce à une approche cytomique de ces graines en développement nous avons, en plus, démontré l'effet de l'auxine sur la mise en place de l'endoreduplication. En effet, celle-ci est l'empreinte cytogénétique de la transition entre les phases de division cellulaire et d'accumulation de substances de réserve lors du développement de la graine. Dans son ensemble, ce travail a démontré que l'auxine module la transition entre le cycle mitotique et les endocycles chez les graines en développement de *M. truncatula* en favorisant la continuité des divisions cellulaires tout en prolongeant simultanément l'endoreduplication.

Mot clés: développement de la graine, Auxine, *In vitro*, remplissage de la graine, *Medicago truncatula*, DOF, Facteur de Transcription, *In silico*, Endoréduplication, Transformation génétique, Cytométrie en flux.

Abstract

Legumes are not only indispensible for sustainable agriculture but are also a rich source of protein in food and feed for humans and animals, respectively. However, major proteins stored in legume seeds are poor in sulfur-containing amino acids, and may be accompanied by anti-nutritional factors causing low protein digestibility problems. In this regard, *Medicago truncatula* serves as a model legume to study legume seed development especially the phase of seed storage protein accumulation. As developing legume seeds are complex structures, a thorough knowledge of the morphogenesis of the seed and the characterization of regulatory mechanisms underlying the embryo development and seed filling of legumes is essential. Mutant studies have identified a DOF1147 (DNA-binding with One Finger) transcription factor belonging to the Zn-Finger family which was expressed in the endosperm at the transition period between embryogenesis and seed filling phase. During my PhD work, a number of transgene constructs were successfully generated for expression analysis of *DOF1147* gene as well as the DOF1147 protein. A successful transformation protocol was also established for stable genetic transformation of *M. truncatula*. Subcellular localization studies have demonstrated that DOF1147 is a nuclear protein. A phylogenetic tree revealed different groups of DOF transcription factors with conserved domains in their protein sequence. *In silico* promoter analysis of putative target genes of DOF1147 identified *cis*-regulatory elements of various transcription factors along with auxin responsive elements (AuxREs) suggesting a possible role of auxin during seed development. A study of *in vitro* seed development under different hormone regimes has demonstrated the positive effect of auxin on kinetics of seed development in terms of gain in seed fresh weight and size, with NAA having a stronger effect than IBA. Using the cytomic approach, we further demonstrated the effect of auxin on the onset of endoreduplication in such seeds, which is the cytogenetic imprint of the transition between the cell division phase and the accumulation of storage products phase during seed development. As a whole, this work highlighted that the auxin treatments modulate the transition between mitotic cycles and endocycles in *M. truncatula* developing seeds by favouring sustained cell divisions while simultaneously prolonging endoreduplication.

Key words: Auxin, DOF, Endoreduplication, Flow cytometry, Genetic transformation, *In vitro*, *In silico*, Legumes, *Medicago truncatula*, Seed development, Transcription factor.

Table of Contents

Chapter 1: General Introduction **1**
1.1 Legume Family 2
1.1.1 Characteristics of legumes 3
1.1.1.1 Impact on soil and environment 3
1.1.1.2 Impact as food and feed 4
1.1.2 Protein sector in Europe 5
1.1.3 Legume status in France 7
1.1.4 Constraints for lower production 7
1.1.5 Need for legume improvement 8
1.1.6 Need for a model plant 9
1.2 Medicago truncatula as a model legume 9
1.2.1 Biological and agronomic interests 10
1.2.2 Synteny with legumes 10
1.2.3 Availability of genomic and bioinformatics tools 11
1.3 Seed development 12
1.3.1 What is seed development 12
1.3.1.1 Cell division phase 13
1.3.1.2 Seed filling and maturation phase 13
1.3.1.3 Dehydration phase 14
1.3.1.4 Proteomics of seed development 15
1.4 Transcriptional control of Seed development 16
1.4.1 Plant transcription factors 16
1.4.2 Transcriptional regulation of seed development 17
1.4.3 Role of master regulators during SSP accumulation 18
1.4.4 Regulatory network of seed maturation 19
1.4.5 Repressors of maturation-related genes 20
1.4.6 DOF transcription factors 21
1.4.6.1 Structure of DOF proteins 21
1.4.6.2 Role of DOF in seed filling 22
1.5 Hormonal control of seed development 24
1.5.1 Role of Auxins 24
1.5.2 Role of Abscisic acid 27
1.5.3 Role of Gibberellins 27
1.5.4 Role of Cytokinins 28
1.6 Stable genetic transformation and functional genomics 28

1.6.1 Gene transfer methods 29
1.6.1.1 *Agrobacterium-mediated* gene transfer 29
1.6.1.1.1 Agrobacterium rhizogenes 29
1.6.1.1.2 Agrobacterium tumefaciens 29
1.6.1.2 Direct gene transfer 30
1.6.1.2.1 Electroporation 30
1.6.1.2.2 Biolistics 30
1.6.2 Gene transfer in *Medicago truncatula* 30
1.7 Flow cytometry in plant breeding 33
1.8 Rationale of the thesis 34
1.9 Scheme of work 36
Chapter 2: Materials and Methods 37
2.1 Materials 38
2.1.1 Plant material and growth conditions 38
2.1.2 Cloning vectors, bacterial strains and primers 39
2.1.2.1 Plant expression vectors 39
2.1.2.1.1 PMDC 83 vector 39
2.1.2.1.2 PMDC 107 vector 40
2.1.2.1.3 pDONR221 vector 41
2.1.2.2 *E coli* strains 41
2.1.2.3 *Agrobacterium tumefaciens* strains 42
2.1.2.3.1 LBA4404: 42
2.1.2.3.2 EHA105: 42
2.1.3 Laboratory chemicals and consumables 43
2.1.4 Solutions, media, buffers and antibiotics 43
2.1.4.1 Stock Solutions 43
2.1.4.2 Media 44
2.1.4.3 Buffers 44
2.1.4.4 Antibiotics and selection agents 45
2.2 Methods 45
2.2.1 *In silico* and bioinformatic analysis 45
2.2.1.1 Sequence data collection 45
2.2.1.2 Identification of Dof family members from *M. truncatula* genome 46
2.2.1.3 Sorting of identified DOF proteins 46
2.2.1.4 Mapping of *MtDof* genes on chromosomes 46
2.2.1.5 Comparison of conserved residues in the Dof DNA-binding domain 47
2.2.1.6 Construction of Phylogenetic tree 47

2.2.1.7 Identification of conserved domains and duplications in Dof proteins 47
2.2.1.8 Analysis of cis-regulatory elements in putative target gene promoters...... 47
2.2.1.9 *In silico* prediction of DOF1147 subcellular localization........................... 48
2.2.2 Preparation of Constructs 48
2.2.2.1 Primers designing... 49
2.2.2.1.1 Calculation of annealing temperature by temperature gradient PCR..... 49
2.2.2.1.2 Designing of AttB-PCR primers.. 49
2.2.2.1.3 Amplifying attB-PCR Products... 50
2.2.2.1.4 Agarose Gel Electrophoresis 51
2.2.2.1.5 Separation and purification of amplified attB-PCR Product 51
2.2.2.2 BP Recombination Reaction 52
2.2.2.2.1 Transformation of entry clone into E. Coli... 52
2.2.2.2.2 Entry clone verification by PCR.. 53
2.2.2.2.3 Plasmid preparation from E.coli colonies... 53
2.2.2.2.4 Sequencing of entry clones... 53
2.2.2.3 LR Recombination Reaction.. 54
2.2.2.3.1 Expression clone verification and sequencing....................................... 55
2.2.3 Manipulations with *Agrobacterium tumefaciens* 55
2.2.3.1 Preparation of competent *A. tumefaciens*.. 55
2.2.3.2 Transformation of *Agrobacteria* by heat shock 56
2.2.3.3 Transformation of electro-competent *Agrobacteria* 56
2.2.4 Stable transformation of *Medicago truncatula* 57
2.2.4.1 Preparation of *Agrobacterium* culture 57
2.2.4.2 Preparation of explants 57
2.2.4.3 Inoculation of explants.. 57
2.2.4.4 Co-cultivation of explants.. 58
2.2.4.5 Callus induction and somatic embryogenesis 58
2.2.5 Analysis of putative transformants... 59
2.2.5.1 Flow cytometric analysis of putative transformants 59
2.2.5.2 Molecular verification of putative transformants 59
2.2.5.2.1 Genomic DNA extraction.. 59
2.2.5.2.2 Verification of transformants by Polymerase Chain Reaction 60
2.2.5.3 Germination and transgene segregation test in the progeny 60
2.2.6 Transient expression assays 61
2.2.6.1 Transient expression in *Nicotiana tabacum* through agro-infiltration........ 61
2.2.7 Microscopic studies.. 61
2.2.7.1 Laser scanning confocal microscopy 61

2.2.7.2 Fluorescence microscopy ... 62

Chapter 3: Genome wide identification of DOF transcription factors in *M. truncatula* and *in silico* promoter analysis ... 63

3.1 Introduction ... 64

3.2 In-silico subcellular localization of DOF1147 ... 65

3.3 Genome wide identification of MtDof family members ... 65

3.4 New annotations of *M. truncatula Dof* genes ... 66

3.5 Mapping MtDOF genes on *M. truncatula* chromosomes ... 67

3.6 Comparison of DNA-binding Dof domains of MtDofs ... 69

3.7 Phylogenetic analysis of Dof proteins ... 70

3.8 Analysis of conserved domains and gene duplications ... 75

3.9 *In silico* motif analysis of putative target gene promoters ... 76

Chapter 4: *Agrobacterium-tumefaciens* mediated transformation of *M. truncatula* and expression studies of DOF1147 ... 81

4.1 Introduction ... 82

4.2 Results ... 83

4.2.1 Use of Gateway cloning technology for preparation of constructs ... 83

4.2.1.1 Preparation of promoter reporter construct (*pdof::gfp)* ... 83

4.2.1.2 Preparation of translational fusion construct (*p35S::dof-gfp)* ... 84

4.2.2 Transformation of expression clones into *A. tumefaciens* ... 85

4.3 Genetic transformation of *Medicago truncatula* ... 85

4.3.1 Generation of *pdof::gfp* cassette containing plants ... 88

4.3.1.1 Verification of *pdof::gfp* cassette at molecular level ... 88

4.3.2 Generation of *35S::dof-gfp* cassette containing plants ... 89

4.3.2.1 Verification of *35S::dof-gfp* cassette at molecular level ... 90

4.3.3 Comparison of transformation efficiency ... 90

4.3.4 Flow cytometric characterization of transformants ... 90

4.3.5 Phenotyping of transformants ... 92

4.3.6 Transgene segregation in the progeny ... 93

4.4 Laser scanning confocal microscopy ... 93

4.4.1 Subcellular localization of Dof1147 ... 93

4.4.2 DOF1147 subcellular localization studies using Agroinfiltration ... 95

4.4.3 Expression of *Dof1147* during seed development ... 96

Chapter 5: Role of auxin in *in vitro* seed development in *M. truncatula* ... 98

5.1 Introduction ... 99

5.2 *In vitro* auxin treatment promotes cell division and delays endoreduplication in developing seeds of the model legume species *Medicago truncatula*. ... 100

5.2.1 Abstract 100
5.2.2 Introduction 101
5.2.3 Materials and methods 103
5.2.3.1 *In vitro* culture 103
5.2.3.2 Determination of seed size, fresh weight and viability 104
5.2.3.3 Flow cytometry 105
5.2.4 Results 106
5.2.4.1 Development on auxin-free medium (MS130) 106
5.2.4.2 Effect of auxin on seed size 107
5.2.4.3 Effect of auxin on seed fresh weight 108
5.2.4.4 Flow cytometry of *in vitro* developing immature seeds 110
5.2.5 Discussion 112
Chapter 6: General Discussion 111
6.1 Construction of plant transformation vectors using Gateway technology 123
6.2 Stable genetic transformation of *M. truncatula* 124
6.3 Genome-wide identification of MtDOF TFs and their phylogenetic analysis 126
6.4 Expression of *Dof1147* in the seed endosperm 128
6.5 Subcellular localization of Dof1147 protein 129
6.6 Role of auxin in seed development 131
Conclusions and Perspectives 136
References 139

List of Figures

Figure 1. Schematic evolutionary tree of legume family. 3
Figure 2. Comparison of seed import and export by France from 1998-2011 7
Figure 3. Statistics of European Legume production 8
Figure 4. Evolutionary and taxonomic relationships of Legumes 11
Figure 5. Comparison of seed proteomes of pea, *M. truncatula* and *A. thaliana* 15
Figure 6. Frequency analyses of different families of transcription factor 17
Figure 7. Tissue specific expression of master regulators 18
Figure 8. Proposed model of regulatory network of TFs and hormones 20
Figure 9. A proposed structure of the Dof DNA Binding Domain 22
Figure 10. Model for regulation of cereal seed storage protein genes 23
Figure 11. Cell-based auxin signalling model 26
Figure 12. Detailed restriction map of Gateway vector pMDC83 39
Figure 13. Detailed restriction map of Gateway vector pMDC107 40
Figure 14. Detailed restriction map of donor vector pDONR 41
Figure 15. BP recombination reaction pathway 52
Figure 16. LR recombination reaction pathway 54
Figure 17. The localization of 22 *M. truncatula Dof* genes on chromosomes 69
Figure 18. Dof DNA-binding domain comparison 70
Figure 19. Phylogenetic relationship among Dofs of different plant species 74
Figure 20. Over-represented cis-regulatory elements in the promoters 78
Figure 21. Schematic representation of various MEME-based motifs 79
Figure 22. Schematic map of the T-DNA region of constructs 84
Figure 23. PCR verification of transformed *A. tumefaciens* colonies 85
Figure 24. Different steps of *M. truncatula* regeneration after transformation 87
Figure 25. PCR amplification of transgene fragment 89
Figure 26. Flow cytometry profiles of transgenic *M. truncatula* plants 91
Figure 27. Phenotypic comparison of pods and seeds 92
Figure 28. Subcellular localization of DOF1147 protein in *M. truncatula* leaves 94
Figure 29. Nuclear localization of DOF1147 protein in tobacco leaf epidermal cells 96
Figure 30. Cytological studied of DOF1147 expression in *M. truncatula* seed 97
Figure 31. Surface area of immature *Medicago truncatula* seeds 108
Figure 32. Fresh weight of immature *Medicago truncatula* seeds 109
Figure 33. Flow cytometry profiles of developing *Medicago truncatula* seeds 111
Figure 34. Role of Auxin in mitotic cell cycle 134

List of Tables

Table 1. Seed composition of different grain legumes and cereals 6
Table 2. Detail of antibiotics and selection agents 45
Table 3. Primers used for Gateway cloning 50
Table 4. The subcellular localization analysis of Dof1147 protein 65
Table 5. New annotation and comparison of *M. truncatula Dof* genes 68
Table 6. Multilevel consensus sequences for the MEME defined domains..................................74
Table 7. Classification and presence of identified motifs in different gene promoters.................77
Table 8. Detail of *M. truncatula* transformation with different constructs 88
Table 9. Comparison of germination and transgene segregation 93

List of abbreviations

°C	degrees Celsius
µsec	micro seconds
2,4-D	2,4-Dichlorophenoxyacetic acid
4-Cl-IAA	4 chloro Indole Acetic acid
AA(s)	Amino acid(s)
A. rhizogenes	*Agrobacterium rhizogenes*
A. thaliana	*Arabidopsis thaliana*
A. tumefaciens	*Agrobacterium tumefaciens*
ABA	Abscisic acid
ABI3	abscisic acid insensitive 3
ABRE(s)	ABA-responsive element(s)
ASIL1	arabidopsis 6b-interacting protein1-like1
AtDof	Dof proteins of *A. thaliana*
ATP	Adenosine triphosphate
AuxRE(s)	Auxin responsive element(s)
BAP	6-benzylaminopurine
bHLH	Basic helix-loop-helix
bp	Base pair
bZIP	Basic Leucine Zipper
CaMV	Cauliflower mosaic virus
CBF	CCAAT-binding factor
CDS	Coding sequence
cf.	*confrontum*
ChIP	Chromatin immunoprecipitation
CKs	Cytokinins
cm	centimeter(s)
CTAB	Cetyl trimethylammonium bromide
DAP	Days after pollination
DAPI	4, 6-diamidino-2-phenylindole
DBD	DNA binding domain
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DOF	DNA with one finder
E. coli	*Escherichia coli*

e.g.	*exempli gratia*
EB	Ethidium bromide
EDTA	Ethylenediaminetetraacetic acid
EMS	Ethyl methylsulfonate
EST(s)	Expressed sequence tag(s)
et al.	*et alia*
EU	European Union
F. wt.	Fresh weight
FDA	fluorescein diacetate
Fig.	figure
FUS3	FUSCA 3
g	gram(s)
GA	Gibberellic acid
h	hour(s)
H_2SO_4	Sulphuric acid
HCl	Hydrochloric acid
HIS	high-level expression of sugar-inducible gene
i.e.	*id est*
IAA	Indole-3yl-Acetic acid
IBA	Indole-3yl-butyric acid
kb	Kilo base pair
kDa	Kilo Dalton
KV	kilo Volt
L1L	leafy cotyledon 1 like
LB	Lysogeny broth
LEC1	leafy cotyledon 1
M	Molar
M. truncatula	*Medicago truncatula*
MALDI-TOF	Matrix-assisted laser-desorption ionization time of flight
MES	2-N-morpholinoethane sulfonic acid
mg	milli gram(s)
min	minute(s)
miRNA(s)	micro RNA(s)
mM	milli molar
mRNA	messenger RNA
MS	Murashige and Skoog basal medium

MtDof	Dof proteins of *M. truncatula*
MYB TF	Myeloblastosis transcription factor
NAA	α–Naphthaleneacetic acid
NCBI	National Center for Biotechnology Information
NLS	Nuclear localization signal
nos	nopaline synthase
O2	Opaque 2
ORF	Open reading frame
P box	Prolamin box
P. sativum	*Pisum sativum*
PBF	Prolamin binding factor
PCR	polymerase chain reaction
PsDof	Dof proteins of pea
psi	pounds per square inch
QTL	Quantitative trait loci
Ri	root inducing (plasmid)
RNA	Ribonucleic acid
rpm	Revolutions per minute
sec	seconds
SSP	Seed storage protein
SSR	Simple sequence repeat
T-DNA	transfer DNA
TE	Tris EDTA
TF(s)	Transcription factor(s)
Ti	Tumor inducing (plasmid)
TILLING	Targeting induced local lesions in genomes
UV	Ultraviolet
V/cm	Volt(s) per centimeter
vir	virulence gene
VP1	Viviparous 1
Z. mays	*Zea mays*
Zn-finger	Zinc-finger

Chapter 1: General Introduction

Introduction

Socio-Economic context

There is a growing demand for food and feed since the world population is increasing, while the area of cultivated crops remains constant. As a result of food deficits, nearly one billion people do not get enough to eat and over 700 million are chronically malnourished. Over the next 30 years, the world's population will probably double, so that the global food supply would need to double just to stay even (http://www.euroseeds.org/). The world needs fundamental changes to the global food system to feed the expanding population, and the huge challenge to balance population increase and food supply makes the need for crop improvement indispensable to fulfil these requirements.

Crop seeds are a major food source, directly or indirectly consumed by human beings as well as their domesticated animals (Grant *et al.,* 1983). Seeds from cereals (wheat, rice, maize, millet, sorghum, etc.) and legumes (pea, soybean, groundnut, chickpea, cowpea, etc.) alone provide about 70% of human food. Seed proteins directly provide more than half of the global intake of dietary proteins in humans. Because of the importance of seeds in the diet, the global seed market has doubled in size over the last couple of decades. According to the International Seed Foundation, to keep pace with population growth it is estimated that in the next 4 decades food production must equal the amount of food produced over the past 12,000 years (http://www.worldseed.org/isf/agriculture_under_pressure.html).

1.1 Legume Family

As legumes are a rich source of protein for food, feed and fodder, they can be used to cope with the challenges of rising food demands. The legume family, *Leguminoseae / Fabaceae*, comprises of over 650 genera and 18,000 species distributed in three subfamilies (Graham *et al.,* 2003). It is the third largest family of flowering plants and is second only to cereals in agricultural importance and worldwide production (Young *et al.,* 2003). Since legumes have been exploited in agricultural production since the very beginning of the civilizations, they are extremely diverse but significant components of nearly all terrestrial biomes, on all continents (except Antarctica).

Based on their flower structure, the legume family is normally divided into three sub-families i.e. i) *Papilionoideae*, ii) *Caesalpinioideae* and iii) *Mimosoideae* (Fig. 1). The

subfamily *Papilionoideae* includes the economically most important crops e.g. pea, soybean, alfalfa, chickpea etc. as shown in Fig. 1.

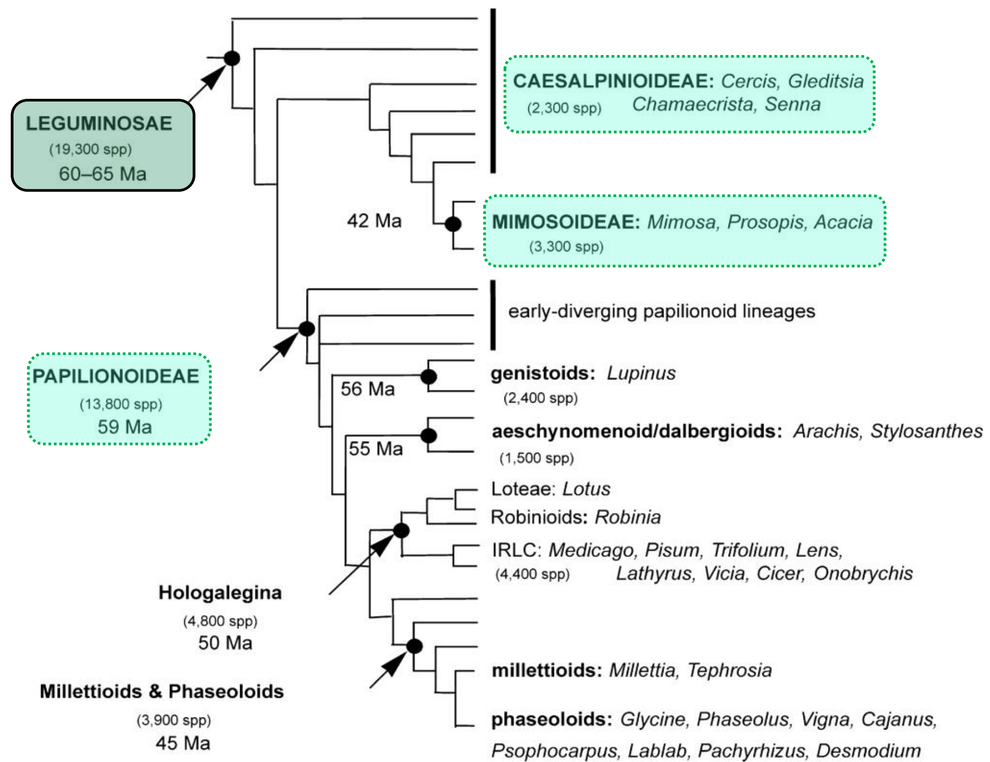


Figure 1. Schematic evolutionary tree of legume family. The three subfamilies (*Caesalpinioideae*, *Mimosoideae*, and *Papilionoideae*) shown by dotted rectangles and major subclades identified by recent molecular phylogenetic studies are shown in boldface and their positions are indicated by black circles (Adapted from Gepts *et al.,* 2005).

1.1.1 Characteristics of legumes

The family *Leguminoseae* constitutes one of the most distinct groups of land plants which are characterized by their typical flower structure, seed anatomy and their unique ability to fix and utilize the atmospheric nitrogen. Globally, the legumes have undeniable impacts serving as food and feed, as well as on soil and environment and will be described below.

1.1.1.1 Impact on soil and environment

Legumes have a unique ability to fix and utilize the atmospheric nitrogen through their root nodules and hence playing a critical role in global nitrogen cycle. Their root nodules house the symbiotic bacteria of the genus *Rhizobium* which reduce atmospheric nitrogen in

that micro-environment through enzymatic activity of nitrogenase into ammonium (NH4), a rich source of nitrogen which can readily be used by the plants (Mylona *et al.,* 1995).

This nitrogen fixation process enables legumes to produce additional proteins even in the absence of supplementary nitrogen fertilization and to survive and compete effectively in poor nitrogen conditions. The economic importance of the *Leguminosae* is likely to increase as human pressure places greater demand on marginal land. About 60% of the nitrogen in the shoots of grain legumes is considered to come from symbiotic nitrogen fixation (Schmid *et al.,* 2000). The quantity of biologically fixed N_2 each year by legumes varies greatly from zero to several hundred kg N_2/ha depending upon several factors including legume species and cultivar, soil texture and structure, pH, soil nitrate-N_2 level, temperature and water regimes, availability of other nutrients, and crop (especially harvest) management (Heichel, 1987). For instance, a crop rotation with legumes reduces energy consumption and saves greenhouse gas emissions; by breaking the annual cycle of cereals, they may enhance soil structure and improve air-water relations in the soil by improving soil organic matter content, and reduce the build-up of cereal weeds and pests and the corresponding need for pesticides (Power, 1987). Some legume species may also serve as cover crops and green manure for soil remediation and restoration and to reduce erosion. Beside these agricultural benefits, a wide range of natural products are also extracted from them to synthesize flavours, drugs, poisons and dyes.

Legumes, like many other plant species, can also associate with endomycorrhiza-forming fungi to establish a symbiotic relationship. Developed mycelium in the soil penetrates the root cells to get food in the form of sugars from the photosynthesis mechanism of plants (Gianinazzi-Pearson *et al.,* 2007). In return, the plant benefits from enhanced water, essential inorganic mineral and nutrient intake through the mycelium extending the roots' impact on plant nutrition.

1.1.1.2 Impact as food and feed

Being rich in proteins, grain legumes are an efficient source of protein for human consumption (e.g. beans, lentils, peas, chickpeas, faba beans) (Iqbal *et al.,* 2006) and animal feed as (peas, horse beans, soybeans, lupins) (Lallés, 1993) or fodder (alfalfa, clover) (Table 1). Legumes have historically been utilized mainly as whole seeds. However, in recent years, interest has grown in the utilization of legumes as flour, concentrates and isolates (Doxastakis,

2000). Some species having seeds rich in protein and oil are classified as oilseeds (soybeans, peanuts), whilst those with seeds rich in protein are classified as protein legumes (peas, faba) or pulses (beans, lentils, chickpeas).

Legumes provide about one third of all dietary protein nitrogen and one-third of processed vegetable oil for human consumption (Graham and Vance, 2003). They provide a substitute to meat, eggs and other protein sources that are high in cholesterol and saturated fats. Grain legumes contain no cholesterol and little fat, are composed of many complex carbohydrates and contain both soluble and insoluble fiber. Legumes usually contain more lysine, an essential amino acid, than cereals, but are low in cysteine and methionine (Allan *et al.,* 2000).

1.1.2 Protein sector in Europe

In Europe, about 110 million hectares of the agricultural land (1-5%) are cultivated with legumes, out of 1.5 billion hectares worldwide, where it represents 15-25% of the arable crop area. Europe is, though, one of the most productive areas for world agriculture on a per hectare basis.. In the E.U., there were about 1.124 million ha and a production of 2.793 million tons in 2011, distributed mainly in five countries: France, UK, Spain, Germany and Italy, which account for around 86% of the EU grain legume production.

Grain legumes cultivated in Europe (peas, soybeans, faba beans, lupins, chickpeas, lentils and *Phaseolus* beans) represent a strategic resource for environmentally-friendly farming, the diversification of production, and developing sources of vegetable proteins for European cattle-rearing. However, partly for historical reasons, grain legumes are underused in Europe, which imports 70% to 75% of its plant protein needs. The annual consumption of legume seed in some parts of the world is 5.4-14.4 Kg/capita while it is 2.5 kg in Europe. At the European level, about 65-85% of pea and 45-55% of faba bean seed is utilized in the animal feed industry.

Table1. Seed composition of different grain legumes and cereals

		Proximate composition (%)				Amino acids (% protein)				Vitamins (mg, %)			Minerals (mg, %)						
		Protein[a]	Fat[a]	Fibre[a]	Ash[a]	Lysine	Tryptophan	Threonine	Methionine + Cystein	Thiamine	Riboflavin	Niacin	K	P	Na	Mg	Zn	Mn	Fe
Legumes	**Pea**	24	1.1	6	3.5	7.3	0.9	3.8	2.3	-	-	-	758	302.9		84.06	3.97		3.75
	Fababean	29	1.7	9.3	4	6.5	0.8	3.6	2	0.52	0.286	2.52	1503	373.3	11.6	214.7	3.35	4.59	6.66
	Soybean	39	18.7	3.8	5.1	6.2	1.3	4	3.1	0.87	0.330	2.35	1820	477	6.9	284.5	4.48	5.43	8.66
	Chickpea	22	5.7	4	2.7	6.8	0.8	3.4	2	0.51	0.228	1.72	1044	365.7	22.7	202.7	3.54	2.14	6.23
	White lupin	38	9.5	13	3.9	4.9	0.7	3.7	2.4	-	-	-	-	-	-	-	-	-	-
	Lentils	24.7	1	4.1	2.6	-	-	-	-	0.54	0.238	2.3	970	408.5	16.6	180.7	3.51	1.31	8.07
	Peanut	22.7	44.5	2.9	2.2	-	-	-	-	0.90	0.183	15.44	786.6	460.4	34.4	268.3	5.28	2.99	5.92
Cereals	**Wheat**	14.3	1.9	3.4	1.8	0.50	0.20	0.40	0.40	9.9[b]	3.1[b]	48.3[b]	-	-	-	3740[a]	2610[e]	-	5410[a]
	Rice	12.5	3.4	2.2	2	0.30	0.10	0.30	0.30	4.62[b]	1.54[b]	48.4[b]	1.5-3.7[g]	0.35[a]	0.05[a]	0.19[a]	15.4[a]	10.80[a]	50[a]
	Corn	10.2	4.3	2.3	1.2	0.20	0.10	0.40	0.20	3.8[d]	1.4[d]	28[d]	0.32-0.72[g]	0.26-0.75[g]	0.01	1.00[g]	12-30[b]	0.7-54[b]	1-100[b]
	Barley	13	2.1	5.6	2.7	0.60	0.20	0.40	0.40	-	-	-	5070[b]	5630[b]	254[b]	410b	23.6[b]	18.9[b]	36.7[b]
	Oat	12.0	5.1	12.4	3.6	0.40	0.20	0.40	0.40	0.67[d]	0.11[d]	0.80[d]	370[g]	450[g]	4[g]	141[g]	3[g]	-	3.81[g]
	Sorghum	12.5	3.4	2.2	2	0.30	0.10	0.30	0.30	4.62[b]	1.54[b]	48.4[b]	0.38[a]	0.35[a]	0.05[a]	0.19[a]	15.4[a]	10.8[a]	50[a]

[a] Percentage, [b] µg/g, [c] U/g, [d] mg/kg, [e] µg/100g, [f] mg/100g, [g] mg/g, Modified from Harmankaya *et al.,* (2010)

1.1.3 Legume status in France

France is the leading European seed producer, the second largest exporter worldwide and its total turnover amounted to approximately 2.7 billion euros in 2010/2011 (www.gnis.fr). The French seed export is mainly restricted to the European Union (75%), Germany being the main customer. The trade balance has been positive for almost the last fifteen years and is steadily increasing (Fig. 2).

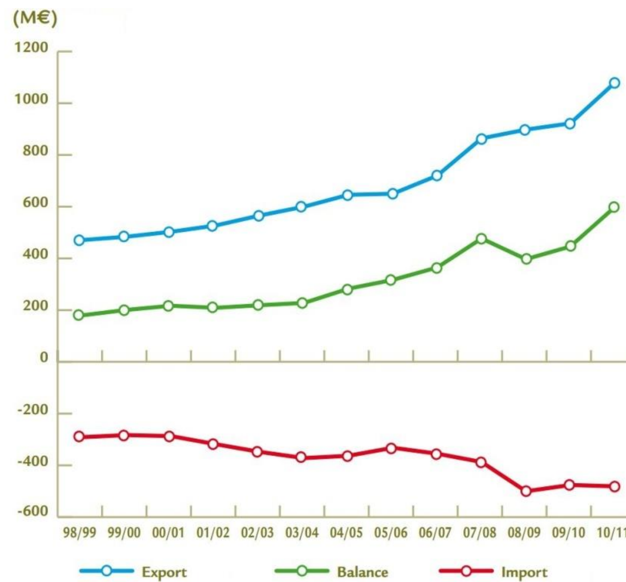


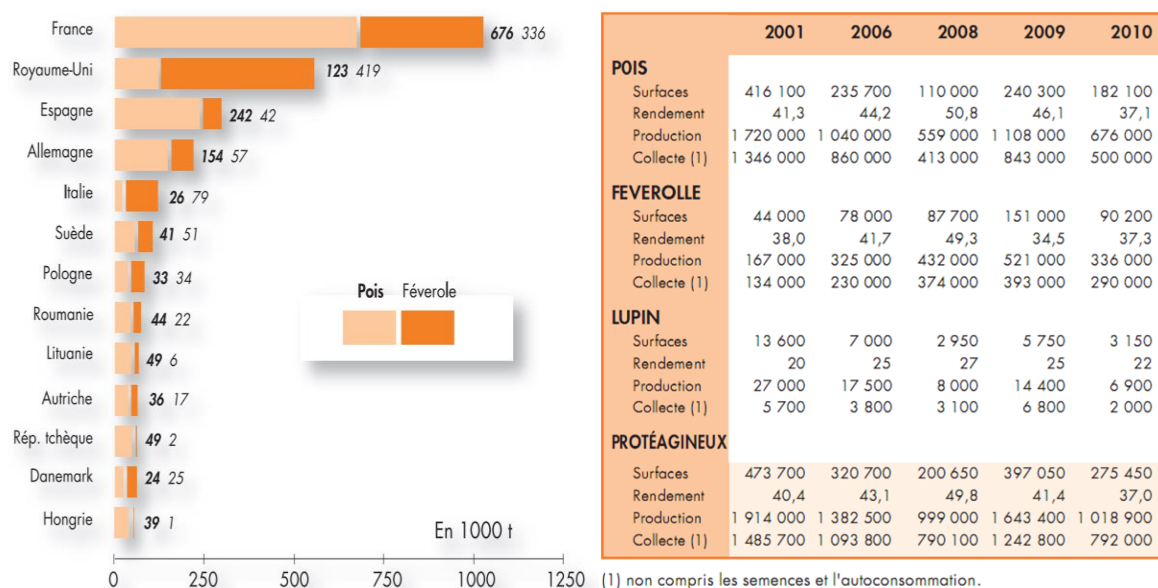
Figure 2. Comparison of seed import and export by France from 1998-2011. (www.gnis.fr)

In contrast to the generally healthy seed business in France, in recent years, the total area and production of major legumes has decreased significantly. In 2010, a 30% decrease in area and 35% decrease in production of grain legumes were recorded in France (Fig. 3B). Compared with 2010, 9% reduction in pea cultivated area was observed in 2011 at the European level. Although France is the largest producer of pea in the European Union, contributing 40-50% of European pea production, and amongst the leading producers of pea and faba bean worldwide, a dramatic decrease in pea cultivated area and production by 29% and 39%, respectively, was recorded in 2011.

1.1.4 Constraints for lower production

Grain legumes have a reputation of high yield variability, due to low tolerance to water stress and lodging in some species, late maturity for others, and variability of the seed quality. The farmers are sometimes hesitant to cultivate legumes because some pests are also

reported to be enhanced by including certain legume cultivars in a cropping system (Power, 1987). Among other reasons, the widespread availability and extensive use of synthetic nitrogen fertilizer in many countries has also been reported as a major factor for a decrease in the cultivation of legumes (Power, 1987). Moreover, farmers prefer to grow the major cash crops on fertile land while grain legumes are grown on the marginal lands, which causes low yields.



	2001	2006	2008	2009	2010
POIS					
Surfaces	416 100	235 700	110 000	240 300	182 100
Rendement	41,3	44,2	50,8	46,1	37,1
Production	1 720 000	1 040 000	559 000	1 108 000	676 000
Collecte (1)	1 346 000	860 000	413 000	843 000	500 000
FEVEROLLE					
Surfaces	44 000	78 000	87 700	151 000	90 200
Rendement	38,0	41,7	49,3	34,5	37,3
Production	167 000	325 000	432 000	521 000	336 000
Collecte (1)	134 000	230 000	374 000	393 000	290 000
LUPIN					
Surfaces	13 600	7 000	2 950	5 750	3 150
Rendement	20	25	27	25	22
Production	27 000	17 500	8 000	14 400	6 900
Collecte (1)	5 700	3 800	3 100	6 800	2 000
PROTÉAGINEUX					
Surfaces	473 700	320 700	200 650	397 050	275 450
Rendement	40,4	43,1	49,8	41,4	37,0
Production	1 914 000	1 382 500	999 000	1 643 400	1 018 900
Collecte (1)	1 485 700	1 093 800	790 100	1 242 800	792 000

(1) non compris les semences et l'autoconsommation.

Figure 3. Statistics of European Legume production. A) Top producers of pea and Faba bean in European Union, B) Area and production of major grain legumes in France (http://www.prolea.com)

1.1.5 Need for legume improvement

Despite having better lysine contents than cereals, the main legume seed storage proteins are deficient in other essential amino acids, notably, they are poor in sulphur-containing amino acids namely methionine and cysteine, which are required for the optimal growth of monogastric animals (pigs, poultry) (Allan *et al.,* 2000, Vigeolas *et al.,* 2008). Moreover, they have a low protein digestibility due to presence of certain anti-nutritional components (Gupta, 1987). For example, protease inhibitors in legume seeds, although protecting them against predator attack, have a major impact on nutritional value as they limit protein digestion by inhibiting pancreatic serine proteases (Guillamon *et al.,* 2008). Nutritionists recommend combining legumes with another protein source such as a cereal grain in the same meal, to balance the amino acid levels (Cordain, 1999). So there is a need to improve these limiting components in legumes.

1.1.6 Need for a model plant

The discovery of the existence of a strong similarity in the organization of plant genomes (syntenic conservation) within the same botanical family suggests that future progress in the study of one species genome should be helpful for other similar species (Gale and Devos, 1998). But the phylogenetic distance between plant families makes it necessary to study one specific plant, which has similar characteristics or is related to an agriculturally important crop family. Thus, rice was chosen as a model species to study cereals (Shimamoto and Kyozuka, 2002), while *Arabidopsis thaliana* was chosen as a model plant for the mustard family *Brassicaceae* (Meinke *et al.,* 1998).

However, genetic analysis of legume specific traits was not possible in non-legume models (Barker *et al.,* 1990), and the selection of a model legume plant was essential (Handberg and Stougaard, 1992). Major legumes are not suitable as models due to features such as complex ploidy levels, large genomes and/or lack of reliable and efficient regeneration/transformation systems. So two legume species, *Medicago truncatula* and *Lotus japonicus* (Barker *et al.,* 1990; Handberg and Stougaard, 1992) having relatively small genomes (450-500 Mb) (Udvardi, 2002) have become popular models for molecular and genetic studies of legumes.

1.2 Medicago truncatula as a model legume

Medicago truncatula Gaertn., a close relative of an important forage legume *Medicago sativa*, is an annual diploid (2n=16) legume species with Mediterranean origin. It is strictly autogamous in nature, with selfing rates reported in the range of 97.6–99.3% (Ellwood, *et al.,* 2006), so an increased level of homozygosity is expected within genomes. It has a small diploid genome (500-600 Mbp/1C, Blondon *et al.,* 1994; Ochatt, 2008) and comprises numerous ecotypes possessing considerable genetic variability (Bonnin *et al.,* 1996). It has a short life cycle but with prolific seed production. Successful *Agrobacterium tumefaciens*-mediated transformation followed by regeneration via somatic embryogenesis (Thomas *et al.,* 1992; Chabaud *et al.,* 1996) has also been reported in *M. truncatula*. The above described traits make *M. truncatula* a suitable model system for biological studies on various aspects of legume molecular genetics and physiology (Cook, 1999). *M. truncatula* offers several benefits at the biological, agronomic and genomic levels which are described below.

1.2.1 Biological and agronomic interests

M. truncatula, like other species and sub-species of *Medicago* is native to Mediterranean region and has a high agronomical potential (Prosperi *et al.,* 1993). It contain traits of agronomic interest, such as grazing tolerance (rooting ability and growth), resistance to drought, salinity and disease (Rubio *et al.,* 2002; Gálvez *et al.,* 2005). It is also cultivated to avoid soil erosion, improve soil fertility and as a source of winter forage (Gepts *et al.,* 2005); thus, it would be desirable to introduce these traits to alfalfa, the most widely grown forage legume crop but having a tetraploid genome (Yang *et al.,* 2008). Some novel genes associated with heavy metal toxicity, resistance and tolerance have also been identified in *M. truncatula* through QTL analysis (Narasimhamoorthy *et al.,* 2007; Chandran *et al.,* 2008 a,b). When transferred into susceptible alfalfa plants, the *RCT1* gene from *M. truncatula* has also been reported to induce broad-spectrum anthracnose resistance in transgenic alfalfa (Yang *et al.,* 2008). Quite recently, the characterization of *M. truncatula* insertional mutants led to the identification of the *SGR* gene, conferring on *M. truncatula* plants the property of staying green for a longer period of time (Zhou *et al.,* 2011). The introduction of this gene into transgenic alfalfa lines allowed them to retain more than 50% of chlorophylls during senescence and increased their crude protein content. So, *M. truncatula* can be used as resource of the above described traits that could potentially be introduced in other legumes as well for their genetic improvement.

The cotyledons of *M. truncatula* seeds are rich in protein (35-45%), the most abundant proteins being homologous to the globulin storage proteins of other legumes. It has also been successfully demonstrated that seed development and the main seed storage protein (SSP) accumulation in *M. truncatula* follows a similar pattern to that observed in other major grain legumes (Djemel *et al.,* 2005; Gallardo *et al.,* 2006) and, hence, confirms the validity of *M. truncatula* as a model plant for analysis of legume seed filling.

1.2.2 Synteny with legumes

A high degree of synteny at the genome level has been found between *Medicago* and other cool-season legumes like pea (Kalo at al. 2004), lentil, chickpea, peanut and soybean (Gepts *et al.,* 2005), contrary to *L. japonicus* that belongs to the *Loteae* and is more distant from cultivated cool season legumes (Ané *et al.,* 2008). A phylogenetic tree showing the taxonomic relationships of different clades of *Papilionoideae* subfamily is shown in Fig. 4.

A good level of genome conservation has been observed between the model legumes *Medicago truncatula, Medicago sativa* and *Lotus japonicus* (Choi *et al.,* 2004; Young *et al.,* 2011). Studies of marker transferability have demonstrated that 80% of *M. truncatula* SSR markers were applicable to alfalfa (Julier *et al.,* 2003), and 25% of SSRs gave a high degree of amplification in pea, faba bean and chickpea.

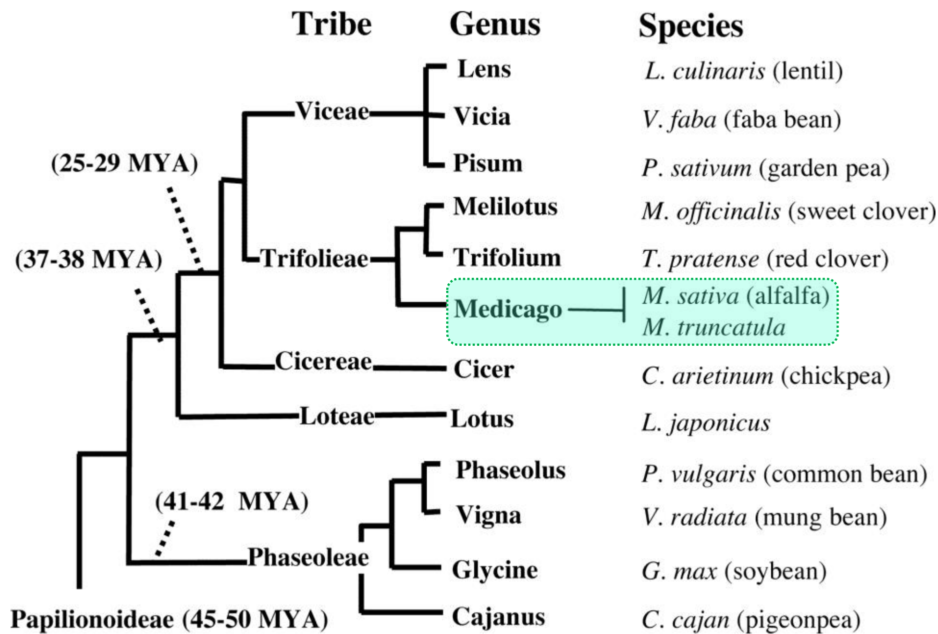


Figure 4. Evolutionary and taxonomic relationships of Legumes. Major clades of crop legume subfamily Papilionoideae. MYA:million years ago (Adapted from Choi *et al.,* 2004)

1.2.3 Availability of genomic and bioinformatics tools

A complete genome sequence is an essential pre-requisite for detailed functional genomics. During the 1990s, studies of the symbiotic interactions of *M. truncatula* with rhizobium bacteria (Gamas *et al.,* 1996) or with arbuscular mycorrhizal fungi (Liu *et al.,* 1998) resulted in the preparation of cDNA libraries. Then, sequencing was boosted by the start of the *M. truncatula* Sequencing Consortium in 2000. The fourth assembly version of *M. truncatula* genome and SNPs i.e. Mt3.5.2, were published in recently (Young *et al.,* 2011) and available at (ftp://ftp.jcvi.org/pub/data/m_truncatula/Mt3.5/Assembly/Mt3.5.2/)

The molecular markers developed for *M. truncatula* could be used in comparative genetic studies with other legumes. Comparative genetic mapping is particularly essential for species with large and complex organization. A sequence-based genetic map of *M. truncatula*,

showing polymorphism of EST-based SSR markers has been developed. Comparative genetic and cytogenetic analysis of these markers has shown marker colinearity with *M. sativa* (Choi *et al.,* 2004). So it is likely that gene order in alfalfa is analogous to that identified in the *M. truncatula* genome. Thus the knowledge achieved from comparative mapping of legumes may have considerable utility to unlock basic and applied agronomic questions in other legumes. The current availability of the mutant *M. truncatula* populations resource for reverse genetics, the TILLING (targeting induced local lesions in genomes) and *Tnt1* retrotransposon-tagged populations, are also important tools for further studies concerning legume biology and functional genomics ((Le-Signor *et al.,* 2009; Tadege *et al.,* 2008; Cheng *et al.,* 2010). A *M. truncatula* gene expression Atlas (Benedito *et al.*, 2008) was also made available at webserver (http://mtgea.noble.org/v2/) hosted by Samuel Roberts Noble Foundation, USA (He *et al.*, 2009). The web server provides the opportunity to explore gene expression data from 156 Affymetrix GeneChip® *Medicago* genome arrays in 64 different experiments.

In conclusion, the genome sequencing of *M. truncatula*, the wide availability of ESTs and the cloning of an increasing number of genes involved in various biological processes open new research corridors for ecological, evolutionary and functional genetics. This knowledge when combined with QTL mapping, will allow inspecting allelic diversity in targeted genomic regions. We are employing *M. truncatula* for studying legume seed development.

1.3 Seed development

Seeds are not only the predecessors but also the successors of plants that supply nutrients to nourish the plantlets in their juvenile stages (heterotrophic stage, before photosynthesis begins).

1.3.1 What is seed development

Generally, seeds comprise three distinct components i.e. embryo, endosperm, and seed coat. Seed development is a genetically programmed process that commences immediately after the double-fertilization, a novel event in which two haploid sperm cells from the male gametophyte enter the female gametophyte through the micropylar end via the growing pollen tube, and fertilize the haploid egg and diploid central cell respectively (Gehring *et al.,* 2004). The fertilization events give rise to the diploid embryo and triploid endosperm. The endosperm expands rapidly to occupy most of the post-fertilization embryo sac and nourishes

the embryo in the early phase of development (Gehring *et al.,* 2004). Changes in maternal tissues are also triggered after fertilization. The seed is developed from the ovary while the ovule integuments give rise to the protective seed coat. During seed development, the seed coat plays a vital role in protecting the embryo as well as transferring the nutrients from the maternal plant to the developing embryo (Borisjuk *et al.,* 2004).

In genetics language, seed development is the post-fertilization output of a montage of numerous gene expression patterns occurring in parallel in different seed tissues along with the production of various phytohormones (Le *et al.,* 2007). Developing legume seeds are complex structures containing the embryo and several other tissues including the seed coat, endosperm and suspensor. The seed development amongst all legumes follows a somewhat similar pattern after the double fertilization process (Weber *et al.,* 2005) and can be divided into three key phases i.e. cell division, seed filling and maturation, and dehydration.

1.3.1.1 Cell division phase

The Cell division phase is characterized by multiple and rapid cellular divisions and cell differentiation which establish the total number of cotyledonary cells and basic body plan of the embryo (Baud *et al.,* 2002). This phase, which is also associated with the synthesis of soluble carbohydrates and starch (Munier-Jolain and Ney, 1998; Domoney *et al.,* 2006), has significance in determining the final seed weight (Guldan and Brun, 1985). The duration of the phase of embryogenesis has proved to be essential for determination of the potential of accumulation of reserves in the body through the number of cotyledonary cells and of the maximum size they can reach (Ochatt *et al.,* 2008).

1.3.1.2 Seed filling and maturation phase

During the second, maturation, phase, embryo morphogenesis and cell cycle activities are arrested (Borisjuk *et al.,* 1995, Weber *et al.,* 1998) while the embryo goes through a state of cellular expansion and differentiation (Raz *et al.,* 2001). There is a linear increase in dry weight of seed due to an accumulation of nitrogen storage (proteins) and energy storage compounds (lipids and/or starch, depending on the legume) in great quantities in the cells and a concomitant decrease in relative water contents of the seed (Weber *et al.,* 2005; Fait *et al.,* 2006; Angeles-Núnez and Niessen, 2010). As a result, the dry weight of seed increases quickly (Weber *et al.,* 1998, 2005). In the economically and nutritionally important legumes, storage protein accumulation occurs during this phase (Gallardo *et al.,* 2003, 2007). In most

dicotyledonous species including most legumes, by maturation the endosperm is absorbed by the embryo during development and is present in only a thin layer in the mature seed (Goldberg *et al.,* 1994).

1.3.1.3 Dehydration phase

The third and last phase of seed development corresponds to the end of maturation and dehydration, during which the dry weight of the seed remains constant whereas it loses much water (West *et al.,* 1994). The gradual accumulation of sucrose and some oligosaccharides have also been reported at the end of the maturation phase (Baud *et al.,* 2002; Focks and Benning, 1998). Conversely, accumulation of starch and hexoses is only transient during seed development and dry seed contains very small amounts of them (Santos-Mendoza *et al.,* 2008). At the end of the maturation phase, the embryo becomes metabolically inactive and desiccation tolerant (Santos-Mendoza *et al.,* 2008). The duration of the phase of embryogenesis has proved to be essential for determination of the potential of accumulation of reserves in the body through the number of cotyledonary cells and the maximum size they can reach (Ochatt *et al.,* 2008).

Seed quality depends upon the strict regulation of embryo morphogenesis, maturation and germination (Gallardo *et al.,* 2008). A better comprehension of the physiology and cytology as of the genetic control of the duration of each phase of the development of the seed would make it possible to improve their quality. Many genes are active during seed development. During the studies conducted on the gene expression during embryogenesis in plants, the majority of work was carried out on the model plants *Arabidopsis thaliana* and *Medicago truncatula*. Several genes were identified, their roles were described and they were grouped in classes according to their properties. Some genes encode transcription factors.

During the last decade, many efforts by plant breeders and geneticists were directed toward the improvement of seed quality. As a result, a significant progress has been made towards a better understanding of the mechanisms of regulation of legume seed development at the genomic, proteomic, transcriptomic and post genomic level, and has been summarized in a number of overviews (Borisjuk *et al.,* 2003; Hills, 2004, Gallardo *et al.,* 2008; Thompson *et al.,* 2009). Previously in our research unit, it has been demonstrated that seed developmental transitions, including the phase of accumulation of seed storage protein are mainly regulated at the transcriptional level in the case of the model legume, *Medicago truncatula* (Gallardo *et al.,* 2007; Verdier *et al.,* 2008).

1.3.1.4 Proteomics of seed development

Gallardo *et al.* (2003) carried out a proteomic analysis in *M. truncatula* using MALDI-TOF and mass spectrometry and identified 84 proteins (Fig. 5). They described the accumulation of major seed storage proteins i.e. vicilins, legumins and convcilins after 14, 16 and 18 DAP (days after pollination) respectively during the seed filling stage of developing *M. truncatula* seeds. Previously, some of these proteins were shown to accumulate during seed development in legumes (e.g. legumins, vicilins, convicillins, and lipoxygenases; cf. Fig. 5), confirming the validity of *M. truncatula* as a model for analysis of legume seed filling. Moreover, proteins associated with cell expansion have been reported to accumulate during the reserve deposition phase, thus validating the cell expansion by accumulation of seed storage proteins at this stage (Weber *et al.,* 2005; Gallardo *et al.,* 2007). A cytological approach using *in situ* hybridization to study histo-differentiation at mid-embryogenesis in *M. truncatula* has been employed by Abirached-Darmency *et al.* (2005). The pattern of mRNA accumulation indicated that vicilin and legumin A gene expression was just beginning in 12–14 DAP seeds, validating the previous 2-D protein electrophoresis data (Gallardo *et al.,* 2003).

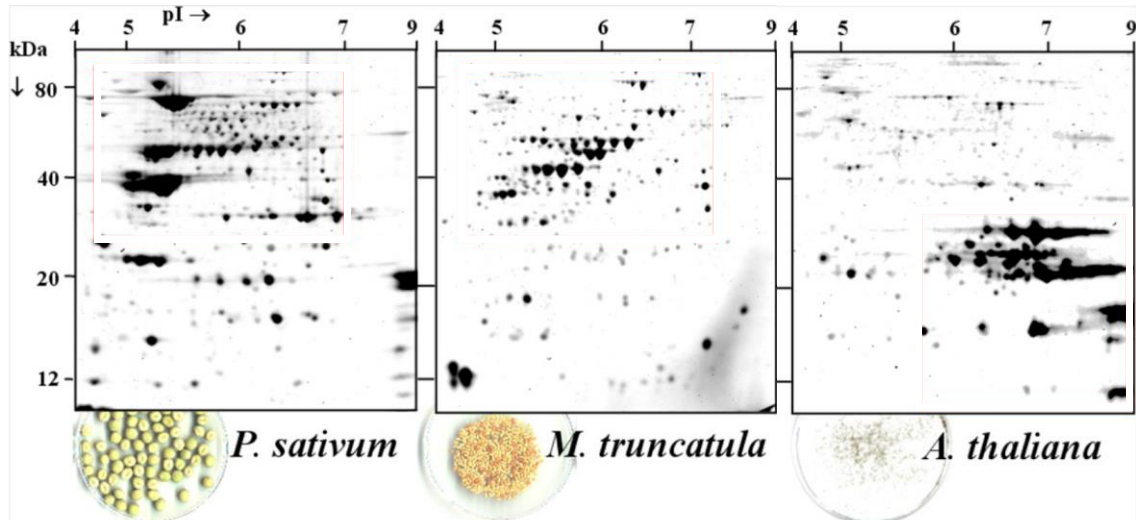


Figure 5. Comparison of seed proteomes of pea, *M. truncatula* and *A. thaliana*. The figure represents the 2-D gels of total soluble proteins separated according to size by SDS-PAGE from dry mature seeds of pea, *M. truncatula* and *A. thaliana*. (Adapted from Gallardo et al. 2008)

Previously in our lab, a study of the kinetics of accumulation of the various protein fractions of seeds and embryos, cultivated in presence and absence of nitrogen confirmed the

role of tissues surrounding the embryo in the synthesis of storage proteins (Gallardo *et al.,* 2006). This result is in agreement with those obtained by somatic embryos of *M. truncatula* by Imin *et al.* (2005), who compared the proteomes of somatic embryos from 2HA and Jemalong lines.

1.4 Transcriptional control of Seed development

1.4.1 Plant transcription factors

Transcription, a process of copying genetic information from DNA to complementary RNA is an essential and genetically regulated program in all biological life forms. Transcription in eukaryotes is coordinated by various enzymes and proteins e.g. RNA polymerase, general transcription factors, activators, co-activators, repressors etc. Transcription starts with the formation of transcription pre-initiation complex during which RNA polymerase binds to the regulatory sequence in the promoter region usually present at upstream of the start codon. It is then followed by elongation and termination steps. This flow of genetic information from DNA to mRNA is sometimes controlled by special proteins called transcription factors (TFs).

TFs perceive specific DNA sequences (often called *cis*-regulatory elements) in the promoter regions of target genes and bind through unique DNA Binding Domains, to modulate their expression through activation or repression (Riechmann and Ratcliffe, 2000). Therefore they are responsible for controlling the temporal and spatial expression patterns of target genes and hence have an imperative function in all living organisms. TFs are classified into various families on the basis of their signature DBDs, some of which (e.g. DOF) are plant-specific, possibly participating in kingdom-specific processes whilst others (e.g. MYB or bHLH) are found in all eukaryotes (Riechmann, 2002; Coulson and Ouzounis, 2003). Being sessile in nature, plant TF families are more diverse compared with fungi and animals (Shiu *et al.,* 2005) and thus appear to need a larger number of transcription factors.

Generally, 6-7 % of the genes in green plant genomes have been estimated to encode TFs (Gong *et al.,* 2004; Udvardi *et al.,* 2007). TFs play a critical role in controlling plant developmental processes so there is a great interest in unravelling their functions. But unfortunately, the functions of only a few legume TF genes are known. For instance, less than 1% of TF genes in the model legumes *M. truncatula* and *Lotus japonicus* have been

genetically characterized (Udvardi *et al.,* 2007). Therefore a tremendous effort is still required to characterize the role of TF genes in legumes.

1.4.2 Transcriptional regulation of seed development

During the last couple of decades, various genetic and molecular studies have been carried out to dissect the regulatory mechanisms involved in seed development. To date, various transcription factors have been identified in *A. thaliana* that are implicated in the control of seed development (Santos-Mendoza *et al.,* 2008; North *et al.,* 2010).

In order to get insight into legume seed development, Gallardo and colleages (2007) compared different stages of seed development at the proteomic and transcriptomic level in the three main seed tissues, i.e. seed coat, endosperm and embryo. The results revealed that seed developmental transitions, including the phase of accumulation of seed storage protein, are mainly regulated at the transcriptional level in the model legume *M. truncatula*. An overview of *M. truncatula* transcription factor genes expressed during different seed development stages was provided by Verdier *et al.* (2008) through their expression profiling. Among 169 differentially expressed TFs, 41 were associated with late embryogenesis processes, 80 were preferentially expressed during storage compound accumulation, and 48

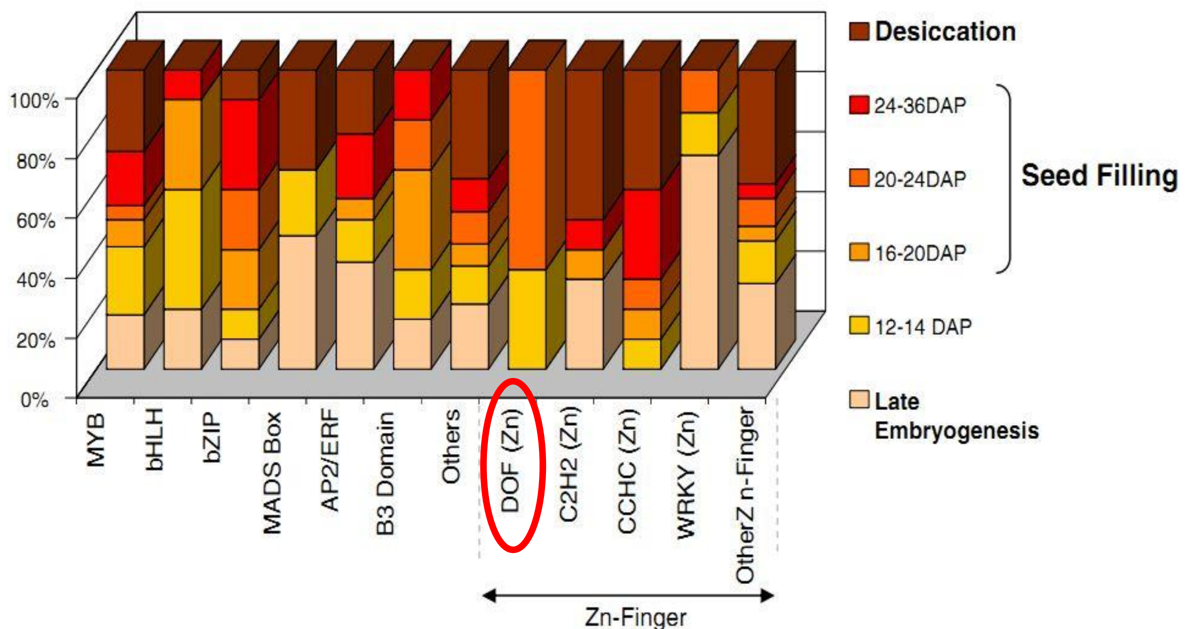


Figure 6. Frequency analyses of different families of transcription factor classes expressed at different stages of seed development in *M. truncatula*. DOF is encircled in red (Adapted from Verdier et al., 2008)

were desiccation specific (Fig. 6). These TF supervise the spatial and temporal regulation of gene expression which is a pre-requisite for flawless seed maturation.

1.4.3 Role of master regulators during SSP accumulation

Genetic and molecular studies carried out on mutants of abnormal phenotypes have revealed that four genes i.e. *LEC1 (LEAFY COT-YLEDON 1)*, *LEC2 (LEAFY COT-YLEDON 2)*, *FUS3 (FUSCA 3)* and *ABI3 (ABSCISIC ACID INSENSITIVE 3)* encoding TFs act in concert with ABA (abscisic acid) as master regulators of seed storage protein synthesis during the maturation phase of seed development (reviewed in Santos-Mendoza *et al.,* 2008; Suzuki and McCarty, 2008, North *et al.,* 2010). *ABI3*, *FUS3* and *LEC2* encode proteins that belong to the plant specific B3 TFs possessing a VP1/ABI3-like conserved B3 DNA binding domain (Luerssen *et al.,* 1998; Stone *et al.,* 2001), suggesting a similar regulatory mechanism on seed maturation specific genes. *LEC1* encodes a protein that is homologous to the HAP3 subunit of CCAAT-binding factor (CBF), a family of heteromeric transcription factors found in all eukaryotes (Edwards *et al.,* 1998; Lotan *et al.,* 1998).

	LEC2	*FUS3*	*ABI3*	*LEC1*
	pLEC2::GUS	*pFUS::FUS GFP*	ISH	ISH
	Santos *et al.* unpubl. To *et al.*, 2006	Gazzarini *et al.*, 2004	To *et al.*, 2006	Lotan *et al.*, 1998
Late heart				
Torpedo				
U shape				

Figure 7. Tissue specific expression of master regulators (Adapted from Santos-Mendoza *et al.,* 2008)

Studies carried out on *lec1*, *lec2*, *fus3* and *abi3* mutants have revealed that these TFs are particularly implicated in regulation of SSPs gene expression during seed filling phase along with other aspects of seed development such as cotyledon identity, fatty acid

metabolism, the accumulation of oil, chlorophyll and anthocyanins, and the acquisition of desiccation tolerance and dormancy (Castle *et al.,* 1994; Baumlein *et al.,* 1994; Parcy *et al.,* 1997; Baud *et al.,* 2002, 2007; To *et al.,* 2006). Protein localization and expression studies of these TFs revealed their expression in various tissues of developing seed (Fig. 7).

The B3 domain TFs have been shown *in vitro* to bind with Sph/RY motif (CATGCA) present in the promoter regions of 2S albumin and various other maturation-specific genes, and activates their expression (Reidt *et al.,* 2000; Kroj *et al.,* 2003; Monke *et al.,* 2004), while *LEC1* attach to CAAT box present upstream of the maturation-related genes (Vicente-Carbajosa and Carbonero, 2005). The promoter regions of many of *ABI3*-regulated maturation-specific genes are enriched with G-box-like and RY-like elements (Monke et al. 2012), and are hence possibly co-regulated by other B3 domain proteins i.e. FUS3 and LEC2.

1.4.4 Regulatory network of seed maturation

Gene expression studies of master regulators in *abi3*, *fus3*, *lec1*, and *lec2* single mutation event as well as the additive effects of double and triple mutations revealed a pleiotropic, overlapping and partially redundant gene regulatory network linking these master regulators with phytohormones during seed development (To *et al.,* 2006; Santos-Mendoza *et al.,* 2008), as shown in Fig. 8. *LEC2* and *FUS3* have been shown to function downstream of *LEC1* (Meinke *et al.,* 1994). While *ABI3* and *FUS3* positively regulate themselves; each other as well as both are positively regulated by *LEC1* and *LEC2* in the cotyledons (Parcy *et al.,* 1997; Kagaya *et al.,* 2005; To *et al.,* 2006).

Overexpression studies have shown that *FUS3* acts as an important link in hormone signalling pathways and controls important steps in hormone biosynthesis during seed development (Nambara *et al.,* 2000, Gazzarrini *et al.,* 2004). *FUS3* mediates ABA biosynthesis while repressing GA to avoid precocious germination of immature seeds and thus helps in dormancy acquisition (Gazzarrini *et al.,* 2004; Curaba *et al.,* 2004), . In return, ABA and GA also control the level of *FUS3* protein by interacting with the responsive elements present in the C-terminal region of *FUS3* (Lu *et al.,* 2010). *ABI3* acts in ABA-dependent manner as ABI3 protein contains ABA-responsive elements (ABRE) which are the sites for its regulation.

1.4.5 Repressors of maturation-related genes

The spatio-temporal expression of seed maturation specific genes is essential for normal seed development. Various repressors have been reported to limit and control spatio-temporal expression of these genes (Gao *et al.,* 2009). For example, *ASIL1* (*ARABIDOPSIS 6B-INTERACTING PROTEIN1-LIKE1*), a repressor binds to a GT element, adjacent to the G-box and RY repeats upstream of the 2S albumin (Gao *et al.,* 2009). The seedlings of *asil1* mutants exhibited gene expression resembling with that of late embryogenesis i.e.with *LEC1*, *LEC2* and SSP expression.

Quite recently, microRNAs (miRNAs) have been reported to be implicated in the regulation of the seed maturation program by preventing precocious gene expression (Nodine and Bartel, 2010; Willmann *et al.,* 2011). While characterizing the heterochronic embryos mutant for *DICER-LIKE1* alleles, Willmann and coworkers (2011) demonstrated that miRNAs function partly by repressing *LEC2* and *FUS3* as well as regulating *ASIL1, ASIL2* and histone deacetylase *HDA6/SIL1* genes to control the maturation program early in

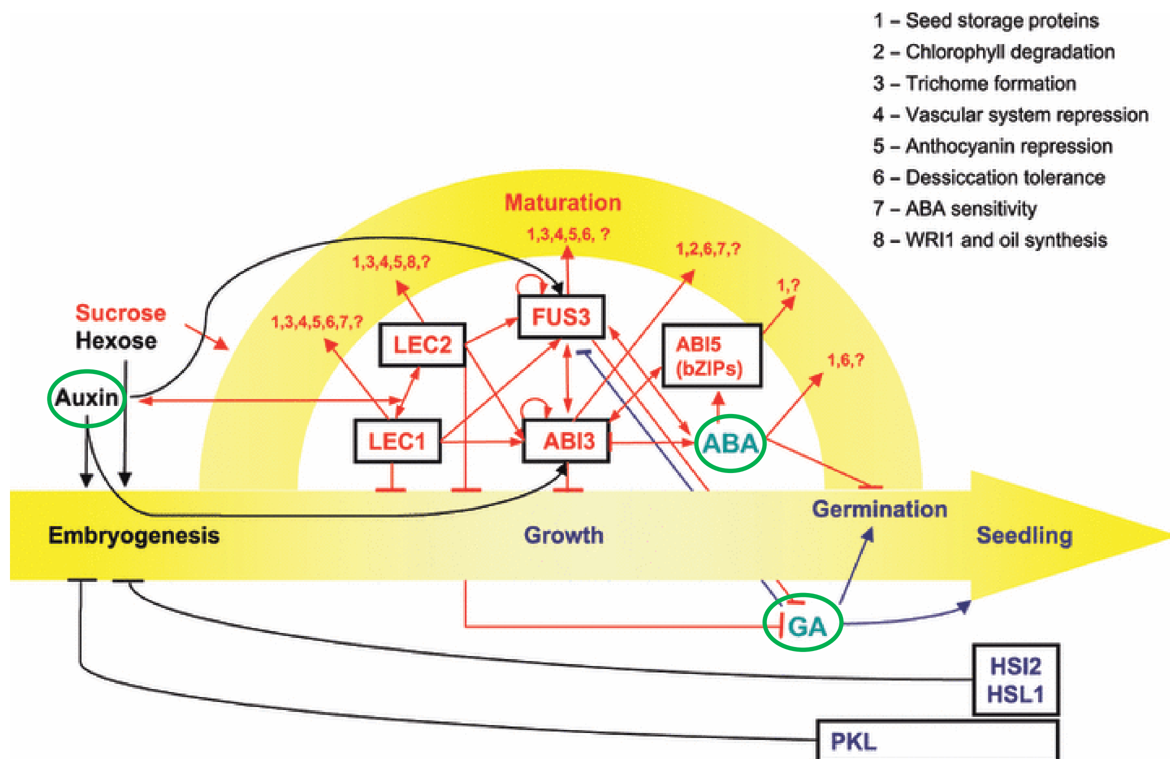


Figure 8. Proposed model of regulatory network of TFs and hormones implicated in the course of seed development in *A. thaliana*. Arrows and T bars show positive and negative regulation, respectively. Seed maturation-specific factors are shown in red; cell growth and differentiation-related factors in blue; phytohormones are encircled in green. The numbers indicate the various targets of the regulators. (Adapted from Santos-Mendoza *et al.,* 2008)

embryogenesis. *ASIL1*, *ASIL2*, and *HDA6/SIL1* have also been demonstrated to redundantly inhibit the maturation related genes during early embryogenesis (Gao *et al.,* 2009).

PICKLE (*PKL*) that encodes a protein with CHD3-chromatin remodeling factor has been speculated as mediator of expression of genes associated with embryonic identity e.g. LEC (Ogas *et al.,* 1999; Rider *et al.,* 2003). Chromatin immunoprecipitation (ChIP) studies confirmed its role as a repressor by showing the presence of PKL protein at *LEC1*, *LEC2* and *FUS3* loci during germination (Zhang *et al.,* 2012). While *HIGH-LEVEL EXPRESSION OF SUGAR-INDUCIBLE GENE 2 (HSI2)* and *HSL1*, encoding a novel class of B3 domain protein, have been demonstrated to repress the sugar-inducible expression of seed maturation-related genes (Tsukagoshi *et al.,* 2005, 2007).

1.4.6 DOF transcription factors

Previously in our unit, Verdier *et al.* (2008) carried out gene expression profiling of numerous TFs implicated during legume seed filling. Among different transcription factors identified, a TF called DOF (DNA with One Finger) belonging to Zn-Finger family was especially active during the seed filling stage of seed development (Verdier *et al.,* 2008). Its transcript is expressed in endosperm at the transition period between embryogenesis and seed filling i.e. 10-12 dap (Fig. 2) and it reaches its peak at the start of seed filling stage. Verdier and Thompson (2008) suggested DOF as possible regulator of SSP synthesis along with B3 domain, bZIP and MYB family members as DOF TFs are known to regulate storage protein accumulation in cereal endosperm in association with bZIP factors (Yamamoto *et al.,* 2006).

1.4.6.1 Structure of DOF proteins

DOF proteins are plant kingdom-specific transcription factors belonging to C2C2 Zinc finger family. They usually contain two principal domains: the N-terminal highly conserved DBD called DOF domain (Yanagisawa, 2000) and the variable C-terminal domain, which can include activation domains (Kang and Singh, 2000; Diaz *et al.,* 2002; Yanagisawa, 2002). The DOF domain consists of 52 A. acid residues including one single C2–C2 zinc finger (Umemura *et al.* 2004), so these are called DOF (DNA-binding with one finger), as depicted in Fig. 9. Although DOF proteins are strictly plant-specific, their DBD has been proposed to share a similarity to the Cys2/Cys2 Zn finger DBD of GATA1 as well as of steroid hormone receptors, in the N terminal region (Kisu *et al.,* 1995; Umemura *et al.,* 2004).

DOF proteins have been reported to function as transcriptional activator as well as repressor in distinct biological processes (Yanagisawa, 2002). DOF proteins regulate gene expression through the specific interaction of their DOF DBD with 5'-AAAG-3'/3'-TTTC-5' sequences in its target gene promoters (Yanagisawa and Schmidt, 1999). It has also been shown that four cysteine residues are necessary for the DNA binding of Dof domain (Umemura *et al.,* 2004). Various metals have been reported to obstruct the DNA binding capacity of Dof domain. Shimofurutani *et al.* (1998) demonstrated an increased binding activity when a Cys residue was altered with Ala.

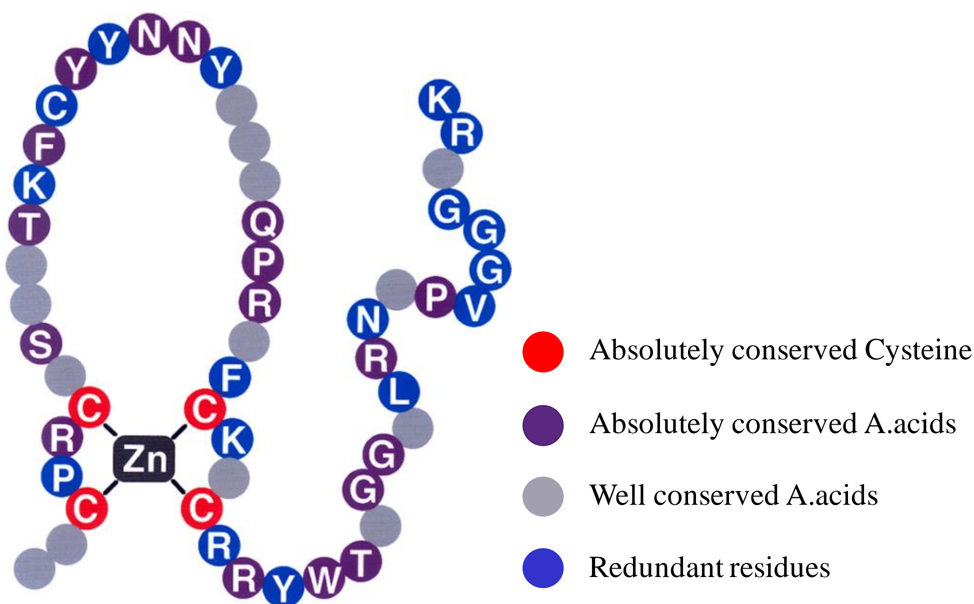


Figure 9. A proposed structure of the Dof DNA BD. Zinc is linked with four highly conserved Cysteine residues to form a finger like structure. (Modified from Yanagisawa, 2004)

1.4.6.2 Role of DOF in seed filling

DOF proteins have been shown to cooperate in the seed filling by directing the regulation of SSP genes in both mono and dicots (Mena *et al.,* 1998; Diaz *et al.,* 2005; Yamamoto *et al.,* 2006; Marzabal *et al.,* 2008). For example, in barley, A Prolamin-box Binding Factor (PBF) encoding a Dof domain protein has been showed to interact and transactivate transcription from the P-box motif in the promoter regions of SSP during seed filling phase (Mena *et al.,* 1998). Any mutation in Dof domain or in P-box elements resulted in binding failure. The maize DOF protein PBF was also reported to transactivate the SSP gene through binding with P-box (Marzábal *et al.,* 2008). DOF has also been reported to

interact with other TFs to direct the synthesis of SSP. Rice PBF, a dof protein has been shown to interact with with RISBZ1, a bZIP factor for regulation of SSP gene expression (Yamamoto *et al.,* 2006), where both the proteins act as transcriptional activators. Barley PBF and SAD (scutellum and aleurone-expressed DOF) proteins are also involved in protein-protein interaction with GAMYB TF (Diaz *et al.,* 2002; Isabel-LaMoneda *et al.,* 2003). The DOF TF along with various other TFs involved in regulation of cereal SSPs by interacting with respective *cis*-regulatory elements in promoter region are depicted in Fig. 10.

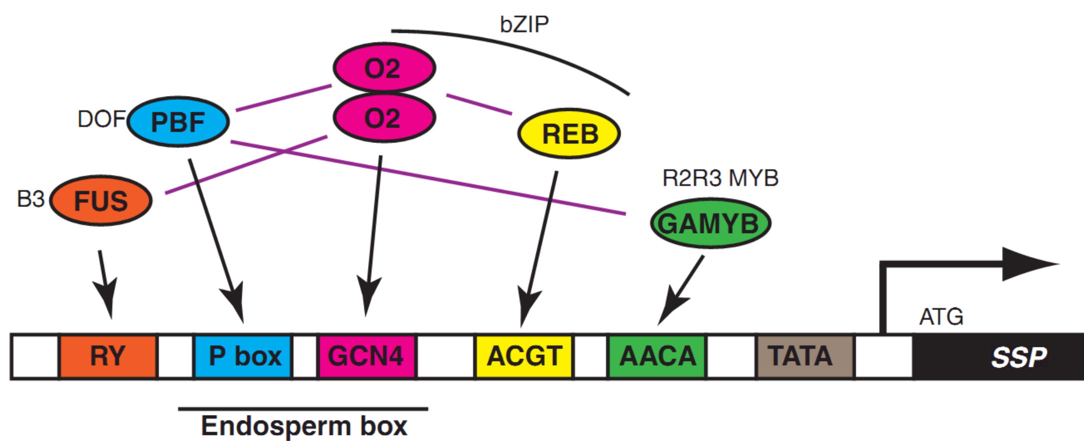


Figure 10. Transcriptional factors and *cis*-regulatory elements for cereal seed storage protein (SSP) gene regulation. Opaque2 (O2)-like basic leucine-zipper (bZip) TF binds to the GCN4 motif, REB binds to ACGT, prolamin box–binding factors (PBF) belonging to DOF family bind to a AAAG mitif in P box, GAMYB binds to the AACA motif, FUS3, B3 TF binds to RY/Sph repeats within SSP promoters. Interactions between TFs are indicated by purple lines. (Adapted from Kawakatsu and Takaiwa, 2010)

Maize Dof2 has been involved in the repression of C4-type Phosphoenolpyuvate Carboxylase (PEPC) promoter (Yanagisawa and Sheen, 1998). Similarly BPBF acts as transcriptional repressor to inhibit transcription from the cathespsin B-like protease gene promoter (Mena *et al.,* 2002). Apart from regulating the expression of SSP genes, DOF proteins have been implicated in diverse plant processes such as seed germination (Mena *et al.,* 2002, Papi *et al.,* 2002, Isabel-LaMoneda *et al.,* 2003), carbon metabolism (Yanagisawa, 2000), secondary metabolism (Skirycz *et al.,* 2006), light mediated gene regulation (Yanagisawa and Sheen, 1998, Shaw *et al.,* 2009), vascular tissue development (Konishi and Yanagisawa, 2007; Guo *et al.,* 2009), legume-rhizobium symbiosis (Kouchi *et al.,* 2004; Asamizu *et al.,* 2008), carbohydrate metabolism (Tanaka *et al.,* 2009) and cell cycle regulation (Skirycz *et al.,* 2008).

1.5 Hormonal control of seed development

During seed development, the role of phytohormones is primordial. They are not only direct regulators of various developmental processes, but are also implicated in the transcriptional control of genes by interacting with various TFs.

1.5.1 Role of Auxins

Indole-3-acetic acid (IAA) is the predominant naturally occuring auxin. It is the best studied phytohormone which is implicated virtually in every aspect of plant growth and development, including seed development (Woodward and Bartel, 2005). Synthesis of IAA in plants has been observed preferentially through a tryptophan (Trp)-dependent pathway, while a Trp-independent pathway has also been reported (Zazimalova and Napier, 2003). As auxins are produced in the proximity of the apical meristems, they play a paramount role during embryo development (Jenik and Barton, 2005; Weijers and Jurgens, 2005). Under the influence of auxin, the diploid zygote goes through a series of cell divisions and differentiations resulting in the formation of the mature embryo (Steinmann *et al.,* 1999; Friml *et al.,* 2003). Auxin interacts with expansin proteins to disrupt hydrogen bonds between adjacent cell wall molecules, resulting in an enhancement of cell expansion (Perrot-Rechenmann, 2010). Quite recently, MAB2 (MACCHI-BOU 2) protein has been suggested to play a vital role in auxin-mediated responses during embryogenesis (Ito *et al.,* 2011), as the embryos of *mab2* mutants showed defective cell divisions leading to abnormal cotyledon formation.

Auxin exerts its physiological effects through transcriptional regulation (Jenik and Barton, 2005; Yamamoto *et al.,* 2010). The spatio-temporal presence of auxin in varying concentrations in different seed tissues is responsible for the “switch on” and “switch off” of genes require for normal embryo development. In *lec* mutants, very few auxin-related phenotypes, i.e. with abnormal development of the suspensor and reduced length of the axis, in early embryogenesis have been observed (Lotan *et al.,* 1998; Stone *et al.,* 2001, 2008). The B3-domain proteins, in contrast, have been shown to take active part in auxin biosynthesis and auxin-mediated gene expression. *FUS3* and *ABI3* have been shown to exhibit an auxin induced expression in loss-of-function mutants (Brady *et al.,* 2003; Gazzarrini *et al.,* 2004). Later on, a *Phaseolus vulgaris ABI3*-like factor was shown through *in vitro* binding assay to interact with auxin responsive elements (AuxREs) present in the auxin responsive genes (Nag

et al., 2005) suggesting that *ABI3* acts downstream to auxin. The expression of auxin-induced genes is greatly increased in plants ectopically expressing *LEC2* (Braybrook *et al.,* 2006). *FUS3* and *LEC2* are also likely to enhance auxin biosunthesis via auxin-biosynthesis genes *YUC4* and *YUC10* during seed development (Yamamoto *et al.,* 2010). Along with *LEC1*, auxin and sucrose have been speculated to play a fundamental role for embryonic cell division and subsequent differentiation (Casson and Lindsey, 2006). Auxin may also modulate the levels of two other phyto-hormones, GA and ABA (Gazzarrini *et al.,* 2004).

Along with cytokinin, auxin has also been shown to play a key role in the induction of cell division and control of cell-cycle progression (Perrot- Rechenmann, 2010). A normal mitotic cell cycle is characterized by theoccurrence of four dictinct phases i.e. S, G1, M and G2 phase that lead to the formation of two daughter cells (Inzé and Veylder, 2006, Perrot-Rechenmann, 2010). On the other hand, during a variant cell cycle often called endocycle the DNA contents of the cell multiplies while the cell does not divide, leading to endoreduplication (Ochatt, 2008). Expression of various mitotic cycle genes has been shown to be downreguled in low auxin levels (Ishida *et al.,* 2010) which in turn reduces the mitotic activity and enhances an entry into endoreduplication, promoting cell expansion (Ishida *et al.,* 2010). On the other hand, increasing auxin levels have been shown to inhibit and delay endocycles, thus prolonging the mitotic phase (Ishida *et al.,* 2010).

Auxin-binding protein 1 (ABP1), PINFORMED protein (PIN) and AUXIN-permeases (AUX) acts as an auxin receptor in plasmalema (Fig. 11). Inside the cell, auxin signaling is accompanied by auxin-mediated transcriptional regulation. AUXIN RESPONSE FACTORS (ARFs), and AUXIN/IAA (AUX/IAAs) are both needed for auxin transcriptional responses. Auxin is involved in the regulation of ARF activity (Ulmasov, 1997), ARFs being DNA-binding transcriptional activators. ARFs identify and bind to AuxREs present upstream of auxin responsive genes (Ulmasov *et al.,* 1999 a,b; Hagen and Guilfoyle, 2002).

The AUX/IAAs proteins are shown to repress the transcriptional activity of ARFs in the absence of auxin (Reed, 2001; Tiwari *et al.,* 2001). The presence of auxin causes removal of AUX/IAAs from their complex with ARF, resulting in regulation of auxin responsive genes by ARFs (Zenser *et al.,* 2001). At the cellular level, a ubiquitination pathway has been reported to play a vital role in auxin signalling. In *A. thaliana*, transport inhibitor response protein 1 (TIR1), an F-box protein, has been identified as a positive regulator of auxin

signalling (Ruegger *et al.,* 1998). It encodes a subunit of Skp1-Cullin1-F-box protein (SCF) ubiquitin ligase complex.

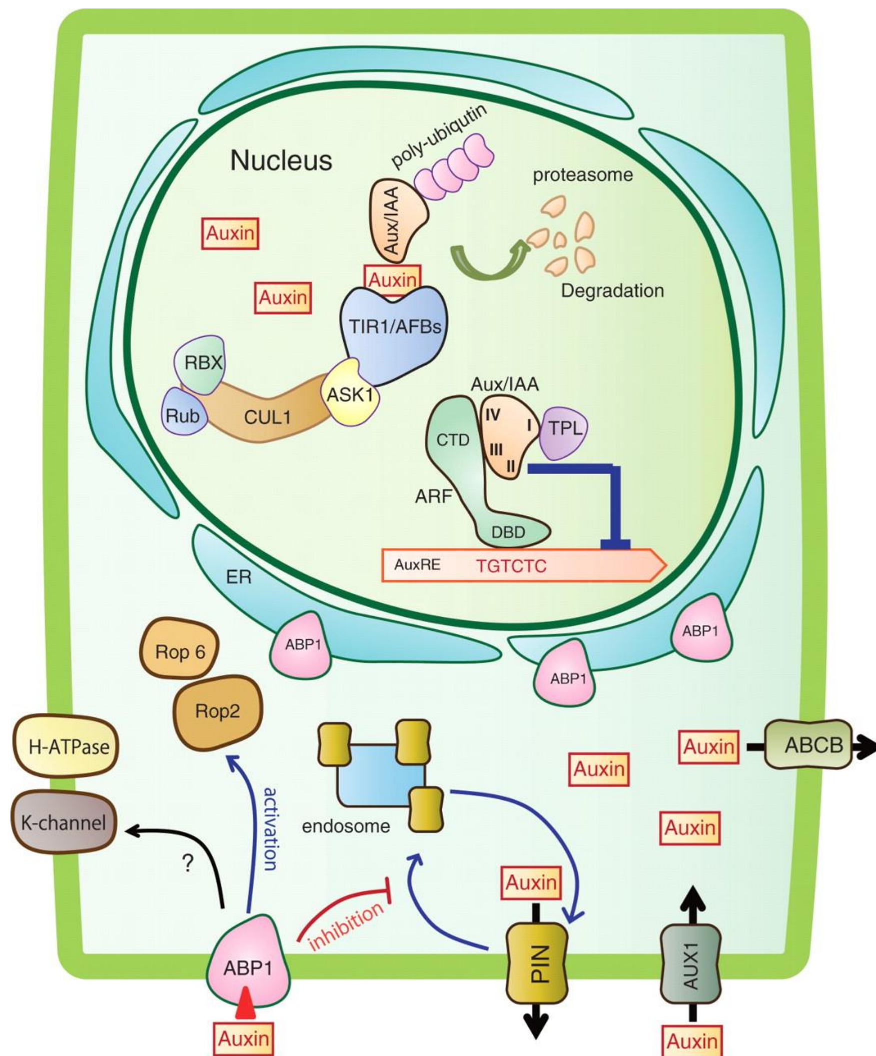


Figure 11. Cell-based auxin signalling model. A ubiquitin ligase complex SCF$^{TIR1/AFB}$ comprising of TIR1/AFB auxin receptor and SCF subunit, is directed to catalyze the ubiquitination of Aux/IAAs transcriptional repressors in the presence of auxin. At the same time, the activity of the ARF TFs is inhibited by Aux/IAA repressors which are also bound to TOPLESS (TPL) proteins. Auxin is normally bound in the small cavity formed between TIR1 proteins and domain II of Aux/IAA repressors. When required, auxin causes the degradation of Aux/IAA repressors in the nucleus through ubiquitinous pathway, resulting in the recovery of ARF TF. ARF TF is then intact with the auxin-reponse elements (AuxRE: TGTCTC) and thus activates the transcription of auxin-responsive genes. ABP1 is basically an endoplasmic reticulum-based protein that acts as an auxin receptor in the plasmalema. The ABP1 signal activates K^+ channels and H-ATP pumps in the cell membrane to induce rapid auxin responses, such as turgor-induced growth. ABP1 is also involved in the auxin-mediated signalling process to block the endocytosis of PIN proteins from the plasma membrane. (Adapted from Hayashi, 2012)

Upon perception of auxin, TIR1 was shown to directly bind with AUX/Ivvs by its conserved domain, causing its SCF^{TIR1}-catalyzed ubiquitination and degradation in the proteosomes (Ramos *et al.,* 2001), and thus releasing ARF protein to activate transcription from its target promoters (Fig. 11). Later on, TIR1 was described to be an auxin receptor (Dharmasiri *et al.,* 2005 a,b; Kepinski and Leyser, 2005). Auxin has also been reported to directly modulate transcription of some AUX/IAAs, a negative feedback loop to regulate auxin activity (Liscum and Reed, 2002; Goda *et al.,* 2004). The auxin signalling pathway has also been shown to be influenced by microRNAs, which control expression of many auxin responsive genes (Jenik and Barton, 2005). The gene families encoding microRNAs, either downregulate (*TIF*, *ARF*) or upregulate (*CUC*) the auxin signal by changing the balance of distinct mRNAs in the auxin pathway (Reinhart *et al.,* 2002; Mallory and Vaucheret, 2004; Mallory *et al.,* 2005).

1.5.2 Role of Abscisic acid

ABA is involved in modulation of a wide spectrum of seed developmental phases. This hormone is indispensable for the progress of seed maturation by preventing precocious germination and reserve mobilization. During early embryogenesis, ABA synthesized from seed maternal tissues (Karssen *et al.,* 1983), promotes embryo growth and avoids its abortion during seed development (Frey *et al.,* 2004). ABA action is via ABA-response-elements (ABRE) present in promoter regions of auxin-induced genes as well as in the promoters of genes belonging to B3 domain family (Ezcurra *et al.,* 1999; Finkelstein and Gibson, 2002). ABA has been implicated in the regulation of SSP genes by interacting with TFs of B3 domain family, LEC1, bZIP and WRKY (Raz *et al.,* 2001; Lara *et al.,* 2003; Nambara and Marion-Poll, 2003, Frey *et al.,* 2004; Zou *et al.,* 2004). Studies on mutants, impaired in ABA synthesis have revealed low SSP accumulation and defective seed phenotypes with non-dormant behaviour (Nambara *et al.,* 1992, To *et al.,* 2006; Seiler *et al.,* 2011).

1.5.3 Role of Gibberellins

Gibberellins (GAs) also play important roles during seed development. A significant decrease in seed fresh and dry weight has been observed in GA-deficient *gib1* mutants (Groot *et al.,* 1987) while Swain *et al.,* (1993, 1997) reported arrest in embryo growth leading to seed abortion in pea *lh-2* mutants having mutations at a gibberellin biosynthesis locus. More recently, seed abortion has also been described in plants overexpressing a GA–inactivating gene (Singh *et al.,* 2002, 2010), demonstrating the importance of active GAs in the endosperm

for normal seed development. Studies have shown that GA biosynthesis during seed development is transcriptionally regulated by various TFs including LEC2, FUS3, MADS (Curaba *et al.,* 2004; Gazzarrini *et al.,* 2004; Zheng *et al.,* 2009).

1.5.4 Role of Cytokinins

During early embryogenesis, elevated levels of CKs have been reported when the embryo undergoes multiple cellular divisions (Morris, 1997), which is consistent with very recent findings of Matsuo *et al.* (2012), who observed the highest transcript level of CK biosynthesis genes in the early embryogenesis. Moreover, the immunolocalization studies of CKs carried out in developing caryopses of maize have shown that CKs are localizes in specific cell types of endosperm and embryo (Rijavec *et al.,* 2011). These reports are in line with the hypothesis that CKs increase the seed sink strength for assimilates transfer (Emery *et al.,* 2000). Nordström and colleagues (2004) have demonstrated that auxin is involved in rapid suppression of cytokinin biosynthesis in the later phases. Various families of TFs are involved in CK signaling pathway during seed development e.g. ARRs (Arabidopsis type-B Response Regulators) and CRFs (cytokinin response factors) etc (Rashotte *et al.,* 2006; Müller and Sheen, 2008). Characterization of loss-of-function mutants demonstrated that CRFs are involved in regulation of development of embryos and cotyledons in a redundant manner (Rashotte *et al.,* 2006; Hwang *et al.,* 2012).

1.6 Stable genetic transformation and functional genomics

Plant transformation may be defined as the sequence of delivery, integration and expression of foreign genes into the plant cells which will ultimately regenerate into a whole plant. This ability to introduce and express or inactivate specific genes in the plant genomes provides a new and powerful experimental tool for validating gene function, particularly in relation with various plant physiology mechanisms and processes that have not been resolved so far using other biochemical approaches. Another non-negligible application of this approach is that of obtaining and transferring genes of novel traits that are not available to a given species due to sexual incompatibility from other plants, from microorganisms or even animals.

1.6.1 Gene transfer methods

Gene delivery systems used to date can be divided into direct gene transfer (mediated by physical or chemical forces for delivery of the gene into plant protoplasts, cells and even tissues) and *Agrobacterium-mediated* gene transfer, where either *A. tumefaciens* or A. rhizogenes are used as vectors for introducing the foreign gene into the plant genome.

1.6.1.1 *Agrobacterium-mediated* gene transfer

In *Agrobacterium*-mediated transformation, the natural infecting capacity of the bacterium is used to transfect genes into plant cells (Chilton, 2001; Somers *et al.,* 2003, Sangwan *et al.,* 2010), and this approach has become a tool widely used in plant breeding for crop improvement. One significant limitation of the *Agrobacterium* gene transfer system is the fact that large groups of commercially important plants and some legumes among them are not hosts for *Agrobacterium* whereby this gene transfer system is not efficient for them. Two strains of *Agrobacterium* are usually used for transformation i.e. *Agrobacterium rhizogenes* and *Agrobacterium tumefaciens.*

1.6.1.1.1 Agrobacterium rhizogenes

A. rhizogenes contains a root inducing (Ri) plasmid that leads to development of transgenic hairy roots from transformed tissues. This system was first described in legumes by Jensen *et al.* (1986) and Petit *et al.* (1987) when they transformed *Lotus corniculatus* to produce transformed nodules. This method is usually employed to study root specific genes, symbiosis related genes as well as for gene overexpression studies (Hansen *et al.,* 1989).

1.6.1.1.2 Agrobacterium tumefaciens

Agrobacterium tumefaciens-mediated transformation is the preferred method to introduce genes into numerous dicotyledonous and some monocotyledonous plants (Komari *et al.,* 1996). A critical step for successful transformation with *A. tumefaciens* is the co-cultivation of plant cells and bacteria to allow gene transfer, generally done by mixing plant cells/tissues with bacteria *in-vitro* for a few days, after which bacteria are removed and cells or organs are regenerated to plants employing adequate culture media. Alternatively, whole plants are dipped in Agrobacterium and subsequently allowed to grow under natural conditions, as in the floral dip transformation method of *Arabidopsis thaliana* (Clough and Bent, 1998). This, unfortunately, has only been applied to a few plant species to date

including the model legume *M. truncatula* (Trieu *et al.,* 2000), and was not reproducible for this species.

1.6.1.2 Direct gene transfer

Direct DNA transfer through physical or chemical means provides an possible alternative to *Agrobacterium,* for example it is the only way to introduce genes into the chloroplast genome (Clarke *et al.,* 2011). Examples of direct gene transfer methods are biolistics and electroporation, but also polyethylene glycol (PEG) treatment.

1.6.1.2.1 Electroporation

Electroporation is a simple and rapid method where plant cells and DNA are put together in a solution and an electric stimulus is used to transfer DNA into the plant cells. When a cell is exposed to an electric field, pores are formed in the plasmalemma through an enhancement of its trans-membrane potential (Cole, 1968; Neumann and Rosenheck, 1973). The pores then allow DNA to enter the cell and nucleus. The method has been used to introduce genes into protoplasts isolated from a range of different species and seems to be a universal method of gene transfer into prokaryotic and eukaryotic cells. Conversely, there are only quite a few examples of electroporation applied for stable gene transfer in other legume species (Ochatt *et al.,* 2005).

1.6.1.2.2 Biolistics

In biolistics, small metal particles are coated with the sequences of interest, and are shot into plant cells. In legumes, the transgene delivery using biolistic approach for successful stable transformation was first demonstrated with soybean (McCabe *et al.,* 1988; Christou *et al.,* 1989). This approach has since been exploited on several occasions in this crop, as reviewed by Dickins *et al.,* (2003). Alternatively, an integrated transformation method by using particle bombardment in combination with *Agrobacterium*-mediated approach (Droste *et al.,* 2000) was also tested.

1.6.2 Gene transfer in *Medicago truncatula*

Successful genetic transformation of *Medicago truncatula* was reported for the first time 20 years ago (Thomas *et al.,* 1992). Various transformation protocols have been developed taking advantage of the superior *in vitro* regeneration and transformation characteristics of a few genotypes, among which Jemalong 2HA, M9-10a and R108-1.

Transformation protocols can be classified according to the nature of the regeneration process: either via embryogenesis or organogenesis. The most efficient protocols currently available use embryogenesis. However, the efficiency of somatic embryogenesis is highly genotype-dependent. Thomas *et al.* (1992) and Rose *et al.* (1999) selected and used genotype 2HA, which belongs to *M. truncatula* ssp. *truncatula*, the same as genotype M9-10a used by several authors (Santos and Fevereiro, 2002; Araujo *et al.,* 2004; Confalonieri *et al.,* 2010), while Trinh *et al.* (1998) worked with genotype R108, sometimes assigned to *M. truncatula* ssp. *tricycla* (Chabaud *et al.,* 2007). Being the genotype used for genome sequencing and for mutagenesis and TILLING programs, Jemalong is also the genotype of choice when requiring genetic crossing of transgenic plants with characterized mutants. However, it is less readily transformed than R108-1.

Hoffmann *et al.* (1997), Trinh *et al.* (1998), Scholte *et al.* (2002), among others, obtained barrel medic transformants using leaves as explants and subsequently regenerating the produced callus into plants through somatic embryogenesis. An efficient transformation procedure for barrel medic has also been developed by Kamate *et al.* (2000), where flower parts were used as explants followed by regeneration through embryogenesis, while Trieu and Harrison (1996), aimed at reducing the tissue culture work involved, developed a new method based on cotyledonary node explants, followed by regeneration of multiple shoots from the pre-existing meristems. Sometime later, Trieu *et al.* (2000) reported two *in planta* procedures for transformation, based on infiltration of flowers and seedlings. Lacking the laborious tissue culture work, these approaches were considered useful for T-DNA insertion analysis. But unfortunately they were unable to reproduce them later on, nor were these results corroborated by other laboratories (Somers *et al.,* 2003).

Recently, genetically modified plants transformed by conventional transformation methods have generated much debate due to the presence of selectable marker genes e.g. herbicide/antibiotic resistance genes that can cause certain environmental problems and concerns in terms of food safety. Therefore, the development of a vector-free genetic transformation system of legumes and hence avoiding the presence of antibiotic resistance genes is very desirable, in order to not only mitigate such issues but also spare the high-cost and time-consuming studies for risk assessment. For the first time in legumes, Scaramelli and colleagues (2009) employed *ipt*-type MAT vector system, based on the use of morphological markers to produce marker-free transgenic *M. truncatula* plants using floral organs as explant source. Although they were able to get gus expressing *ipt*-shooty transgenic lines, typical

characteristic of this transformation system (Ebinuma *et al.,* 2001), they were unable to regenerate *ipt* marker-free transgenic plants. Further molecular investigation carried out revealed the occurrence of DNA rearrangements but yeast Recombinase was unable to excise the undesirable sequences out of genome.

Besides the selectable marker issue, the backbone DNA or T-DNA of transformation vectors being a bacterial/viral in origin is also a great deal of concern as it integrates into the plant genome along with the gene of interest after successful transformation events (Kononov *et al.,* 1997). While in other cases, the integration of complete binary vector sequences, in addition to the T-DNA region into the genomes of transformed plants has also been reported (van der Graaff *et al.,* 1996; Abdal-Aziz *et al.,* 2006; Permyakova *et al.,* 2009). In legumes, various approaches have been exercised to reduce the frequency of transgenic plants containing vector backbone. Ye *et al.* (2008) incorporated four backbone frequency reduction genes that inhibit the development of transgenic plants with integrated vector backbone, and hence permitting the development of only those transgenic plants that lack the backbone. So it is more desirable to limit the amount of this foreign DNA (in the form of T-DNA or plasmid backbone) by transforming plants with genetic elements that are derived from plant origin. The very first demonstration of feasibility of a plant-DNA-based vector system in legumes came from the work of Confalonieri and co-workers (2010), who successfully transformed *M. truncatula* leaf explants, following a protocol from Araujo *et al.* (2004), using plant-based transfer DNA (P-DNA) in the vector backbone. They actually used *M. truncatula*-derived elements in the vector, having sequence homology with the bacterial T-DNA borders (Rommens *et al.,* 2004, 2005). The plant-based right and left borders in the vector were able to efficiently mobilize and integrate the P-DNA into *M. truncatula* genome. This innovative technology can now be safely used in other legumes as well.

Since *M. truncatula* has been proposed as a model especially to study the symbiotic relationship (Barker *et al.,* 1990), development of transgenic hairy roots using *Agrobacterium rhizogenes*, containing root inducing (Ri) plasmid would be preferred to study the endosymbiotic related genes. The first report of use of *A. rhizogenes*-mediated transformation for the production of transgenic *M. truncatula* roots came from the work of Thomas *et al.* (1992), but they were unable to regenerate fertile plants from the root-based callus. Later on, Boisson-Dernier and colleagues (2001) demonstrated the successful production of transgenic *M. truncatula* hairy roots and used this system to explore the symbiotic root colonization system by endomycorrhizal fungi. This transformation system proved its worth in various

studies of root-specific interactions including RNAi-based approaches to study gene functions in root biology (Limpens *et al.,* 2004) and mycorrhizal colonization (Kiirika *et al.,* 2012), functional promoter analyses (Boisson-Dernier *et al.,* 2005), dissection of genes involved in symbiotic relationship and harbouring pathogenic nematodes (Damiani *et al.,* 2012), abiotic stress response of genes in roots (de Zelicourt *et al.,* 2012), root genes associated with heavy metals resistance and tolerance (Chandran *et al.,* 2008 a,b). The silencing of certain genes in *M. truncatula* hairy roots also helped to uncover the signaling pathways of phytohormones during root nodule formation and their symbiotic interactions with *Sinorhizobium meliloti* (Gonzalez-Rizzo *et al.,* 2006, Wasson *et al.,* 2006).

1.7 Flow cytometry in plant breeding

It has previously been discussed that for a successful plant transformation, the transgene sequence delivery, integration and its expression into the plant cells leading to regenerated plants is essential. In the same way, it is also important to characterize the transformed tissues for their regeneration competence, and the primary transformants as well as subsequent progeny at the phenotypic and genotypic level in terms of true-to-typeness and ploidy level (Ochatt *et al.,* 2000, 2010). In recent years, the use of flow cytometry has emerged as an important plus time and money saving tool, not only for such characterizations but also for measurements of relative DNA contents, the extent of endoreduplication and genome size of cell suspension cultures, haploids, somatic hybrids and insertional mutants (Ochatt *et al.,* 2005; Dolezel *et al.,* 2007). Quite recently, it has also been shown as a reliable tool for the determination of transition from cell division phase to seed filling phase, and the onset of endoreduplication with *in vitro* cultured immature developing seeds (Atif *et al.,* 2013) and embryos (Ochatt, 2011). For this purpose, the nuclei from the source tissues are mechanically isolated by chopping in nuclei extraction buffer and then stained with the fluorescent dye DAPI (2-diphenylindole) in the staining buffer. DAPI binds stoichiometrically to the A-T DNA bases and, when passing through the flow cytometer, the nuclei emit an epi-fluorescence upon excitation with a UV laser beam, and the data is recorded by an in-built computer programme of the cytometer that transforms light into a curve, where the intensity of the emitted epi-fluorescence (plotted against the nucleus count) indicates the relative nuclear DNA content (Ochatt, 2008).

1.8 Rationale of the thesis

At the start of this project, various genes encoding putative transcription factors had already been identified through their expression profiling involved in the seed development in the model legume *Medicago truncatula* (Verdier *et al.,* 2008). The expression profiles of a total of 712 TFs helped in the identification of 169 TF genes that were preferentially expressed during different phases of seed development. Among them, the 41 genes expressed at 10 DAP (late embryogenesis) contain principally members of Zn-Finger family, *WRKY* family, *AP2/ERFs*, MADS and MYB factors. The 80 TF genes expressed during the transition and whole seed filling phase contain possible orthologues of regulators of seed storage protein synthesis in other species including B3 domain, bZIP, DOF and MYB family members. Similarly, most of the TFs genes expressed at 36 DAP (Dessication phase) belong to the Zn-Finger and MYB families.

Among these different transcription factors identified, a TF called DOF1147 (DNA with One Finger) belonging to Zn-Finger family was especially active during the seed filling stage of seed development. Its transcript starts expressing at the transition phase, i.e. 10-12 DAP, and it reaches its peak at the mid seed filling stage. As DOF1147 has been expressed in endosperm of developing *M. truncatula* seed, it might control gene expression in mid-term endosperm development as related DOF TFs in cereal endosperm already had the reputation to regulate storage protein synthesis in association with bZIP factors (Yamamoto *et al.,* 2006). The expression of *SSP* genes is regarded as a marker for maturation, and a gradient of storage protein gene expression is generally detected in developing cotyledons or endosperm.

In order to verify this hypothesis of the involvement of *DOF1147* in seed filling, mutants for this gene were isolated from a mutagenized population of *M. truncatula*, created using EMS by TILLING (targeting induced local lesions in genomes), a reverse genetics tool (Le Signor *et al.,* 2009). Among them, mutant EMS109 having stop codon in the DOF DNA binding domain has also been identified. The EMS109 mutant plants exhibit seed abortion in the cell division phase of seed development. These plants only set seeds at the end of their life cycle, and such seeds are significantly fewer in number and smaller in size and hence exhibit a significant decrease in 1000 grain weight as compared to wild types. Upon germination, the majority of the mutant embryos exhibit abnormal and fused cotyledons resulting in seedling mortality. Screening a *Tnt1*-retrotransposon insertional mutant population of *M. truncatula* (Tadege *et al.,* 2008) also resulted in the identification of other mutant alleles of *dof1147*

which are being characterized in our research team. When treated with exogenous auxin, the *dof1147* plants yield better than untreated controls, hence indicating a partial compensation of the *dof1147* mutant defects by auxin (pers.comm. C. Le Signor).

Our research team is investigating several possibilities for obtaining a confirmatory *DOF1147* mutant allele.. The scientific interest of this work is twofold. First, we plan to use the transgenesis approach for *DOF* to confirm or refute the results obtained so far and, secondly, to analyse in more depth the mechanism of action of phytohormones during seed development.

Specifically the objectives of the study were to:

1. Conduct expression studies of promoter and coding sequence of *DOF1147* in *M. truncatula* using a set of transgene constructions
2. Study the effect of phytohormones during *in vitro* seed development of *M. truncatula*

The purpose of this research was to functionally dissect the factors controlling seed filling in legumes and to add to the basic knowledge and understanding of seed development in legumes. This research project was carried out in the UMR1347 Agroécologie, pole GEAPSI, in the team “Developpement et Qualite de la Graine” and lab of “Physiologie Cellulaire, Morphogenèse et Validation (PCMV)”, which hosted me, and focuses on the development and exploitation of in vitro biotechnology approaches for legume breeding, including gene transfer in legumes, somatic hybridisation, the production of double-haploids and insertion mutants, followed by their characterization in terms of phenotype, cell and tissue morphometry, flow cytometry, isoenzymes and molecular markers.

1.9 Scheme of work

The results of this thesis work can be summarized in three parts, shown below.

The first part of this work (Chapter 3) deals with the detailed *in silico* analysis for the genome wide identification of DOF TF family members in *M. truncatula*. A phylogenetic comparison of DOF family members from different plant species was carried out and the evolutionary background of conserved domains among those proteins is also described. Then the promoter sequences of putative target genes of DOF1147 TF were subjected for identification of putative *cis*-regulatory elements targets of transcriptional and hormonal interactions during *M. truncatula* seed development.

The second part (Chapter 4) of my research work deals with the expression studies of *DOF1147* gene as well as its protein in *M. truncatula* using a set of transgene constructions. Different steps for the preparation of transgene constructs using Gateway technology are presented. A detailed protocol for stable transformation of *M. truncatula* with these constructs is described. The microscopic observations of tissues/cells expressing the transgenes are recorded.

The third and last part (Chapter 5) of my thesis research work describes the experiments carried out to investigate the effect of auxins during *in vitro* seed development of *M. truncatula*. For this purpose, an *in vitro* seed culture methodology was employed and the kinetics of seed development was measured during the transition period from the embryo cell division phase to seed maturation phase in *M. truncatula*. Cytomic studies of seed development are also presented. Then the role of auxin especially in cell division, expansion and differentiation during seed development is discussed.

The Higher Education Commission of Pakistan awarded me a fellowship to complete my Ph.D. studies in France while the research grant for this thesis work was partly funded by Eurostar grant PEASTAR E!4770.

Chapter 2: Materials and Methods

Materials and Methods

2.1 Materials

2.1.1 Plant material and growth conditions

- ***Medicago truncatula* genotype A-17:** It is a single-seed descendant line from the cultivar Jemalong. Although it is a genotype difficult to regenerate and transform, it has been used as the primary reference for the genetic, structural and functional genomics studies (Choi *et al.,* 2004; Mun *et al.,* 2006) and in the genome sequencing project (Young *et al.,* 2005). We used genomic DNA of this genotype as a template for the amplification of different fragments
- ***Medicago truncatula* genotype R-108:** This genotype, which was obtained through cell culture selection (Hoffmann *et al.,* 1997), has been reported to exhibit superior *in vitro* regeneration and was thus successfully used for transformation experiments (Trinh *et al.,* 1998). We also used this genotype for *Agrobacterium*-mediated genetic transformation and transgene expression studies.
- ***Nicotiana tabacum* cv. SR1:** Wild type *N. tabacum* cv. SR1 has been widely adopted for transient gene expression studies (Sparks *et al.,* 2006). We employed this genotype for *Agrobacterium* infiltration and the subsequent heterologous expression of proteins in leaf abaxial epidermal cells.

The tobacco plants were grown *in vitro* from seeds in a growth cabinet with a 16 h photoperiodic cycle (warm white fluorescent light with an intensity of 100 $\mu E.m^{-2}.s^{-1}$) at 24 ± 2 °C under sterile conditions, on hormone-free MS medium (Murashige and Skoog, 1962) supplemented with 30 g/l sucrose and solidified with 6 g/l Kalys agar, pH 5.7. Similarly, *M. truncatula* plants were grown *in vitro* either on MS0 medium as described for tobacco, or on SHb10 media (Chabaud *et al.,* 1996) in a growth cabinet with the same culture conditions as above.

2.1.2 Cloning vectors, bacterial strains and primers

2.1.2.1 Plant expression vectors

2.1.2.1.1 PMDC 83 vector

The binary vector pMDC83 has been designed for subcellular localization by expressing a protein in translational fusion (Curtis and Grossniklaus, 2003). For this purpose, the encoded protein can be fused to the N-terminus of the histidine-tagged GFP6, both driven by the dual Cauliflower Mosaic Virus (CaMV) 35S promoter (Odell *et al.,* 1985); and the nopaline synthase *(nos)* terminator of *A. tumefaciens* (Hellens *et al.,* 2000) at the end of T-DNA (Fig. 12).

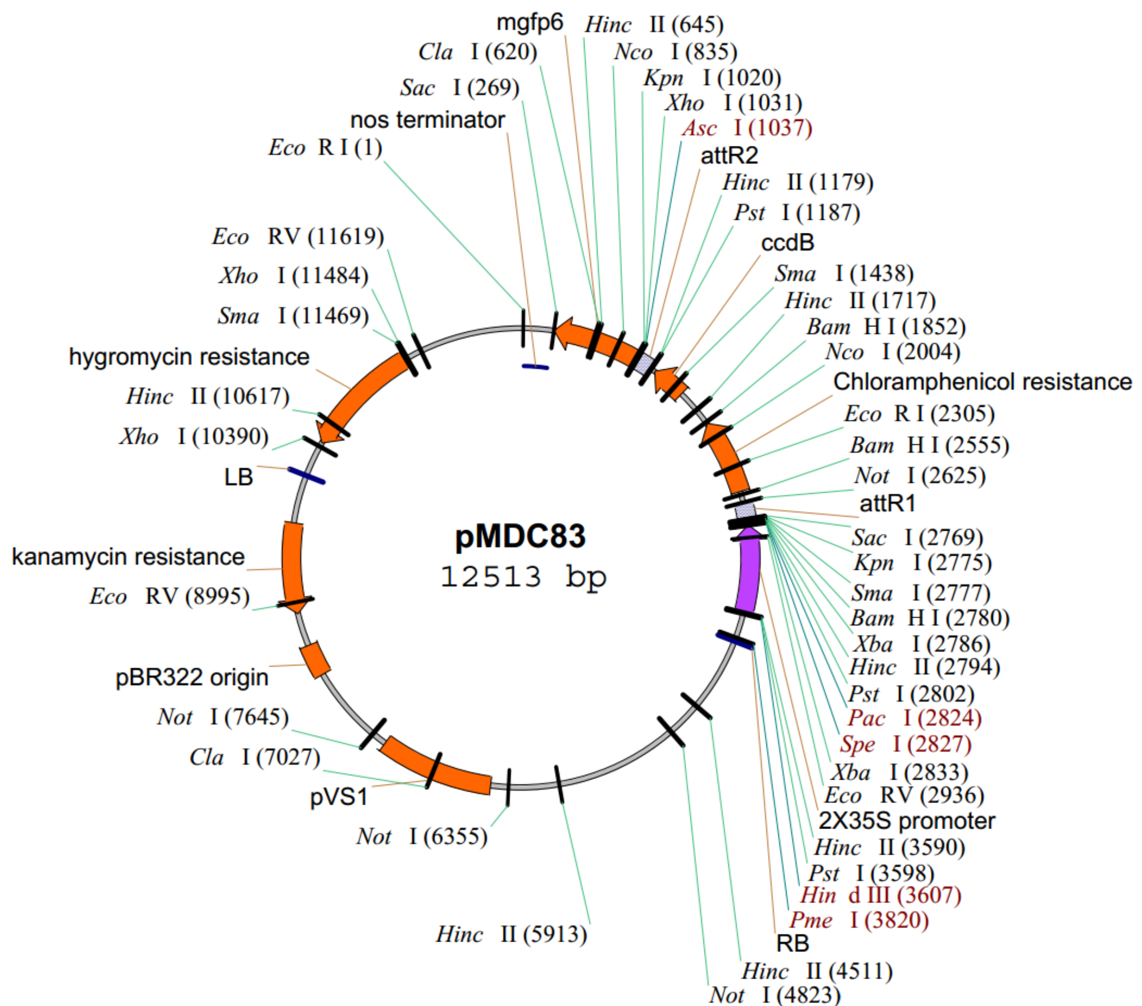


Figure 12. Detailed restriction map of Gateway compatible destination vector pMDC83. It is designed for preparing fusion protein construct.

2.1.2.1.2 PMDC 107 vector

The binary vector pMDC107 has been designed for preparing promoter reporter constructs (Curtis and Grossniklaus, 2003). For this purpose, the promoter sequence can be inserted upstream of a reporter gene, i.e. GFP6, which is followed by the *nos* terminator (Fig. 13). Both of above described expression vectors are adapted for the Gateway® cloning system. The vectors also contain kanamycin resistance gene for selection in *E. coli* and *A. tumefaciens* strains, while hygromycin phosphotransferase is the plant-selectable marker gene. Both the selectable markers are under the transcriptional regulation of the CaMV 35S promoter and *nos* terminator and are adjacent to the left border of the T-DNA (Curtis and Grossniklaus, 2003).

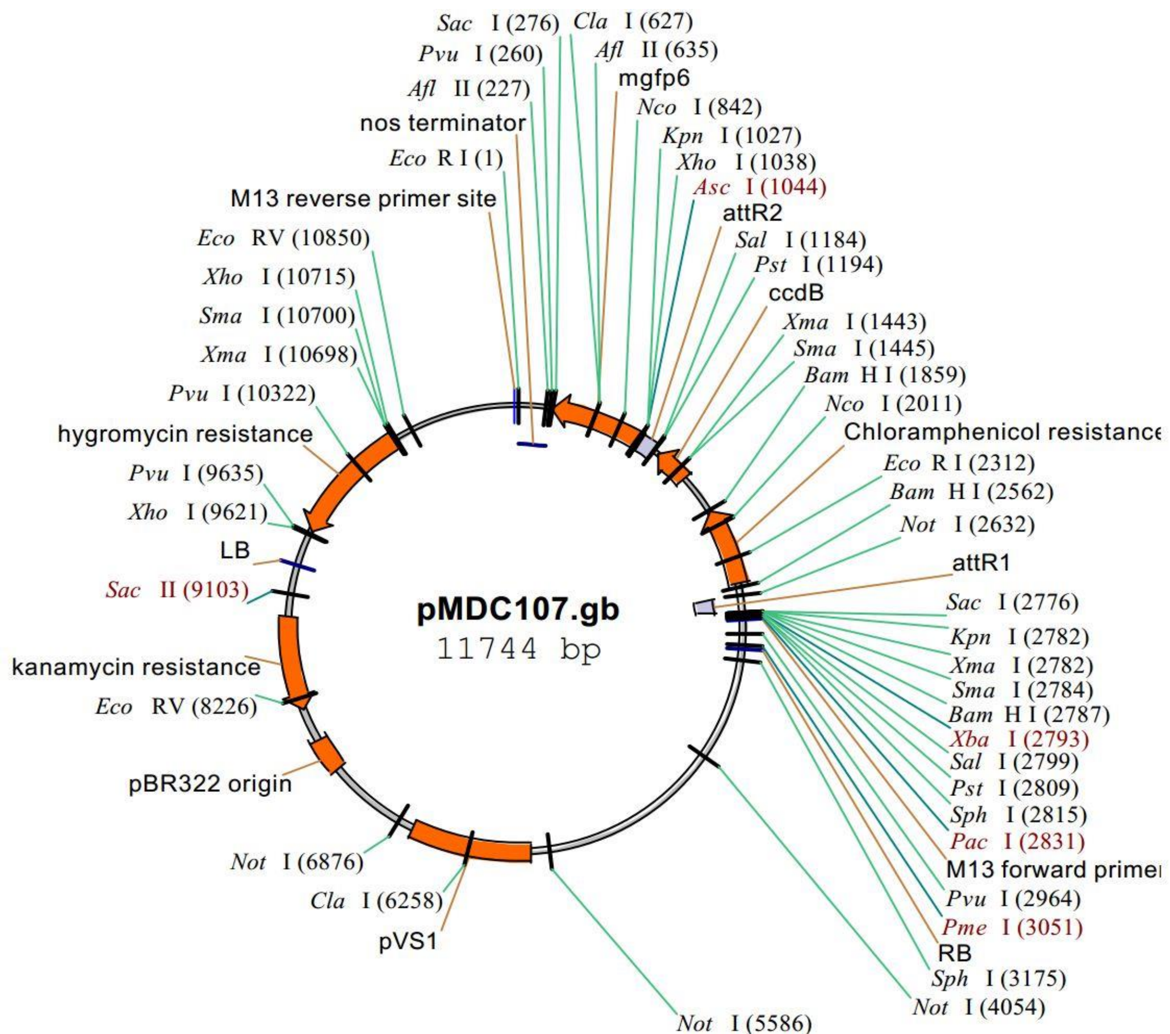


Figure 13. Detailed restriction map of Gateway compatible destination vector pMDC107. It is designed for preparing the promoter reporter construct.

2.1.2.1.3 pDONR221 vector

The pDONR™221 vector is a Gateway®-adapted donor vector designed to generate attL-flanked entry clones containing a cloned sequence following recombination with an attB PCR product. It has a pUC origin for high plasmid yields, kanamycin selection gene and universal M13 sequencing sites (Fig. 14) which are helpful for insert verification by PCR and sequencing. The entry clone can then be used for the LR recombination reaction to transfer the DNA fragment to a destination vector to give rise to the final expression clone.

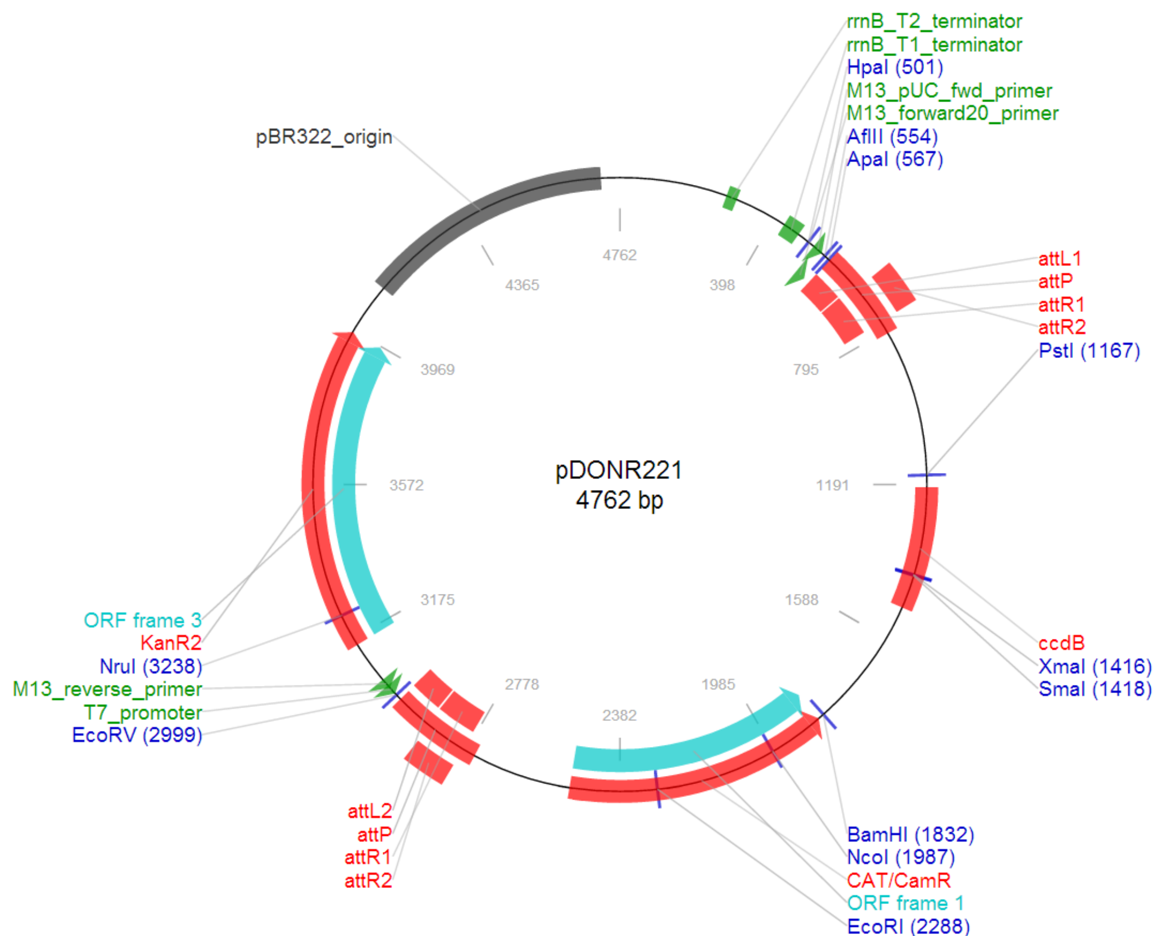


Figure 14. Detailed restriction map of Gateway compatible donor vector pDONR221. It is designed for preparing promoter-reporter constructions.

2.1.2.2 *E coli* strains

- **One Shot TOP10 cells:** These are chemically competent *E. coli* cells and can be used for cloning of non-ccdB-containing plasmids. These cells were used for the

multiplication of plasmids before mini or midipreparations. Their genotype description is as follows: F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG. (Invitrogen ref C404010)

- **One Shot OmniMAX™ 2 T1R :** This *E. coli* strain is an improved chemically competent cell line, suitable for use in the Gateway® Technology, as it offers the highest transformation efficiency of any chemically competent cells. They confer special resistance to T1 and T5 phage infection. Genotype description: F´ *{pro*AB *lac*I^q *lac*Z ΔM15 *T*n10(Tet^R) Δ*(ccdA*B)} *mc*rA Δ*(mrr hs*dRMS -*mc*rBC) Φ 80(*lac*Z)ΔM15Δ *(lac*ZYA-*arg*F)U169 *end*A1 *rec*A1 *su*pE44 *th*i-1 *gy*rA96 *re*lA1 *to*nA *pa*nD. (Invitrogen ref C8540-03)

2.1.2.3 *Agrobacterium tumefaciens* strains

Two *Agrobacterium tumefaciens* strains were used to host the plasmid transformation vectors when producing transgenic plants, LBA4404 and EHA105. *A. tumefaciens* has a tumour inducing (Ti) plasmid containing the disarmed *Vir* virulence genes allowing the transfer of the transfer-DNA (T-DNA) from the plasmid in the nucleus of plant cells where it is integrated into the genome.

2.1.2.3.1 LBA4404:

A. tumefaciens LBA4404 cells contain the disarmed Ti plasmid pAL 4404 which has only the *vir* and *ori* region of the Ti plasmid and confers resistance to streptomycin/spectinomycin (Hoekma, *et al.,* 1983). LBA4404 cells are resistant to rifampicin (Rif). Electro-competent *A. tumefaciens* LBA4404 Cells from Invitrogen has been used for *M. truncatula* transformation during my thesis work.

2.1.2.3.2 EHA105:

It is a disarmed hypervirulent strain of *A. tumefaciens* that is a derivative of EHA101 (Hood *et al.,* 1993) and also possesses Rif^r. This strain has been reported to transform most genotypes of *M truncatula* successfully (*Hellens et al.,* 2000).

2.1.3 Laboratory chemicals and consumables

General laboratory grade chemicals and reagents used during manipulations were purchased from Merck (Darmstadt, Germany), Sigma-Aldrich (Steinheim, Germany) and Duchefa Biochemie (Haarlem, The Nethelands) unless otherwise stated. The vitamins, aminoacids and other tissue culture chemicals, antibiotics and selection agents were obtained from Sigma-Aldrich, Merck, Duchefa Biochemie and Roche unless otherwise stated. Filter papers were obtained from Whatman. The glassware used was from Pyrex and Schott (Germany). The restriction enzymes, dNTPs, DNA polymerases and other molecular biology reagents were bought from Invitrogen (Carlsbad, CA USA), Fermentas Life Sciences (St. Leon-Rot, Germany), Thermo Fisher Scientific (Villebon sur Yvette, France) and New England BioLabs (Génopole d'Evry, France).

2.1.4 Solutions, media, buffers and antibiotics

All the stock solutions, media and buffers were prepared using Milli-Q water (Milli-Q® Integral Water Purification System, Millipore) throughout the study. Water based media and stock solutions were either filter sterilized through a 0.22 µm filter or autoclaved at 121°C for 20 min and 15 psi pressure. Heat labile solutions, on the other hand, were always filter sterilized through a 0.22 µm filter. All media were cooled to room temperature before use or addition of antibiotics.

2.1.4.1 Stock Solutions

- **N6 Macro-salts stock solution (1 L) (10X):** 1.85 g $MgSO_4.7H_2O$, 28.30 g KNO_3, 4.63 g $(NH)_2SO_4$, 1.66 g $CaCl_2.2H_2O$, 4.0 g KH_2PO_4
- **SH Macro-salts stock solution (500 mL) (20X):** 25 g KNO_3, 4 g $MgSO_4.7H_2O$, 2 g $CaCl_2.2H_2O$, 3 g $NH_4H_2PO_4$
- **SH Micro-salts stock solution (100 mL) (1000X):** 1 g $MnSO_4.H_2O$, 500 mg H_3BO_3, 100 mg $ZnSO_4.7H_2O$, 100 mg KI, 10 mg $Na_2MoO_4.2H_2O$, 20 mg $CuSO_4.5H_2O$, 10 mg, $CoCl_2.6H_2O$
- **SH vitamins stock solution (100 mL) (1000X):** 500 mg nicotinic acid, 500 mg thiamine HCl (Vitamin B1), 500 mg pyridoxine HCl (Vitamin B6)

2.1.4.2 Media

- **Luria Bertani (LB) for E.coli culture:** For 1L, 10 g tryptone, 10 g peptone, 5 g yeast extract, 10 g NaCl. For solid medium, 15 g/L Bacto agar was added
- **YM medium for Agrobacterium culture:** For 1L, 0.4 g yeast extract, 10 g mannitol, 0.1 g NaCl, 0.2 g MgSO4.2H2O, 0.5 g K2HPO4.3H2O. Final pH 7.0.
- **YEB medium for Agrobacterium culture:** (Vervliet et al, 1975; Krall et al. 2002): For 1L, 5 g of Bacto beef extract, 1 g of Bacto yeast extract, 5 g of peptone, 5 g of sucrose, 2 ml of MgSo4 from1M stock solution, pH: 7.2. For solid medium, 15 g/L Bacto agar was added before autoclaving.
- **SOC media:** 2% Tryptone, 0.5% Yeast Extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4, 20 mM glucose
- **SH3a medium:** For 1 L, 100 ml N6 Macro-salts, 1 ml SH Micro-salts, 1 ml SH vitamins, 5 ml Fe EDTA, 100 mg myo-inositol, 30 g sucrose, 4 mg 2-4 D, 0.5 mg BAP, pH 5.8, for solid media, 3 g of phytagel was added. When used for cocultivation, 100 µM of acetosyringone was added to the media.
- **SH9 medium:** For 1 L, 100 ml N6 Macro-salts, 1 ml SH Micro-salts, 1 ml SH vitamins, 5 ml Fe EDTA, 100 mg myo-inositol, 20 g sucrose, pH 5.8, for solid media, 7 g of Kalys agar was added.
- **1/2 SH9 medium:** For 1 L, 50 ml N6 Macro-salts, 0.5 ml SH Micro-salts, 0.5 ml SH vitamins, 2.5 ml Fe EDTA, 50 mg myo-inositol, 10 g sucrose, pH 5.8, for solid media, 7 g of kalys agar was added.
- **SHb10 medium:** For 1 L, 50 ml SH Macro-salts, 5 ml SH Micro-salts, 5 ml SH vitamins, 5.5 ml Fe EDTA, 1 g myo-inositol, 10 g sucrose, pH 5.8, for solid media, 6 g of kalys agar was added.
- **infiltration medium:** For 20 mL, 10 ml of 100 mM MES, 2 ml of 20 mM Na_3PO_4, 10 µl of 200 mM Acetosyringone, 0.5% of Glucose.

2.1.4.3 Buffers

- **TAE buffer:** 40 mM Tris-HCl, 1 mM EDTA, pH 8.0, sodium acetate 5mM
- **DNA loading dye:** 10 mM Tris-HCl (pH 7.6), bromophenol blue 0.03% (w/v), xylene cyanol 0.03% (w/v), 60% glycerol (v/v), 60 mM EDTA
- **DNA extraction buffer:** (1.4 M NaCl, 20 mM EDTA, 100 mM Tris-HCl, 2% (w/v) CTAB, pH 8.0, 1% sodium bisulfite added just before use)

2.1.4.4 Antibiotics and selection agents

Milli-Q water was used for the preparation of antibiotics stock solutions and these were filter sterilized through a 0.22 µm filter and kept at -20°C (Table 2). When needed, the antibiotics were added into the media when it had been cooled to about 50°C after autoclaving.

Table 2. Detail of antibiotics and selection agents

Name	Stock conc.	Solvent	Storage °C	Working Conc.
Kanamycin	50mg/ml	Water	-20°C	50mg/L
Rifampicin	50mg/ml	Dimethylsulfoxide	-20°C	100mg/L
Hygromycin	426mg/ml	Water	4°C	10mg/L
Augmentin	100mg/ml	Water	-20°C	100mg/L
Timentin	100mg/ml	Water	-20°C	100mg/L

2.2 Methods

2.2.1 *In silico* and bioinformatic analysis

2.2.1.1 Sequence data collection

About 375 Mb sequence of *M. truncatula* (genome assembly version Mt3.5.2) distributed on 8 pseudomolecules as well as on unfinished and unanchored BAC libraries, based on a recently completed BAC assembly (Young *et al.*, 2011) was downloaded from FTP server of (ftp://ftp.jcvi.org/pub/data/m_truncatula/Mt3.5/Assembly/Mt3.5.2/) Medicago hapmap project website, which covers about 94% of all *M. truncatula* genome. The downloaded genome assembly was imported to desktop version of CLC Sequence Viewer (Build 67103 v6.7.1 downloaded from www.clcbio.com) for further analyses.

For phylogenetic analysis, the complete protein sequences of eight pea Dof (PsDof) TFs were searched and downloaded from NCBI GeneBank® database at web server (http://www.ncbi.nlm.nih.gov/genbank/). The 35 *A. thaliana* Dof (AtDof) protein sequences were kindly provided by Dr. C. Le-Signor (UMRLEG, INRA-Dijon). The protein sequences of Dof proteins belonging to Prolamin Binding Factors (PBF) group from wheat, barley and maize were also searched and downloaded from NCBI GeneBank® database while the only

Dof protein sequence from green alga i.e. *Chlamydomonas reinhardtii* (CrDof1) was downloaded from Moreno-Risueno *et al.* (2007).

2.2.1.2 Identification of Dof family members from *M. truncatula* genome

In order to identify the *M. truncatula* Dof family members, the complete protein sequence of DOF1147 was BLASTed using the default algorithm parameters of BLASTP (Altschul *et al.,* 1990, 1997) against IMGAG protein database from genome assembly Mt3.0 as well as Mt3.5 on the online web server of *Medicago truncatula* Genome Sequence Resources (http://tofu.cfans.umn.edu/advanced_search_page.php?seq). All protein sequences having E value greater than E-^{04} were selected for further analysis.

2.2.1.3 Sorting of identified DOF proteins

As the Dof proteins identified from Mt3.0 and Mt3.5 after BLASTing had different gene model IDs, they were compared for complete protein sequence on the online webserver of *Medicago truncatula* Genome Sequence Resources (http://tofu.cfans.umn.edu/advanced_search_page.php?seq). The sequences of the proteins showing 100% homology from both database i.e. Mt3.0 and Mt3.5 were aligned using built-in CLUSTALW module (v2.1) (Thompson *et al.,* 1997) on CLC Sequences Viewer desktop suite. The proteins showing truncated or no Dof DNA binding domain were rejected for further analyses. The remaining proteins were renamed according to their positions on the *M. truncatula* pseudomolecules.

2.2.1.4 Mapping of *MtDof* genes on chromosomes

The nucleotide sequences of all the renamed MtDof genes were extracted from the *M. truncatula* Resources web server (http://tofu.cfans.umn.edu/advanced_search_page.php?seq). The physical location of each *Dof* gene was identified from genome sequence assembly Mt3.5 using the CLC Sequence Viewer. The identified positions were verified using the Chromosome Visualization tool (CViT), a set of perl and CGI scripts-based visualization program available on the *M. truncatula* Resources web server as well as from the Map Viewer module available at NCBI web server (http://www.ncbi.nlm.nih.gov/mapview/). The resulting precise locations of *MtDof* genes were manually marked on the calibrated *M. truncatula* chromosome bars.

2.2.1.5 Comparison of conserved residues in the Dof DNA-binding domain

The 52 amino acid sequences of Dof DNA-binding domains from all the MtDof proteins were aligned using Constraint-based Multiple Alignment Tool (COBALT) (Papadopoulos and Agarwala, 2007) using default alignment parameters from the NCBI web server (http://www.ncbi.nlm.nih.gov/tools/cobalt/cobalt.cgi?CMD=Web). COBALT computes a multiple protein sequence alignment using conserved domain and local sequence similarity information.

2.2.1.6 Construction of Phylogenetic tree

The complete protein sequences from 22 *M. truncatula* Dofs, 35 *A. thaliana* Dofs, 8 pea Dofs, 3 PBFs from wheat, barley and maize and the only Dof from *C. reinhardtii* were aligned with the CLUSTALW module, configured for the highest accuracy (Thompson *et al.,* 1997). From the resultant multiple sequence alignment, a phylogenetic tree was reconstructed using the neighbour-joining method (Saitou and Nei, 1987) by a built-in module of CLC Sequence Viewer. The reliability for internal branches was assessed using the bootstrapping method (1000 bootstrap replicates).

2.2.1.7 Identification of conserved domains and duplications in Dof proteins

To test whether the Dof proteins sequences had common domains, the deduced protein sequences of the 69 DOF proteins from *M. truncatula*, *A. thaliana*, *P. sativum*, *T. aestivum*, *H. vulgare*, *Z. mays* and *C. reinhardtii* were analyzed by means of the MEME program (Bailey and Elkan, 1994, 1995; Bailey *et al.,* 2006, 2009) available for anlaysis at the online server (http://meme.nbcr.net/meme/cgi-bin/meme.cgi). The program was run with default parameters with the following exceptions: the occurrence of a single motif was set to any number of repetitions, the maximum number of motifs to find was set to 26 and the width of each motif was set to 4-52 amino acid residues. The graphical view of the identified domains was used to explain the phylogenetic relations.

2.2.1.8 Analysis of cis-regulatory elements in putative target gene promoters

In order to identify the presence of cis-regulatory elements in the promoters of putative target genes of Dof1147, the 2.0kb upstream sequence from the transcription start site (TSS) of the selected genes were retrieved from *M. truncatula* genome assembly Mt3.5 using

the CLC Sequence Viewer. For the motif analyses, the dataset was then submitted to online servers of PLACE (Higo *et al.,* 1998, 1999) (http://www.dna.affrc.go.jp/PLACE/), and PlantCARE (Lescot *et al.,* 2002) (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) for the identification of overrepresented sequence motifs. The PLACE and PlantCARE web servers, with motif databases of 469 and 435 different experimentally proved cis-regulatory elements respectively, facilitate the identification of putative motifs in the given promoter sequences.

The sequences were also submitted to the MEME web server (Bailey *et al.,* 1994; Bailey *et al.,* 2006) (http://meme.nbcr.net/meme/cgi-bin/meme.cgi) for the identification of novel motifs shared by all the sequences. The default parameters were selected with following exceptions; the maximum of motifs was selected to be 30, the occurrence of a single motif was set to any number and the motif length 4-12 nucleotides. MEME web server (v4.9.0 Bailey *et al.,* 2009) identifies the common motifs present in the dataset using a complex but reliable algorithm and provides the facility to submit those identified motifs to other modules such as TOMTOM (Tanaka *et al.,* 2011) and MAST (Bailey and Gribskov, 1998) in order to find the specificity and identification of experimentally determined motifs.

2.2.1.9 *In silico* prediction of DOF1147 subcellular localization

The complete protein sequence of DOF1147 was submitted to Cello web server http://cello.life.nctu.edu.tw/ for *in silico* determination of subcellular localization of DOF1147 protein (Yu *et al.,* 2006). In order to identify the nuclear localization sequence (NLS), the protein sequence was then submitted to protein localization prediction tool PSORT accessible at http://psort.hgc.jp/form.html. PSORT uses an artificial intelligence technique for comparison-based determination of NLS in the provided protein sequence. The protein sequence was also submitted to Predictprotein web server (http://www.predictprotein.org/) for determination of important regulatory regions in the protein (Rost *et al.,* 2004).

2.2.2 Preparation of Constructs

In recent years, Gateway® cloning technology has proven an excellent alternative to prepare constructs using restriction enzyme cleavage sites in functional genomics. So, in order to study embryo development and seed filling in *Medicago truncatula*, Gateway® compatible cloning vectors were used for preparing the different constructs. Gateway cloning technology

is based on the site-specific recombination properties of bacteriophage lambda (Landy, 1989). Thus it facilitates to clone the DNA segments through DNA recombination reactions. It ensures the correct orientation of the inserted fragment in a right reading frame (Walhout *et al.,* 2000).

2.2.2.1 Primers designing

The primers were designed manually and their melting temperature was calculated using an algorithm available at: http://www.eurofinsdna.com/home.html. The primers were also blasted against the *M. truncatula* genome sequence to avoid unspecific binding. All the oligonucleotides were synthesized by Eurofins MWG Operon (Ebersberg, Germany). The detail of all the primers used during my research work is given in Table 2.

2.2.2.1.1 Calculation of annealing temperature by temperature gradient PCR

The determination of the optimal annealing temperature is a pre-requisite for exact amplification of desired fragments, and hence to avoid the possible formation of unspecific secondary bands by PCR. The purified genomic DNA extracted from *Medicago truncatula* genotype A17 was used as the DNA template for amplification of all the fragments. The promoter sequence of *DOF1147* was amplified using MtDof2prom-F2 as forward primer and MtDof2prom-R1 as reverse primer (Table 2), while the coding sequence (cds) of *DOF1147* was amplified using MtDof2Cod-F3 as forward primer and MtDof2codfus-R3 as reverse primer (Table 2). The optimized annealing temperatures for all the primers were determined experimentally by temperature gradient PCR. For each primer set, the PCR was normally started at 5°C below the calculated temperature of the primer melting point (Tm). Both, the promoter as well as cds of *DOF1147* showed sequence specific amplification at an annealing temperature of 58°C.

2.2.2.1.2 Designing of AttB-PCR primers

The Gateway cloning takes place by recombination reactions between AttB adapters. AttB1 forward and AttB2 reverse adapters containing four guanine (G) residues at the 5′ end followed by 25 bp attB site (Table 2) are always attached with primers to enable efficient cloning through site specific recombination. So the structure of the Gateway compatible primers would be

AttB1 Forward primer: 5'– GGGG – attb1sequence – gene specific primer sequence – 3'

AttB2 Reverse primer: 5' – GGGG – attb2sequence – gene specific primer sequence – 3'

Table 3. Primers used for Gateway cloning

Primer name	Primer Sequence	Size
MtDof2Prom-F2	TGTAATACTAATGTTTTCTTGACTG	25bp
MtDof2Prom-R1	CTCTCAGTTTTTATTTGAGTGTGCG	25bp
MtDof2cod-F3	ATGGAGCAAGAAAGTGGAGATGGTG	25bp
MtDof2codfus-R3	CTGAGGAGGATTGAACTCTGACAG	24bp
AttB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTN	29bp
AttB2	GGGGACCACTTTGTACAAGAAAGCTGGGTN	29bp
MtDof2Prom-F2- AttB1	**AttB1**-TGTAATACTAATGTTTTCTTGACTG	54bp
MtDof2Prom-R1- AttB1	**AttB2**-CTCTCAGTTTTTATTTGAGTGTGCG	54bp
MtDof2cod-F3-AttB1	**AttB1**-ATGGAGCAAGAAAGTGGAGATGGTG	54bp
MtDof2codfus-R3-AttB2	**AttB2**-CTGAGGAGGATTGAACTCTGACAG	53bp
M13 Fwd	GTAAAACGACGGCCAG	16bp
M13 Rev	CAGGAAACAGCTATGAC	17bp

2.2.2.1.3 Amplifying attB-PCR Products

In order to achieve the gateway specific PCR product, the promoter sequence of *DOF1147* was amplified using MtDof2prom-F2-AttB1 as forward primer and MtDof2prom-R1-AttB2 as reverse primer (Table 2). The cds of *DOF1147* was amplified using MtDof2Cod-F3-AttB1 as forward primer and MtDof2codfus-R3-AttB2 as reverse primer (Table 2). High Fidelity *TAQ* Advantage DNA polymerase was used to amplify the desired sequences. The amplification reaction consists of 3 µl T10X HF PCR buffer, 2µl T10X Advantage buffer, 1.5µl of forward primer (10µM), 1.5µl of reverse primer (10µM), 5µl dNTPs, 5µl of genomic DNA of A17 (80-100ng/ µl), 1 µl *Taq* Advantage DNA polymerase (1 unit/µl) and distilled water to a final volume of 50 µl. The PCR was run on a Biometra T1 Thermocycler (BIOLABO scientific instruments) with the following program: 1 cycle of denaturation at 94°C for 1 min, 35 cycles of each denaturation at 94 °C for 30 seconds,

annealing at 58°C for 30 seconds, elongation at 68°C for 4 min and a final elongation at 68°C for 3 min.

2.2.2.1.4 Agarose Gel Electrophoresis

After PCR completion, agarose gel electrophoresis was carried out for verification, quantification and purification of the amplified fragments. The PCR mixture containing the amplified fragment was loaded on the agarose gel for electrophoresis to remove excessive nucleotides and primers. DNA fragments were separated by horizontal electrophoresis at 10 V/cm using non-denaturing agarose gel [0.8-2% (w/v)] prepared in 1X TAE buffer. 0.2x of DNA loading dye was mixed to the samples. Ethidium bromide (EB) was either added to the agarose gel (final conc. 0.25 µg/ml) or the gel was stained in EB solution (final conc. 1 µg/ml). Ethidium bromide emits fluorescence under UV light when intercalated into DNA or RNA. The bands were visualised and verified under UV by using Gel Doc Quantity One© software (Version 4.5.1, BIO-RAD®) and 1-D Analysis Software. The purified DNA can be used for sequencing, cloning, and restriction enzyme digestion.

2.2.2.1.5 Separation and purification of amplified attB-PCR Product

After gel electrophoresis, if the PCR product was of the expected size, the desired bands of the PCR product from the gel were cut and purified using the Wizard® SV Gel kit and PCR Clean-Up System (Promega). This purification system is designed to extract and purify DNA fragments of 100 bp to 10 kb from standard agarose gels. This system is based on the ability of DNA to bind to the silica membrane in the presence of high concentrations of chaotropic salts such as guanidine isothiocyanate. The binding of the nucleic acid to silica is caused by dehydration and the formation of hydrogen bonds. A high concentration of salts helps in the absorption of nucleic acids on silica.

Briefly, the bands of interest were cut out and dissolved in the presence of guanidine isothiocyanate (Membrane Binding Solution). The mixture was then poured into a column of silica membrane that holds the DNA, then centrifuged at 14000g for 1 min at room temperature. After two successive washings with a membrane wash solution (10 mM potassium acetate pH 5.0, 16.7 µM EDTA pH 8.0), the column was centrifuged at 14000g. The flow-through was discarded and the DNA was then eluted with 50µL of Milli-Q water. The purified PCR-amplfied DNA was then used for the BP recombination reaction.

2.2.2.2 BP Recombination Reaction

The BP recombination reaction enables the easy transfer of attB-PCR product to an attP-containing donor vector to create an entry clone. The 10 µl of BP recombination reaction mixture contained 5 µl AttB-PCR product (50-100 ng), 1 µl pDONR221 (150 ng), 2 µl BP Clonase™ II enzyme mix, and TE buffer (pH 8.0) to make up to the volume. The reaction mixture was vortexed well and incubated at 25°C for 3h. Then 1µl of Proteinase K solution (2 µg/µl in: 10 mM Tris-HCl, pH 7.5, 20 mM $CaCl_2$, 50% glycerol) was added to terminate the reaction and the mixture was vortexed briefly and incubated at 37°C for 10min. The BP recombination reaction resulted in the transfer of AttB-PCR product into donor vector to create the entry clone (Fig. 15).

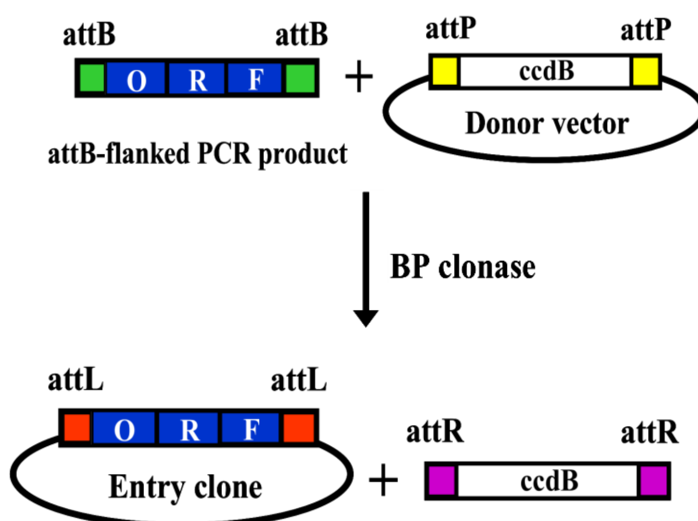


Figure 15. BP recombination reaction pathway

2.2.2.2.1 Transformation of entry clone into E. Coli

The BP recombination reaction product was transformed into OmniMax™ 2-T2^R competent cells (Invitrogen) to get a high transformation efficiency. Briefly, the One Shot® OmniMAX™ chemically competent cells were thawed on ice and about 1µl (50ng) of plasmid DNA was mixed in it gently without pipetting. The vial was then incubated for 30 minutes in ice followed by heat-shock at 42°C for 30 seconds. It was incubated again in ice for 2 minutes and 250 µl of pre-warmed S.O.C. medium was added in it. It was shaken horizontally at 37°C for 1 hour at 225 rpm in a shaking incubator. 20-100 µl of the diluted

transformation mix (1/20th) was spread on pre-warmed plates with LB medium and kanamycin selection and were incubated at 37°C overnight.

2.2.2.2.2 Entry clone verification by PCR

The well-isolated individual colonies of entry clones generated by BP recombination reaction were screened by PCR using M13 standard forward and reverse primers (Table 2). The colonies were dissolved separately in 50 µl of deionized water and 5 µl of this was used for PCR. The composition of PCR reaction mixture was 10 µl T2X buffer (containing dNTPs but without MgCl2), 0.5 µl MgCl2 (25 mM), 1 µl M13 forward primer (10 µM), 1 µl M13 reverse primer (10 µM), 5 µl of colony mixture (described earlier), 0.8 µl *Taq* DNA polymerase (1 unit/ µl) and distilled water to a final volume of 20 µl. The PCR was run on a Biometra T1 Thermocycler (BIOLABO scientific instruments) with the following program: 1 cycle of denaturation at 95°C for 5 mins, 30 cycles each of denaturation at 96 °C for 30 seconds, annealing at 50°C for 15 seconds, elongation at 60°C for 4 min and final cycle of 25°C for 5 min were used during this PCR. The PCR reaction mixture was then run on 2% agarose gel. The colonies of entry clones having the right insert were identified from corresponding bands on the gel for plasmid isolation.

2.2.2.2.3 Plasmid preparation from E.coli colonies

For plasmid isolation, the colonies having the right insert were put in 5ml LB broth with kanamycin selection to start a pre-culture and incubated at 37°C for overnight at 225 rpm. A final culture of 50ml was started from the pre-culture and again incubated overnight at 37°C for overnight at 225 rpm. The plasmid DNA containing entry clones was extracted from this bacterial culture by using PureLink® HiPure Plasmid Midiprep Kit (Invitrogen) according to the manufacturer’s protocol and then sent for sequencing. The DNA was stored at -20°C until use.

2.2.2.2.4 Sequencing of entry clones

In order to verify the correct orientation as well as mutations in the inserted sequence, the entry clones were sent for sequencing. For this purpose, 1-1.5 µg of purifies plasmid DNA of PCR positive entry clones was sent to Eurofins MWG Operon Sequencing Service. After sequencing, the sequences of different clones were compared and the clone having no mutation in the inserted fragment was selected for LR reaction.

2.2.2.3 LR Recombination Reaction

After confirmation of the correct sequence of the entry clones through sequencing, I proceeded with the construction of expression clones by carrying out LR recombination reaction. Before the LR recombination reaction, linearization of entry clones was performed using *Pvu*1 enzyme. The digestion reaction contained 1µg of pENTRY, 2µl of D 10X buffer, 0.2µl of BSA, 1µl of *Pvu*1 enzyme and milli-Q water to make the final volume 20µl. Digestion reaction was incubated overnight at 37°C.

The LR recombination reaction was performed to transfer the gene of interest from the entry clone into an attR-containing destination vector to create an expression clone (Fig. 16). The LR recombination reaction mixture contained: 1-7 µl linearized entry clone (50-150 ng), 1 µl expression vector (150 ng/µl), 2 µl LR Clonase™ II enzyme mix, and TE buffer (pH 8.0) to a final volume of 10 µl. The reaction mixture was mixed well and incubated overnight at 25 °C. Then 2 µl of Proteinase K solution was added to the mixture to terminate the reaction. The reaction mixture was vortexed briefly and incubated at 37°C for 10 min. In short, during LR recombination reaction, the *DOF1147* promoter region is inserted into the destination vector PMDC107 to create an expression clone with orientation *pdof::gfp*. The cds of *DOF1147* was cloned into the destination vector pMDC83 to produce an expression clone containing the cassette: *p35S::dof-gfp*.

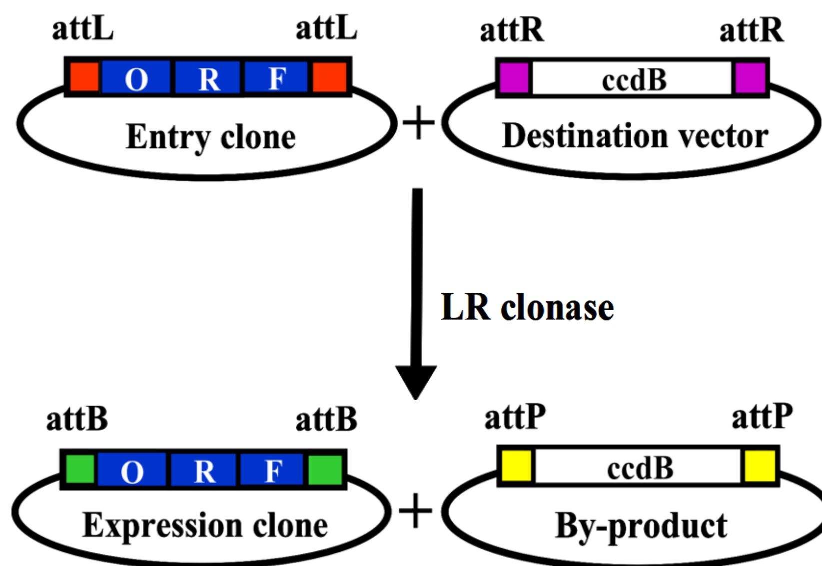


Figure 16. LR recombination reaction pathway

The LR recombination reaction product was transformed into Library Efficiency DH5alpha competent cells (Invitrogen) according to the manufacturer's protocol. 20 µl and 50 µl of each transformation was plated onto LB agar plates containing 50 µg/ml of Kanamycin as selective agent and plates were incubated at 37 °C overnight.

2.2.2.3.1 Expression clone verification and sequencing

Well grown individual colonies of expression clones generated by LR recombination reaction were screened by PCR using vector specific M13 primers as described earlier. Colonies were also double checked with the insert sequence-specific primers (Table 3). The composition of the PCR reaction mixture and running program were as previously stated. Colonies having the correct insert were identified from corresponding bands on the gel. Their plasmid DNAs were prepared by using PureLink® HiPure Plasmid Midiprep Kit (Invitrogen). Extracted DNA was checked for purity and quantified by gel electrophoresis. The correct insertion of the fragment in the expression clone was re-verified with the restriction digestion of corresponding restriction sites of the plasmids, and then sent for sequencing as explained earlier.

2.2.3 Manipulations with *Agrobacterium tumefaciens*

After sequencing, the expression clones having no mutation as well as having the coding sequence in frame were transferred into *Agrobacterium tumefaciens* strains i.e. EHA105 or LBA4404 to enable stable genetic transformation of *M. truncatula*.

2.2.3.1 Preparation of competent *A. tumefaciens*

A fresh culture of *A. tumefaciens* (EHA105) in 10 ml YEB broth was launched from an isolated fresh colony. Then 200 ml of YEB medium was inoculated with it and cultured overnight at 28°C and 200 rpm until its OD_{600} nm had reached between 0.6 and 0.9 which corresponds to the exponential growth phase of culture. The bacterial culture was placed on ice for 15 min and then bacteria were pelleted at 3000 g for 20 min at 4°C. The pellet was washed in ice-cold TE buffer. The pellet was then resuspended in 20 ml of YEB medium. 500 µl aliquots of *Agrobacteria* were frozen in liquid nitrogen and stored immediately at -80°C until further use.

2.2.3.2 Transformation of *Agrobacteria* by heat shock

An aliquot (500 µl) of competent *A. tumefaciens* (EHA105) cells was thawed on ice and 1 µg of salt-free DNA plasmid (pMDC83- *p35S::dof-gfp*) carrying the gene of interest was mixed gently. This cell/DNA mixture was incubated successively on ice, in liquid nitrogen and in a water bath at 37°C for 5 min each. Then, 1 mL of YEB medium was immediately added to the tube and cells were quickly resuspended and incubated at 28°C for 4 hours at 200 rpm. 100-200 µl of culture was spread on a Petri dish containing 20 ml YEB agar medium supplemented with the kanamycin (Table 1). The dishes ware then put at 28°C for 48 hours. Isolated colonies of *Agrobacteria* were verified by PCR using either M13 universal primers or sequence specific primers (Table 2, PCR conditions described earlier) and the culture of correct colonies was glycerol stocked at -80°C for use for stable transformation.

2.2.3.3 Transformation of electro-competent *Agrobacteria*

In order to transform electrocompetent LBA 4404 agrobacterium, 100ng of plasmid DNA (pMDC107- *pdof::gfp*) was added to a pre-chilled microcentrifuge tube. 100ng of PBI121 control DNA was added into another pre chilled microcentrifuge tube to determine the transformation efficiency. ElectroMAX™ *Agrobacterium tumefaciens* LBA 4404 cells were thawed on wet ice, and were mixed by tapping gently. 20 µl of these cells were added to each tube containing plasmid DNA and control plasmid. Cell/DNA mixture was pipetted out into a chilled 1mm electrode gap cuvette. Electroporation was carried out in 1mm cuvette with BTX® ECM® 630 Electro Cell Manipulator at 1.8 kV, 200 Ω, 25 µF. After electroporation, 1 ml of room temperature YM medium was added and gently mixed and the solution was transferred to a 15 ml Falcon tube. The tube was shaken at 225 rpm at 30ºC for 3 hours in the dark. The cells transformed with the control pBI121 DNA were diluted 1:10 with YM Medium and 100 µl of this dilution was spread on pre-warmed YM plates with streptomycin and kanamycin as selection agent (Table 1). 100-200 µl of the cells transformed with *pdof::gfp* expression vector were spread on pre-warmed YM medium plates containing kanamycin as selection agent. Plates were incubated for 48h at 30ºC. Transformation efficiency was calculated and colonies were verified by PCR using either M13 primers or the sequence- specific primers given in the table 2.

2.2.4 Stable transformation of *Medicago truncatula*

2.2.4.1 Preparation of *Agrobacterium* culture

Two days before transformation experiments, an *Agrobacterium* liquid pre-culture was initiated by inoculating a freshly grown isolated single colony of *Agrobacterium* strain EHA105 containing the expression plasmid in 2ml YEB liquid medium supplemented with the appropriate antibiotics for transformation vector selection. It was incubated overnight at 30°C and 200 rpm. One day prior to transformation, 50 ml of YEB medium (supplemented with antibiotics) was inoculated with 2 ml overnight pre-culture and incubated again overnight at 30°C and 200 rpm to attain an OD_{600} of 1.0.

Similarly, pre-culture of LBA4404 was initiated by inoculating a freshly grown isolated single colony in 2ml YM liquid medium supplemented with the appropriate antibiotics for transformation vector selection (described earlier). It was incubated for 36 hours at 30°C and 200 rpm. A day prior to transformation, 50 ml of YM medium (supplemented with antibiotics) was inoculated with 2 ml overnight pre-culture and incubated again overnight at 30°C and 200 rpm to attain an OD_{600} of 1.0.

2.2.4.2 Preparation of explants

Healthy seeds of *M. truncatula* genotype R-108 were scarified either with sand paper or with Sulphuric acid (1N H_2SO_4) for 6-8 min followed by 4-5 rinses with sterile water. The seeds were then disinfected with ethanol 70% for 1 min then soaked in bleach solution (2 % active chlorine) for 15-20 min followed by rinsing three times in sterile water and then put at 4°C for 48 hours for stratification. Individual plants were grown *in vitro* on SHb10 agar media. Leaflets from well expanded leaves were used as explants from 4-5 weeks old *in vitro*-grown plants or alternatively from green house grown plants.

2.2.4.3 Inoculation of explants

The overnight grown liquid cultures of *A. tumefaciens* strains EHA105 and LBA4404 in YEB and YM media respectively were centrifuged at 3000 g for 10 min, and the pellets were resuspended separately in liquid SH3a co-cultivation media to attain an OD_{600} of 0.7-1.0. The explants from green house grown plants were disinfected and pre-cultured on SH3a regeneration media one day prior to transformation. In order to prevent dehydration, the

explants were placed on a drop of *Agrobacterium* culture in a sterile Petri dish and then wounded with a sterile scalpel (3-5 wounds/leaflet). They were then co-cultured with SH3a *Agrobacterium* media in vacuum flasks. Vacuum was applied (650 psi) to the leaf explants for 20 min to facilitate the entry of *Agrobacteria* through wounds in the leaflets. The vacuum should be released slowly in order to avoid cell damage. The flask was then placed on a shaker (50-60 rpm) at room temperature for 1-2 hours to oxygenate the medium and to allow the tissue to recover from the infiltration procedure.

2.2.4.4 Co-cultivation of explants

The infected explants were removed and dried on sterile filter paper to remove excessive bacteria. They were then transferred onto SH3a solid co-cultivation media (supplemented with 100 µM acetosyringone but no antibiotics) and cultured in the dark for 72 hours at 23°C for co-cultivation. This step is most important as the integration of the transgene into the plant genome takes place during this stage. The leaf explants were removed from the co-cultivation medium and rinsed thrice with rinsing media (SH3a + Augmentin + Timintin) in order to remove excessive bacteria that have grown on the explants.

2.2.4.5 Callus induction and somatic embryogenesis

The explants were then transferred to fresh solid SH3a medium supplemented with antibiotics and appropriate selective agent, and placed in the dark in the growth chamber (24°C) for 6 to 7 weeks for callus formation. These developing calli were subcultured to fresh SH3a medium every 2-3 weeks. For induction of somatic embryogenesis followed by plantlet development, the calli were transferred to hormone-free SH9 medium (containing hygromycin) in the light (24°C, 16h photoperiod) while maintaining selection. Green embryos developed from the callus in the light which later on led to plantlet development. Developing shoots were transferred to half-strength SH9 media supplemented with hygromycin for root induction. Plantlets with a few leaves and roots were transplanted for acclimatization into a mini greenhouse covered with a transparent lid to maintain the humidity. The transformed plants were analyzed by PCR in order to prove the integration of the T-DNA into the genomic DNA of the plant. PCR positive plants were transferred to green house until seed set and harvesting.

2.2.5 Analysis of putative transformants

After transformation of *M. truncatula* explants, a high frequency of regenerated plantlets through somatic embryogenesis has normally been observed when using R-108 genotype. So it is highly desirable to screen these putatively transformed plants as early as possible to eliminate the chimeras or escapes, thus saving cost and energy for maintenance of putatively transformed plants. Flow cytometry has emerged as an important tool for verification of true-to-typeness of regenerating plants as well as identification of plants showing undesired somaclonal variations.

2.2.5.1 Flow cytometric analysis of putative transformants

Flow cytometry was used to characterize *M. truncatula* putative transformed plants in terms of true-to-typeness (Ochatt, 2006, 2008). Leaves from Hygromycin resistant plants were taken and nuclei were mechanically isolated by finely cutting and chopping the leaves with a sharp razor blade in CyStain DNA 1 step nuclei extraction and staining buffer (Partec®). This buffer contains 4, 6 diamidino-2-phenylindole (DAPI), a specific fluorochrome which binds to the Adenine and Thymine bases of DNA. The mixture was filtered through a 50 µm mesh. A Partec PA-II flow cytometer supplied with an HBO-100 W mercury lamp and a dichroic mirror was used for flow cytometry analysis. When stained nuclei pass through the UV light, DAPI fluoresces in blue and the in-built computer programme of the cytometer transforms light into a curve, where the intensity of the emitted epi-fluorescence (plotted against the nucleus count) indicates DNA content. Non transformed *M. truncatula* plants were used as controls.

2.2.5.2 Molecular verification of putative transformants

The first step of analysis of putative transformants is to confirm them at DNA level by PCR using primers specific to the transformed gene constructs. Although PCR is very sensitive there are still many factors like quality of DNA that can affect the results.

2.2.5.2.1 Genomic DNA extraction

Extraction of genomic DNA was carried out by modifying the protocol described by Rogers and Bendich (1985). Briefly, approximately 200 mg of leaf material was taken in a 2ml eppendorf tube containing a ball bearing and immediately frozen in liquid nitrogen. Then

grinding of samples was done through vigorous shaking at a frequency of 30Hz for 2 min using Retsch MM301 Ball Mill. Then, 600 µl of DNA extraction buffer was added to the finely pulverized samples and incubated for 20 min at 65°C in order to break the cell walls. Then, 600 µl of chloroform was added and the mixture was vortexed and centrifuged at 10,000 G for 15 min at 10°C. DNA precipitation was carried out by treating the upper aqueous phase with 60µl of 3M sodium acetate (pH5.2) and 600µl of isopropanol respectively followed by centrifugation at 15,000 G for 30 min at 4°C. Finally, the pellet was rinsed with 70% ethanol by centrifuging at 15,000 G for 10 min. The pellet was then air dried and resuspended in 100 µl of TE buffer. Concentration, purity and integrity of isolated DNA were checked by absorbance spectrophotometry and gel electrophoresis. DNA was either stored at 4°C for immediate use or at -20°C.

2.2.5.2.2 Verification of transformants by Polymerase Chain Reaction

The transgenic nature of putative transformants was verified at the molecular level by PCR using DNA extracted from the putative transformants as a template. The plants transformed with *pdof::gfp* construct in LBA4404 were verified using MtDof2prom-F2 and MtDof2prom-R1 primers set while the plants transformed with *p35S::dof-gfp* construct in EHA105 were verified using MtDof2Cod-F3 and MtDof2codfus-R3 primers pair.

2.2.5.3 Germination and transgene segregation test in the progeny

The primary transformants in the greenhouse were allowed to set T1 seeds through self pollination. The mature pods were harvested and the seeds were separated from pods. The segregation analysis was performed on the basis of heritability of hygromycin resistance gene in the progeny of the primary transformants.

In order to perform *in vitro* germination test, some seeds from independent transgenic lines were stratified, disinfected, and put in dark at 4°C for 72 hours to break dormancy. They were then transferred onto MS0 media and put in a growth cabinet at 24°C with 16 hours of photoperiod. After two days, the germinated seedlings were transferred to MS0 media and germination percentage was calculated after one week. The germinated seedlings were then transferred to MS0 supplemented with hygromycin. After two weeks, the seedlings were transferred to fresh media containing hygromycin to select the transgenic plants and rouge out the segregating non trangenics or escapes. The segregation ratios of the transgenic lines were

then determined by counting the surviving plants. Chi square analysis was performed to investigate the inheritance pattern of transgenes and its comparison with Mendelian ratio.

2.2.6 Transient expression assays

Transient gene expression assays using agro-infiltration into tobacco leaves were also conducted to determine the subcellular localization of the fused Dof1147-GFP protein complex.

2.2.6.1 Transient expression in *Nicotiana tabacum* through agro-infiltration

A 1 ml fresh culture of *Agrobacterium* was grown overnight on a shaker at 28°C and 200 rpm. It was then centrifuged at 3000 g for 3 min at room temperature. The pellet was resuspended in infiltration medium to make the final OD_{600} to 0.3, suitable for infiltration. Agro-infiltration was conducted on expanded young leaves from 4-5 week old tobacco plants. The sub-epidermis of leaves was wounded gently with the help of syringe tip. Then 100 µl of bacterial suspension was injected by 1 ml plastic syringe through each wound into intercellular spaces of intact leaves. By following this method, 8-16 spots separated by the veins could be infiltrated in a single tobacco leaf. The plants were then put back in the growth chamber at 22°C with a 16h photoperiod for 24-48 hours. For microscopic observations, a small leaf disc from infiltrated leaves was taken and observed under the epi-fluorescence microscope using a GFP filter under UV light.

2.2.7 Microscopic studies

2.2.7.1 Laser scanning confocal microscopy

A Leica TCS-SP2 AOBS inverted laser scanning confocal microscope (Leica Microsystems, Heidelberg Germany) with adjustable bandwidths of the detected fluorescence wavelengths and equipped with a 3-laser configuration (488/561/633nm) was employed for microscopic studies of *DOF1147* promoter and protein expression. The immature seeds (8-12 DAP) from plants transformed with *pdof::gfp* construct or leaf sections from plants transformed with *p35S::dof-gfp* were mounted on the microscope and observed using a 20X, 40X or 63X (NA 1.4) oil immersion objective.

For observation of green fluorescence, GFP was excited at 488 nm using the inline Argon laser of the microscope, while the emission was detected in a wavelength range of

495–540 nm (see Results) by means of the AOBS-based built-in detectors of the confocal microscope. Scanning was performed using 400 Hz line frequency and each image was the average of four frames. All the images were collected simultaneously in 1024 x 1024 format by Leica Confocal software (LCS) during excitation and emission.

2.2.7.2 Fluorescence microscopy

Transient expression of *GFP* in the infiltrated tobacco leaves was investigated by observing them under UV excitation by using an inverted epi-fluorescence microscope. Excitation and emission wavelengths were controlled by employing GFP specific filters to observe green fluorescence. Image acquisition and treatment were performed with Archimed Pro and Histolab programmes (Microvision, France).

Chapter 3: Genome wide identification of DOF transcription factors in *M. truncatula* and *in silico* promoter analysis

3.1 Introduction

The regulation of different developmental processes in plants requires coordinated expression of thousands of genes. Most of these processes in plants have been reported to be regulated at the transcriptional level. TFs also coordinate these processes in response to hormonal and environmental stimuli. TFs enhance or repress the expression of target genes by binding to specific DNA sequences i.e. *cis*-regulatory elements, present in the promoter regions of target genes.

Publicly available data represents only a small fraction of the full catalog of TF genes encoded by the genomes of different plants. The genome wide identification of genes encoding the TFs belonging to different families is possible through computational methods once a genomic sequence is available. Subsequent comparative phylogenetic studies provide an insight into existing TF families within individual species lineages and may give indications as to their physiological roles in different developmental processes.

Regulatory elements play an essential role in gene function, but their experimental identification and verification continues to be a technical challenge. On the other hand, the development of different bioinformatics tools has been shown to be helpful for their *in silico* prediction. These predicted elements in the target gene promoters can later be used for their experimental verification.

This chapter of my thesis is devoted to the genome wide identification of genes belonging to DOF TF family from the recently published *M. truncatula* genome assembly, their phylogenetic analysis, identification of conserved domains in these proteins and the *in silico* identification of putative *cis*-regulatory elements from the promoter regions of putative target genes of DOF1147.

Results

3.2 In-silico subcellular localization of DOF1147

Predicting protein subcellular localization is an important component for understanding the protein function. The complete sequence of DOF1147 protein was submitted to Cello server (Yu *et al.,* 2006) for the prediction of subcellular localization preference. Cello uses a simple SVM module-based calculation for localization prediction, i.e., classification of DOF1147 based on comparison of the characters listed in Table 4. The DOF1147 protein was predicted to be localized in the nucleus at the subcellular level (Table 4). This prediction is consistent with its role as a transcription factor. The protein sequence was then submitted to web-based PSORT tool for the prediction of its NLS. When compared with functionally described NLS in different TF proteins, the responsible nuclear localisation sequence in Dof1147 could not be identified. This may perhaps be due to the fact that the NLS in the Dof proteins from different plant species have not yet been determined (Krebs *et al.,* 2010). So, it can be speculated that Dof proteins do not contain the conventional NLS in their sequences, whereby a detailed protein functional analysis is required for their prediction.

Table 4 : The subcellular localization analysis of Dof1147 protein. SVM analysis on different composition revealed nuclear localization of Dof1147

SVM	LOCALIZATION	RELIABILITY
Amino Acid Comp.	Nuclear	0.679
N-peptide Comp.	Nuclear	0.642
Partitioned seq. Comp.	Nuclear	0.510
Physico-chemical Comp.	Nuclear	0.461
Neighboring seq. Comp.	Nuclear	0.952
CELLO Prediction:	Nuclear	3.243

3.3 Genome wide identification of MtDof family members

A systematic analysis was carried out to identify non-redundant genes coding TFs belonging to C2-C2 Dof Zn-finger family in the *M. truncatula* genome using the publicly available genomic sequence of *M. truncatula*. A total of 25 *Dof* genes were identified in *M.*

truncatula old genome sequences assembly Mt3.0 (Annexe 1) after BLASTing the DOF1147 protein sequence from the online server of *Medicago truncatula* Genome Sequence Resources.

The presence of *Dof* genes in the latest genome assembly Mt3.5 was also investigated by blast analysis and a total of 25 Dof genes were again identified (Annexe 2), all of these Dof genes having already been annotated as Dof-type except *Medtr7g082600.1* which was marked as an unknown protein. One additional *Dof* gene, i.e. *Medtr2g030030.1*, was recovered by blasting only the Dof DNA binding domain of Dof1147 protein with recent genome assembly Mt3.5. For the final selection, the presence of the Dof domain was considered to be the defining feature and essential for inclusion of a *Dof* gene in the final list. After the manual inspections carried out through multiple sequence alignments of proteins, 4 candidate *Dof* genes were eliminated for further analysis (Annexe 3) because proteins of 2 of these 4 genes, i.e. *Medtr2g030030.1* and *Medtr7g82600.1*, lacked more than 50% of Dof domains and the four cystein residues essential for DNA binding, while the remaining two genes, i.e. *Medtr4g081690.1* and *Medtr4g022220.1*, lacked entire Dof domains despite the fact that these genes have been annotated as Dof-types in the genome annotation assembly.

3.4 New annotations of *M. truncatula Dof* genes

The major problem with the identification of some *Dof* genes in both old and new genome assemblies, i.e. Mt3.0 and Mt3.5, lies in the fact that these assemblies have entirely different gene names as well as gene locations for the same *Dof* genes. So, in order to facilitate future studies and to verify the presence of predicted genes in both the genome assemblies, a sequence comparison was carried out using blast and genome search using CLC Sequence Viewer 6.

The genes from Mt3.0 showing 100% sequence similarity with those from Mt3.5 were identified and were renamed based on their relative position on the chromosomes. A total of 22 *Dof* genes having complete Dof DNA-binding domains and present in both the assemblies were renamed as *MtDof* (*Medicago truncatula Dof*), as shown in Table 5. Our *Dof* gene i.e. *Dof1147* has been renamed as *MtDof2.2* but *Dof1147* would be used throughout the text in the thesis.

3.5 Mapping MtDOF genes on *M. truncatula* chromosomes

All the predicted *MtDOFs* were physically localized on *M. truncatula* chromosomes by CLC Sequence Viewer and Blast program (http://www.medicagohapmap.org/tools/blastform) from Medicago hapmap server and were mapped using online Map Viewer programme from NCBI server (http://www.ncbi.nlm.nih.gov/projects/mapview/). The predicted 22 *MtDof* genes are distributed across seven of eight *M. truncatula* chromosomes. The distribution of *Dof* genes among the 7 chromosomes of *M. truncatula* is shown in Fig. 17.

Compared with other chromosomes, chromosome 2 has the most numbers of *Dof* genes, i.e. five, followed by chromosome 5 with four genes encoding DOF TF. The chromosomes 4 and 7 share three *Dof* genes each while two *Dof* genes are present on each of chromosome 3 and 8. No *Dof* gene was localized on chromosome 1, while two *Dof* genes identified from genomic sequences of BAC libraries that have not yet been attributed to the linkage groups and hence are not placed onto a chromosome.

The physical distribution of *Dof* genes in various other crops has also been reported previously. For example, in the Eudicot model *A. thaliana*, the 36 *Dof* genes were found to be organized among all the five chromosomes with a maximum of nine *Dof* genes on chromosome 1 (Lijavetzky *et al.,* 2003). On the other hand in monocots, 30 rice *Dof* genes were shown to be distributed on 11 out of the 12 chromosomes, with a maximum of 6 *Dof* genes on both chromosomes 1 and 3 (Lijavetzky *et al.,* 2003). Quite recently, 28 genes of *S. bicolor* encoding Dof proteins have been physically localized on 9 chromosomes, with chromosome 1 and 3 showing a distribution of a maximum of 7 and 6 *Dof* genes, respectively (Kushwaha *et al.,* 2011) while no *Dof* gene was reported to be present on chromosome 10.

Table 5 : New annotation and comparison of *M. truncatula Dof* genes in genome assemblies Mt3.0 and Mt3.5. The relative location of *Dof* genes is indicated by start and end position. The name of BAC clones is also written.

Gene new name	Annotation Version 3.5	Start Position (bp)	End Position (bp)	Clone Name	Annotation Version 3.0	Start Position (bp)	End Position (bp)	Clone Name
MtDof2.1	Medtr2g013370.1	3778931	3783848	AC129091.19	Medtr2g015060	4016139	4017481	AC149135.2
MtDof2.2	Medtr2g014060.1	4167122	4168157	AC149801.3	Medtr2g017570	4403997	4405007	AC146855.5
MtDof2.3	Medtr2g014170.1	4205170	4203588	AC149801.3	Medtr2g017650	4441944	4440569	AC146855.5
MtDof2.4	Medtr2g059540.1	18423655	18424544	AC151462.30	Medtr2g074540	17161093	17161772	AC136450.7
MtDof2.5	Medtr2g093220.1	28740812	28743132	AC234096.1	Medtr2g111360	26387661	26389915	AC225206.1
MtDof3.1	Medtr3g077750.1	24743442	24741059	CU062477.6	Medtr3g113950	28735579	28737712	CT954236.4
MtDof3.2	Medtr3g090430.1	30504739	30507030	AC150784.22	Medtr3g125030	31938937	31940032	AC150784.21
MtDof4.1	Medtr4g022370.1	6321773	6324456	AC126019.16	Medtr4g023580	5186896	5188236	AC126019.16
MtDof4.2	Medtr4g082060.1	27681355	27678444	AC140030.6	Medtr4g111370	25504376	25507024	AC202576.10
MtDof5.1	Medtr5g031440.1	13054442	13056859	CR931741.1	Medtr5g032870	13354886	13358741	CR931741.1
MtDof5.2	Medtr5g041380.1	17771949	17773644	CU914136.1	Medtr5g044630	18561316	18562696	CT025840.2
MtDof5.3	Medtr5g041400.1	17776979	17778670	CU914136.1	Medtr5g044650	18566319	18567937	CT025840.2
MtDof5.4	Medtr5g041420.1	17781570	17784480	CU914136.1	Medtr5g044680	1857087	18573780	CT025840.2
MtDof5.5	Medtr5g041530.1	17813535	17816319	CU914136.1	Medtr5g044740	18602937	18605333	CT025840.2
MtDof6.1	Medtr6g012450.1	2931785	2935321	AC139600.16	Medtr6g012960	2585061	2588382	AC166744.12
MtDof7.1	Medtr7g010950.1	2301393	2305397	AC149128.3	Medtr7g010180	2149602	2152776	AC140025.7
MtDof7.2	Medtr7g024670.1	6599803	6597980	AC146757.19	Medtr7g025880	6477938	6479761	AC225096.1
MtDof7.3	Medtr7g086780.1	26404666	26409732	AC158465.3	Medtr7g103580	2328411	23286425	AC137986.27
MtDof8.1	Medtr8g044220.1	11346438	11348897	AY508219.1	Medtr8g047990	10957450	10959909	AC203554.8
MtDof8.2	Medtr8g068210.1	17536849	17537973	AC146709.11	Medtr8g079210	16436932	16437948	AC174360.8
MtDof0.1	AC235757_38.1	157186	155752	AC235757.1	AC235757_43	13474438	13476084	AC235757.1
MtDof0.2	AC235667_13.1	59589	56661	AC235667.4	AC235667_13	20462999	20465450	AC235667.4

3.6 Comparison of DNA-binding Dof domains of MtDofs

The Dof DNA-binding domains of all the DOF proteins present towards the N-terminal region were compared for presence of conserved residues by alignment of complete Dof domain using COBALT from NCBI server (http://www.ncbi.nlm.nih.gov/tools/cobalt/cobalt.cgi). The Dof domain in all the Dof proteins of *M. truncatula* revealed a high level of conservation with 31 out of 52 amino acids being 100% conserved in all 22 proteins (Fig. 18). Another 12 amino acids were very highly conserved with only two substitutions per site.

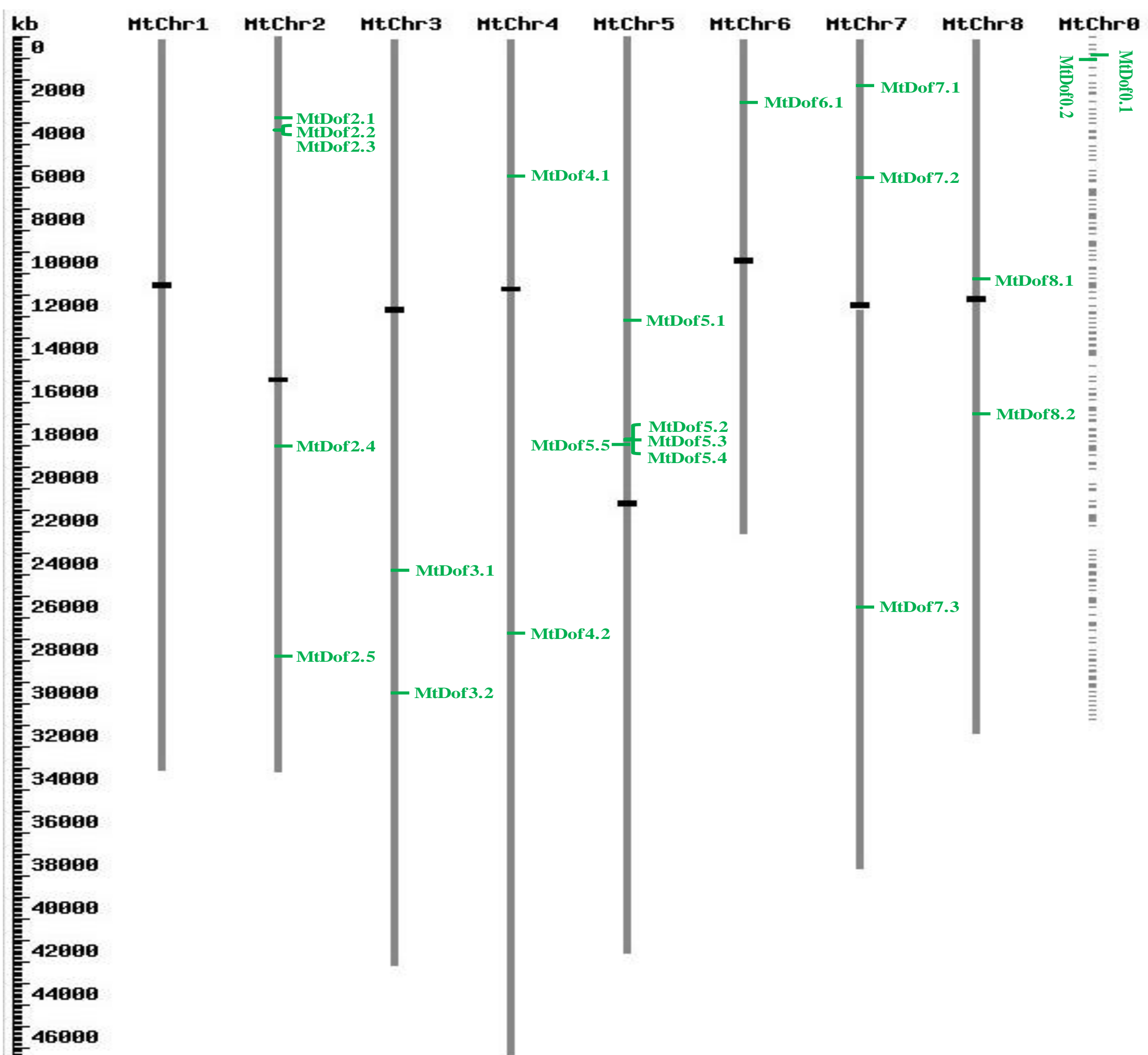


Figure 17 : The localization of 22 *M. truncatula Dof* genes on chromosomes. The chromosome numbers are indicated at the top of each bar. Gene names correspond to those listed in Table 4. The relative position of *MtDofs* and size of chromosome represented using vertical scale. MtChr0 refers to non allocated sequences.

The highly conserved 4 cysteine residues that attach to the Zn^{+} to form Zn-finger are highlighted in Fig. 18.

In Wheat, 23 out of 52 amino acids have been reported to be 100% conserved, while 14 positions in Dof domain were shown to be shared by 2 amino acids (Shaw *et al.,* 2009). Quite recently, Kushwaha and colleagues (2011) have shown the 100% conservation of 25 amino acids in Dof domain of Barley Dof proteins, while 9 amino acids showed variation in only two amino acid residues.

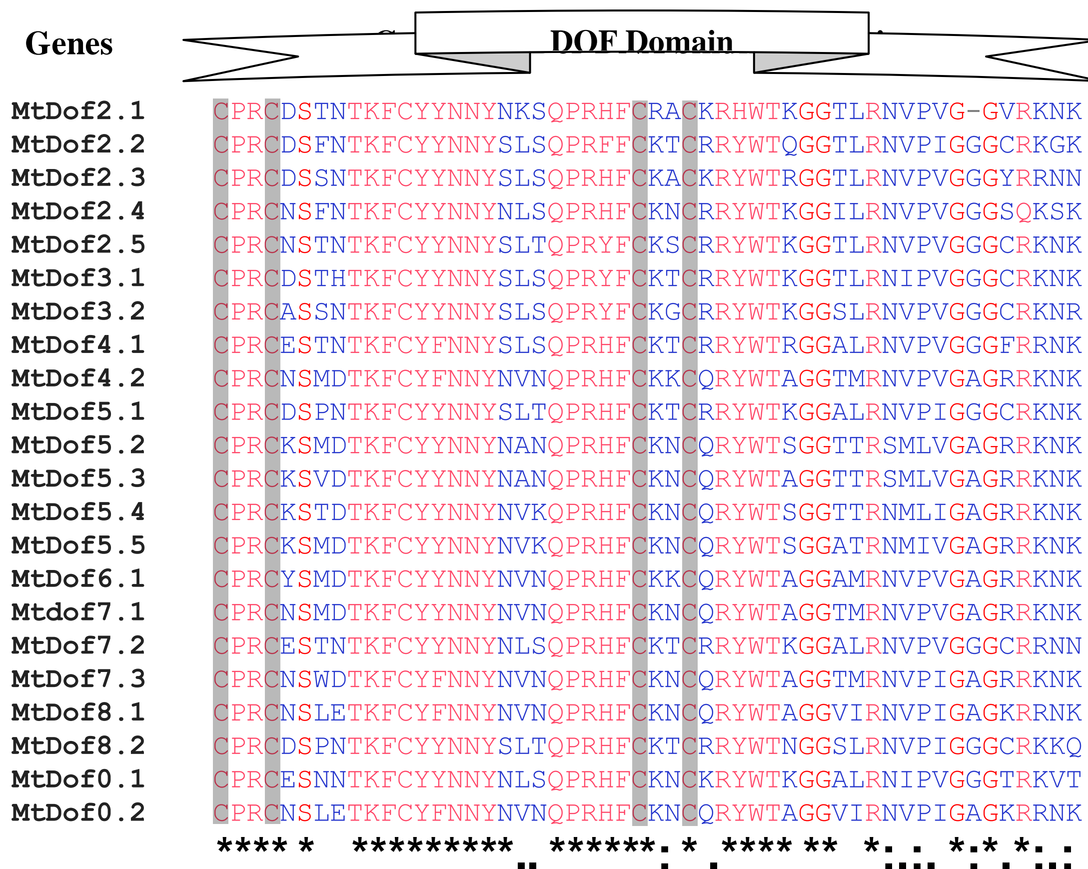


Figure 18 : A comparison of Dof DNA-binding domain among 22 MtDof proteins using COBALT-blast. All the absolutely conserved residues are shown in red while others in blue. The four Cystein core residues are highlighted in grey. The three characters '*', ':' and '.' are used to mark strongly conserved positions according to Clustal X conventions.

3.7 Phylogenetic analysis of Dof proteins

The predicted 22 MtDof proteins were subjected to multiple sequence alignment using ClustalW along with 35 Arabidopsis, 8 pea and 3 PBF family Dof proteins and a phylogenetic

tree was reconstructed using CLC Sequence Viewer software (Version 6.7.1) with neighbour joining method and 1,000 replicates (Fig. 19). Based on the classification of Yanagisawa (2002), a total of seven major groups of Dof proteins were observed in the phylogenetic tree with the participation of MtDofs in all the groups.

The CrDof1 from *Chlamydomonas* being so far the only DOF found in an organism different from land plants displayed a distinct separate position in the phylogenetic tree. CrDof1 with some Dof proteins from moss are believed to be close to the origin of the evolutionary divergence of Dofs in higher plants (Shigyo *et al.,* 2007). Up until now, there is limited information about the relationship between functional diversification and gene multiplication in various crop plants. However, the possibility that members in the same phylogenetic subgroup show redundant, overlapping or related functions cannot be ruled out.

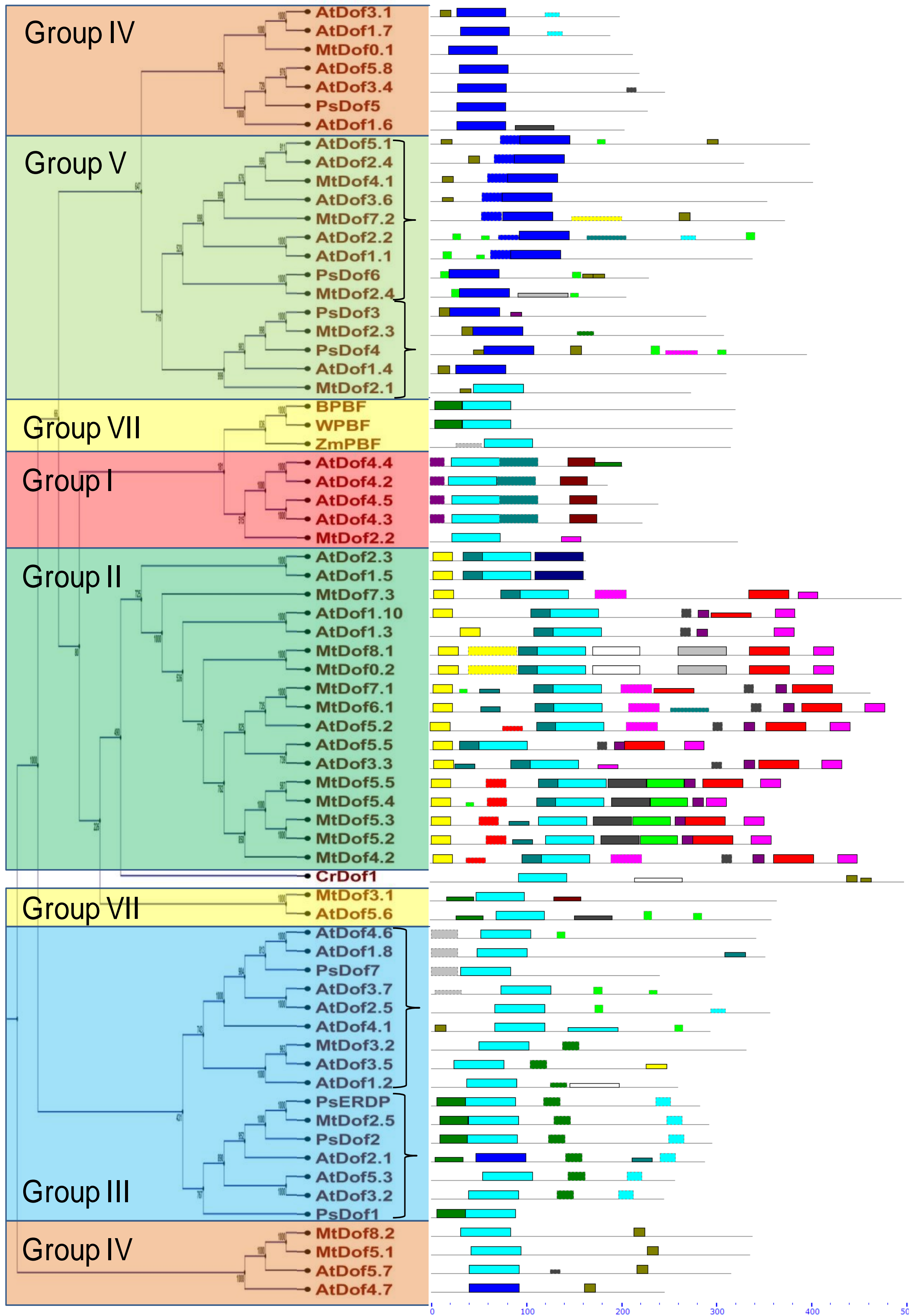
A total of 7 paralogous proteins i.e. 3 from AtDof and 4 from MtDofs have been identified in Group II. Overall, 10 out of 22 MtDof proteins are classified in group II but none of them have yet been characterized for their physiological functions. Among them, MtDof5.2 and MtDof5.3, MtDof8.1 and MtDof0.2 as well as MtDof5.4 and MtDof5.5 are located in pairs close to one another in the phylogenetic tree and in view of their close similarity, probably reflect that they are paralogous. The evolution of these genes seems to be a result of recent duplication events. The sequence similarity among MtDofs located on chromosome 5 in close vicinity suggests that they are tendem repeats. On the other hand, the Arabidopsis group II Dof proteins are encoded by seven Dof genes i.e. AtDof2.3, AtDof1.5 (COG1), AtDof1.10, AtDof1.3, AtDof5.2 (CDF2), AtDof5.5 (CDF1) and AtDof3.3 (CDF3). The physiological functions of most of these genes have already been described. For example AtDof1.5 encoding COG1 is involved in the negative regulation of phytochrome signalling pathways (Park *et al.,* 2003). AtDof5.5, AtDof5.2 and AtDof3.3 encoding CDF1, CDF2 and CDF3 proteins have been proposed to be involved in the regulation of photoperiodic dependent flowering (Imaizumi *et al.,* 2005). Hence, Mt group II Dof domain proteins might be implicated in light regulated mechanisms, although the functional characterization of MtDof domain proteins will be necessary to fully support this possibility.

Group VII comprised of AtDof5.6 and its ortholog MtDof3.1, as well as all orthologs PBF proteins from barley (BPBF), wheat (WPBF) and maize (ZmPBF). In other studies, these PBF proteins have been shown to share a high degree of homology in their protein sequences and have also been reported to form a distinct phylogenetic group (Yanagisawa, 2002). The

PBF proteins are specifically expressed during the grain filling stage and have been demonstrated to regulate the expression of seed storage protein genes in the endosperm of monocots by interacting with signature motifs present in their promoter regions (Vicente-Carbajosa *et al.,* 1997, Mena *et al.,* 1998). AtDof5.6, which encodes the HCA2 protein, has been shown to be expressed in the vascular tissues as well as in seeds and has been implicated in cambium formation and vascular tissue development (Guo *et al.,* 2009) while its ortholog in *M. truncatula* i.e. MtDof3.1 has not yet been characterized.

Group III contains 16 Dof proteins, i.e. 10 from Arabidopsis, 4 from pea and 2 from *M. truncatula*. In this group, three paralogous pairs of AtDof proteins i.e. AtDof5.3 & AtDof3.2, AtDof3.7 & AtDof2.7 and AtDof4.6 & AtDof1.8 have been observed. The similarity in these proteins suggests that they might evolved as a result of gene duplications in *A. thaliana*. The AtDof3.7 and AtDof2.5 genes encode DAG1 and DAG2 proteins which have been involved in light-dependent seed germination (Gualberti *et al.,* 2002). Recently, DAG1 has been shown to negatively regulate GA biosynthetic pathway during seed germination (Gabriele *et al.,* 2010). The pea Dof genes *PsDof1* and *PsERDF* encode elicitor-responsive Dof factors which are involved in the defence mechanism of plants (Seki *et al.,* 2002). The Phylogenetic tree demonstrates that PsERDF pairs with MtDof2.5 which suggests that these are orthologs and have separated as a result of speciation.

Group V is quite heterogeneous, with 6 Dof proteins from Arabidopsis, 5 from *M. truncatula* and 3 from pea. The MtDof2.3 and MtDof2.4 in this group share homology with pea Dof proteins PsDof3 and PsDof6 respectively suggesting that these genes are orthologs and might have evolved in a legume specific manner. Group I contains only one legume Dof protein i.e. MtDof2.2. The *MtDof2.2* encodes Dof protein also named as Dof1147 which has been shown to express in developing seed endosperm during the seed filling phase in *M. truncatula*. Group IV mainly contains the members of Arabidopsis Dof family with I from pea (PsDof5) and one from *M. truncatula* (MtDof0.1). Most Dof proteins in this group have not yet been characterized. Recently, AtDof1.7 from group IV has been shown to regulate the fatty acid metabolic pathway in *A. thaliana* seed (Yin *et al.,* 2011).


Group IV
AtDof3.1
AtDof1.7
MtDof0.1
AtDof5.8
AtDof3.4
PsDof5
AtDof1.6
Group V
AtDof5.1
AtDof2.4
MtDof4.1
AtDof3.6
MtDof7.2
AtDof2.2
AtDof1.1
PsDof6
MtDof2.4
PsDof3
MtDof2.3
PsDof4
AtDof1.4
MtDof2.1
Group VII
BPBF
WPBF
ZmPBF
Group I
AtDof4.4
AtDof4.2
AtDof4.5
AtDof4.3
MtDof2.2
Group II
AtDof2.3
AtDof1.5
MtDof7.3
AtDof1.10
AtDof1.3
MtDof8.1
MtDof0.2
MtDof7.1
MtDof6.1
AtDof5.2
AtDof5.5
AtDof3.3
MtDof5.5
MtDof5.4
MtDof5.3
MtDof5.2
MtDof4.2
CrDof1
Group VII
MtDof3.1
AtDof5.6
AtDof4.6
AtDof1.8
PsDof7
AtDof3.7
AtDof2.5
AtDof4.1
MtDof3.2
AtDof3.5
AtDof1.2
PsERDP
MtDof2.5
PsDof2
AtDof2.1
AtDof5.3
AtDof3.2
PsDof1
Group III
Group IV
MtDof8.2
MtDof5.1
AtDof5.7
AtDof4.7
0
100
200
300
400

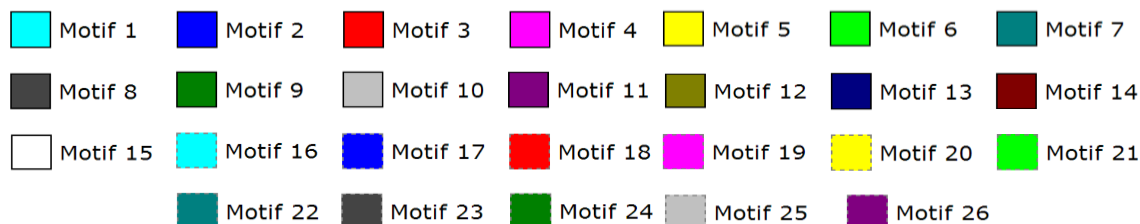


Figure 19 : Phylogenetic relationship among MtDof members with Arabidopsis, pea and PBF Dof proteins. 22 Dof proteins from *M. truncatula*, 35 Dof proteins from *A. thaliana*, 8 Dof proteins from pea, 3 Dof proteins from wheat, barley and maize PBF, and one Dof protein from *Chlamydomonas* used as an outgroup. An unrooted tree is constructed by Neighbour-Joining method using CLC Squence Viewer (version 6.7.1). The resulting groups are indicated as I-VII. All the Dof protein sequences are used for conserved domain and duplication analysis using online MEME server. 26 conserved domains of different lengths are depicted on the protein map next to the phylogenetic tree. While the motif codes are indicated below the tree and the respective sequences of the domains are shown in Table 5.

Table 6: Multilevel consensus sequences for the MEME defined domains observed among different Dof proteins from *Medicago*, *Arabidopsis*, Wheat, Barley, Maize and *Chlamydomonas*.

Motif	Multilevel consensus sequence of conserved domains	Length	E value	No. of sites
1	LPCPRCNSTNTKFCYYNNYSLSQPRYFCKNCRRYWTAGGTLRNVPVGGGCRK	52	6.2e-2236	50
2	LKCPRCDSTNTKFCYYNNYNLSQPRHFCKACRRYWTRGGALRNVPVGGGCRR	52	1.1e-823	19
3	DEGETxEEDSxEKSVLVPKTLRIDDPEEAAKSSIWSTLGIKNE	43	2.6e-182	12
4	SPVLQANPAALSRSLNFHESI	21	6.7e-093	13
5	TKDPAIKLFGRTIPFPTNTDV	21	1.2e-092	17
6	QCFFPQQFPWNPAMCYPVSFQPNIAYYGGCLVPSWSVQPI	40	3.2e-070	4
7	TIEDETTSSSQEKTLKKPDKI	21	2.6e-055	15
8	DASHNCQMSTVLTFGSDSPIMSSTSLAKKMNVGSDEETFDK	41	6.1e-036	4
9	QMSSQSSMENMLGCSKEEQEKKPKPQPEQ	29	5.1e-035	6
10	HYYSAPPWTYPCWNPVAFKSDNITSSPATMMTVEVPMTPSYWGCMPNWVGQM	52	4.2e-030	2
11	PNSPTLGKHSR	11	2.1e-028	12
12	QDLHQQQQQQQ	11	5.0e-027	15
13	RVVVGMLGDGNGVRQVELINGLLVEEWQHAAAAAHGSFRHDFPMKRLRCYSD	52	1.0e-026	2
14	MDRLAFGDESFEQDYYDVGSDDLIVNPLI	29	2.3e-026	4
15	CQVPVTLDAVPVIHIDSKEEVPLSESVETVLNLKGHRKIEMDSSTVKEDIE	51	2.9e-024	2
16	ESRVFWGFPWQMNGG	15	6.5e-024	6
17	IRPGSMADRARMANIPLPEAA	21	3.2e-023	6
18	EHDAKRYKETSRKELTSVLDY	21	1.5e-022	4
19	HHHSLGCNATVLTFGSDSPLCESMASVLNLADK	33	6.4e-020	5
20	HQQSIMNGLEESTNADVEVSSKQVNPQEDANAIPNMYNTGTLKPVTKAVHRT	52	1.1e-018	2
21	NHHHHHLH	8	1.2e-018	13
22	KRAKIDQPSVAQMVSVEIQPGNHQPFFNVQENNDFVGSFGA	41	1.6e-018	4
23	FYPPPAYWGC	10	5.9e-016	8
24	SIDLALAFARLQKQHLG	17	3.8e-013	8
25	MDTAQWPQEIVVKPLAEIVTNTCPKPQ	27	8.9e-013	3
26	MDNLNVFANEDNQVN	15	5.1e-012	4

3.8 Analysis of conserved domains and gene duplications

The complete sequence of Dof proteins used for construction of phylogenetic tree was subjected to MEME server for analysis of conserved domain and duplication events during the evolutionary process. After following the parameters described in M&M (cf. § 2.2.1.7), a total of 26 different conserved domains have been identified in the 69 Dof sequences (Table 6). The identified domains ranged from 8 to 52 residues in length (Table 6).

Some of the groups shown in phylogenetic tree contain single or double conserved domains, while others may contain multiple conserved domains. Group IV and VII contain 2 conserved domains each while I and III contain 4-5 conserved domains of different residues. Maximum number of conserved domains was observed in group II i.e. 10-12 conserved domains.

All Dof proteins in group II, either MtDofs or AtDofs, contain 22 amino acid domain at the extreme N terminal region near the Methionine (Motif 4 in Table 6) apart from Dof binding domain. The MtDofs in this group (i.e. MtDof4.2, MtDof5.2, MtDof5.3, MtDof5.4, MtDof5.5, MtDof6.1 and MtDof7.1) share the 11 amino acids long conserved domain (Motif11: PNSPTLGKHSR) near the C-terminal of the protein. Interestingly, *MtDof5.2*, *MtDof5.3*, *MtDof5.4* and *MtDof5.5* genes are localized on chromosome 5 of *M. truncatula* and, present in close vicinity, are likely to share the same loci, as is evident from Table 6. When further investigated, these proteins contain a highly conserved 22 amino acid sequence EHDAKRYKETSRKELTSVLDY (Motif 18 in Table 6) upstream of the Dof domain, as well as motif 6 and 8. The presence of multiple conserved domains in MtDof5s proteins suggests that the genes encoding these proteins originate from the same duplication events on the same chromosome.

In group I, the AtDof proteins share multiple conserved domains (Motif 14, 22, 26 in Table 6) with each other. These genes encoding these proteins are also localized on the same chromosomes. In *M. truncatula*, 9 Dof proteins i.e. MtDof0.2, MtDof4.2, MtDof5.2, MtDof5.3, MtDof5.5, MtDof6.1, MtDof7.1, MtDof7.3 and MtDof8.1 share the 43 residues conserved domain (motif 3) near the C terminal region of Dof proteins as well as motif 4 at the terminal region of the respective proteins. The presence of conserved domains among different Dof proteins especially in the same group is likely to reflect gene duplication events during evolution and their related functions. Recently, gene expression studies in *A. thaliana*

have revealed that most of the closely related genes in the Dof phylogenetic tree had redundant expression patterns in the different tissues (Moreno-Risueno *et al.,* 2007).

3.9 *In silico* motif analysis of putative target gene promoters

Using Affymetrix, 195 genes have been isolated whose expression has increased during seed development in the EMS109 *dof*-mutants in our research unit (M. Noguero and J. Verdier, pers. comm.). Among them, 34 overexpressed genes were identified having expression more than four times. Based on Affymetrix codes, the sequences of 29 of them were retrieved from *M. truncatula* genome. Using the CLC Sequence Viewer, the 2kb upstream promoter region from transcription start site of these genes were isolated. All the motifs for attachment of TFs to the promoters are generally present in this 2kb upstream region. The dataset was then subjected to PLACE (Higo *et al.,* 1998, 1999) and PlantCARE database (Lescot *et al.,* 2002) for identification of already experimentally described motifs. The analyses have shown that the promoter regions of all the over as well as under-expressed genes contain multiple copies of the basic elements required for promoter identity i.e. TATA-box and CAAT-box (Casimiro *et al.,* 2008) and recruited by basal transcription apparatus before the start of transcription. In order to evaluate the significance of predicted motifs in the promoter regions, a controlled dataset was prepared by randomly choosing the thirty sequences of 2kb each. These control sequences were also subjected for analysis using PlantCARE and PLACE.

The analysis also revealed the presence of multiple binding sites of various TFs and hormones implicated in key developmental and biological processes. Based on the presence of cis- regulatory elements, the binding sites were divided into four major categories i.e. light responsive elements, elements responsible for endosperm-specific gene expression, hormone responsive elements and stress responsive elements including biotic and abiotic stress (Table 7).

All the promoters contain the over-represented sites of light-responsive elements which may be responsible for temporal expression of these genes. Various cis-regulatory elements have also been identified that direct the expression of genes in endosperm. Almost all the gene promoters have Skn_1 motif required for endospermic expression of these genes during seed development except for *Medtr4g107010.1* for which no such element was found. 12 of 29 genes also contain endosperm-specific GCN4 motifs while 6 of 29 genes contain

additional RY elements which are characteristic for endosperm-specific expression of genes during seed development.

Table 7 : Classification and presence of identified motifs in different gene promoters

Annotated genes	Light responsive elements	Elements for endosperm specific gene expression	Hormone responsive elements	Stress responsive elements
Medtr6g012830.1	LS7, Box I, Box 4, AE-box, G-box, G-Box, I-box, GA-motif, GT1 motif	Skn-1_motif, O2-site, GCN4_motif,	ABRE, ERE, TGA-box, EIRE,	CGTCA-motif, MBS, CCAAT-box HSE
Medtr6g012670.1	Box I, Box 4, G-box, G-Box, I-box, ATCT motif, CATT motif, GA-motif, GT1-motif	Skn-1_motif	ABRE, ERE, TGA-element	MBS, TCA-element, Box-W1, CCAAT-box, ARE
AC233577_31.1	Box I, MNF1, Box 4, G-box, I-box, ATCT motif, GT1-motif, TCT-motif	Skn-1_motif, GCN4_motif,	ERE,	CGTCA-motif, MBS, TCA-element, Box-W1, GC-motif, ARE
Medtr5g007230.1	Sp1, Box 4, G-box, G-Box, ATCT motif, GA-motif, GT1-motif	Skn-1_motif, O2-site, GCN4_motif	ABRE	CGTCA-motif, MBS, TCA-element, HSE
Medtr3g099460.1	Sp1, Box1, ATCT motif,GA-motif, GT1-motif	Skn-1_motif, RY-element, O2-site,	P-box	CGTCA-motif, MBS, TCA-element,
Medtr7g014620.1	MNF1, Box 4, G-box, G-Box, GA-motif, TCT-motif	Skn-1_motif	ABRE	CGTCA-motif, MBS, HSE, ARE
Medtr5g071560.1	Sp1, Box1, Box 4, G-box, G-Box, TCT-motif	Skn-1_motif	ABRE, P-box	CGTCA-motif, MBS, HSE, ARE
Medtr7g023560.1	MNF1, SP1, Box I, Box II, Box 4, G-box, G-Box, CATT-motif, GT1-motif	Skn-1_motif, GCN4_motif,	ABRE	MBS, TCA-element, HSE, ARE
Medtr4g107010.1	SP1, Box I, Box II, Box 4, G-box, G-Box, I-bocx, GA-motif, GT1-motif	NILL	ABRE, ERE	CGTCA-motif, C-repeat/DRE, TCA-element, Box-W1, ARE
Medtr1g083580.1	SP1, Box I, Box 4, AE-box, I-box, ATCT-motif, CATT-motif, TCT-motif	Skn-1_motif, O2-site, GCN4_motif	ERE, P-box,	HSE
Medtr2g098790.1	Box I, Box 4, ATCT-motif, GT1-motif	Skn-1_motif		CGTCA-motif, HSE
Medtr4g020450.1	NF1, SP1, Box I, Box 4, AE-box, G-box, I-box, ATCC-motif, CATT-motif, GT1-motif, TGG-motif,	Skn-1_motif, RY-element, O2-site,	ERE, P-box, TGA-element	CGTCA-motif, MBS, TCA-element, CCAAT-box, HSE, ARE
Medtr3g065350.1	Box 4, G-box, G-Box, I-box, CATT-motif, GA-motif, GT1-motif	Skn-1_motif, O2-site, GCN4_motif	NILL	MBS, HSE, ARE
Medtr2g099950.1	SP1, Box I, G-box, G-Box, GT-motif	Skn-1_motif, RY-element, GCN4_motif	ABRE, ERE, P-box	CGTCA-motif, MBS, TCA-element, Box-W1, ARE
Medtr1g106420.1	SP1, Box I, Box 4, G-box, G-Box, ATCT-motif, GA-motif, GT1-motif,	Skn-1_motif, RY-element,	ABRE, ERE, TGA-element, AuxRR-core	CGTCA-motif, MBS, ARE
Medtr5g071360.1	Box I, Box 4, G-box, G-Box, GA-motif, TCT-motif,	Skn-1_motif	ABRE, ERE, TGA-element, AuxRR-core	CGTCA-motif, MBS, TCA-element, HSE, ARE
AC235753_1.1	SP1, Box 4, AE-motif, G-box, G-Box, I-box, ATCC-motif, GA-motif, GTGGC-motif,	GCN4_motif	P-box	CGTCA-motif, MBS, HSE, ARE
Medtr4g060780.1	SP1, Box 4, G-box, G-Box, I-box, ATCT-motif, GT1-motif,	Skn-1_motif, O2-site, GCN4_motif	ABRE	MBS, TCA-element,
Medtr7g075540.1	SP1, Box 4, G-box, G-Box, I-box, ATCT-motif, GT1-motif	Skn-1_motif	ABRE, ERE, P-box	MBS, TCA-element,
Medtr3g046650.1	Box I, Box 4, G-box, GT1-motif	Skn-1_motif	ERE	MBS, CCAAT-box, ARE
Medtr1g102390.1	SP1, Box I, Box 4, G-box, G-Box, I-box, ATCT-motif, TCT-motif	Skn-1_motif	ABRE, AuxRR-core	Box-W1, HSE, ARE
mtr.46259.1.s1	SP1, Box I, AE-box, G-box, G-Box, I-box, ATCT-motif, TCT-motif	Skn-1_motif, GCN4_motif	ERE	TCA-element, Box-W1, HSE, ARE
Medtr8g102350.1	SP1, Box I, Box 4, G-box, G-Box, I-box, ATCT-motif, CATT-motif	Skn-1_motif, O2-site, GCN4_motif	NILL	TCA-element,
mtr.20847.1.s1	SP1, Box I, AE-box, ATCT-motif	Skn-1_motif, RY-element, O2-site, GCN4_motif	ERE P-box	MBS, TCA-element, Box-W1, CCAAT-box, HSE, ARE
Medtr8g093440.1	Box I, Box 4, G-box, G-Box, I-box, GA-motif,	Skn-1_motif, O2-site,	ABRE	CGTCA-motif, TCA-element, HSE, ARE
Medtr3g098910.1	SP1, Box I, AE-box, ATCT-motif, CATT-motif	RY-element, GCN4_motif	ERE	MBS, TCA-element, Box-W1, HSE, ARE
AC235677_4.1	SP1, Box I, G-box, G-Box, I-box, ATCT-motif, CATT-motif, GT1-motif, TCT-motif	Skn-1_motif, O2-site,	ABRE, P-box	CGTCA-motif, MBS, TCA-element, Box-W1, HSE, ARE
Medtr7g065150.1	Box I, Box 4, G-Box, G-box, AE-box, I-box, ATCT-motif, TCT-motif, ATCC-motif, GA-motif, GT1-motif	Skn-1_motif, O2-site,	ABRE	MBS, TCA-element, CCAAT-box, HSE, ARE
Medtr7g085220.1	MNF&, Box I, G-Box, G-box, ATCC-motif, ATCT-motif, CATT-motif, TCT-motif	Skn-1_motif	ABRE, TGA-element, AuxRR-core	CGTCA-motif, MBS, HSE, ARE

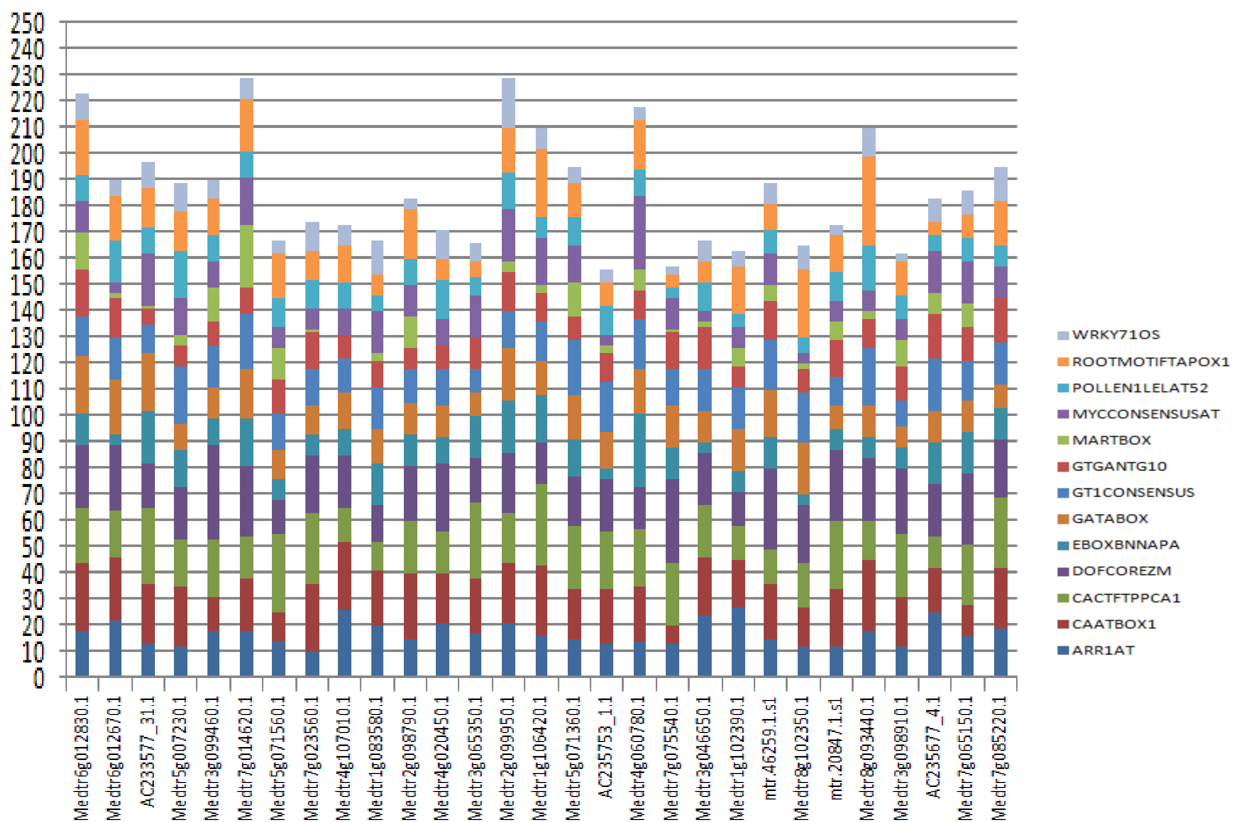


Figure 20 : The presentation of 13 over-represented cis-regulatory elements found in the promoters of genes overexpressed in EMS109 *dof* mutants. The motif frequency of these over-represented motifs varied from 10-30. The gene names are indicated on the X-axis while motif frequency is on the Y-axis. The colour code of motifs is indicated on the righthand side of graph. The AAAT motif encoded by DOFCOREZM element has the maximum frequency in the putative target gene promoters.

A large number of hormone responsive signatures have also been identified in the promoter regions of 29 overexpressed genes (Table 7). Among them, ABRE motifs which are involved in the abscisic acid responsiveness have the highest frequency in the promoters. This highlights the importance of ABA-regulated gene expression during seed development. Various P box and GARE-motif specific for Gibberellin responsiveness, TGA-elements and AuxRR-core motifs involved in auxin responsiveness (Hagen and Guilfoyle, 2002) have also been identified in our dataset. The presence of these motifs in the promoter regions is consistent with the major role of phytohormones during plant seed development (Gutierrez *et al.*, 2007) especially in *M. truncatula*.

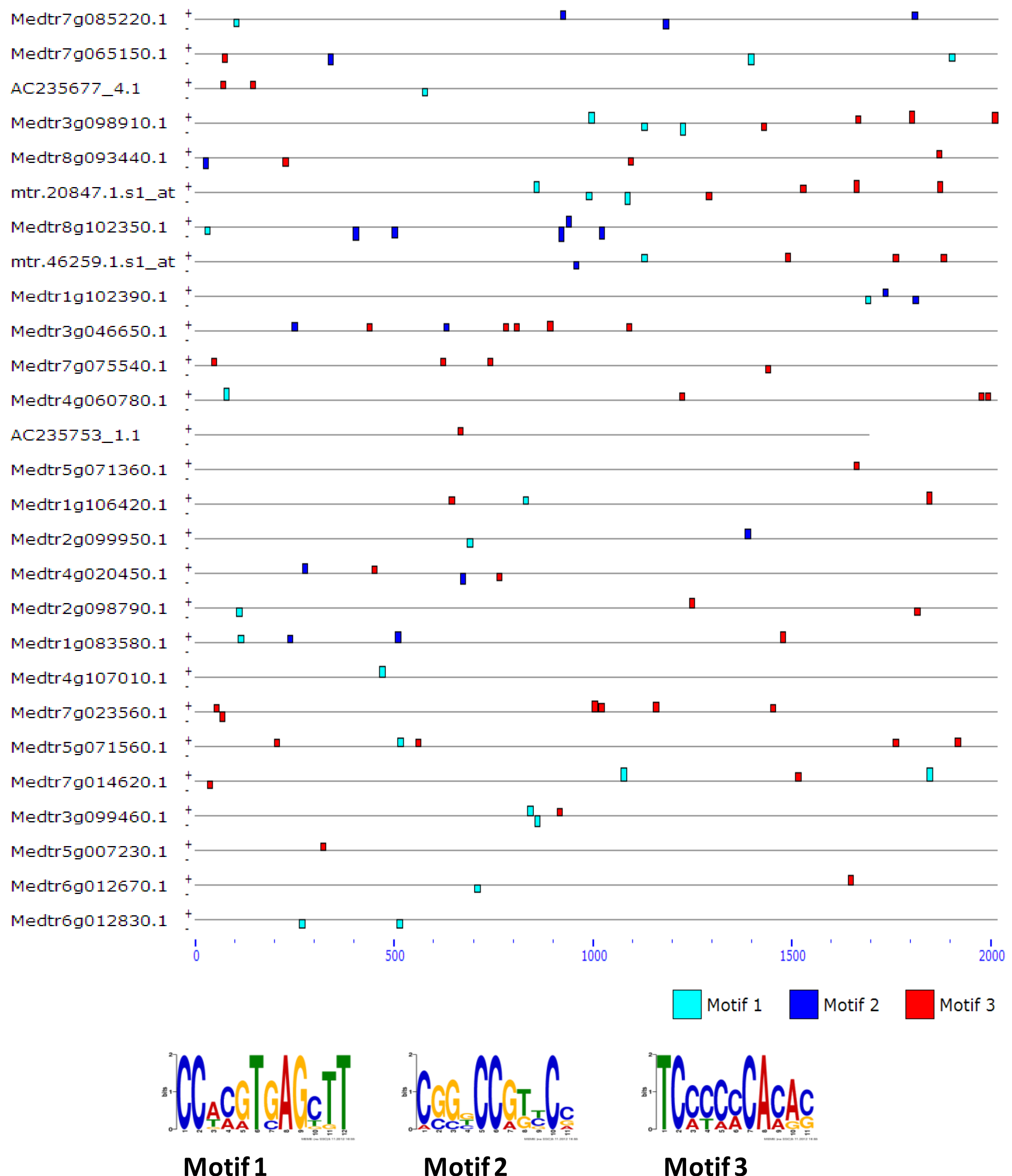


Figure 21: Schematic representation of various MEME-based motifs in 2KB promoter region of genes overexpressed in EMS109 *dof* mutants. The presence of motifs on + or – strand is also indicated. The motif codes and respective sequences are indicated below the figure.

Various motifs that act as responsive elements during cold, dehydration or heat stress have also been identified from the promoters' dataset. Among them are also the CCAAT-box motifs that are binding sites of MYB transcription factors during biotic stress, as well as Box-W1 motif that act as fungal elicitor-responsive element (Table 7). Apart from these motifs, a large quantity of motifs from our promoters' dataset has also been identified using PLACE database. The analysis conducted through PLACE has shown the over-presentation of DOFCOREZM that encodes AAAG motif which is the core recognition site required for DNA binding of Dof proteins. The presence of these *cis*-regulatory elements in the promoter region of over and under-expressed genes suggests important roles of Dof TFs in hierarchical regulatory networks controlling *M. truncatula* seed development. The13 motifs that are over-represented in the promoters are depicted in Fig. 20. The frequency of these motifs ranged from 10-30 with the Dof motif having the highest frequencies i.e. 25-30 sites per sequence (Fig. 20).

MEME motif discovery tool was employed for the identification of novel motifs from the promoters. Three novel motifs of 12 nucleotides have been identified by MEME (Fig. 21). In order to ascertain the presence of already known cis regulatory elements, the consensus sequences of these novel motifs were directed to TOMTOM server. The analysis revealed binding sites of Zn-Dof i.e. AAAG (Yanagisawa and Schmidt, 1999), bZip (Martínez-García *et al.,* 1998), Myb (Hosoda *et al.,* 2002), MADS (Huang *et al.,* 1995), HMG (Webster *et al.,* 1997) and AP2 MBD-like (Niu *et al.,* 2004), ABI4 (Niu *et al.,* 2004) transcription factors (Annexe 4) whereby some of these TFs have already been reported to play an active role in different responses and developmental stages. MEME-based analysis of the control dataset revealed some consensus sequences. When submitted to TOMTOM server for identification of motifs in these sequences, no already known *cis*-regulatory elements was identified.

In this study, multiple Dof binding sites have been reported in the promoter regions of 29 genes which are overexpressed in the EMS109 *dof* mutants. It can thus be speculated that *Dof1147* might be implicated in the regulation of these genes by repressing their expression in developing seeds of a normal plant. Once the DOF1147 protein is truncated due to the presence of a stop codon in the Dof DNA-binding domain, it allows the overexpression of otherwise repressed genes.

Chapter 4: *Agrobacterium-tumefaciens* mediated transformation of *M. truncatula* and expression studies of DOF1147

4.1 Introduction

Plant genetic transformation has been reported as a powerful experimental tool for investigating/validating gene function particularly in relation with plant physiology mechanisms and processes that have not been resolved so far using other biochemical approaches.

In our research unit, various *dof1147* mutants have been identified from a mutagenized population of *M. truncatula*, created using EMS by TILLING (LeSignor *et al.,* 2009). Among them, a *dof1147* homozygous mutant named EMS 109, having a stop codon mutation in the DOF domain, resulting in a truncated DOF protein, has shown premature seed abortion at 8-10 DAP. Many of the pods on the homozygous mutant plant also abort, whilst the rest are much smaller than those on WT plants. The plants of this EMS109 mutant only set seeds at the end of their life cycle, and such seeds are significantly smaller, with a 75% reduction in weight as compared to wild type, and fewer in number. Upon germination, the majority of the mutant embryos exhibit a particular phenotype of fused cotyledons resulting in seedling mortality. Other *dof1147* mutant alleles identified from a *Tnt1*-retrotransposon insertional mutant population of *M. truncatula* (Tadege *et al.,* 2008) are also being characterized in our research team. As a part of my thesis, we are interested in preparing different transgene constructions to be used for expression studies of the promoter and the coding sequence of *DOF1147* in *M. truncatula*.

In plants, promoter expression or protein subcellular localization studies using GUS or GFP fusion proteins have been widely adopted to investigate the sites of gene expression and the sites of protein targeting. The use of conventional restriction-digestion and ligation reactions to prepare constructs for such studies often poses limitations, due to the presence of unwanted restriction sites within the gene of interest, for exemple, and the Gateway® cloning technology has recently proved to be an excellent alternative to prepare constructs used for functional genomics. It enables the transfer of DNA fragments among different vectors in the presence of especially engineered enzymes while maintaining orientation and reading frame of the inserted gene (Karimi *et al.,* 2002). Various Gateway® compatible cloning vectors are now available to prepare constructs containing the transgene in different orientations (Curtis and Grossniklaus, 2003; Kumar *et al.,* 2012). Such constructs can later be used for stable transformation as well as transient gene expression.

In this work, we have studied the expression patterns of promoter and coding sequences of *DOF1147* in *M. truncatula* using a set of transgene constructions. For this purpose, we employed stable *Agrobacterium tumefaciens*-mediated genetic transformation system as well as transient gene expression by agroinfiltration.

4.2 Results

4.2.1 Use of Gateway cloning technology for preparation of constructs

The correct orientation of the DNA sequence in the constructed plasmid is a prerequisite for further downstream applications. For this purpose, Gateway compatible cloning vectors (Invitrogen) have been used to prepare transcriptional and translational GFP reporter fusion constructs. The genomic DNA from *M. truncatula* genotype A17 was used as a template for amplifications of all the DNA fragments described afterwards.

4.2.1.1 Preparation of promoter reporter construct (*pdof::gfp*)

In order to investigate the expression pattern of *Dof1147* during *M. truncatula* seed development, a promoter::reporter construct was prepared in which the green fluorescent protein *(GFP)* reporter gene was driven by the *DOF1147* promoter. For this purpose, a 1.2 kb upstream promoter sequence of *DOF1147* was amplified with primers containing gateway-specific *AttB* adapters (Table 3) at both ends. The *AttB* adapter-containing PCR amplified sequence was then purified by Wizard® SV Gel and PCR Clean-Up System (Promega). It was then cloned into donor vector i.e. pENTR™ (Fig. 14) by BP recombination reaction to prepare the Entry clone, which was allowed to grow overnight at 37°C on LB media with kanamycin selection. Plasmid DNA from the well isolated colonies was extracted using a Midiprep kit (Invitrogen) and sent for sequencing. The entry clones showing no mutation were selected for an LR recombination reaction. Following the LR reaction, the promoter sequence was finally cloned into the pMDC107 destination vector (Curtis and Grossniklaus, 2003) to give the final expression clone. The sequenced expression construct i.e. *pdof::gfp* contains the *DOF1147* promoter sequence fused to the N-terminal of histidine-tagged GFP6 gene which is followed by the *nos* terminator (Fig. 22).

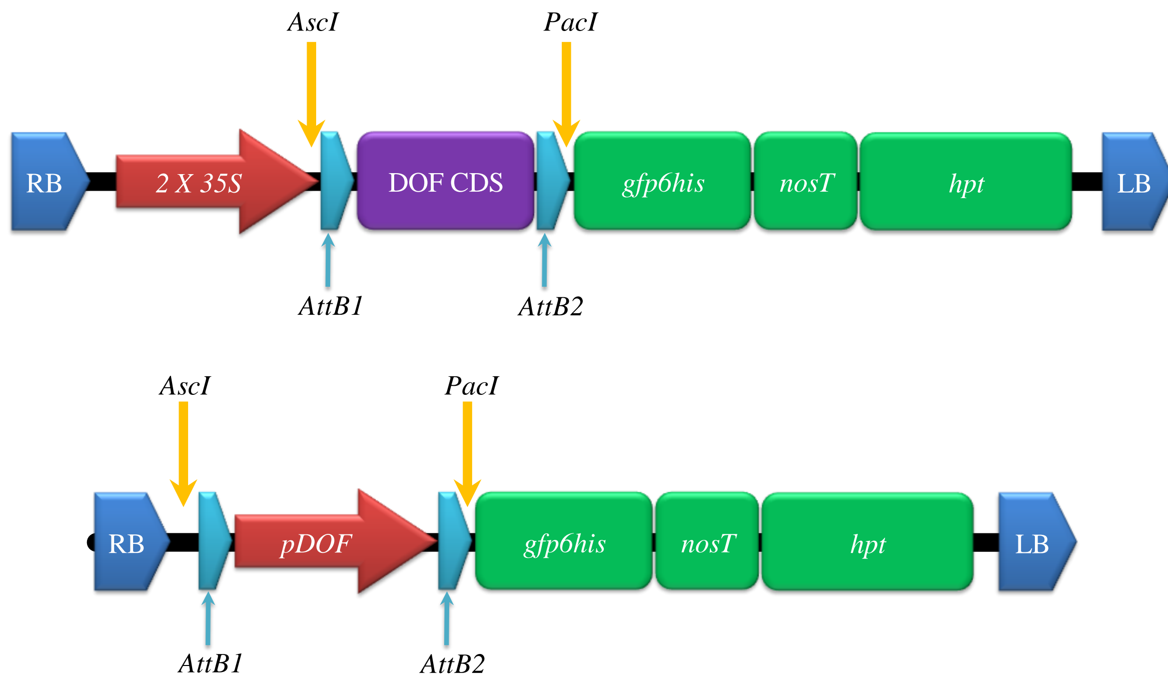


Figure 22. Schematic map of the T-DNA region of constructs: A) T-DNA region of *p35S::dof-gfp* translational fusion construct, B) T-DNA region of *pdof::gfp* promoter reporter construct. *LB* Left border, *RB* right border, *35S* CaMV 35S promoter, *nos*T, nos terminator, *attB1 and AttB2* DNA recombination sequence, *hptII* hygromycin resistance gene, *gfp6his* enhanced green fluorescent protein gene with histidine tag, *AscI* and *PacI*, restriction enzymes.

4.2.1.2 Preparation of translational fusion construct *(p35S::dof-gfp)*

A translational fusion construct was prepared in order to monitor the transcriptional activity of DOF1147 protein by its subcellular localization. Briefly, *DOF1147* coding sequence containing gateway-specific *AttB* adapters at both ends was produced by PCR amplification using the primers given in Table 3. After the purification of PCR product from the gel by employing Wizard® SV Gel and PCR Clean-Up System (Promega), it was then cloned into donor vector i.e. pENTR™ to prepare the Entry clone by BP reaction. The *DOF1147* cds from the sequenced entry clone showing no mutation was then finally cloned into pMDC83 destination vector by LR recombination reaction to give the expression clone. So the final expression construct i.e. *p35S::dof-gfp* contains the *DOF1147* cds fused in frame to the N-terminus of the histidine-tagged GFP6 in expression vector pMDC83 (Curtis and Grossniklaus, 2003) followed by the *nos* terminator, both under the control of dual *CaMV35S* promoter (Fig. 22).

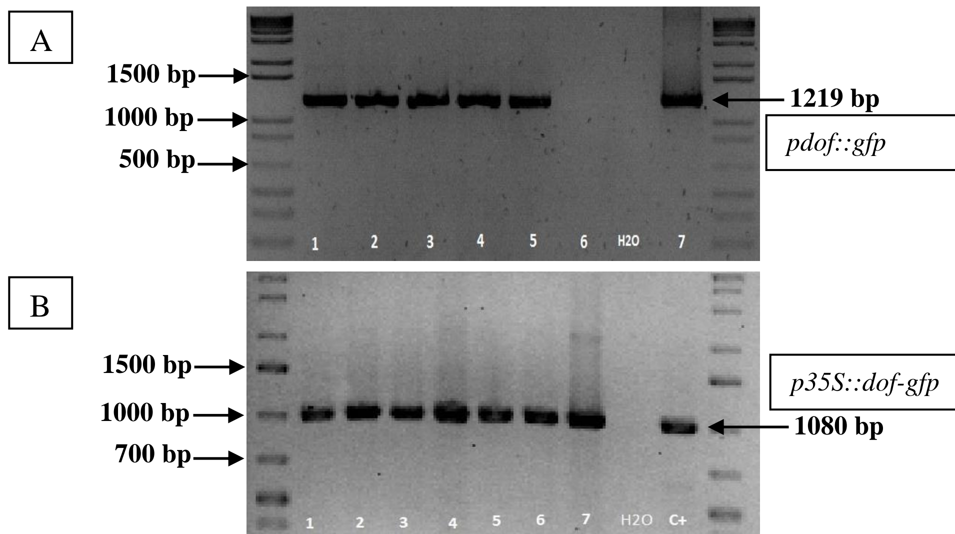


Figure 23. PCR verification of transformed *A. tumefaciens* colonies by using sequence specific primers. (A) 1219 bp amplified from *pdof::gfp* construct in LBA4404 (B) 1080 bp fragment amplified from *p35S::dof-gfp* construct in EHA105. Number indicates different colonies, C+ is positive control i.e. plasmids pMDC107 (A) & pMDC83 (B) and, H_20 is the blank

4.2.2 Transformation of expression clones into *A. tumefaciens*

Based on sequencing data, expression clones having the right insert sequence with no mutation were selected and used for downstream applications. *A. tumefaciens* EHA105 cells transformed with *p35S::dof-gfp* plasmid using heat shock methods were allowed to grow in the dark at 28°C for 48 hours on YEB media with kanamycin selection. On the other hand, *pdof::gfp* plasmid was transformed into *A. tumefaciens* LBA4404 cells through electroporation. The cells were then allowed to grow in the dark at 28°C for 48 hours on YM media with kanamycin selection. Well-isolated transformed colonies were verified for plasmid insertion by PCR (see M&M section 2.2.2.2.2). The PCR positive colonies with the right size (Fig. 23) were selected and used for stable plant transformation and Agroinfiltration.

4.3 Genetic transformation of *Medicago truncatula*

All the transformation experiments were performed with *M. truncatula* genotype R-108, already reported to have an efficient regeneration and transformation capacity. The leaflets were used as explants, which have been shown to produce plants through somatic embryogenesis. The transformation protocol used was modified from Cosson *et al.* (2006). Different stages of *M. truncatula* transformation and regeneration from leaf explants to

greenhouse transferred plants are depicted in the Fig. 24. Healthy leaflets from 4-5 weeks-old *in vitro* grown plants or from green house grown plants were either precultured on SH3a media for 24-36 hours before transformation or used freshly (Fig. 24a).

The explants were wounded and then inoculated with *A. tumefaciens* containing the desired construct. Agrobacterium infiltration into wounds of leaflets was further facilitated by vacuum infiltration. After 3 days of co-cultivation on SH3a co-cultivation medium at 24°C, the leaflets were rinsed to eliminate the overgrown bacteria (Fig. 24b). The transfected explants were then cultured in the dark at 25°C on SH3a (Callus induction media) with selection agent to induce callus formation. It took an average of about 7-8 weeks after transformation to get hygromycin-resistant fragile brown sugar powder-like calli (Fig. 24d). Non-transformed explants used as control were unable to form callus, hence they died and were discarded.

Plant regeneration via somatic embryogenesis was achieved after transferring the callus on hormone free SH9 (embryo induction media) in the light (16h photoperiod) at 24°C, with constant selection for hygromycin resistance (Fig. 24e-j). Some contaminated calli showing the bacterial overgrowth were unable to develop green embryos and hence were discarded. Similarly the fungal infected calli were also discared to reduce the risk of contamination to healthy calli. After about 8-10 weeks of culture, the developing green embryos gave rise to hygromycin-resistant small root-less plantlets. On an average basis, about 4-8 plantlets were observed per calli. 1-2 plantlets per calli were selected to root on ½ SH9 (rooting media with hygromycin) (Fig. 24k) while the still green calli were again subcultured on SH9 media. The plantlets were transferred to a mini-greenhouse (Fig. 24 l,m) and then to the greenhouse until seed harvest (Fig. 24o). The non transformed explants serving as control also developed callus, somatic embryos and plantlets but only on media without hygromycin. After rooting, they were also transferred to the greenhouse.

Figure 24. **Different steps of *M. truncatula* regeneration after transformation** (A) Leaf explants of *M. truncatula* R108 from 6 week old *in vitro* grown plants. (B) Transfected explants after 3 d of co-culture on SH3a without antibiotics. (C) Callus formation from explants after two weeks on SH3a callusing medium in dark. (D) Six-week-old calli on SH3a medium. (E) Friable Calli after 2 wk on SH9 medium at light. Differentiating embryos appear as green spots. (F,G) Calli after 4 wk on SH9 medium under light. (H,I) Development of embryos into plantlets after six wk on SH9 medium. At this stage, both embryos and the first plantlets are present on the plant. (J) The plantlets that developed on SH9 medium are transferred onto ½ SH9 medium to allow rooting. (K) Growth of plantlets after 3 wk on ½SH9 medium. (L,M) Plantlets transferred to the mini greenhouse. Plants maintained under high humidity with a transparent lid. (N) Three-week-old plants after transfer to the mini greenhouse. The lid was removed and plants were watered with nutrient solution. (O) Transgenic plants 5 weeks after greenhouse transfer.

4.3.1 Generation of *pdof::gfp* cassette containing plants

The leaflets from R-108 were used as source of explants for transformation by the LBA4404 strain of *A. tumefaciens* containing the promoter expression construct *pdof::gfp*. After the callus induction phase, 412 hygromycin-resistant fragile calli i.e. (92% of total explants) were recovered from SH3a and transferred to SH9 regeneration media for development of mature embryos and ultimately the plantlets. The non transformed explants were able to develop some callus-like-structures after long-term culturing them on SH3a media, but they died later under hygromycin selection. As shown in Table 8, mature embryos from 55% of total calli i.e. 226/450 developed to higromycin-resistant putatively transformed plantlets. When transferred to ½ SH9 with kanamycin, about 46% of these plantlets i.e. 104/226 (23% of initial explants) regenerated roots. Then only 52 of 104 rooted hygromycin-resistant plantlets were selected randomly from independent transformation events, and transferred to mini-greenhouse for acclimatization, and subsequently to greenhouse for further molecular and cytological analyses until maturity and harvesting.

Table 8. Detail of *M. truncatula* transformation with different constructs

Agrobacterium strain	Constructs	Initial explants	Callus	Hygromycin resistant Plantlets	Rooted plants	Green house	PCR +ve
LBA4404	***pdof::gfp***	450	412	226	104	52	46
EHA105	***p35S::dof-gfp***	325	112	41	33	26	25

4.3.1.1 Verification of *pdof::gfp* cassette at molecular level

As shown in Fig. 22, *pdof::gfp* construct contains *hpt* as a selectable marker gene that confers resistance against hygromycin. Although the plants in the greenhouse were previously selected on media containing hygromycin, verification of their transgenic nature at the molecular level was essential. So, genomic DNA was extracted from those putatively transformed plants through a modified protocol of Rogers and Bendich (1985). A segment of 1219 bp was amplified from the T-DNA region of the construct which has been integrated into the genome. Based on PCR of these putative transformants, 46 of 52 (88.4%) plants in the greenhouse were shown to contain the gene of interest, amplifying the transgene fragment,

while no amplifications were observed when genomic DNA from non-transformed R-108 was used as a control template (Fig. 25b).

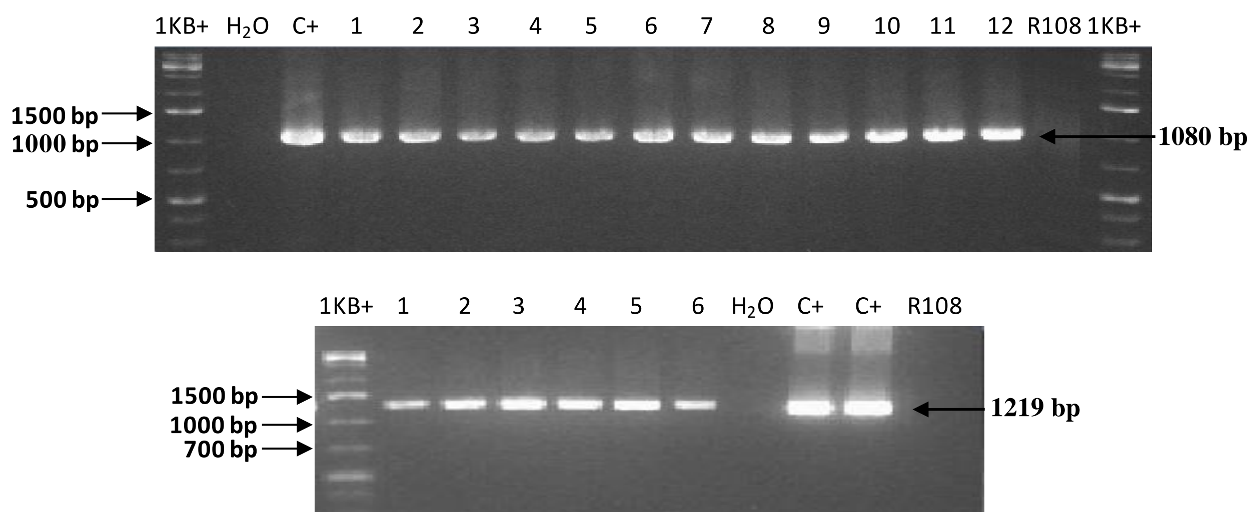


Figure 25. PCR amplification of transgene fragment from genomic DNA of transformants. (A) 1080 bp fragment amplified from gemomic DNA of plants transformed with *p35S::dof-gfp* construct in EHA105. (B) 1219 bp amplified from gemomic DNA of plants transformed with *pdof::gfp* construct in LBA4404. 1KB+ is DNA ladder. Numbers indicates independent lines, C+ is the positive control i.e. construct plasmid, H_20 is the blanc for PCR control

4.3.2 Generation of *35S::dof-gfp* cassette containing plants

As shown in Table 8, a total of 325 explants has been cocultivated with *A. tumefaciens* strain EHA105 containing *35S::dof-gfp*, a translational fusion construct for transformation. About 35% of them i.e. 112/325, were able to develop callus and somatic embryos within 6 weeks on SH3 callus induction media with Hygromycin selection, while non transformed control plants were unable to develop somatic embryos on this media. Various cultures showed bacterial overgrowth so they were discarded. When transferred on SH9 media with hygromycin, about 37% of total calli i.e. 41, gave rise to multiple plantlets from mature embryos. After transferring them on rooting media with hygromycin, about 80% of these plantlets i.e. 33/41 (10% of initial explants) regenerated roots. Then only 26 of 33 rooted hygromycin-resistant plantlets (from independent transformation events) were transferred to mini-greenhouse for acclimatization, and subsequently to the greenhouse for the rest of their lifecycle.

4.3.2.1 Verification of *35S::dof-gfp* cassette at molecular level

Genomic DNA from the putatively transformed greenhouse-grown plants was extracted purified and quantified. The PCR from genomic DNA amplified a 1080 bp segment from the T-DNA region of the construct integrated within the genome. As shown in Table 1, 25 of 26 (96%) greenhouse plants were proven to contain the gene of interest amplifying the transgene fragment, while no amplifications were observed when genomic DNA from non-transformed R-108 was used as a control template (Fig. 25).

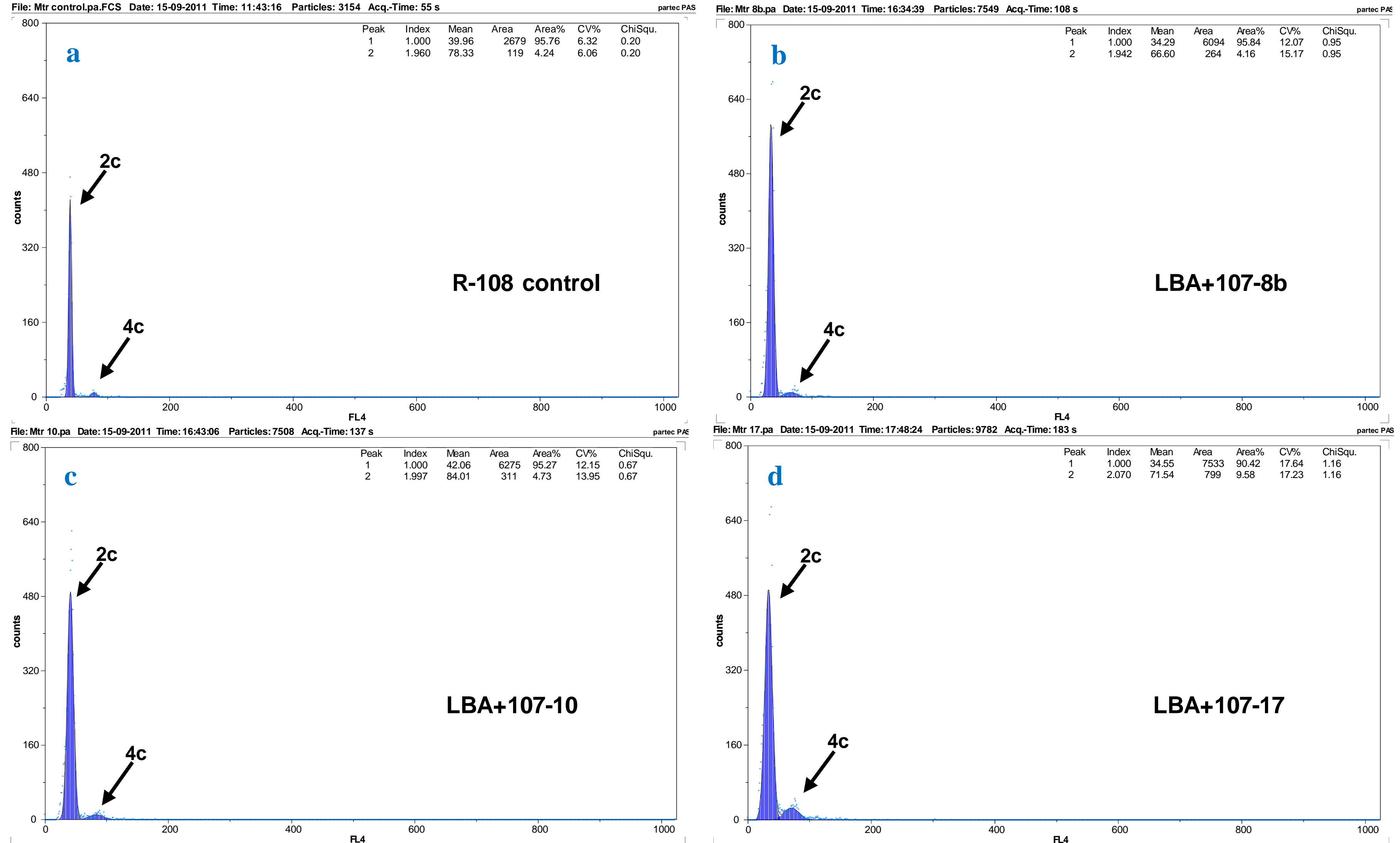
4.3.3 Comparison of transformation efficiency

The *M. truncatula* transformation experiments were carried out using two different strains of *A. tumefaciens* i.e. LBA4404 and EHA105. Both of these strains vary in their ability to transfer the transgene cassette efficiently into the plant genome. The transformation efficiency was calculated on the basis of hygromycin-resistant plantlets producing calli. The results showed that transformation frequency differs with respect to the strains employed. For example, when using LBA4404 strain as a transformation vector, the calli from about 50% of the initial explants were able to develop hygromycin resistant plantlets while 88% of the tested green house plants were PCR positive.

On the other hand, the employment of EHA105 strain as a transformation vector has resulted in production of hygromycin-resistant plantlets from the calli which are about 12% of the initial explants population. Moreover, 96% of the analyzed green house plants were transgenic in nature as verified by PCR. The decreased percentage of plantlet formation may be due to the fact that fungal infection as well as bacterial overgrowth has been observed on the developing calli from EHA105-based transformation event, and hence they were discarded thus reducing final transformation efficiency.

4.3.4 Flow cytometric characterization of transformants

Flow cytometry has emerged as an important as well as time and money saving tool to check the true-to-typeness of the transformants (Ochatt, 2008). In our lab, it has also been employed in characterization of insertional mutants, regeneration competence of calluses (Ochatt *et al.,* 2005, 2010), developing embryos and cell suspension cultures, for the taxonomic distinction among landraces within a species, and to assess the development of *in vitro* cultured immature seeds.



Peak	Index	Mean	Area	Area%	CV%	ChiSqu.
1	1.000	39.96	2679	95.76	6.32	0.20
2	1.960	78.33	119	4.24	6.06	0.20

Peak	Index	Mean	Area	Area%	CV%	ChiSqu.
1	1.000	34.29	6094	95.84	12.07	0.95
2	1.942	66.60	264	4.16	15.17	0.95

Peak	Index	Mean	Area	Area%	CV%	ChiSqu.
1	1.000	42.06	6275	95.27	12.15	0.67
2	1.997	84.01	311	4.73	13.95	0.67

Peak	Index	Mean	Area	Area%	CV%	ChiSqu.
1	1.000	34.55	7533	90.42	17.64	1.16
2	2.070	71.54	799	9.58	17.23	1.16

Figure 26: Representative flow cytometry profiles of developing *M. truncatula* plants before transferring them to greenhouse

a) Non-transformed plant used as a control , b,c,d) plants transformed with *pdof::gfp* construct

Thus, in our experiments flow cytometric analysis was employed to characterize putative transformants prior to their transfer to greenhouse. The leaf tissues were utilized for cytometric analysis. The flow cytometry profiles at this stage show the typical two-peak distribution of a normal dividing tissue in the putative transformants (Fig. 26b-d) as well as non-transformed control plants (Fig. 26a). These two peaks correspond to the nuclei that are distributed into two subpopulations, those at G0/G1 and those at G2/M. When checked for transgene insertion later on by PCR, these plants proved to be transgenic in nature. Only a few plants showed a disturbed ploidy level, and were later on proved by PCR to be non transgenic suggesting they might be chimaeric.

4.3.5 Phenotyping of transformants

A successful genetic transformation event is characterized by normal growing transformants showing no significant differences compared to the non-transformed plants (provided the transgene studied is not implicated in growth responses). So in our case, once in the greenhouse, the phenotype of the transformed *M. truncatula* plants was also observed. The growth of the transformants was comparable to the non-transformed control plants, although the plants which developed very late from the mature embryos, when transferred to the greenhouse, showed reduced seed set. Compared with non-transformed R-108 plants, the transformants exhibited normal pods and seeds (Fig. 27). Moreover, the germination percentage of T1 seeds was quite high i.e. >85% which is comparable to seeds from control plants.

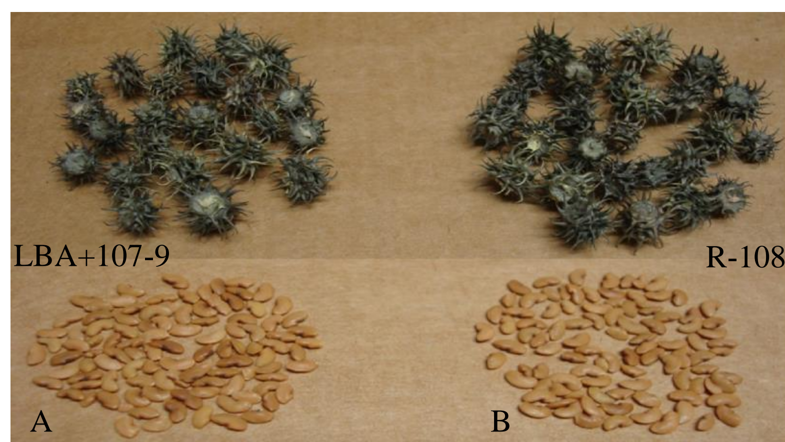


Figure 27. Phenotypic comparison of pods and seeds. A) from plant transformed with *pdof::gfp* construct, B) from non-transformed control plant

4.3.6 Transgene segregation in the progeny

The segregation analysis was performed on the basis of heritability of hygromycin resistance gene in the progeny of the primary transformants. For this purpose, the primary transformants in the greenhouse were allowed to set T1 seeds through self pollination. The mature pods were harvested and the seeds were separated from pods. The *in vitro* germination test has shown that the germination percentage of T1 seeds from four independent transformed lines is quite high, ranging from 85-100% (Table 9). The germinated seedlings were then transferred to MS0 media supplemented with hygromycin to select the transgenic plants and eliminate the segregating non transgenics or escapes. After two successive subcultures on this media after 2 weeks, the number of surviving plants was counted. The percentage of hygromycin-resistant surviving plants ranges from 82% to 95%.

Based on the Chi square analysis for Mendelian 3:1 segregating ratio (Table 9), three out of four lines showed a single locus integration for *hpt* gene, while line i.e. EHA+83-18b showed a ratio of 4:0. The Mendelian inheritance of the transgene could not be studied on a molecular basis, and it was in any case beyond the scope of this project.

Table 9. Comparison of germination and transgene segregation

Transgenic line	Total no. of seeds	Germination %	Hyg^r plants	Dead plantlets	Percentage of Hyg^r plants	x^2 value for Mendelian 3:1
EHA+83-11	24	91%	18	4	81%	0.546
EHA+83-12	23	86%	17	3	85%	1.067
EHA+83-18a	23	100	19	4	82%	0.362
EHA+83-18b	24	87%	20	1	95%	4.587 for 3:1 0.048 for 4:0

4.4 Laser scanning confocal microscopy

4.4.1 Subcellular localization of Dof1147

To determine its subcellular localization, the coding sequence of *DOF1147* was fused in frame to the 5' terminus of the GFP reporter gene under the control of the cauliflower

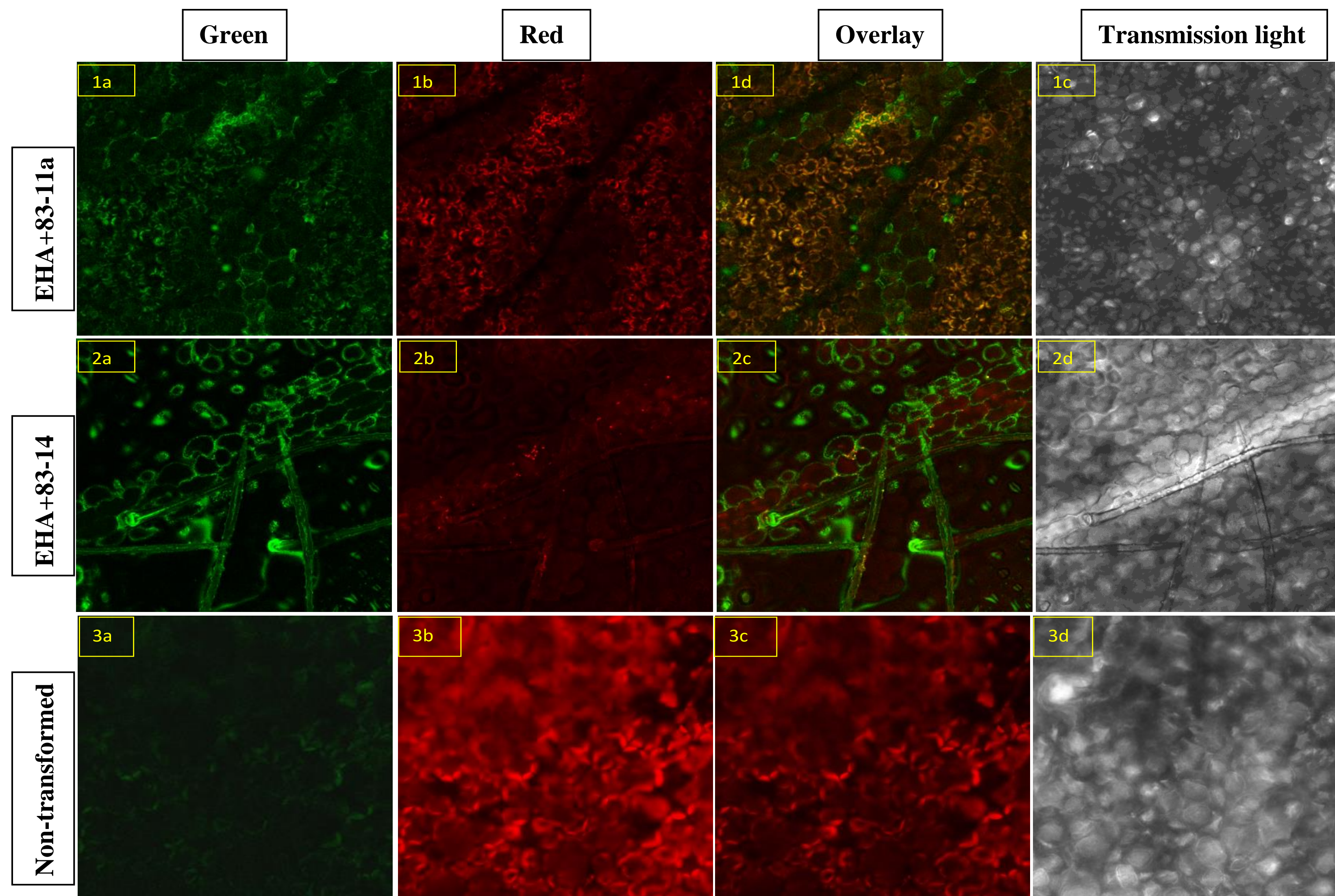


Figure 28. Subcellular localization of DOF1147 protein as observed in leaves of *M. truncatula* plants transformed with translation fusion construct *p35S::dof-gfp*. Two fields (1,2) from independently transformed lines while 3) from non-transformed controls. a) with green excitation, b) with red excitation, c) overlay image, d) under transmission light

mosaic virus dual 35S promoter to form *p35S::dof-gfp*. The recombinant plasmid was inserted into *A. tumefaciens* strain EHA 105 and was then co-cultivated with *M. truncatula* leaf explants. Transgenic plants were regenerated through somatic embryogenesis. Transformed plants which had been confirmed by PCR on leaf sections were selected for microscopic studies. We used a Leica TCS-SP2-AOBS inverted laser scanning confocal microscope (Leica Microsystems, Heidelberg Germany) which has adjustable bandwidths of the detected fluorescence wavelengths and is equipped with a 3-laser configuration (488/561/633nm).

When leaf sections of the transformed plants were observed under laser scanning confocal microscope, GFP was excited with 488 nm excitation light provided by the inline Argon laser of microscope. The emission band for GFP detected a range between 495–540 nm. The images of leaf sections acquired using Leica Confocal software (LCS) are shown in Fig. 7. The green fluorescence from fused GFP-DOF1147 protein was localized in the nucleus of some leaf spongy mesophyll cells (Fig. 28-1). This nuclear localization is consistent with the presumed transcription factor function of DOF1147. Some transient flourescence signals have also been observed in the cell membrane of the cells in leaves of some transgenic plants (Fig. 28-2). Conversly, no fluorescence was observed in leaf cells from non-transformed control plants (Fig. 28-3).

4.4.2 DOF1147 subcellular localization studies using Agroinfiltration

Agroinfiltration in tobacco leaves has been widely used for the heterologous expression and localization of a protein fused to a detectable marker. So in order to investigate the subcellular localization of DOF1147 protein, healthy tobacco leaves from 3-4 week-old plants were infiltrated with the *A. tumefaciens* strain EHA105 containing the *p35S::dof-gfp* construct. The transient localization of the DOF1147 protein in the leaves was investigated after 36 hours of agroinfiltration.

When tobacco leaves were observed under epi-fluorescence microscope using ultraviolet light, a distinct nuclear localization signal was observed by GFP fluorescence in the nucleus of the lower epidermal cells of infiltrated tobacco leaves (Fig. 29-1) while no fluorescence was observed in sub-epidermal cells of non-transformed leaves. The localization of fused DOF1147-GFP protein complex into the nucleus is consistent with the transcription function of genes belonging to the C2-C2 zinc-finger family.

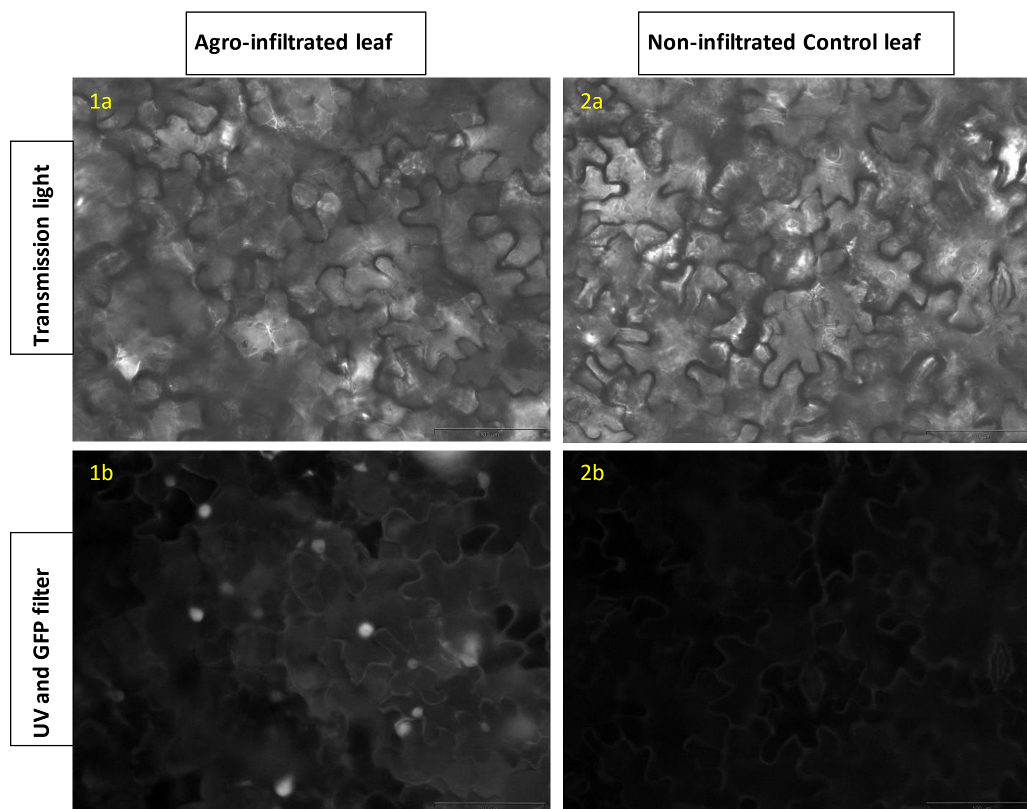


Figure 29. Localization of Dof1147 protein in tobacco leaf epidermal cells. a) nucleus fluorescing under UV. b) Respective field under transmission light

4.4.3 Expression of *Dof1147* during seed development

In order to investigate the site of expression of Dof1147 during *M. truncatula* seed development, a *pDOF::GFP* construct was prepared in which the green fluorescent protein reporter gene was driven by the *DOF1147* promoter. *M. truncatula* plants were then transformed with this construct and selected against hygromycin. Developing immature seeds from transformed plants were extracted from pods at 8, 10 and 12 DAPs based on the *DOF1147* gene expression profile. When the seeds were placed under laser scanning inverted confocal microscope, GFP was excited with 488 nm light provided by Argon laser of microscope. The emission band for GFP was detected in the range 495–540 nm. The images of seeds acquired using LCS software are shown in Fig. 9. The GFP expression driven by *DOF1147* promoter in the endosperm of developing immature seeds was not clearly visible (Fig. 30-1) and was comparable to the seeds from non-transformed plants (Fig. 30-1). This lack of GFP florescence could be attributed to the weakness of the *DOF1147* promoter. On the other hand, GFP fluorescence has also been reported to fade with time.

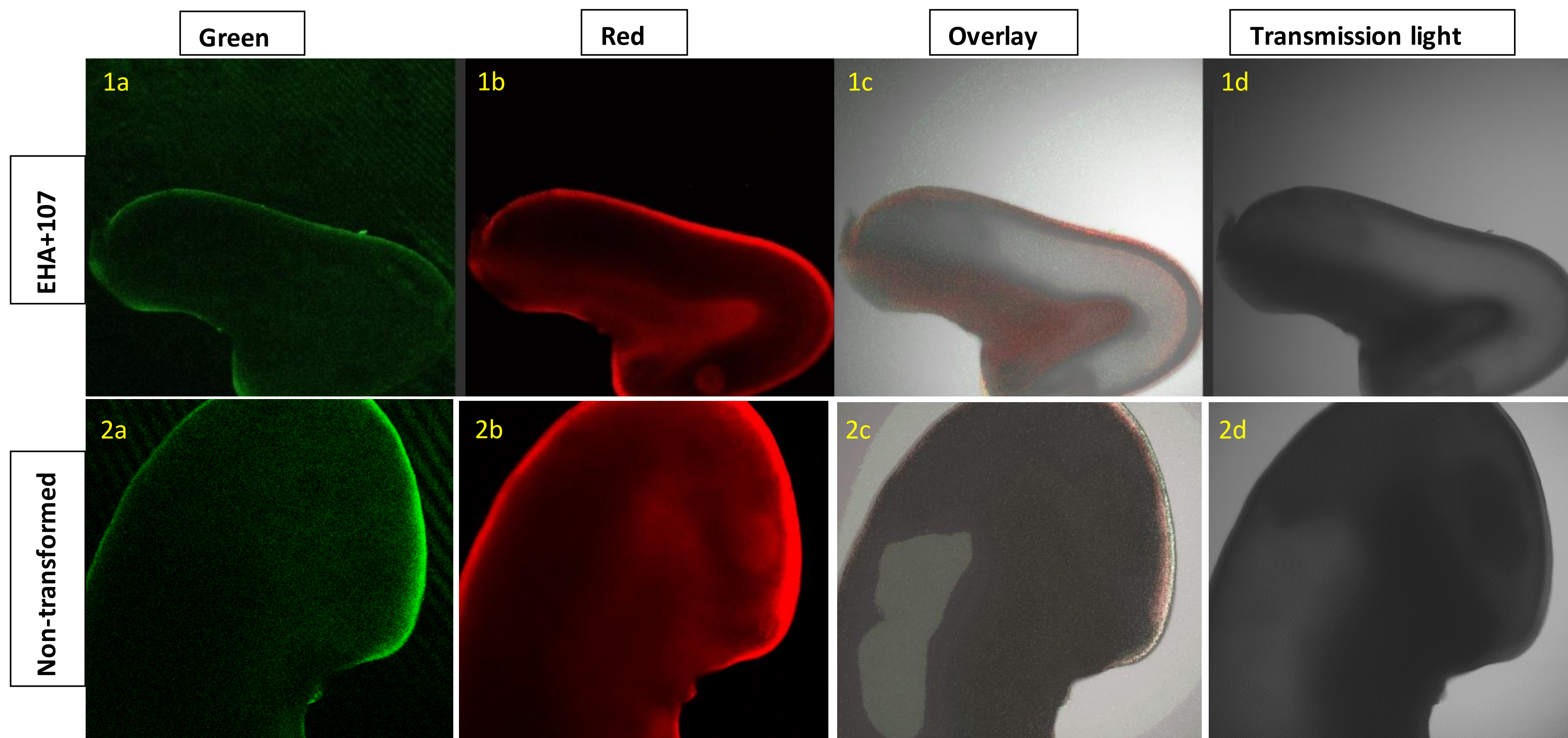


Figure 30. Cytological study of *DOF1147* promoter exression in *M. truncatula* developing seed. Immature developing seed of *M. truncatula* of 1) transformed plants and 2) non-transformed control at 8 DAP under confocal microscope. a) under green excitation, b) red excitation, c) Overlay of all the images, d) transmission light

Chapter 5: Role of auxin in *in vitro* seed development in *M. truncatula*

5.1 Introduction

The phytohormone auxin has been implicated in the regulation of diverse processes, including cell division, expansion and differentiation, the modulation of the mitotic and endocycle, the formation and maintenance of meristems, vascular tissue differentiation, phototropism, and endosperm and embryo growth in developing seeds (Chapman and Estelle 2009; Davies, 2010).

In our research unit, studies were carried out on plants of the EMS109 *dof1147* mutant identified from a TILLING population. These mutants showed defective seed development, mostly leading to seed abortion at the early cell division phase. Even if the plants developed seeds, they were significantly reduced in size and number. Importantly in this context, the *in silico* promoter analysis of DOF1147 putative target genes which I have carried out has identified auxin-responsive elements thus supporting a possible role of auxins during seed development in *M. truncatula*. On the basis of these preliminary results, we decided to investigate further the role of an *in vitro* auxin treatment on the seed development particularly during the embryogenesis-seed filling transition and the seed filling phase itself. For this purpose, a reliable technique for studying *in vitro* seed development developed in our laboratory (Gallardo *et al.,* 2006; Ochatt, 2011) was employed to compare the effects of the naturally occurring auxin IBA and the synthetic auxin NAA on the in vitro cultured immature *M. truncatula* seeds.

The effect of these auxins on immature in vitro seed development was determined by measuring seed fresh weight and seed size. The effects on the cell endocycle leading to endoreduplication were also investigated. The possible role of auxin in seed development especially with reference to the transition from the cell division phase to the seed reserve accumulation phase was also discussed. The bulk of the results obtained were published in the journal Physiologia Plantarum, as presented in the attached paper (DOI: 10.1111/j.1399-3054.2012.1719.x online since November 9, 2012).

5.2 *In vitro* auxin treatment promotes cell division and delays endoreduplication in developing seeds of the model legume species *Medicago truncatula*.

Rana Muhammad Atif[1], Florian Boulisset[1], Catherine Conreux[1], Richard Thompson[1], Sergio J. Ochatt[1]*

[1]INRA CR de Dijon, UMR1347 Agroécologie, BP 86510, F-21065 Dijon Cedex, FRANCE
*Corresponding author.
E-mail address: ochatt@dijon.inra.fr
Physiologia Plantarum (2012) DOI: 10.1111/j.1399-3054.2012.01719.x

5.2.1 Abstract

The role of auxins in the morphogenesis of immature seeds of *Medicago truncatula* was studied, focusing on the transition from the embryo cell division phase to seed maturation. We analyzed seed development *in vitro*, by flow cytometry, and through the determination of the kinetics of seed fresh weight and size. Thus, seeds were harvested at 8, 10 and 12 days after pollination and cultured *in vitro* on a medium either without auxin or supplemented with indole-3-butyric acid (IBA) or naphthalene acetic acid (NAA) at 1 mgl^{-1}. All parameters studied were determined every 2 days from the start of *in vitro* culture. The results showed that both auxins increased the weight and size of seeds with NAA having a stronger effect than IBA. We further demonstrated that the auxin treatments modulate the transition between mitotic cycles and endocycles in *M. truncatula* developing seed by favoring sustained cell divisions while simultaneously prolonging endoreduplication, which is known to be the cytogenetical imprint of the transition from the cell division phase to the storage protein accumulation phase during seed development.

Abbreviations

4-Cl-IAA	4-chloro-3-indoleactic acid,	ARF	auxin-response factor
DAP	days after pollination,	FDA	fluorescein diacetate
IAA	indole-3-acetic acid,	IBA	indole-3-butyric acid
MS	Murashige and Skoog,	NAA	naphthalene acetic acid
PIN	PINFORMED proteins,		

5.2.2 Introduction

Seed development in plants involves a double fertilization event yielding two zygotic products: the embryo that will give rise to the daughter plant and the endosperm that transmits nutrients from the maternal tissues to the embryo during embryogenesis and germination (Lopes and Larkins 1993). The third major seed component, the seed coat, differentiates after fertilization from maternally derived tissues including the integuments, and while in most monocots a persistent endosperm accounts for most of the mature seed, in most dicots the endosperm is only transient and is successively replaced by the growing embryo. As a result, while seed size in monocots is often attributable to the extent of endosperm growth (Reddy and Daynard 1983, Chojecki *et al*. 1986), in dicot seeds such as those of legumes it is the cotyledon cell number and size that is directly linked to final seed size (Munier-Jolain and Ney 1998). However, in pea, the extent of mitosis in cotyledons is correlated with the extent of invertase activity in the seed coat, reflecting the importance of the mother plant in determining final seed size (Weber *et al*. 1996). Further, in the model Arabidopsis thaliana, the extent of endosperm proliferation is correlated with seed weight and embryo size, although the mature seed contains only a single layer of endosperm cells (Scott *et al*. 1998, Garcia *et al*. 2003). Hence, the embryo is not the only organ that determines final seed size. Despite their importance, much still remains to be learnt about the genetic, physiological and environmental determinism of the final size and weight of seeds (Schruff *et al*. 2006).

Legumes are an important source of proteins for food and feed. Their capacity to fix nitrogen through symbiosis with Rhizobium also renders them environmentally friendly and permits them to produce abundant protein-rich seeds even in the absence of added nitrogen fertilizers. *Medicago truncatula* is an annual diploid ($2n = 16$) autogamous legume species originated in the Mediterranean basin with a small genome size of 500 Mb/1C (Blondon *et al*. 1994), which is phylogenetically close to the most widely cultivated legumes in Europe, pea and faba bean (Kaló *et al*. 2004, Aubert *et al*. 2006), and has been used extensively in research and yielded important advances in terms of mutagenesis, genomics and also at the plant physiology level (Holdsworth *et al*. 2008, Verdier and Thompson 2008, Verdier *et al*. 2008)

An *in vitro* method for seed filling, developed with immature seeds of *M. truncatula*, allowed us to demonstrate that the accumulation of storage proteins in embryos cultured *in vitro* on a Murashige and Skoog (MS) medium corresponded to that occurring *in planta* (Gallardo *et al*.

2006). Later (Ochatt 2011), we also examined the effects of varying sulfur levels in the medium on the accumulation of storage proteins richer in this element whose content is generally poor in legume seeds as compared to cereals. We subsequently assessed the effects of the growth regulators gibberellins and abscisic acid on embryo development and seed filling using the same technique (Ochatt 2011).

The role of auxin is decisive in plant development and in the plant's responses to changes in its micro and macro environment. Auxin primarily acts on plant development through transcriptional regulation. The AUX/IAA repressor proteins bind to ARF (auxin-response factor) transcription factors to inhibit their function. Upon its perception in the nucleus, auxin enhances degradation of the AUX/IAA repressors hence releasing ARF transcriptional regulators from inhibition (Chapman and Estelle 2009). The ARFs induce early auxin-responsive gene expression by binding to auxin-responsive elements in the promoter regions of the target genes, which elicit different cellular and developmental responses (Hayashi 2012). At the cellular level, auxin is implicated in cell division, expansion and differentiation (Chapman and Estelle 2009).

Many studies have been undertaken to assess the effects of hormones on the development of embryos as on seed filling in the past, but the vast majority of them focused on ABA and on the latter stages of development (Quatrano *et al.* 1997, Finkelstein *et al.* 2002, Nambara and Marion-Poll 2005), while the role of auxins during this crucial segment of plant growth has received much less attention (Forestan *et al.* 2010, 2012, and references therein), and has been mostly restricted to the auxin-mediated control of cell patterning within the embryo (Möller and Weijers 2009). Several endogenous auxins have been identified in plants, i.e. indole-3-acetic acid (IAA), indole-3-propionic acid, indole-3-butyric acid (IBA), phenylacetic acid and 4-chloro-3-indoleactic acid (4-Cl-IAA) (Schneider *et al.* 1985). A significantly higher amount of 4-Cl-IAA has been observed, as compared to other auxins, in developing seeds of some key legumes, e.g. lentil (Engvild *et al.* 1980), faba bean (Pless *et al.* 1984) and pea (Nadeau 2009). In faba bean, Pless *et al.* (1984) reported that a higher amount of 4-C1-IAA during seed development was correlated with an increase in fresh weight. The 4-C1-IAA appears to influence seed development in several ways. In pea fruit, 4-C1-IAA stimulates pericarp growth (Reinecke 1999) as well as gibberellin biosynthesis (Ozga *et al.* 2009), while Nadeau (2009, see above) suggested an involvement of 4-Cl-IAA in the repression of transcript levels of the putative auxin receptor PsAFB6A in the pericarp by mimicking the presence of seeds.

Endoreduplication is a variant cell cycle, characterized by replication of DNA in the cell nucleus but without going through karyokinesis or cytokinesis (Ochatt 2008). So, after multiple endocycles, DNA contents and hence the ploidy level of nuclei increase many folds, e.g. up to 32C in *A. thaliana* (Galbraith *et al*. 1991), 96C to 192C in *Zea mays* (Larkins *et al*. 2001), 128C in *Pisum sativum* (Lemontey *et al*. 2000) and 128C in *M. truncatula* (Ochatt 2011). A normal cell cycle comprises of four distinct phases, i.e. G1 which acts as a checkpoint to ensure the presence of adequate material to move to the next phase, a step of DNA biosynthesis and chromosome duplication (S phase), G2 phase that ensures that DNA is duplicated accurately and the mitosis/M phase with segregation of the duplicated chromosomes and subsequent cytokinesis (Inzé and De-Veylder 2006, Perrot-Rechenmann 2010). In plants, the phenomenon of endoreduplication has been reported to operate in trichomes (Traas *et al*. 1998), seed development (Kowles and Phillips 1985, Lemontey *et al*. 2000) and fruit formation (Joubès *et al*. 1999). The precise physiological function of endoreduplication is still debatable.

In this work, we have studied the *in vitro* action of the auxins, naphthalene acetic acid (NAA) and IBA, on dissected immature seeds of *M. truncatula* during the transition between embryogenesis and maturation. We examined the status of the cell cycle and endoreduplication, and investigated the temporal and spatial effects on fresh weight and size of seeds.

5.2.3 Materials and methods

5.2.3.1 *In vitro* culture

Medicago truncatula genotype R108-1, a highly regenerating and transformable line derived from cultivar Jemalong, was used throughout. Plants were maintained in a culture chamber under a 16-h light photoperiod of 220 µmol m^{-2} s^{-1}, at 22°C during daytime and 19°C overnight, with a relative humidity of 60–70%.

For harvest of immature seeds, flowers were identified morphologically the day following pollination (when petal size is maximum but the pistil is not apparent) and were labeled with the date for the determination of the number of days after pollination (DAP). Pods were harvested from the mother plants at 8, 10 and 12 DAP, i.e. 2 and 4 days before and up to the end of embryo morphogenesis, which occurs at 12 DAP (Gallardo *et al*. 2003). They were

disinfected by immersion for 1 min in 70° ethanol followed by 15 min in a calcium hypochlorite solution (37 g l^{-1}) and were finally rinsed four times in sterile water prior to their culture onto the different media tested.

All media were based on Murashige and Skoog (1962) basal medium supplemented with 6 g l^{-1} agar, 130 g l^{-1} sucrose and without growth regulators or supplemented with NAA or IBA at 1 mg l^{-1} each (i.e. 5.37 µM NAA and 4.92 µM IBA). The pH of all media was adjusted to 5.6 before autoclaving for 30 min at 112°C. Media were then dispensed as 2 ml aliquots into 5 × 5 multiwell plastic dishes and stored at 4°C in the dark until used. At least 25 seeds per time point were cultured on each tested medium and all experiments were repeated three times. Seeds harvested at 8, 10 and 12 DAP *in planta* served as controls for each age, their development on the various treatments being followed at different intervals.

5.2.3.2 Determination of seed size, fresh weight and viability

For all treatments, the seed size and fresh weight were measured *in planta* at the three stages of harvest, i.e. 8, 10 and 12 DAP, and then every two days following the start of *in vitro* culture up to a stage equivalent to 20 DAP.

Fresh weight was determined immediately after extraction of seeds from the pods prior to their culture, with all seeds within a pod weighed together, and the mean fresh weight was established. Subsequently, at the intervals above after the onset of culture, seeds were weighed individually. Results were always expressed as the means (mg) ± standard error from experiments repeated at least three times per developmental stage and culture medium.

Measurements were performed on images acquired using the program ArchimedPro (Microvision, Evry, France) and treated with program Histolab (Microvision) following procedures as described elsewhere (Ochatt *et al* 2008, Ochatt and Moessner 2010). Results for size were expressed as the surface area in mm^2 and for fresh weight in mg, as the means ± standard error from 25 replicates per experiment, with experiments repeated at least three times per developmental stage and culture medium. Data were subjected to analysis of variance using the Statistica software Program (version 6.0), the difference between means being evaluated with Newman–Keul's multiple range test (Brunning and Kintz 1977).

Viability of seeds was evaluated with FDA (fluorescein diacetate) (Widholm 1972) under UV, using an OMRB Leica photonic microscope (Leica Microsystems CMS GmbH, Wetzlar, Germany). The FDA solution employed was prepared by diluting 60 µl of a stock solution of 5 mg FDA ml^{-1} acetone in 8 ml mannitol (9%, w/v). Then, two droplets (approximately 50 µl) of this solution were added to each slide. Once excited with UV, living cells fluoresce yellow-green, confirming membrane integrity and the conservation of enzymatic activity. Viability is expressed as a percentage (of the total number of cells counted in a sample). A minimum of 100 cells per microscopic field was counted and at least three independent counts were performed for each sample analyzed.

5.2.3.3 Flow cytometry

Flow cytometry assessments were carried out on seeds at 8, 10 and 12 DAP, harvested directly from the mother plants for the controls, and on seeds that had been cultured on media MS130, NAA and IBA taken from *in vitro* cultures at 2 days intervals as above.

All assessments were performed using a Partec PAS II flow cytometer, equipped with a mercury HBO 100W lamp (Partec GmbH, Munster, Germany), a dichroic mirror (TK 420) and an in-built program for the treatment of data (Flomax, Partec GmbH). All plant tissues (leaves of mother plant, seeds and embryos) were prepared following methods as described in Ochatt (2006, 2008), and were stained with DAPI (4′,6-diamidino-2-phenylindole), which specifically binds to the adenine and thymine bases of DNA, were excited under UV (at 372 nm) and emit (at 456 nm) a fluorescence proportional to DNA content per nucleus. DAPI was chosen instead of propidium iodide because cell cycle assessments and not genome size determinations were envisaged.

Briefly, nuclei were mechanically isolated by chopping the tissues with a razor blade in 400 µl of nuclei extraction buffer and 1.6 ml of staining buffer (Partec GmbH). The suspension was filtered through a 50 µm mesh Celltrics® (Partec, GmbH), and DAPI was added to the filtrate to a final concentration of 1 µg ml^{-1}. When stained nuclei pass through UV light DAPI fluoresces in blue and the cytometer transforms light into a curve, where the intensity of the emitted epifluorescence is indicated in the abscissa and the ordinates indicate the number of objects counted (Ochatt 2008).

The parameters for flow cytometry readings were as described previously for *M. truncatula* (Ochatt 2006). For peak and cell cycle analysis, the studied tissues were compared with leaves taken from mother plants of *M. truncatula* A17 cultured on MS0 medium (blank), while the cytometer was calibrated prior to the analyses with leaves of the pea cultivars Victor, Cameor or Frisson (*P. sativum*), routinely used as a standard in our laboratory (Ochatt 2008, 2011).

5.2.4 Results

5.2.4.1 Development on auxin-free medium (MS130)

We have studied the development of immature seeds beginning at 8 DAP on an auxin-free medium (MS130). Controls were harvested *in planta* from 8 to 12 DAP, when the seed was 3 ± 0.12 mm and the embryo 1.5 ± 0.05 mm in length. Previous studies (data not shown) had revealed that, at 20 DAP, they both measured 5 ± 0.20 mm (the embryo occupying the embryo sac at the expense of the endosperm).

When analyzing the rate of early germination (defined as bipolar growth of the embryo axis resulting in plantlet development) this was nil on MS130 medium, confirming data previously reported (Gallardo *et al*. 2006, Ochatt 2011), and reflecting the inhibition of germination *in vitro* by a high osmolarity medium.

At 8 DAP, there is a significant although non-progressive increase of the seed size after 2, 6, 8 and 10 days of culture on MS130 medium, when compared to the size of seeds *in planta* at 8 DAP. Such irregularity observed in seed size can be attributed to a dehydration process taking place when individual seeds were taken from the multiwell plates for weighing. At 10 DAP, the size of seeds cultured on medium MS130 only increases significantly compared to the one *in planta* after 4 and 8 days of culture. Finally, for seeds harvested at 12 DAP, culture on medium MS130 results in a significantly increased size compared to the original one, starting already after 2 days of culture.

With respect to seed fresh weight, the same trend as for seed size above is generally observed, and immature seeds cultured at 8 DAP exhibit a discrete increase of fresh weight as from after 6 days of culture on the auxin-free medium, while this increase is observed as from 2 days of culture for the 10 and 12 DAP seeds, nonlinear for the former but significant and more pronounced for the 12 DAP seeds.

5.2.4.2 Effect of auxin on seed size

Preliminary experiments performed to test the effect of the natural auxin IAA on immature seeds *in vitro* (data not shown) revealed that 1 mg l−1 of auxin (i.e. about 5 μM depending on the specific auxin used) was best and also that the optimum developmental stages to observe such effects were between the end of the cell division phase and the onset of the storage product accumulation phase of seed development, corresponding to 8 and 12 DAP, respectively. However, IAA turned out to be not strong enough an auxin to significantly affect the size or weight of immature seeds. Thus, the study was pursued using seeds at those developmental stages but treated with two stronger auxins.

Irrespective of the duration of *in vitro* culture (2, 4, 6, 8 or 10 days), we observed a stimulatory effect of auxins on the size of seeds of *M. truncatula* (Fig. 31). This effect was stronger with NAA than with IBA, a superior efficiency that had already been reported for NAA in other contexts (Katekar 1999).

Interestingly, for 8 DAP seeds, the stimulatory effect of NAA on seed size was consistently significant with increases beyond the initial area at 8 DAP *in planta* at all time points of culture but the size of seeds was larger than on the auxin-free medium MS130 only after 2 and 10 days of culture (8 DAP + 2 and 8 DAP +10 in Fig. 31), while at 8 DAP + 6 and 8 DAP + 8 the seeds on MS130 were the largest. IBA had little effect on the size of seeds harvested and cultured at 8 DAP and only at 8 DAP + 2 it exceeded that of seeds on the auxin-free medium. Moreover, the size of some seeds cultured on IBA even diminished compared to that at the onset of culture, probably due to desiccation.

For 10 and 12 DAP seeds, there was a net increase of seed size after 2, 4 and 8 days of culture on the medium with NAA, contrasting with results on IBA, where seed size also increased but in a much less significant fashion under the same culture durations. Furthermore, at 10 DAP, seeds on the NAA medium exhibited a consistently larger size than those on the auxin-free medium at all culture times while on IBA seeds had size increases that were comparable to that of seeds on MS130 medium. A similar trend was observed for the immature seeds cultured from 12 DAP, with NAA inducing a significantly larger increase of the seed size compared both to media MS130 and IBA but, at this stage, after the same culture durations the seeds on IBA were larger than those on the auxin-free medium.

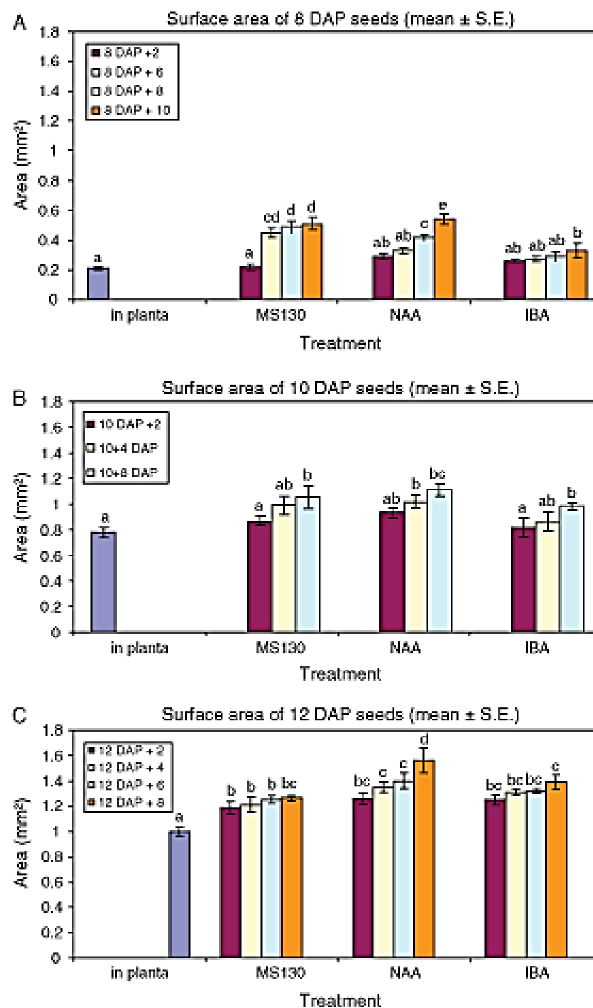


Figure 31. Surface area of immature seeds of *Medicago truncatula* (mean ± se) *in planta* at 8 DAP, 10 DAP or 12 DAP and cultured *in vitro* on media MS130, NAA or IBA. Results were pooled from two (A and B) or three (C) independent experiments.

5.2.4.3 Effect of auxin on seed fresh weight

Whether at 8, 10 or 12 DAP, seeds grown on NAA medium always had equal or greater fresh weight than those observed either on MS130 or on the IBA medium (Fig. 32), and they were also significantly heavier than on IAA in our preliminary studies (not shown). In this respect, however, the inductive effect of auxin toward seed weight was often more pronounced in the later stages of seed development (10 and 12 DAP) analyzed.

Thus, at 8 DAP, the fresh weight of seeds increased after 2, 6 8 and 10 days of culture on NAA compared to its values *in planta* and at those times of culture on both media MS130 or IBA. Comparatively, on IBA, the 8 DAP seeds exhibited a fresh weight that remained more or less stable during culture and were comparable to that observed for seeds on the auxin-free

medium throughout. At 10 DAP, seeds on NAA exhibited a constant increase of their fresh weight that was significantly higher at each time point than those observed on both MS130 and IBA media, which were generally comparable, as also observed in terms of seed size. Finally, for 12 DAP seeds both auxins induced a significantly higher fresh weight than the auxin-free medium and at all times of culture duration assessed. This stimulatory effect of auxin was significantly stronger for NAA compared to IBA although seeds on both media followed a similar trend.

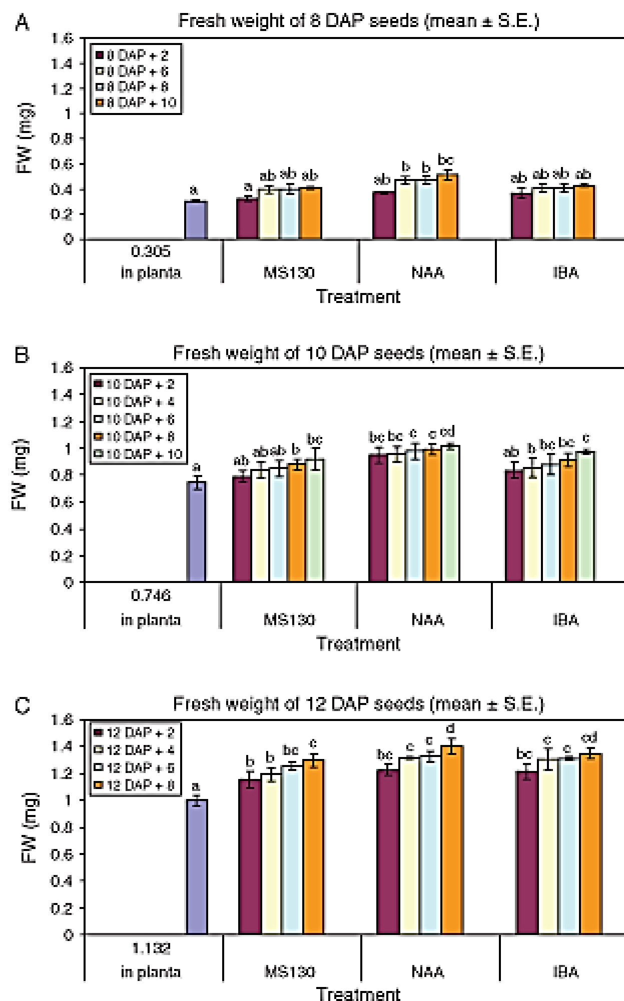


Figure 32. Fresh weight of immature seeds of *Medicago truncatula* (mean ± SE) *in planta* at 8 DAP (A), 10 DAP (B) or 12 DAP (C) and cultured *in vitro* on media MS130, NAA or IBA. Results were pooled from two (A and B) or three (C) independent experiments.

It is noteworthy that, whatever the evolution of seed size or fresh weight on the various media and developmental stages tested, the addition of hormones to the medium had no deleterious effect on the viability of cells within the immature seeds cultured.

5.2.4.4 Flow cytometry of *in vitro* developing immature seeds

Fig. 33 shows representative flow cytometry profiles of immature seeds at different stages of development both *in planta* and cultured on the three media tested *in vitro*. These flow cytometry analyses showed that auxins disrupt the onset of endoreduplication. From this perspective, the action of auxin differed depending on the seed developmental stage, but also on the type of auxin used, NAA or IBA.

At an early stage (8 DAP), the lack of auxins in medium MS130 induced an extension of the cell division phase. The flow cytometry profiles at this stage show the typical two-peak distribution of a dividing tissue, corresponding to nuclei that are distributed into two subpopulations, those at G0/G1 and those at G2/M (Fig. 33D). This status corresponds with that observed *in planta* at 8 DAP (Fig. 33A), and could explain the increase in seed weight and size observed at this stage, that would thus result from the larger number of cells per seed. Conversely, auxins seemed to promote more an expansion of cells, via an enhancement of endoreduplication, rather than cell division as was clearly observed by the appearance of multiple (and in any case more than two) peaks of a decreasing intensity in the flow cytometry profiles of nuclei from seeds cultured on the auxin containing media from 8 DAP onwards. In this respect, such phenomenon was stronger on a medium with NAA (Fig. 33G) than when IBA was added; both seeds at 8 DAP + 2 exhibited profiles with four peaks and a larger proportion of nuclei in the population analyzed in the S phase for NAA medium (not shown) while flow cytometry profiles of seeds cultured on the medium containing IBA were comprised of only three peaks (Fig. 33J) and had less nuclei in the S phase. This trend was maintained later in culture for the seeds harvested at 8 DAP.

At 10 DAP, the flow cytometry analyses gave comparable results and, after 2 days of culture (i.e. 10 DAP + 2), they included three peaks for media MS130, indicating a transition from the cell division phase to endoreduplication, but four peaks for NAA and IBA media with endoreduplication clearly established (comparable to what is observed for 12 DAP seeds *in planta*). When culture duration increased (e.g. for seeds at 10 DAP + 4; Fig. 33E, H, K) this was even magnified, and flow cytometry profiles for such seeds included three peaks for medium MS130 (Fig. 33E) but consistently five peaks on NAA (Fig. 33H) and four on IBA (Fig. 33I).This trend remained unaltered up to the time when endoreduplication decreased and flow cytometry profiles with two peaks only were observed for all media (as also observed *in planta* from around 18 to 20 DAP; Ochatt 2011), except with IBA where even after this stage

profiles with three peaks were still commonplace suggesting that, with this auxin, the onset of endoreduplication is delayed but the phenomenon lasts for longer.

Figure 33. Representative flow cytometry profiles of developing seeds *in planta* and cultured *in vitro* on media without auxin (MS130) or containing NAA or IBA. (A–C) *in planta* immature seeds at 8 DAP (A), 10 DAP (B) and 12 DAP (C); (D–L) immature seeds of different ages cultured *in vitro*, i.e. on MS130 medium (D, E, F) 8 DAP seeds after 2 days culture (D), 10 DAP seeds after 4 days culture (E), 12 DAP seeds after 4 days culture (F); on medium with NAA (G,H,I): 8 DAP seeds after 2 days culture (G), 10 DAP seeds after 4 days culture (H), 12 DAP seeds after 4 days culture (I); on medium with IBA (J, K, L): 8 DAP seeds after 2 days culture (J), 10 DAP seeds after 4 days culture (K), 12 DAP seeds after 4 days culture (L).

At 12 DAP *in planta* (Fig. 33C), the developing seeds are in the transition between the cell division and the storage product accumulation phases and the flow cytometry profiles at this stage comprise of four or five (and sometimes up to seven) peaks. The endoreduplication phenomenon is clearly observable for 12 DAP seeds cultured for 2 days, with or without auxin. In this respect, the addition of auxins to the culture medium resulted in a significant

enhancement of endoreduplication for seeds that have been cultured for four days (12 DAP + 4; Fig. 33F, I, L), and was coupled with flow cytometry profiles with routinely four and as many as five peaks, particularly with NAA (Fig. 33I). As also observed for the 10 DAP seeds on auxin media, the 12 DAP seeds continued to yield cytometry profiles including consistently three and up to four peaks for seeds that had been cultured for as long as 8 days (i.e. 12 DAP + 8) on auxin-containing media, although at this stage *in planta* (equivalent to 20 DAP) endoreduplication would have been completed (Ochatt 2011).

5.2.5 Discussion

It is now accepted that seed maturation begins with a transition phase during which control switches from maternal to filial, as shown in pea (Weber *et al*. 2005). It has also been discussed that the maturation processes are under the control of the concerted action of a considerable number of signaling pathways, which integrate genetic, metabolic and hormonal signals (Gazzarini and McCourt 2003, Gutierrrez *et al*. 2007, Verdier and Thompson 2008, Wu *et al*. 2011). In this context, sugars in general and the sucrose/glucose ratio in particular are major components of the signaling pathway that triggers the onset of the transition phase (Borisjuk *et al*. 2003, Weber *et al*. 2005), via a dramatic decrease of free hexose levels in the embryo parallel to a strong and transient increase in sucrose uptake (Gibson 2005). As a result, sucrose signals have been postulated to control storage and differentiation processes thanks to their potential to act on transcriptional and post-transcriptional regulation, through a regulation of metabolic enzymes and a modulation of gene expression and activity (Brocard-Gilford *et al*. 2003, Baud *et al*. 2005, Gibson 2005). Here, we have used a very high concentration of sucrose in the culture medium as in Gallardo *et al*. (2006), since this was shown to preclude precocious germination in *M. truncatula*, and permitted a better analysis of the effects of auxins on immature seed development without the interference of germination in the process.

The role of auxins in seed development is not very well-known beyond their effect on embryo differentiation (Möller and Weijers 2009) and on the fate of the seed-transfer cells (Forestan *et al*. 2010). The scanty literature available on the role of auxins in seed development indicates that they are localized asymmetrically in different seed tissues and they participate in the formation of apical-basal axis during embryo development (Benkova *et al*. 2003, Blilou *et al*. 2005, De Smet and Jurgens 2007). Auxins are normally passively imported into cells but

AUX/LAX permeases have been reported to be involved in their active uptake (Terasaka *et al*. 2005). On the other hand, their export outside the cells is mediated by PINFORMED (PIN) efflux carriers present in the cell membrane (Petrasek *et al*. 2006) as well as a group of ABCB transport proteins (Geisler and Murphy 2006). The heterogeneous cellular distribution of PIN proteins regulates the direction of cell-to-cell auxin flow during seed development (Benkova *et al*. 2003, Friml *et al*. 2003, De Smet and Jurgens 2007). Thus during the early seed development phase in Arabidopsis, these PIN proteins moderate an apical-basal auxin gradient to initiate the formation of apical embryo architecture (Friml *et al*. 2003, De Smet and Jurgens 2007). A model for auxin transport mediated by ZmPIN1 was recently formulated by Forestan *et al*. (2010), who analyzed the related fluxes of auxin during embryogenesis and endosperm development in maize, and compared it with the one existing for the eudicot Arabidopsis, where ZmPIN1-mediated transport of auxin and auxin accumulation were related to cellular domain differentiation during seed formation.

In this study, we report the action of auxins on the development of dissected immature seeds of *M. truncatula*. Seeds were cultured *in vitro* on media containing NAA or IBA during the transition period toward maturation, and the evolution of fresh weight and size of seeds, and the status of cell cycle was examined by flow cytometry. The choice of the developmental status at which immature seeds were harvested and cultured *in vitro* in this research was based on previous studies of histo-differentiation at mid-embryogenesis in *M. truncatula*. Abirached-Darmency *et al*. (2005) indicated that vicilin and legumin A gene expression was just beginning in 12–14 DAP seeds, thus validating two dimensional protein electrophoresis data (Gallardo *et al*. 2003). On the other hand, more recent studies of the kinetics of accumulation of the various protein fractions of seeds and embryos cultivated *in vitro* in presence and absence of nitrogen by Gallardo *et al*. (2006) confirmed the role of tissues surrounding the embryo in the synthesis of storage proteins.

The results obtained in presence of auxins clearly underline their stimulatory effect on the two parameters chosen in this study, i.e. fresh weight and size of seeds. It is also clear that the developmental stage at which immature seeds are harvested and cultured as well as the type of auxin added to the culture medium modulate these parameters differentially. In this context, the differential response was correlated with the strength of activity of the particular auxin employed (Katekar 1999) and, thus, IAA a natural endogenous auxin with a weak activity had practically no visible influence on seed weight or size, and irrespective of the developmental

stage at which seeds were studied (not shown). IBA and NAA stimulated both the fresh weight gain and the size of seeds, with NAA (of a stronger activity, Katekar (1999) having a larger effect. One hypothesis about this increase of cell size and volume is that, upon initial perception of auxin at the outer side of the cell plasma membrane, ABP1 induces activation of the H+ pump ATPase resulting in acidification of the extracellular space. This acidification prompts the action of cell-wall-modifying enzymes such as the expansins, which play a critical role in cell-wall loosening and thereby assist cell expansion (Perrot-Rechenmann 2010). Application of exogenous auxins is reported to increase the expansin transcript level, resulting in an increased cell size (Hutchison *et al*. 1999). The possibility that endogenous auxins coming from other seed tissues rather than from other tissues of the mother plant cannot, however, be completely ruled out and further studies with isolated immature embryos cultured with exogenous auxin might be envisaged to answer this.

Our flow cytometry analyses of seeds cultured at all three developmental stages (8, 10 and 12 DAP) yielded profiles including a number of peaks of decreasing intensity, reflecting endoreduplication (Ochatt *et al*. 2000, Ochatt 2006, 2008). There is therefore an arrest of divisions; phases G1, S and G2 do take place but with very few mitoses, most nuclei in the second peak being in fact in phase G2 rather than M of the cycle. This phenomenon, *in planta*, occurs for developing seeds only from 10 DAP onwards and up to 18–20 DAP, when endoreduplication is replaced by normal two-peak profiles (Ochatt 2011). In maize endosperm cells, cell expansion following mitosis is simultaneous with the endoreduplication phase (Mambelli and Setter 1998), and in a range of tissues from different species cell volume within the tissue is correlated with the extent of nuclear endoreduplication (Melaragno *et al*. 1993, Ochatt 2008). High auxin levels have been shown to inhibit and delay endocycles, prolonging the mitotic phase in Arabidopsis root meristems (Ishida *et al*. 2010). Conversely, low auxin conditions downregulated the expression of several mitotic cycle genes rapidly, reduced the mitotic activities and triggered an entry into endoreduplication, promoting cell expansion (Ishida *et al*. 2010). Thus, entry and exit from endoreduplication might be linked processes where auxin signaling plays a major role.

Interestingly, our data also showed that auxins had a strong and significant effect on both the onset and the duration of endoreduplication in the developing seeds, with NAA provoking a stronger enhancement of the phenomenon compared to IBA, IAA (not shown) or the auxin-free culture conditions (medium MS130). This effect varied with the developmental stage at

which the immature seeds were harvested and cultured, and was stronger at 10 and 12 DAP than at 8 DAP. More importantly, our data underline that flow cytometry can be used as a reliable marker of the kinetics of embryo developmental stage in its transition toward seed maturation and also to date the beginning of storage product accumulation.

These studies have shown an onset of endoreduplication in the seed cells during the transition between the cell division and maturation phases, as also observed in maize (Graffi and Larkins 1995). Endoreduplication was observed in our control *in planta* seeds from 10 DAP, although our results also show that cell division continues despite endoreduplication. Thus, while division frequency is reduced the accumulation of storage compounds commences. Hence, it seems more a matter of transition periods than distinct fixed phases. Previously it was shown that the permissive auxin signal for entry into the cell cycle is mediated by ABP1. An arrest in cell cycle was observed when ABP1 was inactivated functionally, providing evidence that ABP1 plays a critical role in modulation of the cell cycle by acting at both the G1/S and G2/M checkpoints (David *et al*. 2007). Ishida *et al*. (2010) also demonstrated that, in A. thaliana, auxin is involved in the modulation of the switch from mitotic cycle to endocycle in cells of root meristems and hence is strongly linked with the developmental transition from cell proliferation to cell differentiation. Despite auxin being the key regulator, several other growth regulators mediate endoreduplication. Thus, the cytokinins have been reported to work antagonistically with auxins by downregulating the expression of auxin transport genes, during transition from cell division to endoreduplication (Dello Ioio *et al*. 2008, Ruzicka *et al*. 2009, Ishida *et al*. 2010). Gibberellin and ethylene have been shown to play a vital role to control the frequency of endoreduplication in the A. thaliana hypocotyl (Gendreau *et al*. 1999), with an overall positive effect of gibberellin on endoreduplication while ethylene acted distinctly rather than globally. Similar effects of gibberellin when opposed to ABA have been observed in developing seeds of *M. truncatula* (Ochatt 2011).

Endoreduplication is a progressive phenomenon in storage accumulating organs during the transition between cell division and maturation phases. Controlling endoreduplication is of considerable interest as, in plants with a small genome, this alternative to a cell cycle is a means of increasing the genetic and metabolic capacities. The results shown here with the developing *M. truncatula* seeds show that auxin plays a paramount role in its onset.

Acknowledgements

We acknowledge financial support from HEC (Higher Education Commission, Pakistan) to R. M. A., this research was partially financed by grant E!4770 PEASTAR (Eurostars, E.U.).

References

Abirached-Darmency M, Abdel-Gawwad MR, Verdeil JL, Thompson R (2005) *In situ* expression of two storage protein genes in relation to histo-differentiation at mid-embryogenesis in *Medicago truncatula* and *Pisum sativum* seeds. J Exp Bot 56: 2019–2028

Aubert G, Morin J, Jacquin F, Loridon K, Quillet MC, Petit A, Rameau C, Lejeune-Hénaut I, Huguet T, Burstin J (2006) Functional mapping in pea, as an aid to the candidate gene approach and for investigating the synteny with the model species *Medicago truncatula*. Theor Appl Genet 112: 1024–1041

Baud S, Wuillème S, Lemoine R, Kronenberger J, Caboche M, Lepiniec L, Rochat C (2005) The AtSUC5 sucrose transporter specifically expressed in the endosperm is involved in early seed development in Arabidopsis. Plant J 43: 824–836

Benkova E, Michniewicz M, Sauer M, Teichmann T, Seifertova D, Jurgens G, Friml J (2003) Local, efflux-dependent auxin gradients as a common module for plant organ formation. Cell 115: 591–602

Blilou I, Xu J, Wildwater M, Willemsen V, Paponov I,Friml J, Heidstra R, Aida M, Palme K, Scheres B (2005) The PIN auxin efflux facilitator network controls growth and patterning in Arabidopsis roots. Nature 433: 39–44

Blondon F, Marie D, Brown S, Kondorosi A (1994) Genome size and base composition in *Medicago sativa* and *M. truncatula* species. Genome 37: 264–270

Borisjuk L, Rolletschek H, Walenta S, Panitz R,Wobus U, Weber H (2003) Energy status and ist control on embryogenesis of legumes: ATP distribution within *Vicia faba* embryos is developmentally regulated and correlated with photosynthetic capacity. Plant J 36: 318–329

Brocard-Gilford IM, Lynch TJ, Finkelstein RR (2003) Regulatory networks in seeds integrating developmental, abscisic acid, sugar, and light signalling. Plant Physiol 131: 78–92

Brunning JL, Kintz JL (1977) Computational Handbook of Statistics, 2nd Edn. Scott. Foresman, Glenview, CA

Chapman EJ, Estelle M (2009) Mechanism of auxin-regulated gene expression in plants. Annu Rev Genet 43: 265–285

Chen JG, Ullah H, Young JC, Sussman MR, Jones AM (2001) ABP1 is required for organized cell elongation and division in Arabidopsis embryogenesis. Genes Dev 15: 902–911

Chojecki AJS, Bayliss MW, Gale MD (1986) Cell production and DNA accumulation in the wheat endosperm, and their association with grain weight. Ann Bot 58: 809–817

David KM, Couch D, Braun N, Brown S, Grosclaude J, Perrot-Rechenmann C (2007). The auxin binding protein 1 is essential for the control of cell cycle. Plant J 50: 197–206

Dello Ioio R, Nakamura K, Moubayidin L, Perilli S, Taniguchi M, Morita MT, Aoyama T, Costantino P, Sabatini S (2008) A genetic framework for the control of cell division and differentiation in the root meristem. Science 322: 1380–1384

De Smet I, Jurgens G (2007). Patterning the axis in plants – auxin in control. Curr Opin Genet Dev

17: 337–343

Engvild KC, Egsgaard H, Larsen E (1980) Determination of 4-chloroindole-3-acetic acid methyl ester in Lathyrus, Vicia and Pisum by gas chromatography-mass spectrometry. Physiol Plant 48: 499–503

Finkelstein RR, Gampala SSL, Rock CD (2002) Abscisic acid signalling in seeds and seedlings. Plant Cell 14: 15–45

Forestan C, Meda S, Varotto S (2010) ZmPIN1-mediated auxin transport is related to cellular differentiation during maize embryogenesis and endosperm development. Plant Physiol 152: 1373–1390

Forestan C, Farinati S, Varotto S (2012) The maize gene PIN family of auxin transporters. Frontiers in Plant Science 3: article 16, 23 pp; doi: 10.3389/fpls.2012.00016

Friml J, Vieten A, Sauer M, Weijers D, Schwarz H, Hamann T, Offringa R, Jurgens G (2003) Efflux dependent auxin gradients establish the apical-basal axis of Arabidopsis. Nature 426: 147–153

Galbraith DW, Harkins KR, Knapp S (1991) Systemic endopolyploidy in *Arabidopsis thaliana*. Plant Physiol 96: 985–989

Gallardo K, Le Signor C, Vandekerckhove J, Thompson R, Burstin J (2003) Proteomics of *Medicago truncatula* seed development establishes the time frame of diverse metabolic processes related to reserve accumulation. Plant Physiol 133: 664–682

Gallardo K, Kurt C, Thompson R, Ochatt S (2006) In vitro culture of immature *M. truncatula* grains under conditions permitting embryo development comparable to that observed in vivo. Plant Sci 170: 1052–1058

Garcia D, Saingery V, Chambrier P, Mayer U, Jürgens G, Berger F (2003) Arabidopsis *haiku* mutants reveal new controls of seed size by endosperm. Plant Physiol 131: 1661–1670

Gazzarini P, McCourt P (2003) Cross-talk in plant hormone signalling: What arabidopsis mutants are telling us. Ann Bot 91: 605–612

Geisler M, Murphy AS (2006) The ABC of auxin transport: the role of P-glycoproteins in plant development. FEBS Lett 580: 1094–1102

Gendreau E, Orbovic V, Hofte H, Traas J (1999) Gibberellin and ethylene control endoreduplication levels in the *Arabidopsis thaliana* hypocotyl. Planta 209: 513–516

Gibson SI (2005) Control of plant development and gene expression by sugar signalling. Curr Opin Plant Biol 8: 93–102

Graffi G, Larkins BA (1995) Endoreduplication in maize endosperm: involvement of M phase promoting factor inhibition and induction of S phase-related kinases. Science 269: 1262–1264

Gutierrrez L, Van Wuytswinkel O, Castelain M, Bellini C (2007) Combined networks regulating seed maturation. Trends Plant Sci 12: 294–300

Hayashi K-I (2012) The interaction and integration of Auxin signaling components. Plant Cell Physiol 53: 965–975

Holdsworth MJ, Finch-Savage WE, Grappin P, Job D (2008) Post-genomics dissection of seed dormancy and germination. Trends Plant Sci 13: 7–13

Hutchison KW, Singer PB, Diaz-Sala C, Greenwood MS (1999) Expansins are conserved in conifers and expressed in response to exogenous auxin. Plant Physiol 120: 827–832

Inzé D, De Veylder L (2006) Cell cycle regulation in plant development. Annu Rev Genet 40: 77–105

Ishida T, Adachi S, Yoshimura M, Shimizu K, Umeda M, Sugimoto K (2010) Auxin modulates the transition from the mitotic cycle to the endocycle in Arabidopsis. Development 137: 63–71

Jones AM, Im K-H, Savka MA, Wu MJ, DeWitt NG, Shillito R, Binns AN (1998) Auxin-dependent cell expansion mediated by overexpressed auxin-binding protein 1. Science 282: 1114–1117

Joubès J, Phan T-H, Just D, Rothan C, Bergounioux C, Raymond P, Chevalier C (1999) Molecular and biochemical characterization of the involvement of cyclin-dependent kinase CDKA during the early development of tomato fruit. Plant Physiol. 121: 857–869

Kaló P, Seres A, Taylor SA, Jakab J, Kevei Z, Kereszt A, Endre G, Ellis THN, Kiss GB (2004) Comparative mapping between *Medicago sativa* and *Pisum sativum*. Mol Gen Gen 272: 235–246

Katekar GF (1999) Structure-activity relationships of plant growth regulators. In: Hooykaas PJJ, Hall MA, Libbenga KR (eds) Biochemistry and molecular biology of plant hormones. Elsevier Science BV, Amsterdam, pp. 89–111

Kowles RV, Phillips RL. 1985. DNA amplification patterns in maize endosperm nuclei during kernel development. Proc Natl Acad Sci USA 82: 7010–7014

Larkins BA, Dilkes BP, Dante RA, Coelho CM, Woo YM, Liu Y (2001) Investigating the hows and whys of DNA endoreduplication. J Exp Bot 52: 183–192

Lemontey C, Mousset-Declas C, Munier-Jolain N, Boutin JP (2000) Maternal genotype influences pea seed size by controlling both mitotic activity during early embryogenesis and final endoreduplication level/cotyledon cell size in mature seed. J Exp Bot 51: 167–175

Lopes MA, Larkins BA (1993) Endosperm origin, development, and function. Plant Cell 5: 1383–1399

Mambelli S, Setter TL (1998) Inhibition of maize endosperm cell division and endoreduplication by exogenously applied abscisic acid. Physiol Plant 104: 266–272

Melaragno JE, Mehrotra B, Coleman AW (1993) Relationship between endopolyploidy and cell size in epidermal tissue of Arabidopsis. Plant Cell 5: 1661–1668

Möller B, Weijers D (2009) Auxin control of embryo patterning. Cold Spring Harb Perspect Biol

Munier-Jolain NG and Ney B (1998) Seed growth rate in grain legumes II. Seed growth rate depends on cotyledon cell number. J Exp Bot 49: 1971–1976

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15: 473–497

Nadeau D (2009) Hormone metabolism and action in developing pea fruit. PhD Thesis. University of Alberta, Edmonton. Canada

Nambara E, Marion-Poll A (2005) Abscisic acid biosynthesis and catabolism. Ann Rev Plant Biol 56: 165–185

Ochatt SJ (2006) Flow cytometry (ploidy determination, cell cycle analysis, DNA content per nucleus). In: *Medicago truncatula* handbook, Chapter 2.2.7. Online version November 2006, http://www.noble.org/MedicagoHandbook/ ISBN number 0-9754303-1-9

Ochatt SJ (2008) Flow cytometry in plant breeding. CytometryA 73: 581–598

Ochatt SJ (2011) Immature seeds and embryos of *Medicago truncatula* cultured in vitro. In: Thorpe TA, Yeung EC (eds) Plant Embryo Culture: Methods and Protocols; Methods in Molecular Biology 710. Springer Protocols, Humana Press, New York, pp. 39–52

Ochatt SJ, Moessner A (2010) Rounding up plant cells. Int J Plant Biol 2010; 1: e8; doi: 10.4081/pb.2010.e8

Ochatt SJ, Mousset-Déclas C, Rancillac M (2000) Fertile pea plants regenerate from protoplasts when calluses have not undergone endoreduplication. Plant Sci 156: 177–183

Ochatt S, Muilu R, Ribalta F (2008) Cell morphometry and osmolarity as early indicators of the onset of embryogenesis from cell suspension cultures of grain legumes and model systems. Plant Biosystems 142: 480–486

Ozga JA, Reinecke DM, Ayele BT, Ngo P, Nadeau C and Wickramarathna AD (2009) Developmental and hormonal regulation of gibberellin biosynthesis and catabolism in pea fruit. Plant Physiol 150: 448–462

Perrot-Rechenmann C (2010). Cellular responses to auxin: Division versus expansion. Cold Spring Harb. Perspect. Biol. 2: a001446

Petrasek J, Mravec J, Bouchard R, Blakeslee JJ, Abas M, Seifertova D, Wisniewska J, Tadele Z, Kubes M, Covanova M, *et al* (2006) PIN proteins perform a rate-limiting function in cellular auxin efflux. Science 312: 914–918

Pless T, Bottger M, Hedden P, Graebe J (1984) Occurrence of 4-C1-indoleacetic acid in broad beans and correlation of its levels with seed development. Plant Physiol 74: 320–323

Quatrano RS, Bartels D, Ho THD, Pages M (1997) New insights into ABA-mediated processes. Plant Cell 9: 470–475

Reddy VM, Daynard TB (1983) Endosperm characteristics associated with rate of grain filling and kernel size in corn. Maydica 28: 339–355

Reinecke DM (1999) 4-Chloroindole-3-acetic acid and plant growth. Plant Growth Regul 27: 3–13

Ruzicka K, Simaskova M, Duclercq J, Petrasek J, Zazimalova E, Simon S, Friml J, Van Montagu MC, Benkova E (2009) Cytokinin regulates root meristem activity via modulation of the polar auxin transport. Proc Natl Acad Sci USA 106: 4284–4289

Schneider E, Kazakoff C, Wightman F (1985) Gas chromatography-mass spectrometry evidence for several endogenous auxins in pea seedlings organs. Planta 165: 232–241

Schruff MC, Spielman M, Tiwari S, Adams S, Fenby N, Scott RJ (2006) The AUXIN RESPONSE FACTOR 2 gene of Arabidopsis links auxin signalling, cell division, and the size of seeds and other organs. Development 133: 251–261

Scott RJ, Spielman M, Bailey J, Dickinson HG (1998) Parent-of-origin effects on seed development in *Arabidopsis thaliana*. Development 125: 3329–3341

Terasaka K, Blakeslee JJ, Titapiwatanakun B, Peer WA, Bandyopadhyay A, Makam SN, Lee OR, Richards EL, Murphy AS, Sato F, et al (2005) PGP4, an ATP binding cassette P-glycoprotein, catalyzes auxin transport in *Arabidopsis thaliana* roots. Plant Cell 17: 2922–2939

Traas J, Hulskamp M, Gendreau E, Hofte H (1998) Endoreduplication and development: Rule without dividing? Curr Opin Plant Biol 1: 498–503

Verdier J, Thompson RD (2008) Transcriptional Regulation of Storage Protein Synthesis During Dicotyledon Seed Filling. Plant Cell Physiol 49: 1263–1271

Verdier J, Kakar K, Gallardo K, Le Signor C,Aubert G, Schlereth A, Town CD, Udvardi MK, Thompson R (2008) Gene expression profiling of M. truncatula transcription factors identifies putative regulators of grain legume seed filling. Plant Mol Biol 67: 567–580

Weber H, Borisjuk L, Wobus U (1996) Controlling seed development and seed size in *Vicia faba*: A role for seed coat-associated invertases and carbohydrate state. Plant J 10: 823–834

Weber H, Borisjuk L, Wobus U (2005) Molecular physiology of legume seed development. Ann Rev Plant Biol 56: 253–279

Widholm JM (1972) The use of fluorescein diacetate and phenosaphranine for determining the viability of cultured cells. Stain Technol 47: 189–194

Wu H-M, Hazak O, Cheung AY, Yalovsky S (2011) RAC/ROP GTPases and auxin signalling. The Plant Cell 23: 1208–1218

Chapter 6: General Discussion

General Discussion

Seeds act not only as a carrier of genetic information from ancestors to the progenies but, in grain legumes, they also serve as an important source of food and feed. Seed development is an important process that determines the quantity and quality of different food reserves within the seed, which in turn not only determines the nutritional value of seeds for food or feed but also provides a rich resource to nourish the plantlet after germination.

Recent studies have revealed that seed development in *M. truncatula*, a model legume is transcriptionally regulated and that some of the TFs involved are expressed in specific tissues (Gallardo *et al.,* 2007; Verdier *et al.,* 2008). Among them, the *DOF1147* gene, belonging to a plant-specific Zn-finger family, has been shown to express in an endosperm-specific manner at the transition period between embryogenesis and seed filling. A mutation in its DNA-binding domain resulted in a drastic decrease in fresh seed weight, seed size and seed harvest index. These studies have also revealed the possible role of DOF1147 protein in hormonal signalling. As the storage protein synthesis in legume seeds takes place in the embryonic tissue, thus *DOF1147* may be involved in controlling gene expression in mid-term endosperm development by interacting indirectly to control the key processes during seed filling.

The members of the Dof family of TFs have been demonstrated to play a paramount role in mediating the expression of seed storage protein genes, in the developing endosperm of monocots by trancriptional activation as well as repression (Mena *et al.,* 1998; Yanagisawa, 2004). In cereals, *DOF* was shown to regulate SSP synthesis in the endosperm in association with B3 domain, bZIP and MYB TF family members (Yamamoto *et al.,* 2006). Dof proteins are recruited to the AAAG, *cis*-regulatory element found in promoters of target genes especially in the prolamin box of SSP gene promoters like glutelin (Yanagisawa *et al.,* 2002).

During the first part of my thesis, I carried out a thorough search for the identification of *DOF* genes in the *M. truncatula* genome. After the phylogenetic comparison of identified members of DOF TF family, the protein sequences have been observed for conserved

domains to determine the important evolutionary events in DOF protein genetic history. A motif-based search of putative target genes identified through microarray from *dof* mutants has revealed the presence of characteristic motifs for transcriptional and hormonal interaction.

In the second part, the expression pattern of *DOF1147* gene as well as its protein was then studied in detail by employing a transgenesis approach. Various transgene constructions were prepared using Gateway cloning technology. These constructs were then employed for stable transformation of *M. truncatula* as well as for transient gene expression. The microscopic observations of tissues/cells expressing the transgenes were recorded. The expression analyses using *CaM35S* promoter have demonstrated that DOF1147 protein is localized in the nucleus. On the other hand, the promoter strength of *DOF1147* was found to be too weak to drive the expression of GFP.

In the third part, an *in vitro* immature seed culture strategy was employed to analyse in more depth the mechanism of action of phytohormones during seed development. We focused on the role of auxin on the transition from the embryo cell division phase to seed maturation in *M. truncatula*. The *in vitro* seed development was analysed by flow cytometry, and through the determination of the kinetics of seed fresh weight and seed size.

The results obtained during my thesis as presented in the previous chapters are discussed in detail as follows.

6.1 Construction of plant transformation vectors using Gateway technology

The selection of plant expression vectors is very important not only for the preparation of desired constructs but also for the integration of the T-DNA into the plant genome. For example, the use of pBin19-derived vectors has been reported to give complex and abnormal T-DNA integration patterns in *M. truncatula* leading to formation of concatamers (Scholte *et al.,* 2002). In the last decade, the use of Gateway-compatible vectors has revolutionized the high throughput functional analysis of plant genes (Karimi *et al.,* 2007).

We employed the Gateway compatible destination vectors PMDC107 and pMDC83 for cloning of desired sequences. The backbone of both these vectors is derived from the pCambia series of binary vectors for *Agrobacterium*-mediated plant transformation (Curtis

and Grossniklaus, 2003). Moreover, both of these vectors confer hygromycin resistance which has been reported to be preferred as a selection agent in *M. truncatula* transformation (Cosson *et al.,* 2006; Atif *et al.,* 2013; Ochatt *et al.,* 2013) rather than kanamycin, which was shown to be problematic (as will be discussed in detail below).

Our results showed that the use of Gateway cloning technology ensured the correct orientation and reading frame of transgene cassette in the final construct. Although the optimization of parameter of BP and LR recombination reactions sometimes requires time, but once optimized, the constructs can be prepared in significantly less time as compared to conventional cloning techniques.

6.2 Stable genetic transformation of *M. truncatula*

Being a model legume plant, *M. truncatula* has been exploited in various genetic and physiological studies, especially for legume-specific characters. Thus, the development of an efficient regeneration and transformation system is imperative for this model crop. Somatic embryogenesis and direct organogenesis have been described for plant regeneration employing a wide range of tissues including leaflets, floral organs, cotyledons and in vitro-developed somatic embryos (Chabaud *et al.,* 1996; Trieu and Harison, 1996; Trinh *et al.,* 1998, Iantcheva *et al.,* 2005, Wright *et al.,* 2006). These studies have revealed that the regeneration in *M. truncatula* is highly source tissue and genotype dependent (Atif *et al.,* 2013). Among different genotypes described, R-108 has been demonstrated to exhibit an excellent ability to regenerate through somatic embryogenesis from leaf explants (Hoffman *et al.,* 1997; Trinh *et al.,* 1998; Ochatt *et al.,* 2013).

In our studies, the successful transformation of *M. truncatula* R-108 leaf explants and plant regeneration through somatic embryogenesis has been achieved. We also compared the efficiency of two *A. tumefaciens* strains i.e. LBA4404 and EHA105 to transform the R-108. LBA4404 has previously been reported to be inefficient for transformation of R-108 leaf explants (reviewed in Somers *et al.,* 2003, and Atif *et al.,* 2013) but our results demonstrated higher transformation efficiency with LBA4404 as compared to EHA105. About 50% of the initial explants were able to develop hygromycin resistant plantlets when transformed with LBA4404. The higher transformation efficiency with LBA4404 in our case may be attributed to the modification made to the protocol of Cosson *et al.,* (2006). We actually pre-cultured the

explants on SH3a regeneration media for 24-36 hours prior to inoculation of *Agrobacteria*. This might enhance their ability to withstand the sudden shock of transformation and render them more susceptible to *Agrobacterium* and thereby enhancing their transformation competence. LBA4404 has also been employed previously for transformation of *M. truncatula* using explants other than leaflets (Iantcheva *et al.,* 2005; Crane *et al.,* 2006).

Explant pre-culture on regeneration media prior to transformation has already been reported to increase transformation efficiency of different plant species including various seed legumes in *A. tumefaciens*-mediated transformation (Kar *et al.,* 1996; Karthikeyan *et al.,* 1996; Somers *et al.,* 2003; Barik *et al.,* 2005; Atif *et al.,* 2013). Lawrence and Koundal (2001) reported increased transformation efficiency when embryo axis explants were pre-cultured even one week before inoculation. Explant pre-culture has also been shown to increase transgene transfer efficacy in biolistics-mediated transformation (Gulati *et al.,* 2002). When observed for transient GUS expression after transformation, the pre-cultured explant tissues exhibited increased expression compared to non pre-cultured tissues (Sonia *et al.,* 2007). The increased transformation reported for pre-cultured explants could, probably, be due to cytokinin-induced active divisions of the cells which are in contact with the BAP-containing regeneration media (Sangwan *et al.,* 1992, 2010), leading to rapid DNA synthesis in host cells which appears to be required for T-DNA integration in the host genome (Schroeder *et al.,* 1993). Therefore, in conclusion, it can be postulated that certain cells become metabolically disposed to receive or integrate foreign DNA during pre-culture (Sonia *et al.,* 2007).

The use of acetosyringone (100 µM) in co-cultivation liquid and solid media in our experiments could also have contributed to the higher transformation efficiency with LBA4404 although we have not studied the results for medium without acetosyringone. Despite the use of acetosyringone being described as not needed for R108 transformation (Cosson *et al.,* 2006), its addition to the co-cultivation medium has been shown to increase transformation frequency both in other *M. truncatula* genotypes (Trieu and Harrison, 1996) and in other legumes (Sonia *et al.,* 2007, Švábová and Griga, 2008). The use of acetosyringone has also been shown to increase transformation potential of *Agrobacterium* strains with moderately virulent *vir* region in various other plant species (Sheikoleslam and Weeks, 1987; Godwin *et al.,* 1991; Wordragen and Dons, 1992).

Our results have shown a low transformation efficiency when EHA105 was employed which may be due to the use of leaflets from greenhouse-grown plants. Although they were disinfected before use, still some of them showed fungal contamination and so were eliminated. Despite the hypervirulent strain EHA105 having been reported in several studies to be highly efficient in transforming R-108 leaf explants with transgene constructs (Trinh *et al.,* 1998; Chabaud *et al.,* 2003; Araujo *et al.,* 2004), it has also been shown sometimes to cause bacterial overgrowth problems on the inoculated tissues (Chabaud *et al.,* 1996) and, most recently, to be not as efficient as other strains such as C58pMP90 that is also derived from the pCambia series of binary vectors for *Agrobacterium*-mediated plant transformation (Ochatt *et al.,* 2013). Similar bacterial overgrowth was observed on the developing calli as well as on the developing embryos in our case. Javot and Harrison (2007) also reported difficulty in eliminating EHA105 cells after their overgrowth, leading to death of plant tissues. Despite of all these problems, among the available mature plants tested about 96% showed transgene amplification by PCR. So, it can be concluded that if the explants are handled carefully to decrease the overgrowth risk after inoculation, higher transformation efficiency may be achieved.

These studies have shown hygromycin to be an efficient selection agent during R-108 transformation. Regardless of the *Agrobacterium* strain used, the recovery of transgenic plantlets was improved and was coupled with recovery of very few escapes as demonstrated by PCR verification. Previously, kanamycin has been widely employed for *in vitro* selection of transgenic *M. truncatula* plants (Thomas *et al.,* 1992; Chabaud *et al.,* 1996; Wang *et al.,* 1996; Kamaté *et al.,* 2000), but it led to the survival of a high percentage of escape plants (d'Erfurth *et al.,* 2003). The survival of escapes might be attributed to the natural resistance of legumes to aminoglycoside antibiotics like kanamycin (Christou, 1994).

6.3 Genome-wide identification of MtDOF TFs and their phylogenetic analysis

Transcriptional regulation is a fundamental mechanism underlying gene expression in all plant tissues. The TFs are generally classified into various gene families on the basis of presence of specific DNA-binding domains. The determination of phylogenetic relationship within different members of a specific family among different species serves as a clue for understanding their evolutionary process. The sequence homology-based classification of

certain protein family members in different groups and subgroups is also helpful in elucidation of their physiological roles and functions in the host species.

During the course of my thesis research, a comprehensive search has been conducted to identify the members of the Dof TF family in the *M. truncatula* genome database. A thorough overview of this TF family in *M. truncatula* has also been demonstrated, including a comparison of encoding proteins through multiple sequence alignment, phylogenetic relationship with other dicots i.e. *A. thaliana* and pea, as well as presence of conserved domains.

Multiple sequence alignment of Dof DNA-binding domains from newly annotated MtDof proteins has indicated the conservation of maximum number of domain residues as compared to Dof domain residues for monocots such as wheat and barley (Shaw *et al.,* 2009; Kushwaha *et al.,* 2011). The amino acids present in the Dof domain have been reported to play an essential role in the loop-like structure formed by the Zn-finger in the translated protein (Yanagisawa, 2004), and are important for the determination of protein function. The conservation of these residues in Dof domain of MtDof protein family could point out to the lesser frequency of natural mutation events during the evolutionary process (Shigyo *et al.,* 2007). Thus, this family seems to be evolutionarily more conserved in model legumes than in cereals.

The present comparative phylogenetic study of DOF TFs identified the distribution of MtDofs in six major groups Dof proteins of pea and Arabidopsis. Similar groups for AtDofs alone were also observed in some phylogenetic studies (Yanagisawa, 2002; Kushwaha *et al.,* 2011) while other studies divided the Dof protein family into four groups on the basis of the identification of major clusters of paralogous and orthologous genes (Lijavetzky *et al.,* 2003; Hernando-Amado *et al.,* 2012). Our results also confirmed the presence of such paralogous and orthologus proteins especially in *M. truncatula* and *A. thaliana*.

On the other hand, when comparing Dof domains from moss, green algae, rice and arbidopsis, Shigyo *et al.* (2007) classified the resulting phylogenetic tree into three major groups i.e. A, B and C. Group A contained domains from moss, the only domain from green algae and some domains from the two angiosperms i.e. rice and Arabidopsis, group B contained domains only from moss while group C showing all the domains from the rice and

Arabidopsis. Our Group II, containing most of the Dof proteins from *M. truncatula* and the remaining ones from *A. thaliana*, corresponds exactly to the group A of Shigyo *et al.* (2007), and the Dof proteins in this group are the most primitive ones sharing homology with the Dof domains of moss and green algae i.e. *Chlamydomonas*. From the comparison of both these trees, it can be speculated that the ancestral Dofs of the identified groups might have been multiplied independently of the multiplication of other ancestral genes (Moreno-Risueno *et al.,* 2007; Shigyo *et al.,* 2007).

When comparing the common domains identified in the respective Dof proteins, our results suggested the occurrence of critical evolutionary processes by multiple gene duplications resulting in paralogous genes during the evolution of Dof family in *M. truncatula*, *A. thaliana*, pea and PBF (Yanagisawa, 2002; Lijavetzky *et al.,* 2003). The results from conserved domain analysis of all the Dof proteins also highlighted the importance of using the complete protein sequence for the construction of phylogenetic tree rather than just Dof DNA binding domain for their functional classification. This is so because the C terminal of proteins may contain important functional domains that determine the role of that protein. For example, all three OBP proteins in *A. thaliana*, i.e. OBP1, OBP2 and OBP3 encoded by *AtDof* genes *AtDof3.4*, *AtDof1.1* and *AtDof3.6*, respectively, show a complete homology of 52 amino acids in their Dof domain but they exhibited completely different tissue-specific expression patterns (Kang and Singh, 2000). When examined thoroughly, it was revealed that these proteins contain transcriptional activation domains in their C-terminal region which are responsible for their distinct functions (Kang and Singh, 2000).

6.4 Expression of *Dof1147* in the seed endosperm

The cytological determination of the site of expression of a certain gene can help in studying its function in detail. So in order to determine the site of expression of *DOF1147* gene, the *M. truncatula* plants were stably transformed with the construct containing *pDof::gfp* expression cassette. The immature developing seeds from transgenic plants at 8, 10 and 12 DAP were collected and cytological observations were made to investigate the site of expression of *DOF1147* in the endosperm of the developing seed. The observations made were not conclusive.

Tissue specific transcriptional profiling data suggests that *DOF1147* has been expressed in the endosperm of developing *M. truncatula* seeds at 8-12 DAP (Verdier *et al.,* 2008). The endosperm specific expression of genes depends upon the availability of specific motifs in their promoter regions. The two core motifs i.e. an essential GCN4 element (TGASTCA) and an enhancer –300 core element (TGHAAARK) have been demonstrated to make up the bifactorial endosperm box which is responsible for driving the endosperm-specific expression of plant promoters (Quayle and Feix 1992; Wu *et al.,* 1998). Moreover, it has been demonstrated that the *cis*-regulatory elements AACA, ACGT and prolamin-box are required for quantitative regulation of the rice gluten gene in the endosperm of developing seeds (Wu *et al.,* 2000).

Based upon these findings, the 1.2 kb *DOF1147* promoter region used for the construct preparation was submitted for *in silico* analysis of motifs required for its expression. The results revealed the presence of three enhancer –300 core elements, two ACGT elements, two ACGT motifs and one prolamin-box element, but no GCN4 elements have been detected. This might be the possible reason for the lack of detection of GFP signals in the seed endosperm by *DOF1147* promoter because all the other elements found have been demonstrated to be unable to confer seed-specific expression without GCN4 element (Wu *et al.,* 2000). Therefore, a longer promoter fragment e.g. 2.0 kb, may be preferable for promoter expression studies.

Although the use of GFP has proven a useful visually screenable marker for plant transformation (Haseloff *et al.,* 1997), some problems of GFP fading have also been reported in transformed *M. truncatula* plants (Kamaté *et al.,* 2000; Zhou *et al.,* 2005). Moreover, the expression profiling results have indicated the native *Dof1147* promoter to be a weak promoter (Verdier *et al.,* 2008), which may barely drive detectable fluorescence from the GFP marker protein. So all the above described factors might have contributed to the failure of detection of tissue-specific *DOF1147* expression in plants transformed with this *pDof::gfp* construct.

6.5 Subcellular localization of Dof1147 protein

The transcription factors are generally involved in the transcriptional regulation of their target genes by interacting with the *cis*-regulatory elements mainly present in the

upstream regions of target genes. So, the import of transcription factors from the cytoplasm into the nucleus is mandatory for this post-translational regulation (Schwechheimer *et al.,* 1998). TF proteins having small sizes between 40–60 kDa can be imported into the nucleus by diffusion, while most of the TFs possess a special nuclear localization signal which is recruited by transporter proteins acting as a shuttle between the cytoplasm and the nucleus for their nuclear transport (Dingwall and Laskey, 1991; Raikhel, 1992; Laskey and Dingwall, 1993).

In our studies, DOF1147 protein belonging putatively to Zn-finger C2-C2 Dof type TF family has been predicted *in-silico* to localize in the nucleus. The experimental determination for subcellular localization of DOF1147 protein was carried out using a translational fusion protein construct containing N-terminal fused Histidine-tagged GFP with *DOF1147* promoter. The construct when used for transient gene expression has demonstrated the localization of DOF1147 protein in the nuclei of epidermal cells of tobacco leaves. Similar nuclear localization has been confirmed in leaf sections of stably transformed *M. truncatula* plants by accumulated green fluorescence signals. These studies are consistent with DOF1147 being a *bona fide* transcription factor.

During subcellular localization of DOF1147, the same GFP signals were also observed in cytoplasm and cell membranes in some transgenic lines of *M. truncatula*. In order to investigate this phenomenon, the DOF1147 protein sequence was submitted to Predictprotein online server (Rost *et al.,* 2004) for further analysis. This analysis identified the presence of 14 putative N-myristoylation sites in DOF1147 protein. The N-terminal myristoylation has been reported to play a vital role in membrane targeting and signalling (Podell and Gribskov, 2004). A couple of DOF proteins from each *A. thaliana* and rice, as well as some other proteins from TF families such as bZIP, Myb and WRKY have also been shown to contain N-terminal myristoylation sites (Podell and Gribskov, 2004). As these TF proteins must be localized in the nucleus to perform their functional role, it could be possible that they reside temporarily in cytosolic locations that require myristoylation. According to this assumption, the transient localization of DOF1147 protein in the cell membrane could be explained. On the other hand, cytoplasmic as well as nuclear localization of AtDof1.7 protein have been reported (Krebs *et al.,* 2010), and may be due to mutations in the NLS of DOF protein.

Although DOF1147 protein was predicted and later on experimentally proved to be nuclearly localized, the *in silico* prediction tools were unable to identify a nuclear localization sequence in DOF1147 protein as well as in other DOF proteins (Krebs *et al.,* 2010). Thus, it seems that DOF proteins contain a quite different, yet uncharacterized, NLS which is essentially different from the known NLSs, thus precluding its recognition by the available localization prediction tools (Krebs *et al.,* 2010). When studying the conserved residues in Dof domains of AtDof proteins, Krebs *et al.* (2010) reported the presence of highly conserved novel bipartite NLS in all the AtDof proteins. Both the basic regions of novel bipartite NLS i.e. KXCRR and RKXKR were separated by a unique 17 residues long linker. The functional validation of this novel NLS from AtDof4.7 and AtDof1.7 demonstrated the nuclear localized expression. Some proteins also harboured an extra monopartite NLS in the C terminal e.g. AtDof4.7 capable of nuclear localization independent of bipartite NLS (Krebs *et al.,* 2010).

The literature research has demonstrated that NLS from only a few DOF proteins has been identified experimentally (Krebs *et al.,* 2010). So, it is thought that DOF proteins possess a distinct nuclear localization signal not shared by other TFs. Usually a typical NLS consists of one or more short sequences of the basic amino acids lysine or arginine in the primary structure of the protein (Dingwall and Laskey, 1991). Different nuclear localized proteins may also share the same NLS (Raikhel, 1992). The maize transcription factor DOF1 has been experimentally demonstrated to contain a KKKPASKKRR amino acid sequence in the center of the protein that serves as an NLS (Yanagisawa *et al.,* 2001). This identified region is absent from all the other known DOF proteins, indicating that other motifs for nuclear targeting might exist in DOF transcription factors.

6.6 Role of auxin in seed development

Seed developmental being a critical process in the lifecycle of legumes like that of other angiosperms, provides a means to fulfil the nutritional requirements for humankind, their domestic animals as well as for the plants themselves. During the last decade, extensive studies have pointed out the key processes and factors controlling seed development in legumes (Weber *et al.,* 2005). Seed development is essentially a combination of cell division, expansion and differentiation events that are under the control of inherited developmental programmes modified by hormonal and environmental stimuli. It has now been functionally established that transitions between seed developmental phases are controlled by the

concerted action of different transcription factors and phytohormonal signalling pathways (Gutierrez *et al*. 2007; Verdier *et al.,* 2008). Among the different plant hormones, auxin is implicated virtually in every aspect of plant growth and development, including seed development (Woodward and Bartel, 2005). The role of auxins is much less well known in seeds than in tissues such as roots, even though seeds accumulate higher levels of auxin than any other plant tissue. On the basis of some preliminary results obtained from auxin treatments of *dof* mutants (CV. Le Signor, pers. comm), we investigated the role of an *in vitro* auxin treatment on the seed development in *M. truncatula*.

The results obtained (Ch. 5 of this thesis) demonstrated the overall stimulatory effects of the exogenously applied auxins on both the fresh weight and size of seeds with NAA more efficient to induce these effects than IBA. The key role of phytohormones in the alteration of cell division activities or in the initiation of differentiation is well endorsed by experimental data (Dudits *et al.,* 2011). For example, for auxin, both local biosynthesis and intercellular transport have been shown to contribute to the establishment and maintenance of active auxin gradients in the tissues (Vanneste and Friml, 2009; Sundberg and Ostergaard, 2009). At the cellular level in roots, auxin is perceived differentially in a tissue- and dose-dependent manner (Tromas *et al.,* 2009) and hence it has been reported to modulate cell division, expansion and differentiation (Stals and Inzé, 2001; Davies, 2010). The intercellular transport and signalling of auxin is accomplished mainly by transcriptional regulation-mediated pathways, while some transcription-independent mechanisms have also been reported for auxin-signalling in the cells. For example, for the former one, various proteins have been reported for auxin uptake in the root cells including AUX/LAX permeases (Terasaka *et al*. 2005) and ABP1 protein (Tromas *et al.,* 2009), while for the latter one, ABP1 protein present in the plasmalema is presumably an auxin receptor (Tromas *et al.,* 2009).

Once in the cell, the availability of auxins has been reported to enhance the transcription of drivers of the cell cycle i.e. the cyclin-dependent kinases (CDKs) and mitotic cyclin mRNA levels in roots which in turn can result in DNA replication and cell division (Hemerly *et al.,* 1993; Doerner *et al.,* 1996). Hence in our case, the increased cell divisions might be responsible for increased seed fresh weight and size. However, this is not the only mechanism but various parallel processes have also been held responsible for auxin-induced cell division. Several proteins have been reported to be responsible for auxin-mediated spatial and temporal regulation of cell division in plants (Hemerly *et al.,* 1993; Bhalerao and Bennett,

2003). For example, auxin is reported to promote pericycle cell division from G2 phase of cell cycle, which might be achieved by SCF^{TIR1}-induced degradation of negative regulators of the cell cycle (Gray *et al.,* 1999).

In our study, a delay in the onset of endoreduplication in the cells of developing seeds was also demonstrated. The onset of endoreduplication serves as a marker of the initiation of cell differentiation as well accumulation of protein reserves in the developing seed (Ochatt, 2006, 2008; 2011). The cell cycle-modulated expression pattern of auxin-induced genes has revealed that a discrete set of genes are expressed at different phases of the cell cycle i.e. S phase, or M phase, to control induction of mitotic- or endocycle (Breyne *et al.,* 2002). In line with our results with developing immature seeds of *M. truncatula*, high auxin levels have been demonstrated to inhibit and delay endocycles, prolonging the mitotic phase in Arabidopsis root meristems (Ishida *et al.,* 2010). Conversely, low auxin conditions rapidly downregulated the expression of several mitotic cycle genes, reduced mitotic activity and triggered entry into endoreduplication, promoting cell expansion (Ishida *et al.,* 2010). At the molecular level, auxin-induced downregulation of CDK activities at the G2-to-M transition has resulted in the cells transferring from the mitotic cell cycle to an endocycle and thereby leading to endoreduplication (Edgar and Orr-Weaver, 2001; Lee *et al.,* 2009). Inzé and De Veylder (2006) suggested downregulation of the expression of CDKB, specific only for the G2-to-M transition. E2FB proteins have been reported as one of the key targets for auxin, which determine whether cells proliferate through a normal mitotic cell cycle, or whether they exit the cell cycle, enlarge, and endoreduplicate their DNA (Magyar *et al.,* 2005). An EBP1-like protein (the ErbB3 epidermal growth factor receptor binding protein) that controls cell size is also known to be regulated in an auxin-dependent manner (Horvàth *et al.,* 2006). Thus, entry and exit from endoreduplication might be linked processes where auxin signalling plays a major role. A complete proposed mechanism of action of above described factors for the auxin-induced cell cycle regulation is elaborated in Fig. 31.

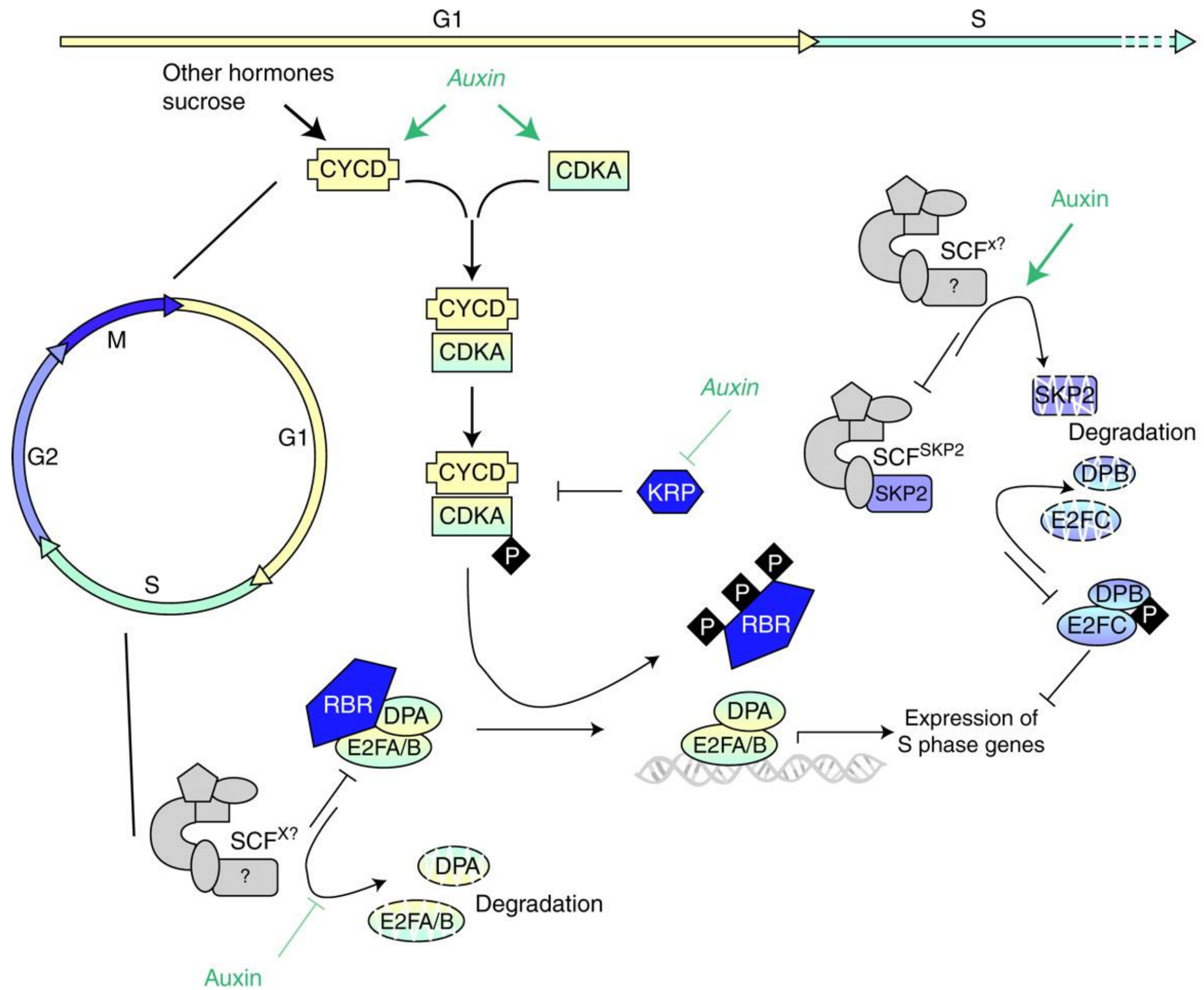


Figure 31. Role of Auxin in mitotic cell cycle. The cell cycle is divided into four phases: DNA replication (S), mitosis (M), and two Gap phases (G1 and G2, between M/S, and S/M, respectively). A normal cell cycle starts with G1, during which auxin induces expression of D-cyclins and cyclin-dependent kinase CDKA. CDK inhibitors (KRP) are reported to block the functioning of CDKA/CYCD complex while phosphorylation activates this complex. Auxin is shown to reduce the expression of inhibitors i.e. KRPs. CDK/CYCD complex in activated form provokes phosphorylation of retinoblastoma-related protein (RBR) transcriptional repressor, thus promoting expression of genes essential for the starting of the S phase under the control of the E2FA/B and DPA complex. At the end of S phase, the expression of S phase-specific genes is repressed by E2Fc and DPB. Auxin-induced E3 ubiquitine ligase SCFSKP2 pathway is involved in the degradation of these proteins. The genes active during different phases are indicated by colour codes. Auxin-induced factors are shown with green arrow bars, while auxin-repressed factors with T bars (Adapted from Perrot-Rechenmann, 2010)

Quite recently, characterization of the defective endosperm18 *(de18)* maize mutant has demonstrated its defect in auxin biosynthesis during endosperm development, leading to a significantly decreased seed dry mass due to a lower total cell number and reduced cell volume (Bernardi *et al.,* 2012). Flow cytometry profiles have also exhibited a low level of endoreduplication in the mutant endosperm cells (Bernardi *et al.,* 2012). These findings are in accordance with our results that auxin not only controls cell division but also the onset of endoreduplication in the developing seeds. So it is believed that at transcriptional level, the reprogramming of a variety of genes directly or indirectly associated with auxin signalling process is an essential step for modulation of the above studied traits during seed development. So, in order to dissect the role of auxin during seed development, it would be interesting to exercise a combined proteomic and transcriptomic approach to identify the target genes in different seed tissues. The promoter-reporter constructs could also be employed to explore the effects of auxin treatment on tissue-specific reporter expression during seed development.

At the transcriptional level, auxin-induced transcription factors i.e. ARF regulate the expression of auxin response genes by specifically binding to AuxR elements (TGTCTC) in the promoters of these genes (Ulmasov, 1997; Guilfoyle and Hagen, 2007; Hayashi, 2012). Our *in silico* analysis carried out on the promoters of selected putative target genes has also revealed the presence of AuxRE. Moreover in recent studies, the post transcriptional regulation of auxin-regulated genes especially ARFs by miRNAs has also been reported to play a role in seed development. In soybean, miRNAs are reported to silence ARF transcription factors in a tissue-specific manner in the seed coat (Zabala *et al.,* 2012), while Curaba and colleagues (2012) identified 7 *ARF* genes potentially regulated by conserved families of miRNAs and tasiRNAs during barley seed development. Since the *M. truncatula* genome encodes 24 *ARF* genes (Libault *et al.,* 2009), the transcriptional profiling of these genes and of the corresponding mutants and under different developmental or environmental conditions would be helpful to reveal the diverse transcriptional responses to auxin during *M. truncatula* seed development.

Conclusions and Perspectives

Conclusions and Perspectives

Genetic, cytological and environmental factors all contribute to the determination of balanced seed development. In this thesis research work, we combined different approaches including *in silico* motif analysis of promoter sequences, stable genetic transformation of *Medicago truncatula* and *in vitro* seed development studies to dissect the important contributors of seed development in the model legume, *Medicago truncatula*.

The use of an *in silico* approach for the identification of motifs in the promoter regions of putative target genes has enabled us to identify various previously described *cis*-regulatory elements. These elements are putative binding sites of various transcription factors and hormone signaling pathways implicated during seed development. These interactions can be experimentally determined by transient gene expression using reporter and expression constructs. Putative target genes of DOF1147 having motifs in their promoter regions could be verified by qRT-PCR, comparing WT and mutant lines, and by DNA-binding assays using the purified protein.

We further demonstrated an efficient genetic transformation protocol employing two *Agrobacterium tumefaciens* strains, LBA4404 and EHA105, to obtain a high percentage of stably transformed, phenotypically normal, *M. truncatula* plants. The transgene integration pattern and copy number in the primary transformants and their progeny could be verified by Southern blotting. This protocol highlighted that the physiological status of the explant as well as cocultivation conditions were important factors that considerably influenced the early events in T-DNA transfer. This transformation protocol can reliably be utilized for other functional genomics purposes.

On the basis of the labour-intensive analysis of *in vitro* seed development, it has been established that auxin concentration level in different tissues of developing seed, together with other local factors, contributes to cell differentiation and specification of the cell fate during seed development. Further studies with isolated immature embryos cultured with exogenous auxin might be envisaged to investigate the possibility of other seed tissues as sources of endogenous auxins. It would also be interesting to conduct a proteomic approach to study the

effect of different auxin concentrations on the proteome in the component seed tissues during the course of seed development.

What still remains to be established is how these relatively short auxin signalling cascades mediate so many different cellular responses, including cell proliferation, cell expansion and cell differentiation in phase specific manner during seed development. On the basis of available literature, it can be hypothesized that auxin-induced cellular responses during seed development involve multiple divergent and intersecting signal transduction pathways that presumably involve the activation or repression of different sets of genes and subsequent cellular activities. The use of a transcriptomic approach would help to identify the target genes involved in the auxin signaling pathway in different seed tissues.

Taking in account of the whole thesis work, our basic knowledge about the transcriptional and hormonal factors controlling seed development in model legume *M. truncatula* has been imrpoved. Our data not only elaborated the role of auxin in seed development but also provided new insights into the subsequent responses in seed development at the cytomic level in legumes.

References

References

Abdal-Aziz, S. A., Pliego-Alfaro, F., Quesada, M. A, & Mercado, J. A. (2006) Evidence of frequent integration of non-T-DNA vector backbone sequences in transgenic strawberry plant. *J Biosci & Bioeng*, 101(6): 508–10.

Abirached-Darmency, M., Abdel-Gawwad, M. R., Conejero, G., Verdeil, J. L. & Thompson, R. (2005) In situ expression of two storage protein genes in relation to histo-differentiation at mid-embryogenesis in *Medicago truncatula* and *Pisum sativum* seeds. *J Exp Bot*, 56: 2019-2028.

Allan, G. L., Parkinson, S., Booth, M. A., Stone, D. A. J., Rowland, S. J., Frances, J. & Warner-Smith, R. (2000) Replacement of fish meal in diets for Australian silver perch, *Bidyanus bidyanus*: I. Digestibility of alternative ingredients. *Aquaculture* 186: 293-310.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) Basic local alignment search tool. *J Mol Biol*, 215:403–410.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.,* 25: 3389-3402.

Ané J-M., Zhu, H. & Frugoli, J. (2008) Recent Advances in *Medicago truncatula* Genomics. *Int J Plant Genom,* 2008:256597. *In press* doi:10.1155/2008/256597

Angeles-Núnez, J. G. & Tiessen, A. (2010) Arabidopsis sucrose synthase 2 and 3 modulate metabolic homeostasis and direct carbon towards starch synthesis in developing seeds. *Planta*, 232: 701–18.

Araujo, S. S., Duque, A. S. R. L., dos-Santos D. M. M. F. & Fevereiro, M. P. S. (2004) An efficient transformation method to regenerate a high number of transgenic plants using a new embryogenic line of *Medicago truncatula* cv Jemalong. *Plant Cell Tiss Org Cult*, 78: 123–131.

Asamizu, E., Shimoda, Y., Kouchi, H., Tabata, S. & Sato, S. (2008) A positive regulatory role for *LjERF1* in the nodulation process is revealed by systematic analysis of nodule-associated transcription factors of *Lotus japonicus*. *Plant Physiol*. 147: 2030-2040.

Atif, R. M., Boulisset, F., Conreux, C., Thompson, R. & Ochatt, S. (2013) *In vitro* auxin treatment promotes cell division and delays endoreduplication in developing seeds of the model legume species *Medicago truncatula*. *Physiologia Plantarum, in press;* DOI: 10.1111/j.1399-3054.2012.1719.x (online since November 9, 2012)

Atif, R. M., Patat-Ochatt, E. M., Svabova, L., Ondrej, V., Klenoticova, H., Jacas, L., Griga, M. & Ochatt, S. J. (2013) Gene transfer in legumes. In: U. Lüttge et al (eds), *Progress in Botany* 74, DOI 10.1007/978-3-642-30967-0_2, Springer-Verlag, Berlin, 37-100.

Bailey T.L., Elkan C. (1995) Unsupervised Learning of Multiple Motifs In Biopolymers Using EM. *Mach. Learn.,* 21:51–80.

Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers", Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, pp. 28-36, AAAI Press, Menlo Park, California, 1994.

Bailey, T. L. & Gribskov, M., (1998) Combining evidence using p-values: application to sequence homology searches. *Bioinformatics*, 14(1): 48-54.

Bailey, T. L., Bodén, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W. & Noble, W. S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res,* 37: 202-208.

Bailey, T. L., Williams, N., Misleh, C. & Li, W. W. (2006) MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.,* 34: 369–373.

Barik, D. P., Mohapatra, U. & Chand. P. K. (2005) Transgenic grasspea (*Lathyrus sativus* L.): Factors influencing *Agrobacterium*-mediated transformation and regeneration. *Plant Cell Rep,* 24: 523–531.

Barker, D. G., Bianchi, S., London, F., Dattee, Y., Duc, G., Essad, S., Flament, P., Gallusci, P., Genier, G., Guy, P. et al. (1990) *Medicago truncatula*, a model plant for studying the molecular genetics of the Rhizobiumlegume symbiosis. *Plant Mol Biol*, 8: 40-49.

Baud, S., Boutin, J. P., Miquel, M., Lepiniec, L. & Rochat, C. (2002) An integrated overview of seed development in *Arabidopsis thaliana* ecotype ws. *Plant Physiol Biochem*, 40: 151–60.

Baud, S., Mendoza, M. S., To, A., Harscoet, E., Lepiniec, L. & Dubreucq, B. (2007) Wrinkled1 specifies the regulatory action of leafy cotyledon2 towards fatty acid metabolism during seed maturation in Arabidopsis. *Plant J.,* 50: 825–38.

Baumlein, H., Misera, S., Luerssen, H., Kolle, K., Horstmann, C., Wobus, U. & Mu¨ller, A. (1994) The FUS3 gene of *Arabidopsis thaliana* is a regulator of gene expression during late embryogenesis. *Plant J.,* 6: 367–387.

Benedito, V., Torez-Jerez, I., Murray, J. Andriankaja, I., Allen, I., Kakar, K., Wandrey, M., Verdier, J., Zuber, H., Ott, T., Moreau, S., Niebel, A., Frickey, T., Weiller, G., He, J., Dai, X., Zhao, P., Tang, Y., & Udvardi, Y. (2008) A gene expression atlas of the model legume *Medicago truncatula*. *Plant J.,* 55(3): 504--513.

Bernardi, J., Lanubile, A., Li, Q-B., Kumar, D., Kladnik, A., Cook, S. D., Ross, J. J., Marocco, A. & Chourey, P. S. (2012) Impaired auxin biosynthesis in the defective endosperm18 mutant is due to mutational loss of expression in the ZmYuc1 gene encoding endosperm-specific YUCCA1 protein in maize. Plant Physiol, *in press* http://dx.doi.org/10.1104/pp.112.204743

Bhalerao, R. P. & Bennett, M. J. (2003) The case for morphogens in plants. *Nat Cell Biol,* 5:939–43.

Blondon, F., Marie, D., Brown, S. & Kondorosi, A. (1994) Genomic size and base composition in *Medicago sativa* and *M. truncatula* species. *Genome,* 37: 264-270.

Boisson-Dernier, A., Andriankaja, A., Chabaud, M., Niebel, A., Journet, E. P., Barker, D. & de Carvalho-Neibel, F. (2005) *MtENOD11* gene activation during rhizobial infection and mycorrhizal arbuscule development requires a common AT-rich-containing regulatory sequence. *Mol. Plant-Microbe Interact,* 18: 1269-76.

Boisson-Dernier, A., Chabaud, M., Garcia, F., Be´card, G., Rosenberg, C. & Barker, D. G. (2001) Agrobacterium rhizogenes-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Mol Plant Microbe Interact*, 14: 695–700.

Bonnin, I., Huguet, T., Gherardi, M., Prosperi, J. M. & Olivier, I. (1996) High level of polymorphism and spatial structure in a selfing plant species, *Medicago truncatula* (Leguminosae), shown using RAPD markers. *American J. of Botany*, 83: 843–855.

Borisjuk, L., Rolletschek, H., Radchuk, R., Weschke, W., Wobus, U., & Weber, H. (2004) Seed development and differentiation: a role for metabolic regulation. *Plant biology*, *6*(4): 375–86.

Borisjuk, L., Rolletschek, H., Wobus, U. & Weber, H. (2003) Differentiation of legume cotyledons as related to metabolic gradients and assimilate transport into seeds. *J. Exp. Bot.*, 54: 503-512.

Borisjuk, L., Weber, H., Panitz, R., Manteuffel, R. & Wobus, U. (1995) Embryogenesis of Vicia faba L: Histodifferentiation in relation to starch and storage protein synthesis. *J. Plant Physiol,* 147: 203-218.

Brady, S. M., Sarkar, S. F., Bonetta, D. & McCourt, P. (2003) The abscicic acid insensitive 3 (*ABI3*) gene is modulated by farnesylation and is involved in auxin signaling and lateral root development inArabidopsis. *Plant J.,* 34: 67–75.

Braybrook, S. A., Stone, S. L., Park, S., Bui, A. Q., Le, B. H., Fischer, R. L., Goldberg, R. B., et al. (2006) Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. *Proc Nat Acad USA*, 103(9): 3468–73.

Breyne, P., Dreesen, R., Vandepeole, K., De Veylder, L., Van Breusegem F., Callewaert L., Rombauts S., Raes J., Cannoot B., Engler G., Inzé D. & Zabeau M. (2002) Transcriptome analysis during cell division in plants. *Proc Natl. Acad. Sci. USA* 12: 14825-14830.

Casimiro, A. C., Vinga, S., Freitas, A. T., & Oliveira, A. L. (2008) An analysis of the positional distribution of DNA motifs in promoter regions and its biological relevance. *BMC bioinformatics*, 9: 89.

Casson, S. A, & Lindsey, K. (2006) The turnip mutant of Arabidopsis reveals that LEAFY COTYLEDON1 expression mediates the effects of auxin and sugars to promote embryonic cell identity. Plant Physiol, 142(2): 526–41.

Castle, L. A. & Meinke, D. W. (1994) A FUSCA gene of Arabidopsis encodes a novel protein essential for plant development. *Plant Cell* 6: 25–41.

Castle, L.A. & Meinke, D.W. (1994) A *FUSCA* gene of Arabidopsis encodes a novel protein essential for plant development. *Plant Cell* 6(1): 25-41.

Chabaud, M., de Carvalho, N. F. & Barker, D. G. (2003) Efficient transformation of *Medicago truncatula* cv Jemalong using the hypervirulent Agrobacterium tumefaciensstrain AGL1. *Plant Cell Rep*, 22: 46–51.

Chabaud, M., Larsonneau, C., Marmouget, C. & Huguet, T. (1996) Transformation of barrel medic (*Medicago truncatula* Gaertn.) by Agrobacterium tumefaciens and regeneration via somatic embryogenesis of transgenic plants. *Plant Cell Rep*, 15: 305–310.

Chabaud, M., Ratet, P., Araújo S. S., Duque, A. S. R. L. A., Harrison, M. & Barker, D. G. *Agrobacterium tumefaciens*-Mediated Transformation and in vitro Plant Regeneration of *M. truncatula* . 2007. Available from http://www.noble.org/MedicagoHandbook/

Chabaud, M., Ratet, P., de Sousa Arau´jo, S., Roldao-Lopes, A. S., Duque, A., Harrison, M. & Barker, D. G. (2007) Agrobacterium tumefaciens-mediated transformation andin vitroplant regeneration of *M. truncatula*. In: *Medicago truncatula* handbook. *http://www.noble.org/MedicagoHandbook/pdf/AgrobacteriumTumefaciens*

Chandran, D., Sharopova, N., Ivashuta, S., Gantt, J. S., Vandenbosch, K. A. & Samac, D. A. (2008a) Transcriptome profiling identified novel genes associated with aluminum toxicity, resistance and tolerance in *Medicago truncatula*. *Planta* 228: 151–166.

Chandran, D., Sharopova, N., VandenBosch, K. A., Garvin, D. F. & Samac, D. A. (2008b) Physiological and molecular characterization of aluminum resistance in *Medicago truncatula*. *BMC Plant Biol*., 8: 89.

Cheng, X., Wen, J., Tadege, M., Ratet, P. & Mysore, K. S. (2010) Reverse genetics in *Medicago truncatula* using *Tnt1* insertion mutants. *Methods Mol Biol*, 678: 179-90.

Chilton, M. D. (2001) *Agrobacterium*. A memoir. *Plant Physiol*, 125: 9–14.

Choi, H. K., Kim, D., Uhm, T., Limpens, E., Lim, H., Mun, J. H., Kalo, P., Penmetsa, R.V., Seres, A., Kulikova, O., *et al.,* (2004) A sequence-based genetic map of *Medicago truncatula* and comparison of marker colinearity with *M. sativa*. *Genetics,* 166: 1463–1502.

Christou, P. (1994) The biotechnology of crop legumes. *Euphytica*, 74: 165-185.

Christou, P., Swain, W. F., Yang, N. S. & McCabe, D. E. (1989) Inheritance and expression of foreign genes in transgenic soybean plants. *Proc Natl Acad Sci USA,* 86: 7500–7504.

Clarke, J. L., Daniell, H. & Nugent, J. M. (2011) Chloroplast biotechnology, genomics and evolution: current status, challenges and future directions. *Plant Mol Biol*, 76: 207–209.

Clough, S. J., Bent, A. F. (1998) Floral dip: a simplified method forAgrobacterium-mediated transformation of*Arabidopsis thaliana*. *Plant J.,* 16: 735–743.

Cole, K. S. (1968) Membranes, ions and impulses: a chapter of classical biophysics. University of California Press, Berkeley.

Confalonieri, M., Borghetti, R., Macovei, A., Testoni, C., Carbonera, D., Salema-Fevereiro, M. P., Rommens, C., Swords, K., Piano, E. & Balestrazzi, A. (2010) Backbone-free transformation of barrel medic (*Medicago truncatula*) with a Medicago-derived transfer DNA. *Plant Cell Rep*, 29: 1013–1021.

Cook, D. R. (1999) *Medicago truncatula*-a model in the making. *Curr Opin Plant Biol,* 2: 301-4.

Cordain, L. (1999) The late role of grains and legumes in the human diet, and biochemical evidence of their evolutionary discordance. Internet: http://www.beyondveg.com/cordain-l/grains-leg/grains-legumes-1a.shtml (accessed 26 August 2012).

Cosson, V., Durand, P., d'Erfurth, I., Kondorosi A. & Ratet, P. (2006) *Medicago truncatula* R108-1 (c3) *Agrobacterium* transformation. In Methods in Molecular Biology K. Wang (ed): *Agrobacterium Protocols* (2nd edition). Humana press Inc, Totowa, N.J. USA.

Coulson, R. M. & Ouzounis, C. A. (2003) The phylogenetic diversity of eukaryotic transcription. *Nucleic Acids Res.* 31: 653–660.

Crane, C., Wright, E., Dixon, R. a, & Wang, Z.-Y. (2006). Transgenic *Medicago truncatula* plants obtained from *Agrobacterium tumefaciens*-transformed roots and *Agrobacterium rhizogenes*-transformed hairy roots. *Planta*, *223*(6): 1344–54.

Curaba, J., Moritz, T., Blervaque, R., Parcy, F., Raz, V., Herzog, M. & Vachon, G. (2004) AtGA3ox2, a key gene responsible for bioactive gibberellin biosynthesis, is regulated during embryogenesis by LEAFY COTYLEDON2 and FUSCA3 in Arabidopsis. *Plant Physiol,* 136: 3660-3669.

Curaba, J., Spriggs, A., Taylor, J., Li Z. & Helliwell, C. (2012) miRNA regulation in the early development of barley seed. *BMC Plant Biology,* 12: 120.

Curtis, M. D. & Grossniklaus, U. (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol,* 133: 462–469.

d'Erfurth, I., Cosson, V., Eschstruch, A., Lucas, H., Kondorosi, A. & Ratet, P. (2003) Efficient transposition of the Tnt1tobacco retrotransposon in the model legume *Medicago truncatula*. *Plant J.,* 34: 95-106.

Damiani, I., Baldacci-Cresp, F., Hopkins, J., Andrio, E., Balzergue, S., Lecomte, P., Puppo, A., Abad, P., Favery, B. & Hérouart, D. (2012) Plant genes involved in harbouring symbiotic rhizobia or pathogenic nematodes. *New Phytol.*, 194(2): 511-22.

Davies PJ (ed).2010.Plant hormones: biosynthesis, signal transduction, action! Revised 3rd edn, Dordrecht: Springer.

de Zélicourt, A., Diet, A., Marion, J., Laffont, C., Ariel, F., Moison, M., Zahaf, O., Crespi, M., Gruber, V. & Frugier, F. (2012) Dual involvement of a *Medicago truncatula* NAC transcription factor in root abiotic stress response and symbiotic nodule senescence. *Plant J.*, 70(2): 220-30.

Dharmasiri, N., Dharmasiri, S. & Estelle, M. (2005a) The F-box protein TIR1 is an auxin receptor. *Nature,* 435: 441-445.

Dharmasiri, N., Dharmasiri, S., Weijers, D., Lechner, E., Yamada, M., Hobbie, L., Ehrismann, J. S., Jurgens, G. & Estelle, M. (2005b) Plant development is regulated by a family of auxin receptor F box proteins. *Dev. Cell* 9: 109-119.

Diaz, I., Martinez, M., Isabel-LaMoneda, I., Rubio-Somoza, I., Carbonero, P. (2005) The DOF protein, SAD, interacts with GAMYB in plant nuclei and activates transcription of endosperm-specific genes during barley seed development. *Plant J.,* 42(5): 652–662.

Diaz, I., Vicente-Carbajosa, J., Abraham, Z., Martinez, M., Isabel-La Moneda, I. & Carbonero, P. (2002) The GAMYB protein from barley interacts with the DOF transcription factor BPBF and activates endosperm-specific genes during seed development. *Plant J..,* 29**:** 453-464.

Dickins, R. D., Reddy, M. S. S., Meurer, C. A., Remond, C. T. & Collins, G. B. (2003) Recent advance in soybean transformation. In: Jaiwal PK, Singh PK (eds) Applied genetics of leguminosae biotechnology. *Kluwer Academic Publishers, Great Britain,* pp 3–21

Dingwall, C. & Laskey, R. A. (1991) Nuclear targeting sequences—a consensus? *Trends Biochem. Sci.*, 16: 478–81.

Djemel, N., Guedon, D., Lechevalier, A., Salon, C., Miquel, M., Prosperi, J. M., Rochat, C. & Boutin, J. P. (2005) Development and composition of the seeds of nine genotypes of the *Medicago truncatula* species complex. *Plant Physiol Biochem* 43: 557–566.

Doerner, P., Jorgensen, J. E., You, R. Steppuhn, J. & Lamb, C. (1996) Control of root growth and development by cyclin expression. *Nature,* 380: 520–523.

Dolezel, J. Greilhuber, J. & Suda, J. eds. (2007) Flow cytometry with plant cells. Analysis of genes, chromosomes and genomes. Weinheim: Wiley-VCH Verlag GmbH & Co. KgaA. ISBN 978-3-527-31487-4; 454 pp.

Domoney, C., Duc, G., Ellis, T. H. N., et al (2006) Genetic and genomic analysis of legume flowers and seeds. *Current Opinion in Plant Biology,* 9: 133-141.

Droste, A., Pasquali, G. & Bodanese-Zanettini, M. H. (2000) Integrated bombardment and *Agrobacterium transformation* system: an alternative method for soybean transformation. *Plant Mol Biol Rep*, 18: 51–59.

Dudits, D., Ábrahám, E., Miskolczi, P., Ayaydin, F., Bilgin, M. & Horváth, G. V. (2011) Cell-cycle control as a target for calcium, hormonal and developmental signals: the role of phosphorylation in the retinoblastoma-centred pathway. *Annals of Botany*, 107 : 1193–1202.

Ebinuma, H., Sugita, K., Matsunaga, E., Endo, S., Yamada, K. & Komamine, A. (2001) Systems for the removal of a selection marker and their combination with a positive marker. *Plant Cell Rep*, 20(5): 383–392.

Edgar, B. A. & Orr-Weaver, T. L. (2001) Endoreplication cell cycles: More for less. *Cell,* 105: 297–306.

Edwards, D., Murray, J.A. & Smith, A.G. (1998) Multiple genes encoding the conserved CCAAT-box transcription factor complex are expressed in *Arabidopsis*. *Plant Physiol.* 117(3): 1015-22.

Ellwood, S. R., D'Souza, N. K., Kamphuis, L. G., Burgess, T. I., Nair, R. M. & Oliver, R. P. (2006) SSR analysis of the *Medicago truncatula* SARDI core collection reveals substantial diversity and unusual genotype dispersal throughout the Mediterranean basin. *Theor Appl Genet*: 112: 977-983.

Emery, R. J. N., Ma, Q. & Atkins, C. A., (2000) The forms and sources of cytokinins in developing white lupine seeds and fruits. Plant Physiol, 123: 1593–1604.

Ezcurra, I, Ellerstrom, M., Wycliffe, P., Stalberg, K. & Rask, L. (1999) Interaction between composite elements in the napA promoter: both the B-box ABA-responsive complex and the RY/G complex are necessary for seed-specific expression. *Plant Mol Biol,* 40: 699-709.

Fait, A., Angelovici, R., Less, H., Ohad, I., Urbanczyk-Wochniak, E., Fernie, A. R., et al. (2006) Arabidopsis seed development and germination is associated with temporally distinct metabolic switches. *Plant Physiol.*, 142: 839–54.

Finkelstein, R. R. & Gibson, S. I. (2002) ABA and sugar interactions regulating development: cross-talk or voices in a crowd?. *Current Opinion in Plant Biology*, 5: 26-32.

Focks, N. & Benning, C. (1998) Wrinkled1: a novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. *Plant Physiol.,* 118: 91–01.

Frey, A., Godin, B., Bonnet, M., Sotta, B., & Marion-Poll, A. (2004) Maternal synthesis of abscisic acid controls seed development and yield in Nicotiana plumbaginifolia. *Planta*, *218*(6): 958–64.

Friml, J., Vieten, A., Sauer, M., Weijers, D., Schwarz, H., Hamann, T., Offringa, R., & Jurgens, G. (2003). Efflux-dependent auxin gradients establish the apical-basal axis of Arabidopsis. *Nature*, 426: 147–153.

Gabriele, S., Rizza, A., Martone, J., Circelli, P., Costantino, P. & Vittorioso, P. (2010) The Dof protein DAG1 mediates PIL5 activity on seed germination by negatively regulating GA biosynthetic gene AtGA3ox1. *Plant J..,* 61: 312–323.

Gale, M. D. & Devos, K. M. (1998) Plant comparative genetics after 10 years. *Science* 282: 656-659.

Gallardo, K., Firnhaber, C., Zuber, H., et al (2007) A Combined Proteome and Transcriptome Analysis of Developing *Medicago truncatula* Seeds: Evidence for Metabolic Specialization of Maternal and Filial Tissues. *Mol. Cell Proteomics*, 6: 2165-2179.

Gallardo, K., Kurt, C., Thompson, R. & Ochatt, S. (2006) In vitro culture of immature *M. truncatula* grains under conditions permitting embryo development comparable to that observed in vivo. *Plant Science,* 170**:** 1052-1058.

Gallardo, K., Le Signor, C., Vandekerckhove, J. et al (2003) Proteomics of *Medicago truncatula* **s**eed development establishes the time frame of diverse metabolic processes related to reserve accumulation. *Plant Physiol.*, 133: 664-682.

Gallardo, K., Thompson, R., & Burstin, J. (2008) Reserve accumulation in legume seeds. *Comptes rendus biologies*, 331(10): 755–62.

Gálvez, L., González, E. M. & Arrese-Igor, C. (2005) Evidence for carbon flux shortage and strong carbon/nitrogen interactions in pea nodules at early stages of water stress. *J Exp Bot,* 56: 2551-2561.

Gamas, P., Niebel, F. D. C., Lescure, N. & Cullimore, J. V. (1996) Use of a subtractive hybridization approach to identify new *Medicago truncatula* genes induced during root nodule development. *Mol Plant Microbe Interact,* 9: 233–242.

Gao, M. J., Lydiate, D. J., Li, X., Lui, H., Gjetvaj, B., Hegedus, D. D., & Rozwadowski, K. (2009). Repression of seed maturation genes by a trihelix transcriptional repressor in Arabidopsis seedlings. *The Plant cell*, 21(1): 54–71.

Gazzarrini, S., Tsuchiya, Y., Lumba, S., Okamoto, M. & McCourt, P. (2004) The transcription factor FUSCA3 controls developmental timing in Arabidopsis through the hormones gibberellin and abscisic acid. *Dev. Cell,* 7: 373–385.

Gehring, M., Choi, Y. & Fischer, R. L. (2004) Imprinting and Seed Development. *Plant Cell*, 16: S203–S213

Gepts, P., Beavis, W. D., Brummer, E. C., Shoemaker, R. C., Stalker, H. T., Weeden, N. F., & Young, N. D. (2005) Legumes as a Model Plant Family . Genomics for Food and Feed Report of the Cross-Legume Advances through Genomics Conference. Plant Physiol, 137: 1228–1235.

Gianinazzi-Pearson, V., Séjalon-Delmas, N., Genre, A., Jeandroz, S. & Bonfante, P. (2007) Plants and Arbuscular Mycorrhizal Fungi: Cues and Communication in the Early Steps of Symbiotic Interactions. *Advances in Bot. Res.* 46: 181–219

Goda, H., Sawa, S., Asami, T., Fujioka, S., Shimada, Y. & Yoshida, S. (2004) Comprehensive comparison of auxin-regulated and brassinosteroid-regulated genes in Arabidopsis. *Plant Physiol,* 134: 1555–1573.

Godwin, I., Todd, G., Ford-Lloyd, B. & Newbury, H. J. (1991) The effect of acetosyringone and pH on Agrobacterium-mediated transformation vary according to plant species. *Plant Cell Rep,* 9: 671–675.

Goldberg, R. B., de Paiva, G. & Yadegari, R. (1994) Plant embryogenesis: zygote to seed. *Science* 266: 605–614.

Gong, W., Shen, Y. P., Ma, L. G., Pan, Y., Du, Y. L., Wang, D. H., Yang, J. Y., Hu, L. D., Liu, X. F., Dong, C. X. *et al.,* (2004) Genome-wide ORFeome cloning and analysis of *Arabidopsis* transcription factor genes. *Plant Physiol,* 135: 773–782.

Gonzalez-Rizzo, S., Crespi, M. & Frugier, F. (2006) The *Medicago truncatula* CRE1 cytokinin receptor regulates lateral root development and early symbiotic interaction with Sinorhizobium meliloti. *Plant Cell* 18: 2680–2693.

Graham, P. H. & Vance, C. P. (2003) Legumes: importance and constraints to greater use. *Plant Physiol,*131: 872–877.

Graham, P. H., & Vance, C. P. (2003) Update on Legume Utilization Legumes : Importance and Constraints to Greater Use. Plant Physiol, 131: 872–877.

Grant, G., More, L. J., McKenzie, N. H., & Stewart, J. (1983) A survey of the nutritional and haemagglutination properties of legume seeds generally available in the UK. *British J. Nutrition.,* 50: 207-214.

Gray, W. M., del-Pozo, J. C., Walker, L., Hobbie, L., Risseeuw, E., Banks, T., Crosby, W. L., Yang, M., Ma, H. & Estelle, M. (1999) Identification of an SCF ubiquitin-ligase complex required for auxin response in *Arabidopsis thaliana*. *Genes Dev*., 13: 1678–1691.

Groot, P. C., Bniinsma, J., & Karssen, C. M. (1987) The role of endogenous gibberellin in seed and fruit development of tomato : Studies with a gibberellin-deficient mutant. *Physiol. Plant*, 71: 184-190.

Gualberti, G., Papi, M., Bellucci, L., Ricci, I., Bouchez, D., Camilleri, C., Costantino, P., Vittorioso, P. (2002) Mutations in the Dof zinc finger genes DAG2 and DAG1 influence with opposite effects the germination of Arabidopsis seeds. *Plant Cell,* 14(6): 1253–1263.

Guilfoyle, T. J. & Hagen, G. (2007) Auxin response factors. *Current opinion in plant biology*, 10(5): 453–60.

Guillamon, E., Pedrosa, M. M., Burbano, C., Cuadrado, C., Sanchez, M. C. & Muzquiz, M. (2008) The trypsin inhibitors present in seed of different grain legume species and cultivar. *Food Chemistry,* 107: 68–74.

Gulati, A., Schryer, P. and McHughen, A. (2002) Production of fertile transgenic lentil (*Lens culinaris* Medik) plants using particle bombardment. *In Vitro Cell Dev Biol Plant*, 38: 316–324.

Guldan, S. J., Brun, W. A. (1985) Relationship of Cotyledon Cell Number and Seed Respiration to Soybean Seed Growth. *Crop Sci,* 25: 815-819.

Guo, Y., Qin G., Gu, H. & Qu, L-J. (2009) *HCA2*, a Dof transcription factor gene, regulates interfascicular cambium formation and vascular tissue development in *Arabidopsis*. *Plant Cell*, 21: 3518-3534.

Gupta, Y. P. (1987) Antinutritional and toxic factors in food legumes: a review. *Plant Foods for Human Nutrition,* 37: 201-228.

Gutierrez, L., Van Wuytswinkel, O., Castelain, M. & Bellini, C. (2007). Combined networks regulating seed maturation. *Trends in plant Science*, 12(7): 294–300.

Hagen, G., & Guilfoyle, T. (2002) Auxin-responsive gene expression: genes, promoters and regulatory factors. *Plant molecular biology*, 49(3-4): 373–85.

Handberg, K. & Stougaard, J. (1992) *Lotus japonicus*, and autogamous, diploid legume species for classical and molecular genetics. *Plant J.,* 2: 487-496.

Hansen, J., Jorgensen, J. E., Stougaard, J. & Marcker, K. A. (1989) Hairy roots - a short cut to transgenic root nodules. *Plant Cell Rep.,* 8: 12-15.

Haseloff, J., Siemering, K. R., Prasher, D. C. & Hodge, S. (1997) Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. *Proc Natl Acad Sci USA*, 94: 2122–2127.

Hayashi, K. (2012). The interaction and integration of auxin signaling components. *Plant & cell physiology*, 53(6): 965–75.

He, J., Vagner, A. B., Wang, M., Jeremy, D. M., Zhao, P. X.,Tang, Y. & Udvardi, M. K. (2009) The *Medicago truncatula* gene expression atlas web server. *BMC Bioinformatics,* 10: 441.

Heichel, G. H. (1987). Legumes as a source of nitrogen in conservation tillage. In J. F. Power (ed.) The Role of legumes in conservation tillage systems. Proc. Natl. Conserv. Tillage Conf., Athens, GA. 27-29 April. Soil Conservation Society of America. Ankeny, IA.

Hellens, P., Edwards, E. A., Leyland, N. R., Bean, S. & Mullineaux, P. M. (2000) pGreen, a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol. Biol,* 42: 819–832;

Hellens, R., Mullineaux, P. & Klee, H. (2000) A guide to Agrobacterium binary Ti vectors. *Trends in Plant Sci,* 5: 446-451.

Hemerly, A. S., Ferreira, P. de-Almeida-Engler, J., Van-Montagu, M., Engler, G. & Inzé, D. (1993) cdc2a expression in Arabidopsis is linked with competence for cell division. *Plant Cell,* 5: 1711–1723.

Hernando-Amado, S., González-Calle, V., Cabonero, P. & Barrero-Calle, C. (2012) The family of DOF transcription factors in *Brachypodium distachyon*: phylogenetic comparison with rice and barley DOFs and expression profiling. *BMC Plant Biol,* doi:10.1186/1471-2229-12-202

Higo, K., Ugawa, Y., Iwamoto, M. & Higo, H. (1998) PLACE: A database of plant cis-acting regulatory DNA elements. *Nucleic Acids Res.,* 26: 358-359.

Higo, K., Ugawa, Y., Iwamoto, M. & Korenaga, T. (1999) Plant cis-acting regulatory DNA elements (PLACE) database:1999. *Nucleic Acids Research,* 27 (1): 297-300.

Hills, M. J. (2004) Control of storage-product synthesis in seeds. *Current Opinion in Plant Biol,* 7: 302-308.

Hoekma, A., Hirsch, P. R., Hooykass, P. J. J. & Schilperoort, R. A. (1983) A binary plant vector strategy based on separation of the *vir*-and T-region of *Agrobacterium tumefaciens* Ti plasmid. *Nature,* 303:179-180.

Hoffmann, B., Trinh, T. H., Leung, J., Kondorosi, A., & Kondorosi, E. (1997). A New *Medicago truncatula* Line with Superior *in vitro* Regeneration , Transformation , and Symbiotic Properties Isolated Through Cell Culture Selection. *MPMI*, *10*(3): 307–315.

Hood, E.E., Gelvin, S.B., Melchers, L.S. & Hoekema, A. (1993) New *Agrobacterium* helper plasmids for gene transfer to plants. *Science,* 227: 1229-1231.

Horváth, B. M., Magyar, Z., Zhang, Y. X., Hamburger, A. W., Bakó, L., Visser, R. G., Bachem, C. W., Bögre, L. (2006) EBP1 regulates organ size through cell growth and proliferation in plants. *EMBO J,* 25: 4909–4920.

Hosoda, K., Imamura, A., Katoh, E., Hatta, T., Tachiki, M., Yamada, H., Mizuno, T. & Yamazaki, T. (2002) Molecular structure of the GARP family of plant Myb-related DNA binding motifs of the Arabidopsis response regulators. *Plant Cell.* 14(9): 2015-29.

Huang, H., Tudor, M., Weiss, C. A., Hu, Y. & Ma, H. (1995) The Arabidopsis MADS-box gene AGL3 is widely expressed and encodes a sequence-specific DNA-binding protein. *Plant Mol. Biol.* 28(3): 549-67.

Hwang, I., Sheen, J., & Müller, B. (2012) Cytokinin signaling networks. *Annual review of plant biology*, 63: 353–80.

Iantcheva, A., Vlahova, M. & Atanassov, A. (2005). Genetic transformation of *Medicago truncatula* using system for direct somatic embryogenesis promoted by TDZ. *Biotechnol. & Biotechnol. Eq*. 19(2): 50-56.

Imaizumi, T., Schultz, T. F., Harmon, F.G., Ho, L. A. & Kay, S. A. (2005) FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in *Arabidopsis*. *Science,* 309: 293–297.

Imin, N., Nizamidin, M., Daniher, D. et al (2005) Proteomic Analysis of Somatic Embryogenesis in *Medicago truncatula*. Explant Cultures Grown under 6-Benzylaminopurine and 1-Naphthaleneacetic Acid Treatments. *Plant Physiol* 137: 1250-1260.

Inze, D. & De Veylder, L. (2006) Cell cycle regulation in plant development. *Annu Rev Genet*, 40: 77–105.

Iqbal, A., Khalil, I. A., Ateeq, N. & Khan, M. S. (2006) Nutritional quality of important food legumes. *Food Chemistry,* **97**(2): 331-335.

Isabel-LaMoneda, I., Diaz, I., Martinez, M., Mena, M. & Carbonero, P (2003) SAD: a new DOF protein from barley that activates transcription of a cathepsin B-like thiol protease gene in the aleurone of germinating seeds. *Plant J.,* 33(2): 329–340.

Ishida, T., Adachi, S., Yoshimura, M., Shimizu, K., Umeda, M. 1 Sugimoto, K. (2010) Auxin modulates the transition from the mitotic cycle to the endocycle in Arabidopsis. *Development* 137: 63–71.

Ito, J., Sono, T., Tasaka, M., & Furutani, M. (2011) MACCHI-BOU 2 is required for early embryo patterning and cotyledon organogenesis in Arabidopsis. *Plant & cell physiology*, 52(3): 539–52.

Jenik, P. D. & Barton, M. K. (2005) Surge and destroy: the role of auxin in plant embryogenesis. *Development,* 132: 3577–3585.

Jensen, J. S., Marcker, K. A., Otten, L. & Schell, J. (1986) Nodule-specific expression of a chimaeric soybean leghaemoglobin gene in transgenic Lotus corniculatus. *Nature*, 321: 669-674.

Julier, B., Flajoulot, S., Barre, P., Cardinet, G., Santoni, S., Huguet, T. & Huyghe, C. (2003) Construction of two genetic linkage maps in cultivated tetraploid alfalfa (Medicago sativa) using microsatellite and AFLP markers. *BMC Plant Biology* 3:9-27.

Kagaya, Y. Toyoshima, R., Okuda, R., Usui, H., Yamamoto, A. & Hattori, T. (2005) LEAFY COTYLEDON1 controls seed storage protein genes through its regulation of FUSCA3 and ABSCISIC ACID INSENSITIVE3. *Plant Cell Physiol.* 46: 399–406.

Kalo, P., Seres, A., Taylor, S. A., Jakab, J., Kevei, Z., Kereszt, A., Endre, G., Ellis, T. H. & Kiss, G. B. (2004) Comparative mapping between *Medicago sativa* and *Pisum sativum*. *Mol Genet Genomics,* 272: 235–246.

Kamaté, K., Rodriguez-Llorente, I. D., Scholte, M., Durand, P., Ratet, P., Kondorosi, E., Kondorosi, A. & Trinh, T. H. (2000) Transformation of floral organs with GFP in *Medicago truncatula*. *Plant Cell Rep*, 19: 647–653.

Kang, H. G. & Singh, K. B. (2000) Characterization of salicylic acid-responsive, arabidopsis Dof domain proteins: overexpression of OBP3 leads to growth defects. *Plant J.* 21(4): 329–39.

Kar S, Johnson TM, Nayak P and Sen SK (1996) Efficient transgenic plant regeneration through Agroabcterium-mediated transformation of chickpea (*Cicer arietinum* L.). *Plant Cell Rep*, 16: 32-37.

Karimi, M., Depicker, A. & Hilson, P. (2007) Recombinational Cloning with Plant Gateway Vectors. *Plant Physiol,* 145(4): 1144-1154.

Karimi, M., Inzé, D., and Depicker, A. (2002). GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci*, 7: 193–195.

Karssen, C. M., Brinkhorst-Van der Swan, D. L. C., Breekland, A. D. & Koornneef, M. (1983) Induction of seed dormancy during seed development by endogenous abscisic acid: studies on abscisic acid-deficient genotypes of *Arabidopsis thaliana*L. *Heynh. Planta* 157: 158–165.

Karthikeyan, A. S., Sharma, K. S. & Veluthambi, K. (1996) *Agrobacterium tumefaciens*-mediated transformation of Vigna mungo (L.) Hepper. *Plant Cell Rep*, 15: 328-331.

Kawakatsu, T. & Takaiwa, F. (2010) Cereal seed storage protein synthesis: fundamental processes for recombinant protein production in cereal grains. *Plant Biotech J*, *8*(9): 939–953.

Kepinski, S., & Leyser, O. (2005) The Arabidopsis F-box protein TIR1 is an auxin receptor. *Nature,* 435: 446–451.

Kiirika, L. M., Bergmann, H. F., Schikowsky, C., Wimmer, D., Korte, J., Schmitz, U., Niehaus, K. & Colditz, F. (2012) Silencing of the Rac1 GTPase MtROP9 in *Medicago truncatula* stimulates early mycorrhizal and oomycete root colonizations but negatively affects rhizobial infection. *Plant Physiol.* 159: 501-516.

Kisu, Y., Esaka, M., & Suzuki, M. (1995) Putative zinc binding domain of plant transcription factor AOBP is related to DNA binding domains of steroid hormone receptors and GATA1. *Proc Jpn Acad* 71: 288–292.

Konishi, M., & Yanagisawa, S. (2007) Sequential activation of two Dof transcription factor gene promoters during vascular development in *Arabidopsis thaliana*. Plant Physiol *and biochemistry :* 45(8): 623–9.

Kononov, M. E., Bassuner, B. & Gelvin, S. B. (1997) Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome: evidence for multiple complex patterns of integration. *Plant J.,* 11: 945–957.

Kouchi, H., Shimomura, K., Hata, S., Hirota, A., Wu, G. J., Kumagai, H., Tajima, S., Suganuma, N., Suzuki, A., Aoki, T, et al (2004) Large-scale analysis of gene expression profiles during early stages of root nodule formation in a model legume, *Lotus japonicus*. *DNA Res* 11: 263–274.

Krebs, J., Mueller-Roeber, B. & Ruzicic, S. (2010) A novel bipartite nuclear localization signal with an atypically long linker in DOF transcription factors. *J Plant Physiol*, 167: 583–586.

Kroj, T., Savino, G., Valon, C. et al (2003) Regulation of storage protein gene expression in *Arabidopsis. Development*, 130: 6065-6073.

Kumar, K., Yadav, S., Purayannur, S. & Verma, P. K. (2012) An alternative approach in Gateway cloning when the bacterial antibiotic selection cassettes of the entry clone and destination vector are the same. *Mol Biotechnol,* doi: 10.1007/s12033-012-9549-0.

Kushwaha, H., Gupta, S., Singh, V. K., Rastogi. S. & Yadav, D. (2011) Genome wide identification of Dof transcription factor gene family in sorghum and its comparative phylogenetic analysis with rice and Arabidopsis. *Mol Biol Rep,* 38(8): 5037–5053.

Lallés, J. P. (1993) Nutritional and antinutritional aspects of soyabean and field pea proteins used in veal calf production: a review. *Livestock Production Sci,* 34: 181-202.

Landy, A. (1989) Dynamic, structural, and regulatory aspects of Lambda site-specific recombination. *Ann Rev of Biochem,* 58: 913-949.

Lara, P., Onate-Sanchez, L., Abraham, Z., Ferrandiz, C., Diaz, I., Carbonero, P. & Vicente-Carbajosa, J. (2003) Synergistic activation of seed storage protein gene expression in Arabidopsis by ABI3 and two bZIPs related to OPAQUE2. *J Biol Chem.* 278(23): 21003-11.

Laskey, R. A. & Dingwall, C. (1993) Nuclear shuttling: the default pathway for nuclear proteins? *Cell*, 74: 585–86.

Lawrence, P.K. & Koundal, K.R. (2001) *Agrobacterium tumefaciens* mediated transformation of pigeon pea [*Cajanus cajan* (L.) Millsp.] and molecular analysis of regenerated plants. *Curr. Sci*., 80: 1428-1432.

Le, B. H., Wagmaister, J. A., Kawashima, T., Bui, A. Q., Harada, J. J. & Goldberg R. B. (2007) Using genomics to study legume seed development. *Plant Physiol,* 144: 562–574.

Lee, H. O., Davidson, J. M. & Duronio, R. J. (2009) Endoreplication: polyploidy with purpose. *Genes & development*, 23(21): 2461–77.

Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouze, P. & Rombauts, S. (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Research,* 30: 325–327.

Le-Signor, C., Savois, V., Aubert, G., Verdier, J., Nicolas, M., Pagny, G., Moussy, F., Sanchez, M., Baker, D., Clarke J., & Thompson R. (2009) Optimizing TILLING populations for reverse genetics in *Medicago truncatula*. *Plant Biotechnology Journal,* 7**:** 430-441.

Lijavetzky, D., Carbonero, P. & Vicente-Carbajosa, J. (2003) Genome-wide comparative phylogenetic analysis of the rice and Arabidopsis Dof gene families. *BMC Evol Biol,* 2003: 3:17.

Limpens, E, Ramos, J., Franken, C., Raz, V., Compaan, B., Franssen, H., Bisseling, T. & Geurts, R. (2004) RNA interference in Agrobacterium rhizogenes-transformed roots of Arabidopsis and *Medicago truncatula*. *J Exp Bot,* 55: 983–992.

Liscum, E., & Reed, J. W. (2002) Genetics of Aux/IAA and ARF action in plant growth and development. *Plant molecular biology*, 49(3-4): 387–400.

Liu, H., Trieu, A. T., Blaylock, L. A., & Harrison, M. J. (1998) Cloning and characterization of two phosphate transporters from *M. truncatula* roots: Regulation in response to phosphate and to colonization by arbuscular mycorrhizal (AM) fungi. *Mol. Plant Microbe Interact,* 11: 14-22.

Lotan, T., Ohto, M. A., Yee, K. M. et al (1998) *Arabidopsis* LEAFY COTYLEDON1 Is Sufficient to Induce Embryo Development in Vegetative Cells. *Cell* 93: 1195-1205.

Lu, Q. S., dela Paz, J., Pathmanathan, A., Chiu, R. S., Tsai, A. Y., Gazzarrini, S. (2010) The C-terminal domain of FUSCA3 negatively regulates mRNA and protein levels and mediates sensitivity to the hormones abscisic acid and gibberellic acid in Arabidopsis. *Plant J.*, 64: 100-113.

Luerssen, K., Kirik, V., Herrmann, P., Misera, S. (1998) Fusca3 encodes a proteinwith a conserved vp1/abi3-like b3 domain which is of functional importance for the regulation of seed maturation in *Arabidopsis thaliana*. *Plant J.*, 15: 755–64.

Magyar, Z., De Veylder, L., Atanassova, A., Bako, L., Inze, D. & Bogre, L. (2005) The role of the *Arabidopsis* E2FB transcription factor in regulating auxin-dependent cell division. *Plant Cell,* 17: 2527 – 2541.

Mallory, A. C., & Vaucheret, H. (2004) MicroRNAs: Something important between the genes. *Curr. Opin. Plant Biol.* 7: 120–125.

Mallory, A. C., Bartel, D. P., & Bartel, B. (2005) MicroRNA-Directed Regulation of Arabidopsis AUXIN RESPONSE FACTOR17 Is Essential for Proper Development and Modulates Expression of Early Auxin Response Genes, *Plant Cell*, 17: 1360–1375.

Martínez-García, J. F., Moyano, E., Alcocer, M. J. & Martin, C. (1998) Two bZIP proteins from Antirrhinum flowers preferentially bind a hybrid C-box/G-box motif and help to define a new sub-family of bZIP transcription factors. *Plant J..,* 13(4): 489-505.

Marzabal. P., Gas, E., Fontanet, P., Vicente-Carbajosa, J., Torrent, M. & Ludevid, M. D. (2008) The maize Dof protein PBF activates transcription of gamma-zein during maize seed development. *Plant Mol Biol,* 67(5): 441–454.

Matsuo, S., Kikuchi, K., Fukuda, M., Honda I. & Imanishi, S. (2012) Roles and regulation of cytokinins in tomato fruit development. *J. Exp. Bot*. doi:10.1093/jxb/ers207

McCabe, D. E., Swain, W. F., Martinell, B. J. & Christou, P. (1988) Stable transformation of soybean (Glycine max) by particle acceleration. *Bio/Technology* 6: 923–926.

Meinke, D. W., Cherry, J. M., Dean, C., Rounsley, S. D. & Koornneef, M. (1998) *Arabidopsis thaliana*: a model plant for genome analysis. *Science,* 282: 662–682.

Meinke, D. W., Franzmann, L. H., Nickle, T. C. & Yeung, E. C. (1994) leafy cotyledon mutants of *Arabidopsis*. *Plant Cell,* 6: 1049–1064.

Mena, M., Cejudo, F. J., Isabel-Lamoneda, I., Carbonero, P. (2002) A role for the DOF transcription factor BPBF in the regulation of gibberellin-responsive genes in barley aleurone. *Plant Physiol,* 130(1): 111–119.

Mena, M., Vicente-Carbajosa, J., Schmidt, R. J. & Carbonero, P. (1998) An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm. *Plant J.*, 16(1): 53–62.

Monke, G., Altschmied, L., Tewes, W. R., Mock, H. P., Baumlein, H. & Conrad, U. (2004) Seed-specific transcription factors ABI3 and FUS3: molecular interaction with DNA. *Planta,* 219: 158–166.

Monke, G., Michael, S., Jens, K., Michaela, M., Ivo, G., Urs, H., Astrid, J., Bernd, W., Udo, C., Helmut, B. & Lothar A. (2012) Toward the identification and regulation of the *Arabidopsis thaliana* ABI3 regulon. *Nucleic Acids Res*, 40: 8240-8254.

Moreno-Risueno, M. A., Martinez, M., Vicente-Carbajosa, J. & Carbonero, P. (2007) The family of DOF transcription factors: from green unicellular algae to vascular plants. *Mol Genet Genomics,* 277(4): 379–390.

Morris, R. O. (1997) Hormonal regulation of seed development. In: Larkins BA, Vasil IK, eds.Cellular and molecular biology of plant seed development. Dordrecht, The Netherlands: Kluwer Academic Publishers, 117–148.

Muller, B. & Sheen, J. (2008) Cytokinin and auxin interaction in root stem-cell specification during early embryogenesis. *Nature*, 453: 1094–1097.

Mun, J. H., Kim, D. J., Choi, H. K., Gish, J., Debelle, F., Mudge, J., Denny, R., Endre, G., Saurat, O., Dudez, A. *et al.* (2006) Distribution of Microsatellites in the Genome of *Medicago truncatula*: A Resource of Genetic Markers That Integrate Genetic and Physical Maps. *Genetics*, 172: 2541-2555.

Munier-Julain, N. G. & Ney, B. (1998) Seed growth rate in grain legumes II. Seed growth rate depends on cotyledon cell number. *J. Exp. Bot.,* 49: 1971-1976.

Mylona, P., Pawlowski, K. & Bisseling, T. (1995) Symbiotic nitrogen fixation. *Plant Cell*, 7: 869-885.

Nag, R., Maity, M. K., & Dasgupta, M. (2005). Dual DNA binding property of ABA insensitive 3 like factors targeted to promoters responsive to ABA and auxin. *Plant molecular biology*, 59(5): 821–38.

Nambara, E., & Marion-Poll, A. (2003) ABA action and interactions in seeds. *Trends in plant science*, 8(5): 213–7.

Nambara, E., Hayama, R., Tsuchiya, Y., Nishimura, M., Kawaide, H., Kamiya, Y. & Naito, S. (2000) The role of ABI3 and FUS3 loci in *Arabidopsis thaliana* on phase transition from late embryo development to germination. *Dev Biol*, 220: 412-423.

Nambara, E., Hayama, R., Tsuchiya, Y., Nishimura, M., Kawaide, H., Kamiya, Y. & Naito, S. (2000) The role of ABI3 and FUS3 loci in *Arabidopsis thaliana* on phase transition from late embryo development to germination. *Dev Biol,* 220: 412–423.

Narasimhamoorthy, B., Bouton, J. H., Olsen, K. M. & Sledge, M. K. (2007) Quantitative trait loci and candidate gene mapping of aluminum tolerance in diploid alfalfa. *Theoretic & Appl Genet*, 114(5): 901–13.

Neumann, E. & Rosenheck, K. (1973) An alternate explanation for permeability changes induced by electrical impulses in vesicular membrane. *J Membr Biol,* 14: 193–196.

Niu, X., Helentjaris, T. & Bate, N. J. (2004) Maize ABI4 binds coupling element1 in abscisic acid and sugar response genes. *Plant Cell.* 14(10): 2565-75.

Nodine, M. D. & Bartel, D. P. (2010) MicroRNAs prevent precocious gene ex-pression and enable pattern formation during plant embryogenesis. *Genes Dev* 24: 2678–2692.

Nordstrom, A., Tarkowski, P., Tarkowska, D., Norbaek, R., Astot, C., Dole-zal, K. & Sandberg, G. (2004) Auxin regulation of cytokinin biosyn-thesis in *Arabidopsis thaliana*: a factor of potential

importance for auxin–cytokinin-regulated development. *Proc Natl Acad Sci USA* 101: 8039–8044.

North, H., Baud, S., Debeaujon, I., Dubos, C., Dubreucq, B., Grappin, P., Jullien, M., et al. (2010) Arabidopsis seed secrets unravelled after a decade of genetic and omics-driven research. *Plant J. :*, 61(6): 971–81.

Ochatt, S. J., Delaitre, C., Lionneton, E., Huchette, O., Patat-Ochatt, E. M. & Kahane, R. (2005) One team, PCMV, and one approach, in vitro biotechnology, for one aim, the breeding of quality plants with a wide array of species. In: Ramdane Dris (ed) Crops: growth, quality and biotechnology. WFL Publisher, Helsinki, Finland, pp 1038–1067

Ochatt, S. J., Mousset-Déclas, C. & Rancillac, M. (2000) Fertile pea plants regenerate from protoplasts when calluses have not undergone endoreduplication. *Plant Sci,* 156: 177-183.

Ochatt, S. J., Muilu, R. & Ribalta, F. (2008) Cell morphometry and osmolarity as early indicators of the onset of embryogenesis from cell suspension cultures of grain legumes and model systems. *Plant Biosystems,* 142: 480–486.

Ochatt, S., Jacas, L., Patat-Ochatt, E.M. & Djennane, S. (2013) Flow cytometric analysis and molecular characterization of *Agrobacterium tumefaciens*-mediated transformants of *Medicago truncatula. Plant Cell, Tissue Organ Culture, in press. (PCTOC),DOI: 10.1007/s11240-012-0263-1*

Ochatt, S.J. (2006) Flow cytometry (ploidy determination, cell cycle analysis, DNA content per nucleus). In: *Medicago truncatula handbook*, Chapter 2.2.7. Online version November 2006, http://www.noble.org/MedicagoHandbook/ ISBN number 0-9754303-1-9.

Ochatt, S.J. (2008) Flow cytometry in plant breeding. *CytometryA,* 73: 581-598.

Ochatt, S.J. (2011) Immature seeds and embryos of *Medicago truncatula* cultured *in vitro*. In: Thorpe TA, Yeung EC (eds) Plant Embryo Culture: Methods and Protocols; Methods in Molecular Biology 710. Springer Protocols, Humana Press, New York, pp. 39-52.

Ochatt, S.J., Atif, R.M., Patat-Ochatt E. M., Jacas, L. & Conreux, C. (2010) Competence versus Recalcitrance for in Vitro Regeneration. *Not. Bot. Hort. Agrobot. Cluj,* 38: 102-108.

Odell, J. T., Nagy, F. & Chua, N. H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature,* 313: 810-812.

Ogas, J., Kaufmann, S., Henderson, J. & Somerville, C. (1999) PICKLE is a CHD3 chromatin-remodeling factor that regulatesthe transition from embryonic to vegetative development in Arabidopsis. *Proc Natl Acad Sci U S A.* 96(24): 13839-44.

Papadopoulos, J. S. & Agarwala, R. (2007) COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics,* 23(9): 1073-9.

Papi, M., Sabatini, S., Altamura, M. M., Hennig, L., Scha, E., Costantino, P., Vittorioso, P., et al. (2002) Inactivation of the Phloem-Specific Dof Zinc Finger Gene DAG1 Affects Response to Light and Integrity of the Testa of Arabidopsis Seeds, *Plant Cell*, 128: 411–417.

Parcy, F., Valon, C., Kohara, A., Miséra, S. & Giraudat, J. (1997) The ABSCISIC ACID-INSENSITIVE3, FUSCA3, and LEAFY COTYLEDON1 loci act in concert to control multiple aspects of Arabidopsis seed development. *Plant Cell,* 9: 1265–1277.

Park, H. D., Lim, P. O., Kim, J. S., Cho, D. S., Hong, S. H. & Nam, H. G. (2003) The Arabidopsis *COG1* gene encodes a Dof domain transcription factor and negatively regulates phytochrome signalling. *Plant J.,* 34: 161–171.

Permyakova, N. V., Shumnyi, V. K. & Deineko E. V. (2009) Agrobacterium-mediated transformation of plants: Transfer of vector DNA fragments in the plant genome. *Russian Journal of Genetics,* 45(3): 266-275.

Perrot-Rechenmann, C. (2010) Cellular responses to auxin: division versus expansion. *Cold Spring Harbor perspectives in biology*, *2*(5),

Petit, A., Stougaard, J., KiIhle, A., Marcker K. A. & Tempé, J. (1987) Transformation and regeneration of the legume Lotus comiculatus: a system for molecular studies of symbiotic nitrogen fixation. *Mol. Gen. Genet.,* 207: 245-250.

Podell, S. & Gribskov, M. (2004) Predicting N-terminal myristoylation sites in plant proteins. *BMC Genomics,* 5:37.

Power, J. F. (1987) Legumes : their potential role in agricultural production. *American journal of alternative agriculture*, 2(2): 69-73.

Prosperi, J. M., Angevain, M., Genier, C., Olivieri, I. & Mansat, P. (1993) Sélection de nouvelles légumineuses fourragères pour les zones difficiles méditerranéennes. *Fourrages,* 135: 343-354 ;

Quayle, T. & Feix, G. (1992) Functional analysis of the -300 region of maize genes. *Mol Gen Genet,* 231: 369–374.

Raikhel, N. (1992) Nuclear targeting in plants. *Plant Physiol.,* 100: 1627–32

Ramos, J. A., Zenser, N., Leyser, O. & Callis, J. (2001) Rapid degradation of auxin/indole acetic acid proteins requires conserved amino acids of domain II and is proteasome dependent. *The Plant Cell,* 13: 2349–2360.

Rashotte, A. M., Mason, M. G., Hutchison, C. E., Ferreira, F. J., Schalleret G. E. et al. (2006) A subset of Arabidopsis AP2 transcription factors mediates cytokinin responses in concert with a two-component pathway. *Proc. Natl. Acad. Sci. USA* 103: 11081–11085.

Raz, V., Bergervoet, J. H., & Koornneef, M. (2001). Sequential steps for developmental arrest in Arabidopsis seeds. *Development,* 128(2): 243–52

Reed, J. W. (2001) Roles and activities of Aux/IAA proteins in Arabidopsis. *Trends Plant Sci.,* 6: 420-425.

Reidt, W., Wohlfarth, T., Ellerström, M., Czihal, A., Tewes, A., Ezcurra, I., Rask, L. & Bäumlein, H. (2000) Gene regulation during late embryogenesis: the RY motif of maturation-specific gene promoters is a direct target of the FUS3 gene product. *Plant J.*, 21: 401-408.

Reinhart, B. J., Weinstein, E.G., Rhoades, M. W., Bartel, B., & Bartel, D. P. (2002) MicroRNAs in plants. *Genes Dev.* 16: 1616–1626.

Rider, S. D., Henderson, J. T., Jerome, R. E., Edenberg, H. J., Romero-Severson, J. & Ogas, J. (2003) Coordinate repression of regulators of embryonic identity by PICKLE during germination in Arabidopsis. *Plant J..* 35(1): 33-43.

Riechmann, J. L. & Ratcliffe. O. J. (2000) A genomic perspective on plant transcription factors. *Curr Opin Plant Biol,* 3(5): 423–434.

Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O. J., Samaha, R. R., et al (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science.,* 290: 2105–2110.

Rijavec, T. T., Jain, M. M., Dermastia, M. M., & Chourey P. S. P. S. (2011) Spatial and temporal profiles of cytokinin biosynthesis and accumulation in developing caryopses of maize. *Ann Bot*, 107(7): 1235-45.

Rogers, S. O. & Bendich, A. J. (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol Biol,* 5(2): 69–76.

Rommens, C. M., Bougri, O., Yan, H., Humara, J. M., Owen, J., Swords, K. & Ye, J. (2005) Plant-derived transfer DNAs. *Plant Physiol,* 139: 1338–1349.

Rommens, C. M., Humara, J. M., Ye, J., Yan, H., Richael, C., Zhang, L., Perry, R. & Swords, K. (2004) Crop improvement through modification of the plant's own genome. *Plant Physiol* 135: 421–431.

Rose, R. J., Nolan, K. E. & Bicego, L. (1999) The development of the highly regenerable seed line Jemalong 2HA for transformation of *Medicago truncatula*-implications for regenerability via somatic embryogenesis. *J Plant Physiol,* 155: 788–791.

Rost, B., Yachdav, G. & Liu, J. (2004) The PredictProtein server. *Nucleic Acids Res,* 32: W321–W326.

Rubio, M. C., Gonzalez, E. M., Minchin, F. R., Webb, K. J., Arrese-Igor, C., Ramos, J. & Becana, M. (2002) Effects of water stress on antioxidant enzymes of leaves and nodules of transgenic alfalfa overexpressing superoxide dismutases. *Physiol Plant,* 115: 531-540.

Ruegger, M., Dewey, E., Gray, W.M., Hobbie, L., Turner, J., & Estelle, M. (1998) The TIR1 protein of Arabidopsisfunctions in auxin response and is related to human SKP2 and yeast Grr1p. *Genes Dev.* 12: 198–207.

Saitou, N. & Nei, M. (1987) The neighbor-joining method—a new method for reconstructing phylogenetic trees. *Mol Biol Evol,* 4(4): 406–425.

Sangwan, R. S., Bourgoies, Y., Brow. S., Vasseur, G. & Sangwan-Norreel, B. S. (1992) Characterization of competent cells and early events of Agrobacterium-mediated transformation of*Arabidopsis thaliana*. *Planta*, 188: 439–456.

Sangwan, R. S., Ochatt, S., Nava-Saucedo, J. E. & Sangwan-Norreel, B. S. (2010) T-DNA insertion muta-genesis for gene cloning in plants. In: Shu Q, Nakagawa H, Forster B (eds) Mutation techniques in plants—principles and applications. Springer, Germany, pp 1–19

Santos, D. & Fevereiro, P. (2002) Loss of DNA methylation affects somatic embryogenesis in *Medicago truncatula*. *Plant Cell Tissue Organ Cult,* 70: 155–161.

Santos-Mendoza, M., Dubreucq, B., Baud, S., Parcy, F., Caboche, M., & Lepiniec, L. (2008) Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis. *Plant J.*, 54(4): 608–20.

Scaramelli, L, Balestrazzi, A., Bonadei, M., Piano, E., Carbonera, D. & Confalonieri, M. (2009) Production of transgenic barrel medic (*Medicago truncatula* Gaernt.) using the ipt-type MAT vector system and impairment of Recombinase-mediated excision events. *Plant Cell Rep,* 28: 197–211.

Schmid, M., Neftel, A. & Fuhrer, J. (2000) Nitrous oxide emissions from Swiss agriculture. FAL, the FAL Series, 33: 131.

Scholte, M., d'Erfurth, I., Ripa, S., Mondy, S., Cosson, V., Durand, P., Breda, C., Trinh, T., Rodriguez-Llorente, I., Kondorosi, E., Schultze, M., Kondorosi, A. & Ratet, P. (2002) T-DNA tagging in the model legume *Medicago truncatula* allows efficient gene discovery. *Mol. Breeding,* 10: 203-215.

Schwechheimer, C., Zourelidou, M. & Bevan, M. W. (1998) Plant transcription factor studies. *Annu. Rev. Plant Physiol. Plant Mol. Biol*. 49:127–50

Shaw, L. M., McIntyre, C. L., Gresshoff, P. M. & Xue, G. P. (2009) Members of the Dof transcription factor family in Triticum aestivum are associated with light-mediated gene regulation. *Funct Integr Genomics*, 9(4): 485–498.

Sheikoleslam, S. N. & Weeks D. P. (1987) Acetosyringone promotes high efficiency transformation of Arabidopsis thaliana explants by Agrobacteriura turaefaciens. *Plant Mol Biol,* 8: 291-298.

Shigyo, M., Tabei, N., Yoneyama, T. & Yanagisawa, S. (2007). Evolutionary processes during the formation of the plant-specific Dof transcription factor family. *Plant & cell physiol*, *48*(1) : 179–85.

Shimamoto, K. & Kyozuka, J. (2002) Rice as a model for comparative genomics of plants. *Ann Rev of Plant Biol*, 53, 399–419.

Shimofurutani, N., Kisu, Y., Suzuki, M. & Esaka, M. (1998) Functional analyses of the Dof domain, a zinc finger DNA-binding domain, in a pumpkin DNA-binding pro-tein AOBP. *FEBS Lett*, 430: 251–256.

Shiu, S. H., Shih, M. C. & Li, W. H. (2005) Transcription factor families have much higher expansion rates in plants than in animals. *Plant Physiol.,* 139: 18–26.

Singh, D. P., Filardo, F. F., Storey, R., Jermakow, A. M., Yamaguchi, S. & Swain, S. M. (2010) Overexpression of a gibberellin inactivation gene alters seed development, KNOX gene expression, and plant development in Arabidopsis. *Physiologia Plantarum* 138: 74–90.

Singh, D. P., Jermakow, A. M., & Swain, S. M. (2002) Gibberellins Are Required for Seed Development and Pollen Tube Growth in Arabidopsis. *Plant Cell*, 14: 3133–3147.

Skirycz, A., Radziejwoski, A., Busch, W., Hannah, M. a, Czeszejko, J., Kwaśniewski, M., Zanor, M.-I., et al. (2008) The DOF transcription factor OBP1 is involved in cell cycle regulation in *Arabidopsis thaliana*. *Plant J.,* 56(5), 779–92.

Somers, D. A., Samac, D. A. & Olhoft, P. M. (2003) Recent advances in legume transformation. *Plant Physiol,* 131: 892–899.

Sonia, Saini, R., Singh, R. P. & Jaiwal. P. K. (2007) *Agrobacterium tumefaciens* mediated transfer of *Phaseolus vulgaris* α-amylase inhibitor-1 gene into mungbean *Vigna radiate* (L.) Wilczek using bar as selectable marker. *Plant Cell Rep,* 26: 187–198

Sparkes, I. A., Runions, J., Kearns, A. & Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Prot,* 1: 2019–2025.

Stals, H.& Inzé, D. (2001) When plant cells decide to divide. *Trends in Plant Science,* 6(8): 359–364.

Steinmann, T., Geldner, N., Grebe, M., Mangold, S., Jackson, C., Paris, S., Galweiler, L., Palme, K., and Jurgens, G. (1999) Coordinated polar localization of auxin efflux carrier PIN1 by GNOM ARF GEF. *Science*, 286: 316–318.

Stone, S. L., Braybrook, S. A., Paula, S. L, Kwong, L. W., Meuser, J., Pelletier, J., et al. (2008) Arabidopsis leafy cotyledon2 induces maturation traits and auxin activity: implications for somatic embryogenesis. *Proc Natl Acad Sci USA,* 105: 3151–6.

Stone, S. L., Kwong, L. W., Yee, K. M., Pelletier, J., Lepiniec, L., Fischer, R. L., et al. (2001) Leafy cotyledon2 encodes a b3 domain transcription factor that induces embryo development. *Proc Natl Acad Sci USA,* 98: 11806–11.

Sundberg, E. & Ostergaard, L. (2009) Distinct and dynamic auxin activities during reproductive development. *Cold Spring Harb Perspect Biol,* 1(6): a001628.

Suzuki, M., & McCarty, D. R. (2008) Functional symmetry of the B3 network controlling seed development. *Current opinion in plant biology*, 11(5): 548–53.

Švábová, L. & Griga, M. (2008) The effect of cocultivation treatments on transformation efficiency in pea (*Pisum sativum* L.). *Plant Cell Tiss Org Cult,* 95: 293-304.

Swain, S. M., Reid, J. B., & Kamiya, Y. (1997) Gibberellins are required for embryo and seed development in pea. *Plant J..* 12:1329–1338.

Swain, S. M., Reid, J. B., & Ross, J. J. (1993) Seed development in Pisum: The *lhi* allele reduces gibberellin levels in developing seeds, and increases seed abortion. *Planta* 191: 482–488.

Tadege, M., Wen, J., He, J., Tu, H., Kwak, Y., Eschstruth, A., Cayrel, A., Endre, G., Zhao, P. X., Chabaud, M., Ratet, P. & Mysore, K. S. (2008) Large-scale insertional mutagenesis using the*Tnt1* retrotransposon in the model legume *Medicago truncatula*. *Plant J.*, 54: 335–347.

Tanaka, E., Bailey, T. L., Grant, C. E., Noble, W. S. & Keich, U. (2011) Improved similarity scores for comparing motifs. *Bioinformatics,* 27(12): 1603-1609.

Tanaka, M., Takahata, Y., Nakayama, H., Nakatani, M. & Tahara, M. (2009) Altered carbohydrate metabolism in the storage roots of sweet potato plants overexpressing the SRF1 gene, which encodes a Dof zinc finger transcription factor. *Planta,* 230(4): 737-46.

Thomas, M. R., Rose, R. J. & Nolan, K. E. (1992) Genetic transformation of *Medicago truncatula* using Agrobacterium with genetically modified Ri and disarmed Ti plasmids. *Plant Cell Rep*, 11: 113–117.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res,* 25(24): 4876–4882.

Thompson, R., Burstin, J. & Gallardo, K. (2009) Post-genomics studies of developmental processes in legume seeds. *Plant physiol*, 151(3): 1023–9.

Tiwari, S. B., Wang, X. J., Hagen, G. & Guilfoyle, T. J. (2001) AUX/IAA proteins are active repressors, and their stability and activity are modulated by auxin. *Plant Cell*, 13: 2809–2822.

To, A., Valon, C., Savino, G., Guilleminot, J., Devic, M., Giraudat, J. & Parcy, F. (2006) A network of local and redundant gene regulation governs Arabidopsis seed maturation. *Plant Cell,* 18: 1642–1651.

Trieu, A. T. & Harrison, M. J. (1996) Rapid transformation of*Medicago truncatula*: regeneration via organogenesis. *Plant Cell Rep,* 16: 6–11.

Trieu, A. T., Burleigh, S. H., Kardailsky, I. V., Maldonado-Mendoza, I. E., Versaw, W. K., Blaylock, L. A., Shin, H., Chiou, T. J., Katagi, H., Dewbre, G. R., Weigel, D. & Harrison, M. J. (2000) Transformation of *Medicago truncatula*via infiltration of seedlings or flowering plants withAgrobacterium. *Plant J.,* 22: 531–541.

Trinh, T. H., Ratet, P., Kondorosi, E., Durand, P., Kamate,´ K., Bauer, P. & Kondorosi, A. (1998) Rapid and efficient transformation of diploid*Medicago truncatula*andMedicago sativa ssp. Falcatalines improved in somatic embryogenesis. *Plant Cell Rep,* 17: 345–355.

Tromas, A., Braun, N., Muller, P., Khodus, T., Paponov, I. a, Palme, K., Ljung, K., et al. (2009). The AUXIN BINDING PROTEIN 1 is required for differential auxin responses mediating root growth. *PloS one*, *4*(9): e6648.

Tsukagoshi, H., Morikami, A. & Nakamura, K. (2007) Two B3 domain transcriptional repressors prevent sugar-inducible expression of seed maturation genes in Arabidopsis seedlings. *Proc Natl Acad Sci U S A.,* 104(7): 2543-7.

Tsukagoshi, H., Saijo, T., Shibata, D., Morikami, A. & Nakamura, K. (2005) Analysis of a sugar response mutant of Arabidopsis identified a novel B3 domain protein that functions as an active transcriptional repressor. *Plant Physiol.,* 138(2): 675-85.

Udvardi, M. (2002) Legume genomes and discoveries in symbiosis research. *Genome Biol,* 3: reports 4028.1- reports4028.3.

Udvardi, M. K., Kakar, K., Wandrey, M., Montanari, O., Murray, J., Andriankaja, A., Zhang, J. Y., Benedito, V., Hofer, J. M., Chueng, F., Town, C. D. (2007) Legume transcription factors: global regulators of plant development and response to the environment. *Plant Physiol,* 144(2): 538–549.

Ulmasov, T. (1997) ARF1, a Transcription Factor That Binds to Auxin Response Elements. *Science*, 276(5320): 1865–1868.

Ulmasov, T., Hagen, G. & Guilfoyle, T. J. (1999a) Dimerization and DNA binding of auxin response factors. *Plant J..* 19: 309–319.

Ulmasov, T., Hagen, G. and Guilfoyle, T.J. (1999b) Activation and repression of transcription by auxin-response factors. *Proc. Natl. Acad. Sci. USA,* 96: 5844–5849.

Umemura, Y., Ishiduka, T., Yamamoto, R. & Esaka, M. (2004) The Dof domain, a zinc finger DNA-binding domain conserved only in higher plants, truly functions as a Cys2/Cys2 Zn finger domain. *Plant J.,* 37: 741–749.

Van der Graaff, E., den Dulk-Ras, A. & Hooykaas, P. J. J. (1996) Deviating T-DNA transfer from Agrobacterium tumefaciens to plants. *Plant Mol Biol*, 13: 677–681.

Vanneste, S. & Friml, J. (2009). Auxin: a trigger for change in plant development. *Cell*, 136(6) : 1005–16.

Verdier, J. & Thompson, R. D. (2008) Transcriptional Regulation of Storage Protein Synthesis During Dicotyledon Seed Filling. *Plant Cell Physiol.* 49: 1263-1271.

Verdier, J., Kakar, K., Gallardo, K., Le Signor, C., Aubert, G., Schlereth, A., Town, C. D., et al. (2008) Gene expression profiling of *M. truncatula* transcription factors identifies putative regulators of grain legume seed filling. *Plant molecular biology*, 67(6): 567–80.

Vicente-Carbajosa, J. & Carbonero, P. (2005) Seed maturation: developing an intrusive phase to accomplish a quiescent state. *The Inter J of develop biol*, 49(5-6): 645–51.

Vicente-Carbajosa, J., Moose, S. P., Parsons, R. L. & Schmidt, R. J. (1997) A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. *Proc Natl Acad Sci U S A,* 94(14): 7685–7690.

Vigeolas, H., Chinoy, C., Zuther, E. et al (2008) Combined Metabolomic and Genetic Approaches Reveal a Link between the Polyamine Pathway and Albumin 2 in Developing Pea Seeds. *Plant Physiol.* 146: 74-82.

Walhout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S. & Vidal, M. (2000) GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods Enzymol,* 328: 575-92.

Wang, J., Rose, R. J. & Donaldson, B. I. (1996) *Agrobacterium*-mediated transformation and expression of foreign genes in *Medicago truncatula*. *Austral J Plant Physiol,* 23(3): 265 – 270.

Wasson, A. P., Pellerone, F. I. & Mathesius, U. (2006) Silencing the flavonoid pathway in *Medicago truncatula* inhibits root nodule formation and prevents auxin transport regulation by rhizobia. *Plant Cell,* 18: 1617-1629.

Weber, H., Borisjuk, L. & Wobus, U. (2005) Molecular physiology of legume seed development. *Annu. Rev. Plant Biol.,* 56:253–79.

Weber, H., Heim, U., Golombek, S., Borisjuk, L. & Wobus, U. (1998) Assimilate uptake and the regulation of seed development. *Seed Science Research* 8: 331–345.

Webster, C. I., Packman, L. C., Pwee, K. H. & Gray, J. C. (1997) High mobility group proteins HMG-1 and HMG-I/Y bind to a positive regulatory region of the pea plastocyanin gene promoter. *Plant J..*, 11(4):703-15.

Weijers, D. & Jurgens, G. (2005) Auxin and embryo axis forma-tion: the ends in sight? *Curr. Opin. Plant Biol.* 8: 32–37.

West, M., Yee, K. M., Danao, J., Zimmerman, J. L., Fischer, R. L, Goldberg, R. B, Harada, J. J. (1994) LEAFY COTYLEDON1 Is an Essential Regulator of Late Embryogenesis and Cotyledon Identity in Arabidopsis. *Plant Cell,* 6: 1731-1745.

Willmann, M. R., Mehalick, A. J., Packer, R. L., & Jenik, P. D. (2011) MicroRNAs regulate the timing of embryo maturation in Arabidopsis. *Plant Physiol,* 155(4): 1871–84.

Woodward, A. W. & Bartel, B. (2005) Auxin: regulation, action, and interaction. *Annals of botany*, 95(5): 707–35.

Wordragen, M. F. V. & Dons, H. J. M. (1992) Agrobacterium tumefaciens-mediated transformation of recalcitrant crops. *Plant Mol Biol,* 5: 299–302.

Wright, E., Dixon, R. A. & Wang, Z. Y. (2006) *Medicago truncatula* transformation using cotyledon explants. In K Wang, ed, Agrobacterium Protocols, Vol 343. Human Press, Totowa, NJ, pp 129–136

Wu, C-Y., Suzuki, A., Washida, H. & Takaiwa, F. (1998) The GCN4 motif in a rice glutelin gene is essential for endospermspecific gene expression and is activated by Opaque-2 in transgenic rice plants. *Plant J.,* 14: 673–683.

Wu, C-Y., Washida, H., Onodera, Y., Harada, K. & Takaiwa, F. (2000) Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal cis-element requirements for endosperm-specific gene expression. *Plant J.*, 23: 415–421.

Yamamoto, A., Kagaya, Y., Usui, H., Hobo, T., Takeda, S., & Hattori, T. (2010) Diverse roles and mechanisms of gene regulation by the Arabidopsis seed maturation master regulator FUS3 revealed by microarray analysis. *Plant & cell physiology*, 51(12): 2031–46.

Yamamoto, M. P., Onodera, Y., Touno, S. M., Takaiwa, F. (2006) Synergism between RPBF Dof and RISBZ1 bZIP activators in the regulation of rice seed expression genes. *Plant Physiol,* 141(4): 1694–1707.

Yanagisawa, S. & Schmidt, R. J. (1999) Diversity and similarity among recognition sequences of Dof transcription factors. *Plant J.,* 17(2): 209–214.

Yanagisawa, S. & Sheen, J. (1998) Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression. *Plant Cell,* 10: 75–99.

Yanagisawa, S. (2000) Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. *Plant J.*, 21(3), 281–8.

Yanagisawa, S. (2001) The transcriptional activation domain of the plant-specific Dof1 factor functions in plant, animal, and yeast cells. *Plant & cell physiol*, 42(8): 813–22.

Yanagisawa, S. (2002) The Dof family of plant transcription factors. *Trends Plant Sci.,* 7: 555–560.

Yanagisawa, S. (2004). Dof domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants. *Plant & cell physiol, 45*(4): 386–91.

Yang, S., Gao, M., Xu, C., Gao, J., Deshpande, S., Lin, S., Roe, B. & Zhu, H. (2008). Alfalfa benefits from *Medicago truncatula*: the *RCT1* gene from *M. truncatula* confers broad-spectrum resistance to anthracnose in alfalfa. *Proc Natl Acad Sci USA*, *105*(34): 12164–9.

Ye, X., Williams, E. J., Shen, J., Esser, J. A., Nichols, A. M., Petersen, M. W. & Gilbertson, L. A. (2008) Plant development inhibitory genes in binary vector backbone improve quality event efficiency in soybean transformation. *Transgenic Res*, 17: 827–838.

Yin, M-Z., Guan, M., Xiao, G., Li, X. & Guan C-Y., (2011) RNAi Vector Construction of AtDof1.7 Transcription Factors and Genetic Transformation into *Arabidopsis thaliana*. *Acta Agron Sinca,* 37(7): 1196-1204.

Young, N. D., Debellé, F., Oldroyd, G. E., Geurts, R., Cannon, S. B., Udvardi, M. K., Benedito, V. A., Mayer, K. F., Gouzy, J. et al., (2011) The *Medicago* genome provides insight into evolution of rhizobial symbiosis. *Nature,* 480: 520–524.

Young, N. D., Mudge, J. & Ellis, T. H. N. (2003) Legumes Genomes: more than peas in a pod. *Curr. Opin. Plant Biol.* 6: 199-204.

Young, N. D., Roe, B. A., Town, C. D., Cannon, S. B., Sato, S. & Tabta, S. (2005) Sequencing the genespaces of *Medicago truncatula* and *Lotus japonicus*. *Plant Physiol,* 137: 1174–1181.

Yu, C. S., Chen, Y. C., Lu, C. H. & Hwang J. K. (2006) Prediction of protein subcellular localization. *Proteins,* 64: 643–651.

Zabala, G., Campos, E., Varala, K. K., Bloomfield, S., Jones, S. I., Win, H., Tuteja, J. H., Calla, B., Clough, S. J., Hudson, M. & Vodkin, L. O. (2012) Divergent patterns of endogenous small RNA populations from seed and vegetative tissues of Glycine max. *BMC Plant Biology,* 12: 177.

Zazimalova, E. & Napier, R. M. (2003) Points of regulation for auxin action. *Plant Cell Rep.* 21: 625–34.

Zenser, N., Ellsmore, A., Leasure, C. & Callis, J. (2001) Auxin modulates the degradation rate of Aux/IAA proteins. *Proce Natl Acad Sci USA,* 98: 11795–11800.

Zhang, H., Bishop, B., Ringenberg, W., Muir, W. M., & Ogas, J. (2012) The CHD3 remodeler PICKLE associates with genes enriched for trimethylation of histone H3 lysine 27. *Plant Physiol,* 159(1): 418–32.

Zheng, Y., Ren, N., Wang, H., Stromberg, A. J. & Perry, S. E. (2009) Global identification of targets of the Arabidopsis MADS domain protein AGAMOUS-Like15. *The Plant cell*, 21(9): 2563–77.

Zhou, C., Han, L., Pislariu, C., Nakashima, J., Fu, C., Jiang, Q., Quan, L., Blancaflor, E. B., Tang, Y. J., Bouton, H., Udvardi, M., Xia, G., & Wang, Z. Y. (2011) From Model to Crop: Functional Analysis of a STAY-GREEN Gene in the Model Legume *Medicago truncatula* and Effective Use of the Gene for Alfalfa Improvement. *Plant Physiol,* 157: 1483–1496.

Zhou, X., Carranco, R., Vitha, S. & Hall, T. C. (2005) The dark side of green fluorescent protein. *New Phytol,* 168: 313-322.

Zou, X., Seemann, J. R., Neuman, D. & Shen Q. J. (2004) A *WRKY* gene from creosote bush encodes an activator of the ABA signaling pathway. *J Biol Chem,* 279: 55770–55779.