

**Role of the chromatin modifier
LYN1 gene in
chloroplast development**

Jia He

Submitted for the degree of Doctor of Philosophy

Royal Holloway, University of London

School of Biological Sciences

December 2017

Abstract

Chloroplast development in plants is regulated by light signalling, which is perceived by photoreceptors including phytochromes, to drive gene expression in the nucleus through a complex transcriptional network. In order to enhance our understanding of this process, and by extension of chloroplast development, searches for novel genes through mutant screens were previously carried out. The *lyn1* mutant was identified as a suppressor of a phytochrome-defective mutant of *Arabidopsis*, *hy1*. The *hy1* mutation causes a chloroplast development defect as mutants cannot produce phytochromes to perceive light due to failure of photoreceptor chromophore biosynthesis in plastids.

Multiple assays suggested that *lyn1* can enhance plastid development. These included chloroplast total “coverage” within cauline leaf cells, levels of several plastid-synthesised tetrapyrroles, not only in the light but also in dark conditions, and relative numbers of plastid versus nuclear genome copies. *lyn1* was demonstrated to partly suppress another phytochrome chromophore-defective mutant, but not photoreceptor mutants generically. In contrast to *hy1*, *lyn1 hy1* was found to exhibit a small but detectable amount of photochemically-active phytochrome.

The *LYN1* gene has been identified through a combination of high-resolution mapping and whole-genome sequencing and complementation tests. The *LYN1* gene identity as *JMJ14*, a histone demethylase gene, was confirmed through independent knock-outs, failure of cross-complementation assays, as well as by transgenic complementation of the mutant with wild type *JMJ14*. The *lyn1* mutation was found to reside on a gene predicted to play an indirect role in the regulation of expression of other genes.

Global changes in gene expression in the mutant were assessed by RNA-sequencing. The results were analysed in conjunction with publicly available data on global gene expression during development and on the presence of chromatin marks in whole seedling genomic DNA, to explain how *lyn1* rescues the light-response defect. *lyn1* is proposed to enhance the initiation of cell differentiation, resulting in an enhanced plastid development which, in turn, leads to assembly of a small amount of photoreceptor and rescue of the light response defect.

Table of Contents

Abstract	ii
Table of Contents	iv
List of Figures	ix
List of Tables	xiii
Abbreviations	xiv
Acknowledgements	xviii
Declaration of Authorship	xix
Chapter 1 General Introduction	1
1.1 Photosynthesis and chloroplast evolution.....	1
1.2 Light signalling control of plant development.....	6
1.2.1 Types of photoreceptor	6
1.2.2 Phytochromes' biogenesis	7
1.2.3 Phytochromes' and cryptochromes' nuclear localization.....	8
1.2.4 Light mediates plant morphogenesis	9
1.3 Communication between plastid and nucleus	12
1.3.1 Nucleus-to-plastid anterograde build-up and regulation pathways ...	12
1.3.2 Chloroplast-to-nucleus retrograde signalling pathways	15
1.4 <i>Arabidopsis lyn1</i> mutant	23
1.4.1 Generating <i>lyn1</i> mutation by genetic screens	23
1.4.2 LYN1 is not a light-inactivated repressor of photomorphogenesis ...	24
1.4.3 <i>lyn1</i> suppression of the growth and greening of seedlings	25
1.4.4 <i>lyn1</i> eliminates the <i>gun</i> phenotype of <i>hy1</i>	27
1.4.5 <i>lyn1</i> does not rescue haem oxygenase.....	28
1.4.6 Mapping and cloning of the <i>LYN1</i> gene	29
1.4.7 Current published studies about the role of JM14.....	31
1.5 Hypothesis	37
1.6 Aims	38
Chapter 2 Materials and Methods	40
2.1 Plant materials	40
2.2 <i>Arabidopsis thaliana</i> -related preparations	40
2.2.1 Seedling growth media preparation	40
2.2.2 Seed sterilisation	41
2.2.3 Seed plating	41
2.2.4 Plant soil preparation	41
2.2.5 Seedling and plant growth conditions	41
2.3 <i>Arabidopsis</i> genotyping assays	42
2.3.1 DNA extraction method	42
2.3.2 Primers design	42
2.3.3 Real-time polymerase chain reaction.....	43
2.3.4 Restriction enzyme digestion of dCAPS and CAPS assay	44

3.2.1.2.2 Hypocotyl length and chlorophyll content	77
3.2.1.3 Failure of complementation of <i>lyn1</i> by <i>jmj14-1</i>	80
3.2.1.4 Complementation of <i>lyn1</i> mutants by transferring functional <i>JMJ14</i>	85
3.2.1.4.1 Failure of BP recombination.....	85
3.2.1.4.2 Ligation of <i>JMJ14</i> cDNA into pGEM-T Easy vector	85
3.2.1.4.3 Generation of <i>JMJ14</i> cDNA fragments with Gateway <i>attB</i> sites	87
3.2.1.4.4 Gateway™ vector-related works	88
3.2.1.4.4.1 Transformation into Gateway™ Donor vector p201.....	88
3.2.1.4.4.2 Transformation into Gateway™ Destination vector pB2GW7	91
3.2.1.4.5 Transforming and screening of plants.....	94
3.2.1.4.5.1 <i>Agrobacterium</i> -mediated transformation into plants.....	94
3.2.1.4.5.2 Screening and verification of complemented plants	95
3.2.2 Detecting <i>JMJ14</i> mRNA in <i>jmj14-1</i> and <i>lyn1</i> mutants	98
3.2.3 Structure, expression and evolutionary relationships of <i>JMJ14</i>	100
3.2.4 <i>JMJ14</i> gene expression pattern during <i>Arabidopsis</i> development..	105
3.3 Discussion	106
3.3.1 Confirmation of the identification of <i>lyn1</i>	106
3.3.2 <i>JMJ14</i> protein domains	107
3.3.3 <i>JMJ14</i> evolution	109
3.3.3.1 Current studies of <i>JMJ14</i> paralogues	109
3.3.3.2 Current studies of <i>JMJ14</i> orthologues	111
3.4 Conclusion	113
Chapter 4 Result – Characterisation of the <i>JMJ14</i> Gene.....	114
4.1 Introduction and Methodology	114
4.1.1 <i>jmj14</i> impact on plastid development.....	114
4.1.2 <i>jmj14</i> impact on tetrapyrroles synthesis	117
4.1.3 Assessment of whether <i>lyn1</i> rescues a PΦB synthase defective mutant	118
4.1.4 Assessing whether <i>lyn1</i> suppresses photoreceptor apoprotein mutants, <i>phyA-211</i> , <i>phyB-9</i> and <i>cry1-304</i>	119
4.2 Results	121
4.2.1 Chloroplasts compartment analysis	121
4.2.2 Analysis of plastid to nuclear genome copy number ratio	129
4.2.3 <i>lyn1</i> elevates Pchl _a level	131
4.2.4 <i>lyn1</i> causes a small degree of suppression of <i>hy2-105</i>	139
4.2.4.1 Identification of the <i>lyn1 hy2-105</i> double mutant	139
4.2.4.2 Hypocotyl length	141
4.2.4.3 Chlorophyll content.....	143

4.2.5 <i>lyn1</i> causes a detectable presence of phytochromobilin-containing photochemically active phytochrome in <i>hy1</i> and elevates its level in the WT	145
4.2.6 Confirming whether <i>lyn1</i> suppresses photoreceptor mutants <i>phyA-211</i> , <i>phyB-9</i> and <i>cry1-304</i>	147
4.2.6.1 Identification of <i>lyn1 cry1-304</i> , <i>lyn1 phyB-9</i> and <i>lyn1 phyA-211</i>	148
4.2.6.2 Can <i>lyn1</i> suppress photoreceptor mutants in white light?	154
4.2.6.2.1 Hypocotyl length	155
4.2.6.2.2 Chlorophyll content	157
4.2.6.3 Can <i>lyn1</i> suppress photoreceptor mutants in their diagnostic light wavelength?	159
4.2.6.3.1 Hypocotyl length	160
4.2.6.3.2 Chlorophyll content	165
4.2.7 Overexpression of <i>JMJ14</i>	170
4.3 Discussion	172
4.3.1 <i>jmj14</i> enhances the plastid compartment and plastome to nuclear genome ratio	172
4.3.2 <i>jmj14</i> enhances tetrapyrroles synthesis	174
4.3.3 <i>lyn1</i> partially rescues a PΦB synthase defect mutation, <i>hy2-105</i> ...	176
4.3.4 <i>lyn1</i> does not suppress photoreceptor apoprotein mutants	178
4.3.5 Future work	181
4.4 Conclusion	182
Chapter 5 Result – Global Gene Expression Analysis	183
5.1 Introduction	183
5.1.1 Global gene expression analysis	183
5.1.2 Assessment of global developmental roles through comparative analysis with published transcriptome datasets.	187
5.1.2.1 Gene expression analysis at early stages of differentiation	188
5.1.2.2 Analysis of H3K4me3-marked level of the whole genome	190
5.2 Results	191
5.2.1 Selection of differentially expressed genes and variability of RNA-Seq samples	191
5.2.2 GO term analysis of total 5326 significantly differentially expressed genes	192
5.2.3 Functional gene classification analysis	195
5.2.4 Association of <i>jmj14</i> function with early development.....	202
5.2.5 Association of <i>jmj14</i> function with H3K4me3-marked global genes	210
5.2.6 Association of H3K4me3-marked global genes with early development	217
5.3 Discussion	220
5.3.1 Association of <i>jmj14</i> regulation with H3K4me3 demethylation.....	220
5.3.2 Association of <i>jmj14</i> regulation with early development.....	227

5.3.3 Association of H3K4me3 demethylation with early development....	229
5.3.4 Future work	230
5.4 Conclusion	231
Chapter 6 General Discussion	232
6.1 How does <i>lyn1</i> rescue the light response defect?	232
6.1.1 Is enhanced chloroplast development in <i>lyn1</i> a primary or secondary effect of the mutation?.....	232
6.1.2 How does <i>lyn1</i> increase the extent of plastid development regardless of light?.....	234
6.1.2.1 <i>lyn1</i> boosts general plastid biogenesis during cell entry into differentiation very early in development.....	234
6.1.2.2 <i>lyn1</i> eliminates accumulation of haem.....	237
6.2 Importance of the analysis of epigenetic mechanisms beyond this study	240
6.3 Applications of this work	242
6.3.1 Silencing of <i>JMJ14</i> homologues in rice bundle sheath cells.	242
6.3.2 Stress tolerance and <i>JMJ14</i> homologues	245
6.4 Conclusion	246
References	248
Appendix	283

List of Figures

Figure 1.1 Respiration and photosynthesis maintain energy support for living organisms.....	1
Figure 1.2 Diversity of plastid forms	3
Figure 1.3 Structure and cell types of the <i>Arabidopsis</i> leaf.	5
Figure 1.4 The phytochrome biosynthesis pathway in <i>Arabidopsis</i> plastids and cytoplasm	8
Figure 1.5 Simplified transcriptional networks for skotomorphogenesis and photomorphogenesis.....	11
Figure 1.6 Model for the roles of transcriptional factors in the regulation of chloroplast development	15
Figure 1.7 Summary of chloroplast-to-nucleus retrograde signalling pathways, with a focus on tetrapyrrole synthesis and signals. Note that GUN1 is also involved in monitoring other chloroplasts activities, including translation and protein import.	22
Figure 1.8 <i>lyn1</i> effect of phenotypes of <i>hy1</i> in light and dark	25
Figure 1.9 <i>lyn1</i> suppression of the growth and greening of 20-day plants.....	26
Figure 1.10 <i>lyn1</i> suppression of <i>hy1-1</i>	27
Figure 1.11 Quantification of the expression of light-induced photosynthetic genes in the presence of NF	28
Figure 1.12 Immuno-blotting of phyA apoprotein decay assay	29
Figure 1.13 The region of <i>LYN1</i> on chromosome 4 (coordinates in TAIR10) ...	30
Figure 1.14 <i>JMJ14</i> gene structure and the location of the <i>lyn1</i> and <i>jmj14-1</i>	31
Figure 1.15 Distribution of H3K4me within genes	35
Figure 1.16 A schematic diagram showing the hypothesis of <i>lyn1</i> rescuing light responses in a virtuous cycle.	38
Figure 2.1 Agarose gel images of molecular markers used during the project	45
Figure 2.2 Figure to show FDR-based pairwise multiple comparisons between different genotypes and light conditions	53
Figure 2.3 Designing of <i>attB</i> -adapted <i>JMJ14</i> primers	61
Figure 3.1 The principle of Gateway Cloning used to generate a <i>JMJ14</i> complementation and overexpression construct.....	69
Figure 3.2 Schematic diagram of the cloning strategy for <i>JMJ14</i> complementation and overexpression in <i>Arabidopsis</i>	71
Figure 3.3 Comparison of WT, <i>jmj14-1</i> and <i>lyn1</i> on the phenotypes	73
Figure 3.4 Identification of <i>jmj14-1 hy1</i> by genotyping assays	74
Figure 3.5 <i>jmj14</i> suppression of seedling phenotypes	76
Figure 3.6 <i>jmj14</i> effect of hypocotyl length of 5-day seedlings in different light fluence rate	78

Figure 3. 7 <i>jmj14</i> effect of chlorophyll content of 5-day seedlings in different light fluence rate	80
Figure 3. 8 Confirmation that <i>jmj14-1</i> fails to complement <i>lyn1</i> suppressing hypocotyl length	83
Figure 3. 9 Colony PCR for identification of full-length <i>JMJ14</i> cDNA insertion in the pGEM-T Easy vector	86
Figure 3. 10 Confirmation of correct insertion of <i>JMJ14</i> fragment in pGEM-T Easy vector by PCR	87
Figure 3. 11 Gel image of PCR products of <i>attB</i> -flanked <i>JMJ14</i> and <i>JMJ14</i> amplified from the pGEM-T Easy vector/ <i>JMJ14</i> plasmid template	88
Figure 3. 12 Colony PCR for identification of full-length <i>JMJ14</i> cDNA insertion in the pDONR201 vector	89
Figure 3. 13 Confirmation of correct insertion of <i>JMJ14</i> fragments in the entry clones	90
Figure 3. 14 Colony PCR for identification of full-length <i>JMJ14</i> cDNA insertion in the pB2GW7 vector	92
Figure 3. 15 Confirmation of correct insertion of <i>JMJ14</i> in the plant expression vector	93
Figure 3. 16 Colony PCR confirming correctly inserted <i>JMJ14</i> in the expression vector	94
Figure 3. 17 Comparison of phenotype between <i>lyn1 hy1</i> , T1 and <i>hy1</i>	96
Figure 3. 18 Confirmation of the present of <i>JMJ14</i> inserts in the T1 plants	97
Figure 3. 19 RT-PCR of full-length <i>JMJ14</i> revealed that transcription of <i>JMJ14</i> was eliminated in <i>jmj14-1</i> mutants but not in <i>lyn1</i> mutants	99
Figure 3. 20 Phylogenetic relationship of JmjC-domain-containing proteins from some selected species	103
Figure 3. 21 Image shown the <i>JMJ14</i> gene expression map of <i>Arabidopsis</i> ..	105
Figure 4. 1 <i>Arabidopsis</i> chloroplast genome map	117
Figure 4. 2 <i>hy2-1</i> mutation effect of phytochrome and chlorophyll biosynthesis pathway	119
Figure 4. 3 Observation of chloroplast development under DIC microscopy ..	123
Figure 4. 4 Quantitative analysis of cell index	125
Figure 4. 5 Correlation between total chloroplast area and cell size of all genotypes.....	128
Figure 4. 6 Quantitative analysis of chloroplast density	129
Figure 4. 7 Quantitation of plastome to nuclear genome copy number ratio...	131
Figure 4. 8 Cotyledon greening of (control) plates without light treatment	133
Figure 4. 9 FR-HIR of cotyledon greening of WT and <i>jmj14</i> single mutants by treating with different periods of FR light exposure	135
Figure 4. 10 FR-HIR of cotyledon greening of <i>hy1</i> mutants by treating with different periods of FR light exposure	137
Figure 4. 11 FR-HIR of cotyledon greening rate of the different genotypes	138

Figure 4. 12 <i>HY2</i> mRNA was confirmed as absent in the <i>hy2-105</i> KO plant (Col background) by reverse transcription.....	139
Figure 4. 13 Identification <i>lyn1 hy2-105</i> double mutants.....	140
Figure 4. 14 <i>lyn1</i> suppression of phenotype of <i>hy2-105</i>	141
Figure 4. 15 <i>lyn1</i> effect of hypocotyl length of <i>hy2-105</i>	142
Figure 4. 16 <i>lyn1</i> effect of chlorophyll content of <i>hy2-105</i>	144
Figure 4. 17 Relative different amount of photochemically active phytochrome in 4 genotypes.....	147
Figure 4. 18 Identification <i>lyn1 cry1-304</i> double mutants.....	149
Figure 4. 19 Identification <i>lyn1 phyB-9</i> double mutants	150
Figure 4. 20 Identification <i>lyn1 phyA-211</i> double mutants	151
Figure 4. 21 Germination and segregation of <i>phyA-211</i> and <i>lyn1 phyA-211</i> ..	152
Figure 4. 22 <i>lyn1</i> dCAPs assay of re-identification <i>phyA-211</i> and <i>lyn1 phyA-211</i> mutants	153
Figure 4. 23 Identification <i>phyA-211</i> mutants.....	154
Figure 4. 24 <i>lyn1</i> suppression of hypocotyl length of photoreceptor apoprotein mutants in 50 $\mu\text{mol m}^{-2} \text{s}^{-1}$ white light	156
Figure 4. 25 <i>lyn1</i> suppression of chlorophyll content of photoreceptor apoprotein mutants in 50 $\mu\text{mol m}^{-2} \text{s}^{-1}$ white light	158
Figure 4. 26 Assessment of <i>lyn1</i> suppression of hypocotyl length of <i>hy1</i> over a fluence rate range of white light	161
Figure 4. 27 Assessment of <i>lyn1</i> suppression of hypocotyl length of <i>phyB-9</i> over a fluence rate range of R light.....	162
Figure 4. 28 Assessment of <i>lyn1</i> suppression of hypocotyl length of <i>phyB-9</i> over a fluence rate range of B light	163
Figure 4. 29 Assessment of <i>lyn1</i> suppression of hypocotyl length of <i>phyA-211</i> over a fluence rate range of FR light.....	164
Figure 4. 30 <i>lyn1</i> suppression of chlorophyll level of <i>phyB-9</i> and <i>cry1-304</i> over a fluence rate range	166
Figure 4. 31 <i>lyn1</i> suppression of chlorophyll level of <i>phyA-211</i> , over a range of FR fluence rate light.....	169
Figure 4. 32 Confirming the presence of <i>JMJ14</i> transgene in the T1 <i>JMJ14</i> overexpressed plants	170
Figure 4. 33 WT plant overexpressed with <i>JMJ14</i>	171
Figure 5. 1 Strategy of RNA-Seq analysis.....	186
Figure 5. 2 Sample images to show the RNA samples collected from <i>Arabidopsis</i> at different developmental stages	189
Figure 5. 3 GO term of total 5326 significantly differentially expressed genes	195
Figure 5. 4 Functional gene classification analysis of total 5326 significantly differentially expressed genes.....	196
Figure 5. 5 Cluster 5 of differentially expressed genes reveals enrichment of biological process categories	197

Figure 5. 6 Cluster 1 of differentially expressed genes reveals enrichments of biological function categories	199
Figure 5. 7 <i>lyn1</i> up-regulated subclusters selected from Cluster 1	200
Figure 5. 8 <i>lyn1</i> up-regulated and light-dependent subclusters 1.1 and 1.10 reveal enrichments of biological function categories	201
Figure 5. 9 Functional gene classification of the differentially expressed genes during the early development of 7-day old seedlings	204
Figure 5. 10 Functional gene classification of differentially expressed genes during early development of 17-day young plants	205
Figure 5. 11 Functional gene classification of the differentially expressed genes between photosynthetic vs. non-photosynthetic organs (7/3) of 7-day old seedlings	206
Figure 5. 12 Overrepresentation of genes which are active at a particular developmental stage among <i>lyn1</i> -dependent gene clusters	207
Figure 5. 13 Overrepresentation of genes which are active at a particular developmental stage among light-dependent but <i>lyn1</i> -independent gene clusters	209
Figure 5. 14 (in the following page) Apparent differential extent of H3K4me3 chromatin marking associated with photosynthesis related genes in the DNA of whole seedlings and in the DNA of roots	211
Figure 5. 15 Venn diagram showing the number of genes which are H3K4me3-marked in common or uniquely in genomic DNA from whole seedlings or roots	213
Figure 5. 16 Functional classification analysis of H3K4me3-marked genes ...	215
Figure 5. 17 Over/underrepresentation of H3K4me3-marked genes among all the differentially expressed gene clusters of interest	216
Figure 5. 18 Overrepresentation of H3K4me3-marked genes among genes selected for their increased expression at particular developmental stages	218
Figure 7. 1 <i>lyn1</i> effect of protochlorophyllide level in the dark.	283
Figure 7. 2 Chloroplast compartment quantification in mesophyll cells of rosette and cauline leaves	284
Figure 7. 3 <i>JMJ14</i> location on <i>Arabidopsis</i> genetic map	285
Figure 7. 4 pGEM [®] -T Easy Vector map	286
Figure 7. 5 pDONR [™] 201 Vector map	287
Figure 7. 6 Gateway [™] Destination Vector pB2GW7 vector circle map	288
Figure 7. 7 Schematic domain structure of <i>JMJ14</i> homologues	289
Figure 7. 8 All clusters resulting from the k-means clustering of all differentially expressed genes	290
Figure 7. 9 Sub-clusters resulting from the k-means clustering of differentially expressed genes present in cluster 1	291

List of Tables

Table 2. 1 List of all the biological replicates of each genotype and each light condition included and excluded in the RNA-Seq data analysis	52
Table 2. 2 Chemical agents used for bacterial and transformed plant selection and their preparation details.....	58
Table 3. 1 List of homologues in different species evidenced by the evolutionary relationships of the JMJ14 protein	104
Table 5. 1 Table to show the details of RNA samples for the Affymetrix microarray carried out as the <i>Arabidopsis</i> Gene Atlas	190
Table 7.1 Genotyping and sequencing primer details	292
Table 7.2 Restriction enzyme details	293
Table 7.3 Multiple sequence alignment of JMJ14 homologues.....	294

Abbreviations

AGI	<i>Arabidopsis</i> genome initiative
αKG	α-ketoglutarate
aa	Amino acid
ABI4	ABSCISIC ACID INSENSITIVE 4
Agrobacterium	<i>Agrobacterium tumefaciens</i>
ALA	Aminolevulinic acid
Amp	Ampicillin
APRR	<i>ARABIDOPSIS</i> PSEUDO RESPONSE REGULATOR
Arabidopsis	<i>Arabidopsis thaliana</i>
ARID	AT-rich interaction domain
At	<i>Arabidopsis thaliana</i>
ATX1/2	<i>ARABIDOPSIS</i> TRITHORAX 1/2
B	Blue
BAR	Bio-Array Resource
BB	Bacterial blight
bHLH	basic helix–loop–helix
BLAST	Basic Local Alignment Search Tool
BOP	BLADE-ON-PETIOLE
Bp	Base pair
CA1	Carbonic anhydrase 1
CAB	Chlorophyll a/b-binding proteins of photosystem II (termed LHCb in this thesis)
CAPs	Cleaved Amplified Polymorphic Sequences
CDD	COP10, DDB1 and DET1
cDNA	Complimentary DNA
CGA1	CYTOKININ-RESPONSIVE GATA1
CHLH	H subunit of Mg chelatase
Col Los5	<i>Arabidopsis thaliana</i> Columbia- <i>Lhcb::Ω::GFP</i> reporter transgene 5
conc.	Concentration
COP/DET/FUS	Constitutively-photomorphogenic/de-etiolated/fusca
COP1	Constitutive photomorphogenic 1
CP12	Chloroplast protein 12
CRF2	CYTOKININ RESPONSE FACTOR2
CSN	COP9 signalosome
Cry	Cryptochrome
cue	CAB underexpressed
cue8	<i>Arabidopsis thaliana</i> cue8 mutant
CUL3	Cullin3
Ct	Cycle threshold

dCAPs	Derived Cleaved Amplified Polymorphic Sequences
DET1	DE-ETIOLATED 1
DDB1	DAMAGED DNA-BINDING PROTEIN 1
Dm	<i>Drosophila melanogaster</i>
DIC	Differential interference contrast
DMF	N,N-dimethylformamide
DNA	Deoxyribonucleic acid
E. coli	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
EMS	Ethyl methane sulfonate
EtBr	Ethidium bromide
FAD	Flavin Adenine Dinucleotide
FeChII	Ferrochelataase
FC1	Ferrochelataase 1
FDR	False-discovery rate
FHY1	FAR-RED ELONGATED HYPOCOTYL 1
FHL	FHY1-LIKE
FLC	FLOWERING LOCUS C
FR	Far-red
FR-HIR	Far-red light high irradiance response
FT	FLOWERING LOCUS T
GFP	Green fluorescent protein
GLK	GOLDEN2-LIKE
Gm	<i>Glycine max</i>
GNC	GATA NITRATE-INDUCIBLE CARBON-METABOLISM-INVOLVED
GO	Gene Ontology
GUN	Genome uncoupled
H3K4	Histone H3 lysine 4
HCL	Hydrochloric acid
HEMA1/2/3	Glutamyl-tRNA reductase 1/2/3
HDM	histone demethylase
HFR1	LONG HYPOCOTYL IN FAR-RED
HKMT	Histone lysine methyl-transferase
HO1/2/3/4	Haem oxygenase 1/2/3/4
Hs	<i>Homo sapiens</i>
HY2	P Φ B synthase
HY5	Elongated hypocotyl 5
IPA	Isopropylalcohol
IPTG	Isopropyl- β -d-1-thiogalactopyranoside
IR	Inverted repeats
Jmj	Jumonji
JHDM	Jumonji C (JmjC)-domain-containing histone demethylase

Kan	Kanamycin
KMC	K-means/Medians clustering
KO	Knock out
LAF1	LONG AFTER FAR-RED LIGHT 1
LB	Lithium borate
Ler	Landsberg erecta
LFR	Low fluence responses
LHCB	Light-harvesting chlorophyll-binding proteins of the PSII LHC (formerly termed CAB)
LIN	Lincomycin
LSC	Large single-copy
LSD1	Lysine specific demethylase 1
MeV	MultiExperiment Viewer
MgCl₂	Magnesium chloride
mRNA	Messenger ribonucleic acid
MS	Murashige-Skoog
N_{2(l)}	Liquid nitrogen
NaCl	Sodium chloride
NF	Norflurazon
NMD	Nonsense-mediated decay
OI	<i>Ostreococcus lucimarinus</i>
Os	<i>Oryza sativa</i>
PΦB	Phytochromobilin
p35S	Cauliflower mosaic virus 35S promoter
pB2GW7	Gateway™ pB2GW7 vector
PCA	Principal component analysis
Pchl_{ide}	Protochlorophyllide
PCR	Polymerase chain reaction
PET	Photosynthetic Electron transport
PDV2	PLASTID DIVISION2
P_{fr}	Far red light-absorbing
Pfu	Phusion
pg	Picogram
PhANGs	Photosynthesis-associated nuclear genes
PHD	plant homeodomain
Phot	Phototropin
Phy	Phytochrome apoprotein
Phy	Phytochrome
PIFs	Phytochrome Interacting Factors
PLBs	Prolamellar bodies
POR	NADPH:protochlorophyllide oxidoreductase
PORA	NADPH:protochlorophyllide oxidoreductase A
Pp	<i>Physcomitrella patens</i>

PPR	Plastid-localized pentatricopeptide-repeat
PPT	Phosphinothricin
P_r	Red light-absorbing
PTC	Premature termination codons
PTGS	Posttranscriptional gene silencing
PTM	PHD WITH TRANSMEMBRANE DOMAINS
PTM	Post-translational covalent modification
qg-PCR	Quantitative real-time genomic-PCR
qPCR	Quantitative real-time polymerase chain reaction
R	Red
R-HIR	R light high irradiance responses
Rif	Rifampicin
rbcL	Rubisco large subunit
rpm	Revolution pre minute
RNA	Ribonucleic acid
RNA-Seq	RNA sequencing
ROS	Reactive oxygen species
RPS6	Ribosomal protein S6
RT-PCR	Reverse transcription polymerase chain reaction
SAM	Shoot apical meristem
SDG	SET-domain gene
SDS	Sodium dodecyl sulphate
Sm	<i>Selaginella moellendorffii</i>
SPA	SUPPRESSOR OF PHYA-105
Sp	Spectinomycin
Sc	<i>Saccharomyces cerevisiae</i>
SSC	Small single-copy
T1/2/3	Transformation line - generation 1/2/3
TBE	Tris/Borate/EDTA
TAIR	The <i>Arabidopsis</i> Information Resource
TE	Tris-EDTA
Temp	Temperature
TF	Transcriptional factor
Trx	Trithorax
TSF	TWIN SISTER OF FT
UV	Ultraviolet
UV-B	Ultraviolet-B
V	Volt (s)
VLFR	Very low fluence responses
W	White
WT	Wild type
X-gal	5-bromo-4-chloro-indoyl-β-d-galactopyranoside

Acknowledgements

I would like to express my primary thanks of gratitude to my supervisor, Dr. Enrique López-Juez, for giving me the opportunity to do this PhD. I appreciate for all his guidance, support, patience and tolerance throughout the course of this study. I would like to thank Dr. Paul Devlin for being my advisor and for encouraging me to be a self-confident research scientist. I thank Royal Holloway University of London for giving me the opportunity to study in the U.K. since my Bachelor of Science degree.

Secondly, I would also like to thank my colleagues, Sara and Naresh, who helped me during this process in the lab and exchanged the science knowledge and practical experience. Thanks to the members of Dr. Devoto's lab, Imma and Stacey. Imma gave me a lot of helpful advices on cloning experiment and Stacey shared her experience of using R with me.

Thanks to my friend, Sara, Kelly, Yihan Liu, Hongdan Deng, Yajun Li, who gave me a lot of help and support in life. Especially, to Dr. Hongdan Deng who helped me with designing clear scientific figures and computing skills for optimizing my thesis.

I must thank my uncle, Colin who proofread my thesis when he was busy on running the family business. And I thank Sara who did some proofreading for me by the end when she was busy on her research.

Finally, I must acknowledge my family for their encouragement. Especially, I am extremely grateful to my Mum and grandparents from the bottom of my heart for giving me the moral and financial support to pursue this degree.

Declaration of Authorship

I, Jia He, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed: _____

Date: _____

Chapter 1 General Introduction

1.1 Photosynthesis and chloroplast evolution

Photosynthesis is considered to be the most fundamental mechanism to providing the basic energy source for the biosphere, thanks to photosynthetic organisms. These organisms consume water and atmospheric CO_2 , using sunlight energy, to drive a light-dependent reaction which converts H_2O into O_2 to take the electrons through a photosynthetic electron transport (PET) chain and releases ATP and NADPH used as energy in the Calvin cycle (Figure 1.1). The Calvin cycle is the light-independent reactions, also known as dark reactions that synthesise carbohydrates using carbon from CO_2 (Figure 1.1 shown in green). Carbohydrates are the most abundant existing biomolecule to support living organisms on Earth. Carbohydrates can be catabolized and then go through a citric acid cycle and oxidative phosphorylation to release energy, H_2O and CO_2 (Figure 1.1 shown in red). This effluent CO_2 from respiration is released into the atmosphere for the next photosynthetic reaction. Maintaining atmospheric CO_2 level by photosynthesis and respiration becomes very important for ecological equilibrium.

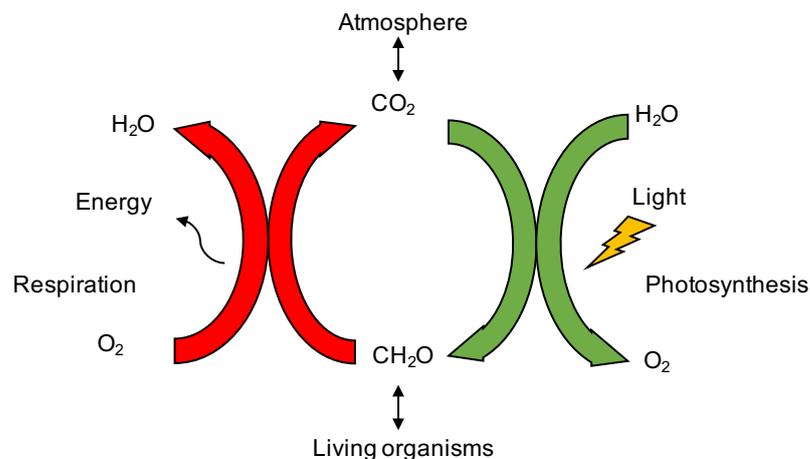


Figure 1.1 Respiration and photosynthesis maintain energy support for living organisms

Through previous research, I have evidence to indicate that oxygenic photosynthesis was prevalent at least 2 billion years ago and maybe even much earlier (Blankenship, 2001). As is known to all, plant chloroplasts are where photosynthesis takes place. Back 300 million years ago, oxygenic photosynthesis was first found in the cyanobacteria. Cyanobacteria have given undisputable evidence that they existed 2.7 billion years ago (Summons et al., 1999). On the evolution process of oxygenic photosynthesis from prokaryotes to eukaryotes it has now been accepted that a single mitochondrion-containing eukaryotic cell engulfed a cyanobacterium by endosymbiosis around 1.2 to 1.5 billion years ago (Dyall et al., 2004). After the engulfment, the cyanobacterium evolved into a chloroplast and this was followed by further diversification (López-Juez and Pyke, 2005)

Chloroplasts are unique organelles present in plants and they are the members of a diverse family, the plastids. In plants, plastids exist in various forms such as chloroplasts, etioplasts, amyloplasts, chromoplasts, elaioplasts and gerontoplasts (López-Juez and Pyke, 2005; Egea et al., 2011) (Figure 1.2). They have diversified in plants and have different functions to play various roles in plant growth and development through biogenesis, varying functional states and metabolic activities, such as embryogenesis, leaf development, gravitropism, temperature response and plant-microbe interactions (Inaba and Ito-Inaba, 2010).

Basic mechanisms in plastid biogenesis (the transformation of a proplastid into a particular type of plastid) and the interconversion between different types of plastid are controlled through their developmental and environmental signals (Jarvis and López-Juez, 2013). Proplastids are converted into etioplasts in the absence of light. Etioplast development allows a very large accumulation of protochlorophyllide (Pchl_{id}), the processed chlorophyll precursor, making it ready to be photoconverted by light. The etioplasts are immature chloroplasts which are developmentally arrested during the transition from proplastid to

chloroplast because of absence of light (Wise and Hooper, 2006). They rapidly differentiate into chloroplasts under the light (Waters and Langdale, 2009).

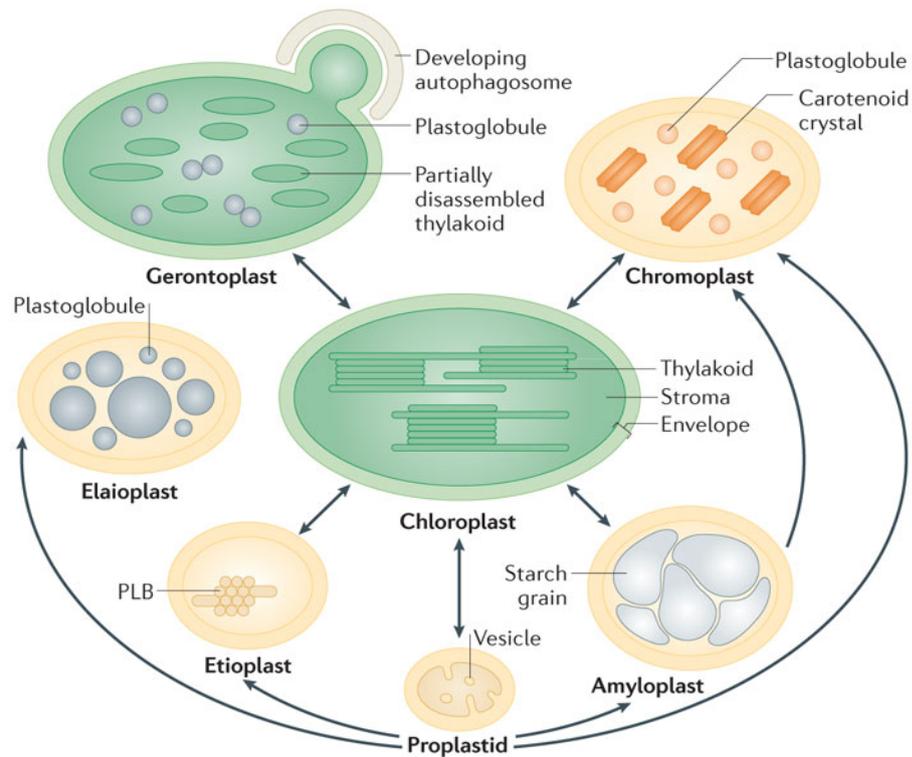


Figure 1.2 Diversity of plastid forms

Plastids exist in different forms and perform different functions in different compartments of the plant. Biosyntheses of different forms of plastid from proplastid and interconversions between different types of plastids are controlled by developmental and environmental signals. The plastid differentiation is controlled by the transcription of nucleus-encoded plastid proteins. The dark-developed etioplasts, which are chloroplast progenitors, containing unique internal membranes called prolamellar bodies (PLBs) are the places where chlorophyll precursors, Pchl_{ide}, are accumulated (Fujii et al., 2017). They can rapidly differentiate into chloroplasts upon illumination, whereas Pchl_{ide} can rapidly convert into chlorophyll by light-dependent NADPH:protochlorophyllide oxidoreductase (POR) to avoid photo-damage. Chloroplasts are photosynthetic plastids. Amyloplasts play important roles in energy storage and gravitropism. Chromoplasts contain carotenoid pigments for facilitating flower pollination and

seed dispersal of fleshy fruit. Elaioplasts play a role in the storage of lipids in lipid droplets as plastoglobules. Gerontoplasts are developed from chloroplasts during senescence, owing to resource recycling. Figure from (Jarvis and López-Juez, 2013)

Arabidopsis leaves are constituted by different types of cells, including epidermal pavement cells, mesophyll parenchyma cells, the bundle sheath cells, etc (Figure 1.3). The mesophyll cells contain the most chloroplasts in leaf mesophyll tissue, and are the main location where carbon fixation and photosynthesis happen in C_3 plants (Edwards and Walker, 1983). Mesophyll tissue consists of palisade cells and spongy cells, both containing a large population of chloroplasts (Figure 1.3 C and D), that enable the leaf to carry out photosynthetic carbon assimilation (Pyke, 2012). Bundle sheath cells are another type of chloroplast-containing cells which surround the vascular bundle in leaf mesophyll tissue (Kinsman and Pyke, 1998) (Figure 1.3 E). They harvest small amounts of light because the majority of light is utilized by mesophyll cells. In C_3 plants, they are assumed to be responsible for only a small fraction of total photosynthesis. Bundle sheath cells have been found to act differently between C_3 and C_4 plants. In C_4 plants, mesophyll cells carry out the light reactions of photosynthesis which convert light energy to ATP and NADPH. They also perform the initial, temporary fixation of CO_2 as a C_4 molecule. This C_4 molecule is shuttled to the bundle sheath cells, where it is de-carboxylated, causing in them an elevated concentration of CO_2 . Bundle sheath cells work cooperatively with mesophyll cells and expend the ATP and NADPH to fix this shuttled CO_2 , now permanently, through the Calvin cycle, classically known as dark reactions (Devi et al., 1995). Interestingly, the cells of stems and petioles that surround the xylem and phloem have been found to show characteristics of C_4 photosynthesis in tobacco, a typical C_3 plant. They contain high activities of C_4 photosynthetic enzymes. The carbon supply for these cells is from neighbouring cells, the vascular system (as in the C_4 photosynthetic pathway) instead of from stomata (as in the C_3 photosynthetic pathway) (Hibberd

and Quick, 2002). The thin cell layers on both surfaces of the leaf are epidermal pavement cells. The epidermal pavement cells are transparent and allow light to pass unimpeded to the mesophyll cells. Although epidermal pavement cells also contain chloroplasts (Pyke, 2009), their chloroplasts are much less fully developed than the mesophyll cell chloroplasts (Chiang et al., 2012; Higa et al., 2014) (Figure 1.3 B). They only contain one-tenth the amount of chloroplasts of the mesophyll cells and the average size of pavement cell chloroplasts is about one-half the size of the mesophyll cell chloroplasts (Barton et al., 2016).

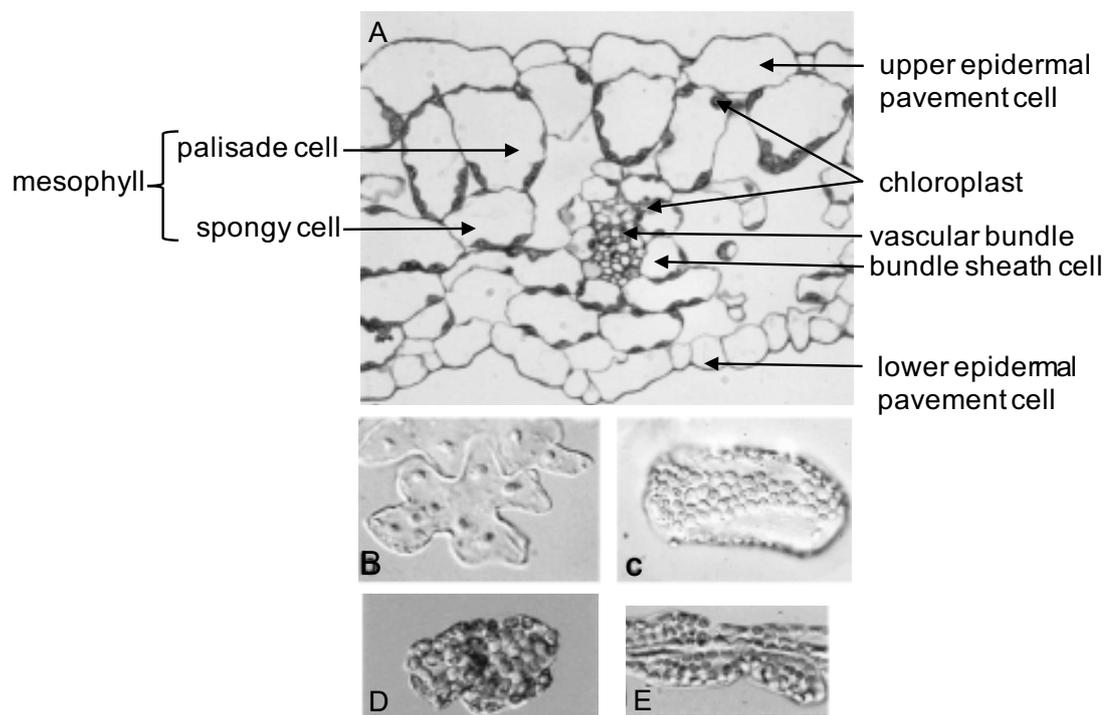


Figure 1. 3 Structure and cell types of the *Arabidopsis* leaf.

A. Transverse section through the first leaf of *Arabidopsis* showing the various cell types;

B. Highly irregular shaped epidermal pavement cell, containing several small chloroplasts;

C. Palisade mesophyll cell showing the top cell surface with many chloroplasts;

D. Spongy mesophyll cell showing the top cell surface with many chloroplasts;

E. Several bundle sheath cells, containing chloroplasts and attached to a vascular bundle. Figure adapted from (Pyke and López-Juez, 1999).

1.2 Light signalling control of plant development

1.2.1 Types of photoreceptor

As mentioned above, light works as a primary energy source in photosynthesis and it is also a very important external factor for plant growth and development. Plants exhibiting different growth stages, which include germination, hypocotyl elongation and cotyledon expansion, are controlled by light signals (Jiao et al., 2007). Plastids, contained in plant cells, also develop during plant growth. Dark-grown seedlings present a skotomorphogenic phenotype. Their cotyledons remain folded and undeveloped, while the hypocotyls rapidly elongate, and etioplasts develop in those cotyledons (Wise and Hooper, 2006; Kobayashi et al., 2012a). Light-grown seedlings present a photomorphogenic phenotype. They exhibit short hypocotyls and open and expanded cotyledons. The shoot apical meristem is activated to produce true leaves, where chloroplasts are developed (Jiao et al., 2007). Light acts as a signal that can be perceived through several photoreceptors in plants. Plant photoreceptors have been classified into four families, namely phytochrome (phy), cryptochrome (cry), phototropin (phot) and an ultraviolet-B (UV-B) receptor called UVR8 (Rizzini et al., 2011; Arsovski et al., 2012). They respond to different types of light. phys respond primarily in red (R) and far-red (FR) light (Li et al., 2011). Cryptochromes and phototropins respond in blue (B) and ultraviolet-A (UV-A) light (Dieterle et al., 2003; Li et al., 2011). In *Arabidopsis*, there are five subtypes of phys, which are phyA, phyB, phyC, phyD and phyE (Sharrock and Quail, 1989). Type I (light labile) phytochrome, phyA, mediates very low fluence responses (VLFR) and far-red high-irradiance responses (FR-HIR). It is abundant and predominant in the etiolated seedlings. The rest of phys, phyB-phyE, are type II (light stable). They mediate low fluence responses (LFR) and R light high irradiance responses (R-HIR) (Nagatani, 2004; Fankhauser and Chen, 2008; Li et al., 2011). In light-grown plants, phyB is the

predominant phytochrome and it is much more abundant than phyC-phyE (Li et al., 2011). phyA have a vital role in controlling deetiolation, while their responses overlap with those of cryptochromes on flowering. In *Arabidopsis*, cryptochrome 1 (cry1) is photostable and mainly responds to high light fluence and cryptochrome 2 (cry2) is photodegradable and responds to low light fluence (Fankhauser and Casal, 2004). Although phototropins display many overlapping responses with phytochromes and cryptochromes on chloroplast movement, stomatal opening and phototropism. These responses, mostly associated with membrane activity, are mainly controlled by phototropins (Briggs and Christie, 2002).

1.2.2 Phytochromes' biogenesis

Phy is synthesized when phytochrome apoprotein is covalently attached to a linear tetrapyrrole chromophore, phytychromobilin (PΦB), in the cytoplasm (Figure 1.4). The PΦB is synthesized from aminolevulinic acid (ALA) in plastids through a series of enzymatic reactions (Muramoto et al., 1999; Emborg et al., 2006). PΦB, chlorophyll and haem biosynthesis share the early pathway. They split at converting protoporphyrin IX into either haem or Mg-protoporphyrin, respectively, thereby directing into the PΦB or chlorophyll biosynthesis pathways (Mochizuki et al., 2010). PΦB is the end product of the phytochrome biosynthesis pathway within plastid for immediately synthesizing phytochrome in the cytoplasm (López-Juez and Pyke, 2005). Haem oxygenase (HY1) and PΦB synthase (HY2) are the two enzymes involved in the last two steps of PΦB biosynthesis pathways within plastid (Muramoto et al., 1999; Li et al., 2011).

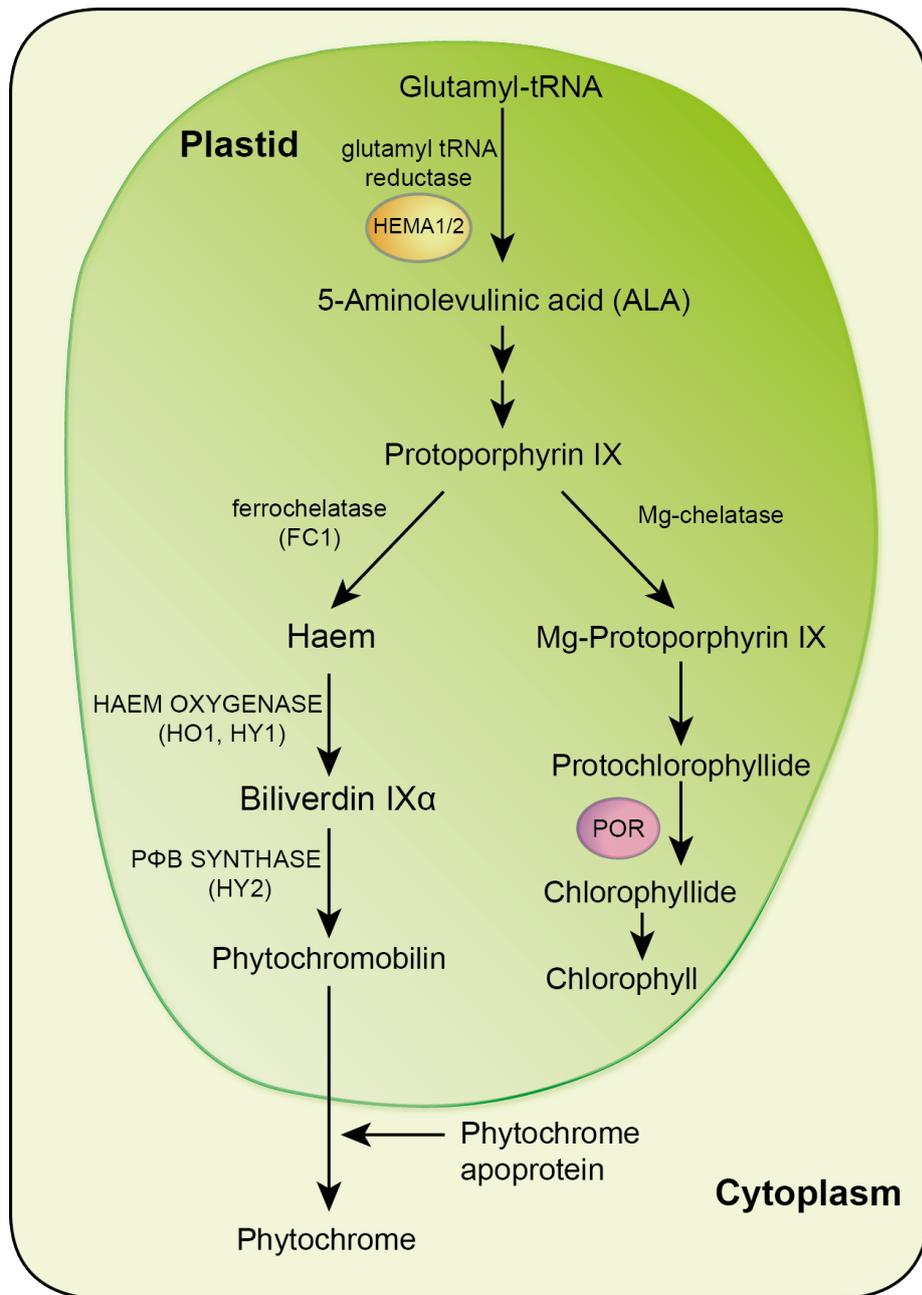


Figure 1.4 The phytochrome biosynthesis pathway in *Arabidopsis* plastids and cytoplasm

1.2.3 Phytochromes' and cryptochromes' nuclear localization

In *Arabidopsis*, phytochromes exist in either of two photo-interconvertible forms: the inactive form (P_r) or active form (P_{fr}) (Li et al., 2011). Phytochromes are synthesized and dimerized in the cytoplasm, accumulating as P_r form in the dark (Nagatani, 2004). Light triggers the conformation change from P_r form to P_{fr} form. The light-

activated Pfr form is imported into the nucleus to drive nuclear gene transcription. phyA subcellular localization is regulated by all light illuminations (FR, R and B) (Kevei et al., 2007). No detectable level of phyA is found in nucleus of dark-grown plants (Kircher et al., 2002). A single light pulse (FR, R or B) causes an immediate importing of phyA. phyB subcellular localization is regulated by continuous R light and to a small extent is regulated by continuous B light, but is completely unaffected by FR light. phyB nuclear localization cannot be induced by a single pulse of R, FR or B light indicating that phyB nuclear localization takes longer periods compared to phyA (Gil et al., 2000). However, phyB has been found to be present in the nucleus in the dark in a small amounts (Kevei et al., 2007). In contrast to single protein translocation of phyB, phyA translocation depends on two small plant-specific proteins, FAR-RED ELONGATED HYPOCOTYL 1 (FHY1) and FHY1-LIKE (FHL). FHY1/FHL act as adaptor proteins to transfer phyA into the nucleus by forming a complex with phyA (Genoud et al., 2008). crys are predominantly nuclear proteins (Lin and Todo, 2005), having a variety of overlapping functions with phys and also cooperating with phys to regulate gene expression through complex transcriptional networks in light-triggered responses (Jiao et al., 2007; Li et al., 2012; Rodrigues et al., 2014).

1.2.4 Light mediates plant morphogenesis

Although light signal-induced transcriptional networks controlling plant morphogenesis are extremely complex, the signalling pathways overall are mainly directed by two negative regulators, which are photomorphogenic/de-etiolated/fusca (COP/DET/FUS) proteins and Phytochrome Interacting Factors (PIFs), and a group of positive regulators, such as LONG HYPOCOTYL 5 (HY5) (Figure 1.5). In the dark, positive transcriptional regulators are inhibited by phosphorylation or degraded by photomorphogenesis negative regulator-caused ubiquitination. Constitutive COP/DET/FUS group proteins are crucial in regulating negatively or positively many of the transcription factors (TFs) that initiate the

photomorphogenesis and skotomorphogenesis (Jiao et al., 2007; Li et al., 2011). The COP/DET/FUS proteins are defined into three groups: the COP1-SPA complexes, the COP9 signalosome (CSN), and the CDD complex (COP10, DDB1 and DET1), all of which are involved in proteasomal degradation of photomorphogenesis positive TFs (Li et al., 2011). COP1-SPA complex is an E3 ubiquitin ligase, formed by interaction of CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1) and SUPPRESSOR OF PHYA-105 (SPA) proteins to degrade several positive TFs in order to inhibit photomorphogenesis in the dark (Jiao et al., 2007; Casal et al., 2014). The CSN is a multiprotein complex (Lau and Deng, 2012) and The CDD complex is formed by DE-ETIOLATED 1 (DET1), DAMAGED DNA-BINDING PROTEIN 1 (DDB1) and COP10 interaction (Waters and Langdale, 2009). They are both essential for regulating activity of E3 ubiquitin ligases (Chen, 2006; Lau and Deng, 2012). CDD directly interacts with COP1 and CSN to promote the E2 ubiquitin enzyme activity in degradation (Fernando and Schroeder, 2016). PIFs, a group of bHLH (basic helix–loop–helix)-type transcription factors, are phyA and phyB-regulated negative regulators of photomorphogenesis and positive regulators of skotomorphogenesis. The inactive phys cannot inhibit the PIFs in the dark, so skotomorphogenesis is promoted by PIFs, whereas photomorphogenesis is repressed (Jiao et al., 2007; López-Juez et al., 2008). Recently, COP/DET/FUS and PIFs have been found to act as cofactors and synergistically repress photomorphogenesis (Xu et al., 2014). The COP/DET/FUS group of proteins can be inhibited by light-activated phys or crys. Once they are inhibited, they are unable to degrade the nuclear localized transcriptional factors (TFs) such as HY5, LONG HYPOCOTYL IN FAR-RED (HFR1), LONG AFTER FAR-RED LIGHT 1 (LAF1), etc (Jiao et al., 2007). HY5 is a well-characterized bZIP (basic leucine zipper) TF playing a major role in promoting photomorphogenesis under various light conditions (Li et al., 2011; Arsoovski et al., 2012). Both phys and crys promote HY5 accumulation in the nucleus. HFR1, a bHLH TF, acts as a suppressor of

hypocotyl elongation in response for phyA-mediated FR and cry1-mediated B light signalling. LAF1 is a Myb transcriptional activator, responsible for transmitting phyA-mediated downstream signalling (Jang et al., 2005; Yang et al., 2005; Casal et al., 2014). HY5, HFR1 and LAF1 can interact with each other to prevent the degradation from COP1 (Casal et al., 2014). Meanwhile, activated phyA cause phosphorylation and degradation of PIFs, resulting in plant photomorphogenesis (Arsovski et al., 2012). According to our knowledge so far, it has been found that PIFs are degraded by BLADE-ON-PETIOLE (BOP) protein, which in association with Cullin3 (CUL3) protein which forms a complex with E3 ubiquitin ligase activity (Zhang et al., 2017). Once plants enter photomorphogenesis, the shoot apical meristem (SAM) is activated to produce leaves and subsequently chloroplasts (López-Juez et al., 2008).

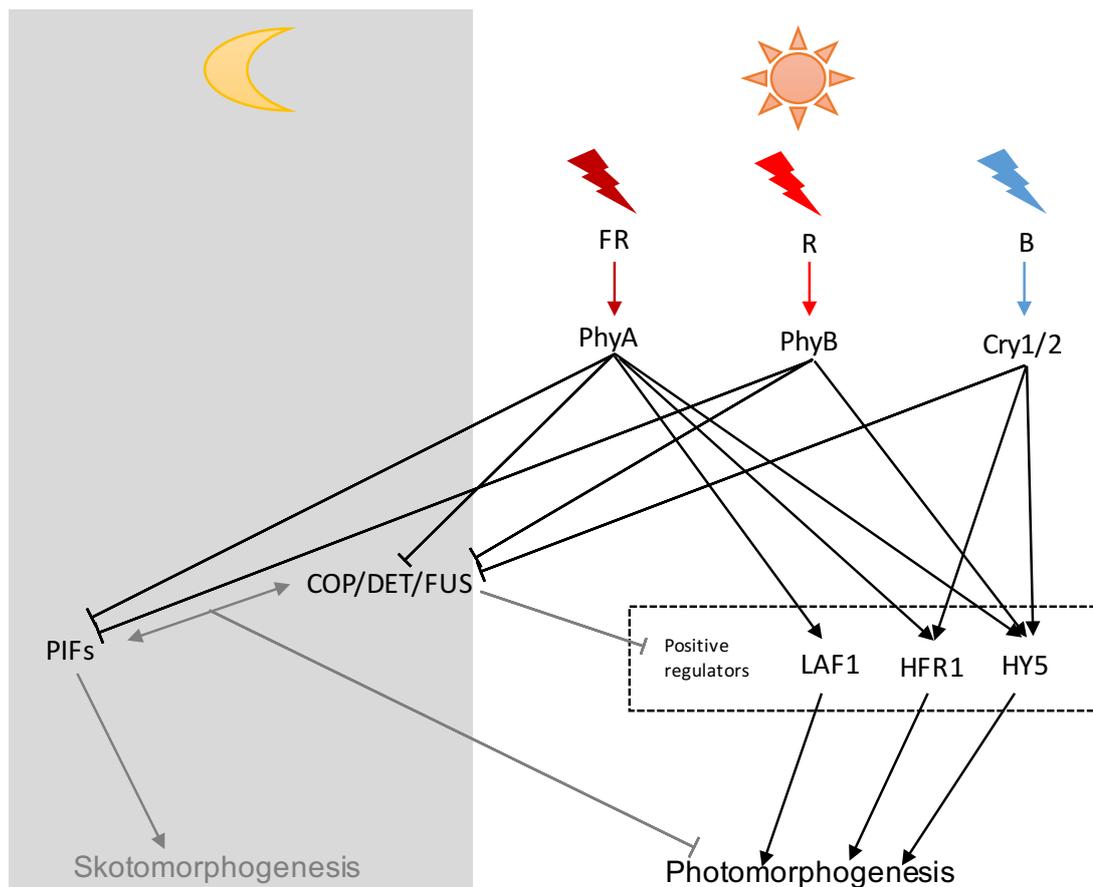


Figure 1.5 Simplified transcriptional networks for skotomorphogenesis and photomorphogenesis

1.3 Communication between plastid and nucleus

1.3.1 Nucleus-to-plastid anterograde build-up and regulation pathways

To build up a mature plastid, about 3000 different proteins are involved. As plastids are evolved from prokaryotes, they are semi-autonomous organelles which retain genes for a small number of polypeptides. However, there are only about 100 genes and the majority of plastid proteins are encoded in nucleus (Inaba and Ito-Inaba, 2010). Moreover, the expression of plastid-encoded genes is also directed by the nucleus. Therefore, plastid biogenesis is dramatically controlled by the nucleus and by the post-translational import of nuclear-encoded plastid proteins into plastids (Kessler and Schnell, 2009). Nuclear-encoded chloroplast genes are rapidly expressed through the transcriptional network. The early light-responsive transcriptional factors express within 1 hour (Jiao et al., 2007). The chloroplast biogenesis is usually completed after 6 to 24 h after light exposure (López-Juez et al., 2008). These nuclear-encoded chloroplast proteins are imported into proplastids, then involved in chloroplast biogenesis and maturation.

Currently, only a few of TFs that regulate chloroplast development have been well studied (Figure 1.6). HY5 is a bZIP transcription factor that binds to a conserved G-box in the promoter region of many light-regulated genes involved in promoting chloroplast biosynthesis genes under a wide spectrum of wavelengths, including FR, R, B, and UV-B (Masuda and Fujita, 2008; Waters and Langdale, 2009; Li et al., 2010, 2011; Kobayashi et al., 2012a). *CHLH* (also called *GUN5*, the H subunit of Mg-chelatase), *GUN4* (required for activating Mg-chelatase) and *CAO* (catalyzes the conversion of chlorophyllide a to chlorophyllide b) which are identified HY5 targeted genes involved in chlorophyll biosynthesis of tetrapyrrole pathway. Importantly, strong evidence has shown

that photosynthesis-associated gene *light-harvesting chlorophyll a/b-binding1* (*Lhcb1/CAB1*) is positively regulated by HY5 (Lee et al., 2007; Waters et al., 2008). LHCB proteins are the subunits of Photosystem II components of PET complexes, located on thylakoid membranes in the chloroplast. They contain protein-bound chlorophyll that optimize light absorption and transfer excitation energy to additional chlorophylls (Waters et al., 2008, 2009; Waters and Langdale, 2009).

LHCB requires a crucial TF - GOLDEN2-LIKE (GLK). The efficient expression of *LHCB* requires HY5 to maximize the GLK function (Kobayashi et al., 2012b). The GLK proteins belong to the GARP superfamily of TFs (Riechmann et al., 2000). *GLKs* gene expression is promoted by R and B light (Waters et al., 2008). In *Arabidopsis*, *GLK1* and *GLK2* are homologues which are largely redundant and functionally equivalent. Water and collaborators showed only *glk1 glk2* double mutant has a significantly defective pigment synthesis and thylakoid stacking. They also showed that GLKs are required for Pchl_{id} synthesis in the dark. Although *GLK1* and *GLK2* both play a same role in retrograde signalling, *GLK1* plays a key role in retrograde signalling. In contrast, *GLK2* has a minor role in retrograde signalling (Martin et al., 2016). However, they also have distinct functions of chloroplast biosynthesis. Although *GLK1* and *GLK2* are both light-dependent genes, *GLK1* is primarily light regulated. By contrast, *GLK2* is regulated by light and also Auxin/Cytokinin signalling (Fitter et al., 2002; Kobayashi et al., 2012a). *GLK2* is more abundantly transcribed compared to *GLK1* in the root (Fitter et al., 2002). Kobayashi and collaborators have found that *GLK2* is a hormone signalling-responsive factor, regulating root greening. The *ARABIDOPSIS* PSEUDO RESPONSE REGULATORS (APRRs) also belong to the GARP superfamily of TFs (Riechmann et al., 2000). They are incompletely studied cytokinin-regulated proteins (Makino et al., 2000; To et al., 2007; Pan et al., 2013). In tomato, APRR2-Like has been found to have a high sequence

similarity to tomato GLK2 (Pan et al., 2013). Therefore, it was proposed to have a function on plastid metabolism and pigment development. A recent study showed that APRR2-Like is a positive transcriptional regulator for mediating plastid compartment size and the levels of chlorophyll in tomato (Pan et al., 2013).

CYTOKININ RESPONSE FACTOR2 (CRF2) which is a cytokinin-responsive TF has been found to control chloroplast growth and division (Okazaki et al., 2009). It is transcriptionally up-regulated by cytokinin and light-independent. CRF2 increases the chloroplast division rate by regulating the PLASTID DIVISION2 (PDV2) protein, components of the division apparatus to change chloroplast size and number, but not other components of the division apparatus (Okazaki et al., 2009; Chiang et al., 2012).

The two positive nucleus-localized TFs, GATA NITRATE-INDUCIBLE CARBON-METABOLISM-INVOLVED (GNC) and CYTOKININ-RESPONSIVE GATA1 (CGA1), have been identified to regulate chloroplast development (Richter et al., 2010). The expression of *GNC* and *CGA1* is light-dependent and also cytokinin-regulated (Chiang et al., 2012). It has been found that PIF3 directly binds to the GNC/CGA1 promoter to repress the gene expression (Richter et al., 2010). It has also been found that GNC/CGA1 has an overlapping function with GLKs in inducing the transformation of chloroplasts from proplastids. GNC/CGA1 also has an overlapping function with CRF2 in chloroplasts growth and division during proliferation. It mediates chloroplast development in multiple cell types and tissues, except flower petals. Its expression is repressed during flower development (Mara and Irish, 2008).

PIFs is a subfamily of phytochrome-interacting bHLH transcription factors (Jiao et al., 2007). PIFs have been found to regulate a number of chloroplast biosynthesis genes and photosynthesis genes, particularly of tetrapyrrole biosynthesis genes. PIF1 and PIF3 both repress *HEMA1*, *GUN5* and *GUN4* that

lead to a reduced Pchlide synthesis in the dark (Stephenson et al., 2009). PIF1 directly targets *PORC* and indirectly targets *PORA*, *PORB*, *haem oxygenase (HO3)*, and *ferrochelatase (FeChII)* genes (Moon et al., 2008; Shin et al., 2009). PIF3 negatively regulates *CAO* (Monte et al., 2004) and two Photosystem I components of PET complexes (*LHCA1* and *PsaE1*) (Shin et al., 2009).

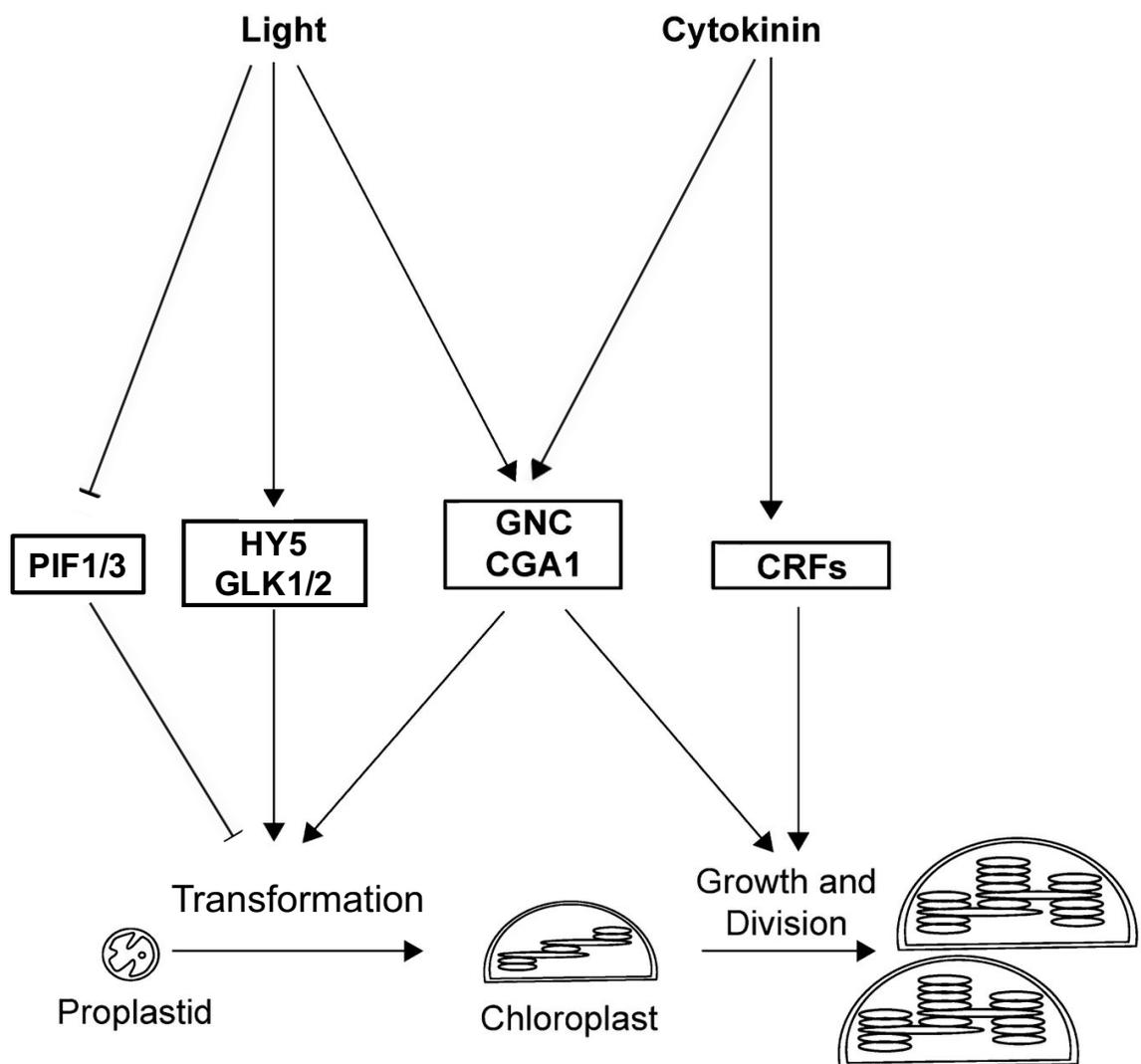


Figure 1. 6 Model for the roles of transcriptional factors in the regulation of chloroplast development

Figure adapted from (Chiang et al., 2012)

1.3.2 Chloroplast-to-nucleus retrograde signalling pathways

Chloroplast biogenesis requires the coordination of “anterograde” build-up and regulation and “retrograde” (from “in reverse”) signalling. Nuclear-encoded plastid

gene expression can be shut down by a negative plastid signalling, when chloroplast biogenesis is inhibited or the chloroplast is damaged by inhibitors or under stress (López-Juez and Pyke, 2005; Waters and Langdale, 2009; Jarvis and López-Juez, 2013). The photobleaching herbicide norflurazon (NF), which is an inhibitor of carotenoid biosynthesis, is commonly used for inhibiting chloroplast function to study retrograde signalling (Nott et al., 2006). Treating plants with NF leads to the loss of expression of hundreds of genes. In WT plants, expression of most photosynthesis-associated nuclear genes (*PhANGs*), such as *LHCB*, *CA1* (*carbonic anhydrase 1*) and *CP12* (chloroplast protein 12), is reduced by treating with NF. Interestingly, the negative effect of retrograde signalling is reduced in a group of *gun* (genome uncoupled) mutants in the presence of NF. The *gun* mutants were termed because of the fact that in them nuclear gene expression is uncoupled from the plastid status and they fail to repress the expression of photosynthesis-associated genes when the plastids are damaged (Susek et al., 1993). In *gun* mutants, either the nucleus fails to respond to negative signalling from the plastid, or negative signalling does not occur, so the *PhANG* expression cannot be consequently shut down. In contrast to WT, loss of GUN function leads to an increased *PhANG* expression with NF treatment. This phenomenon will be demonstrated later (section 1.4.4).

Five *gun* (*gun1~gun5*) mutants were first identified (Susek et al. 1993). GUN2~GUN5 are enzymes involved in different steps of the plastid tetrapyrroles' synthesis pathway (Figure 1.7). GUN2 is the same protein as HY1 (HO1) whereas *HY2* encodes the same protein as *GUN3* (Mochizuki et al., 2010). They are both involved in the haem and PΦB metabolic pathway. *gun2* and *gun3* mutations block the haem catabolic steps resulting in an accumulation of haem. GUN4 and GUN5 are directly involved in regulating Mg-chelatase activity at the first step of chlorophyll biosynthesis pathways. GUN4 acts as an activator increasing Mg-chelatase efficiency (Larkin, 2003; Peter and Grimm, 2009). *GUN5*

encodes CHLH, which is a subunit of Mg-chelatase (Mochizuki et al., 2001). *gun4* and *gun5* block the Mg-protoporphyrin IX biosynthesis in the tetrapyrrole pathway, resulting in reducing the accumulation of Mg-protoporphyrin IX and directing protoporphyrin IX into the haem biosynthesis pathway. More recently, one more *gun* mutant, *gun6-1D*, was identified (Woodson et al., 2011). *GUN6* encodes ferrochelatase 1 (FC1) which is haem synthase (Figure 1.7). In contrast to other *gun* mutants, *gun6-1D* is a dominant mutation that in the presence of NF causes a *gun* phenotype when *GUN6* is overexpressed. *gun6-1D* leads to an overflow of tetrapyrrole biosynthesis towards haem biosynthesis pathway that consequently increases the amount of haem. According to this model, the increased haem level in *gun2~gun6-1D* acts as positive retrograde signalling for chloroplast biogenesis (Woodson et al., 2011). An accumulation of haem in *gun2* and *gun3* and the FC1-derived pool of haem act as a positive signal to increase the *PhANG* expression in the nucleus. Mg-protoporphyrin IX, a mobile signal mediator, was proposed to be the negative regulator. According to the older, alternative model, increasing Mg-protoporphyrin IX represses *PhANG* expression (Strand et al., 2003). However, additional evidence shows that there are difficulties with a model that proposes elevated Mg-proto as a negative signalling molecule leading to the repression of nuclear genes, including that this intermediate is phototoxic. Additionally, there is not always a correlation between the Mg-protoporphyrin IX level and the gene expression level (Moulin et al., 2008). Instead, those authors suggest that the generation of reactive oxygen species (ROS) from phototoxic compounds in the downstream of Mg-protoporphyrin IX biosynthesis pathway is more likely the suppressor of nuclear gene expression.

Although the retrograde signalling mechanism is not clearly understood, it has been proposed that the *PhANG* expression is closely regulated by the activities of positive (HY5 and GLKs) and negative (ABI4) factors (Figure 1.7). The proteolytic cleavage of a chloroplast envelope-bound plant homeodomain (PHD)

transcription factor, PTM (PHD WITH TRANSMEMBRANE DOMAINS), has been found to transfer the signal between chloroplast and nucleus (Sun et al., 2011). The N-terminus of PTM is released by proteolytic cleavage and then imported into the nucleus to activate the *ABI4* (ABSCISIC ACID INSENSITIVE 4) expression. However, a more recent study suggests that seedlings lacking the PTM protein do not show a *gun* mutant phenotype (Page et al., 2017). *ABI4*, which is a nuclear-localized TF, competes with TF HY5 on G-box-binding, in turn, leading to the repression of *LHCB* gene expression (Koussevitzky et al., 2007). Seedlings carrying *abi4* mutation have shown a *gun* phenotypes with NF treatment (Koussevitzky et al., 2007). Additionally, it is proposed that the FC1-derived pool of haem acts as a positive signal of *PhANG* expression by promoting the *GLK1* gene expression (Woodson et al., 2011). Evidence has shown that overexpressing *GLK* induces a *gun* phenotype (Leister and Kleine, 2016).

Unlike GUN2-GUN5, GUN1 is a special GUN protein which works independently from the tetrapyrrole biosynthesis pathway (Mochizuki et al., 2001; Larkin et al., 2003; Koussevitzky et al., 2007; Woodson et al., 2011). The function of GUN1 is still unknown when chloroplasts function normally. Nevertheless, GUN1 has been postulated to act as a signal transducer, sending a signal to the nucleus to repress photosynthesis-associated gene expression in the presence of NF. Strong evidence shows that *gun1* mutation can cause death to a protein import defective mutant, *cue8* (Loudya and López-Juez, personal communication) and equally of *ppi2* (Kakizaki et al., 2012). Without suppressing the nuclear gene expression by GUN1, the nuclear-encoded chloroplast proteins are expressed abundantly and they remain in the cytoplasm. Some of these cytoplasmic free nuclear-encoded chloroplast proteins may be toxic. The free (protein-unbound) tetrapyrroles could also cause ROS in the plastid. In both cases, the *gun1* mutation can lead to cell death.

Although the molecular mechanism by which GUN1 integrates retrograde signalling in the chloroplast is unclear, GUN1 has been proposed by many studies to be a master regulator which integrates multiple chloroplast signals. In contrast to other *gun* mutants, the effects of Lincomycin (LIN) on *LHCB1.2* expression were also suppressed in *gun1* (Gray et al., 2003). LIN is an inhibitor of organellar gene expression that triggers a different type of signalling from chloroplasts. Therefore, GUN1 is expected to act in both signalling pathways. GUN1 is a plastid-localized pentatricopeptide-repeat (PPR) protein which can associate with nucleoids to form transcriptionally active complexes of plastid DNA, RNA and ribosomes (Koussevitzky et al., 2007). This suggests that GUN1 is presumably involved in organellar gene expression. Although the model of GUN1-mediated tetrapyrrole signalling is currently still unclear, the most promising GUN1 signalling mechanism is that it controls plastid gene expression to subsequently affect the tetrapyrrole signalling pathway (Terry and Smith, 2013). The GUN1 function is proposed to be upstream of the haem-regulated signal (Figure 1.7). It may target plastid-transcribed glutamyl-transfer RNA (glutamyl-tRNA) and control the availability of glutamyl-tRNA. Glutamyl-tRNA is the substrate of glutamyl-tRNA reductase which catalyses the initial and rate-limiting step of tetrapyrrole synthesis pathway. Removing GUN1 increases the capacity of the tetrapyrrole synthesis and consequently enhances the haem level. Another possibility is GUN1 restricts the accessing of glutamyl-tRNA to glutamyl-tRNA reductase. In the presence of NF, GUN1 was found to induce ABI4 in order to inhibit *LHCB* expression. Meanwhile, it inhibits *GLK1* expression, in turn, reduces *LHCB* expression (Koussevitzky et al., 2007; Kakizaki et al., 2009). GUN1-mediated plastid signalling suggests that there are at least two different types of retrograde signals. GUN1 has in fact recently been shown to interact chloroplast protein import, protein degradation in the stroma, chloroplast metabolism (photosynthesis and tetrapyrrole biosynthesis) and chloroplast unfolded protein

response process as a kind of “jack of all trades” (Colombo et al., 2016; Tadini et al., 2016).

Tetrapyrrole signal regulating nuclear gene expression is a primary plastid signal during chloroplast biogenesis. Probably, tetrapyrrole is one of the most ancient prosthetic groups which have been found. Tetrapyrrole compounds contain four pyrrole rings and exist in various forms in the organelle, including light absorption (chlorophyll), electron transfer (sirohaem), oxygen binding (haem) (Mochizuki et al., 2010). They act as cofactors that bind to their apoproteins and are transferred throughout all cellular compartments. Chlorophyll, Pchl_{id}, haem, P Φ B and sirohaem are the main tetrapyrrole products synthesized in plastids (Mochizuki et al., 2010). Some tetrapyrroles can cause photo-oxidative damage, and cell death through generated free radicals and ROS, primarily singlet oxygen (Mochizuki et al., 2010). In this case, the accumulation of tetrapyrroles in Mg-protoporphyrinIX biosynthesis pathway acts as a negative retrograde signalling (Terry and Smith, 2013). Pchl_{id}s, tetrapyrrole compounds in Mg-protoporphyrinIX biosynthesis pathway, are essential cofactors and pigment precursors in photosynthetic organisms (Schlicke et al., 2015). The tetrapyrrole character of causing photooxidative damage can be visualized through the “FR light block of greening” experiment (Barnes et al., 1996; McCormac and Terry, 2002). This experiment is carried out by treating seedlings under a continuous FR light, then transferring to white light (Barnes et al., 1996). phyA, a unique subtype of phytochrome, is different from other members as it triggers de-etiolation in FR light rich environments (Rausenberger et al., 2011; Debrieux et al., 2013). It has been determined as the only photoreceptor responsible for controlling the FR-HIR (Reed et al., 1994; Rausenberger et al., 2011). In the dark, Pchl_{id} is synthesised and accumulated in the plastid. Simultaneously, the light-dependent enzyme NADPH:protochlorophyllide oxidoreductase A (PORA) which is positively regulated by DELLAs, also accumulates in the dark and binds to

Pchl_a to catalyse the photo-conversion process after the seedlings are exposed to the light (Cheminant et al., 2011). In general, the accumulated levels of Pchl_a and PORA within seedlings are comparable. All the Pchl_a should be rapidly converted into chlorophyll during initial light irradiation. In contrast, FR-HIR causes degradation of PORA by nuclear-encoded proteases and inhibits POR mRNA expression (APEL, 1981; Santel and Apel, 1981; Paul, 2013). The mechanism of light-activation of protease expression or activity is unknown, although one can speculate that it could be related to the other forms of protein degradation which control photomorphogenesis. The free Pchl_a (not bound to PORA) operates as a photosensitizer upon light exposure to produce ROS, thereby causing photo-oxidative stress (op den Camp et al., 2003; Mochizuki et al., 2010). Photosensitizing molecules (in an excited state) can generate singlet oxygen using light energy to excite triplet molecular oxygen (ground state) (Gollnick, 1968). Pchl_a is accumulated abundantly during continuous FR treatment, without PORA. After seedlings are exposed to white light, the excess accumulated Pchl_a act as a signal to activate two stress responses, growth inhibition and cell death, inhibiting the light-dependent signal by repressing photosynthetic and tetrapyrrole synthetic genes rather than acting as a primary toxin (op den Camp et al., 2003; Wagner et al., 2004). Because singlet oxygen has a short half-life, the molecules that are affected need to be close to where single oxygen is synthesized (Gorman and Rodgers, 1992). Currently, how single oxygen causes the stress responses is not fully understood. It has been speculated that ROS oxidized lipid-related metabolites, such as β -cyclocitral, may be involved in signal transfer (Ramel et al., 2012). Nevertheless, inactivation of EXECUTER 1 and EXECUTER 2 have been found to abrogate stress responses (Lee et al., 2007). *executer 1* can suppress the *flu* mutation which leads to the over-accumulation of free Pchl_a during the light/dark shift, resulting in a WT phenotype of double mutant (op den Camp et al., 2003; Wagner et al., 2004).

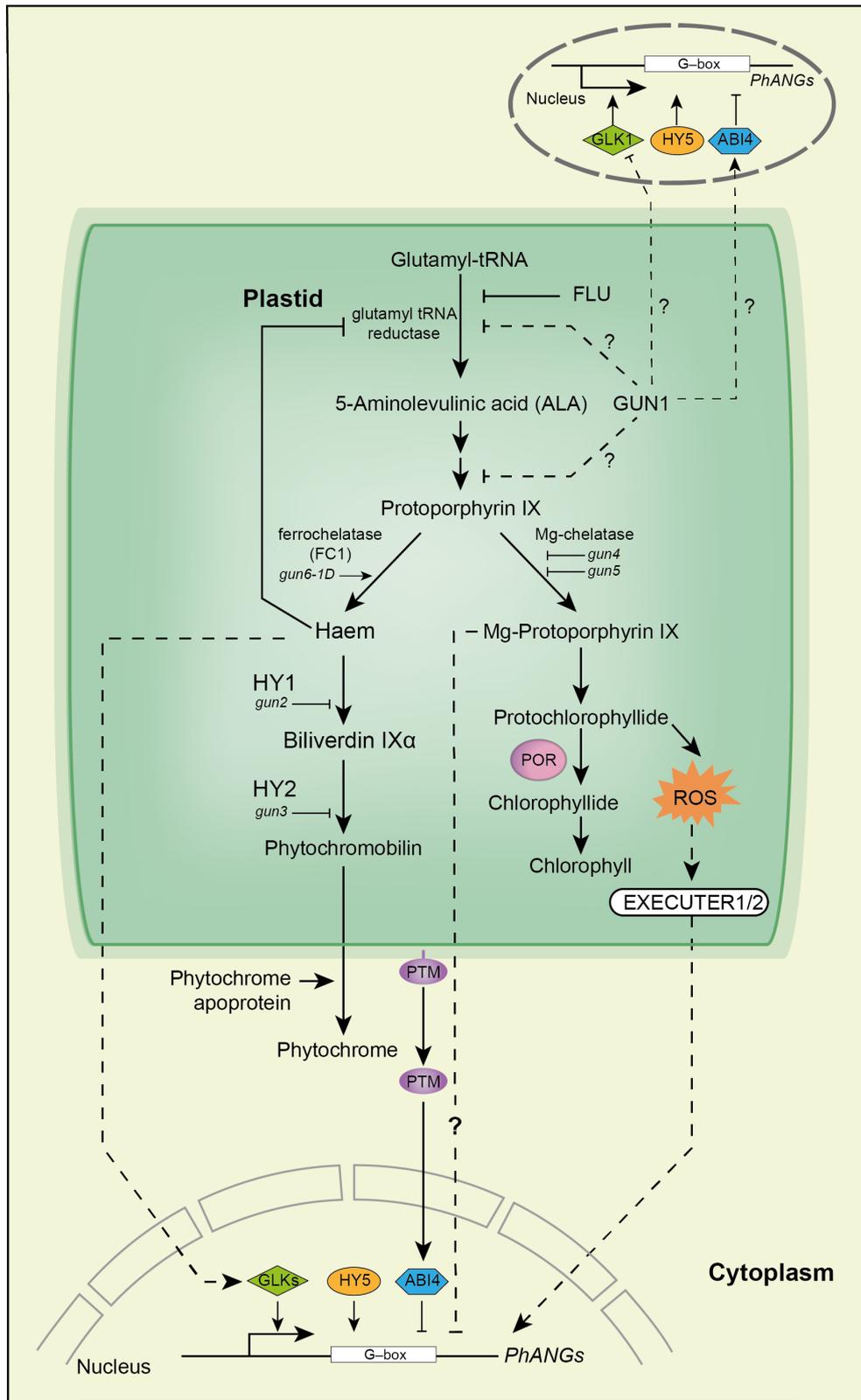


Figure 1. 7 Summary of chloroplast-to-nucleus retrograde signalling pathways, with a focus on tetrapyrrole synthesis and signals. Note that GUN1 is also involved in monitoring other chloroplasts activities, including translation and protein import.

1.4 *Arabidopsis lyn1* mutant

1.4.1 Generating *lyn1* mutation by genetic screens

I am interested in how light regulates the conversion of etioplasts to chloroplasts in order to understand the process of chloroplast development, so a genetic screening was carried out on a light-response defective mutant, *hy1*. *HY1*-null line *hy1*-100 (hereafter *hy1*) carries a splicing mutation in Columbia (Col) ecotype with a single base pair change from G to A at the 3' boundary of the first intron (Davis et al., 1999; Muramoto et al., 1999). In *Arabidopsis*, *HY1* locus encodes the haem oxygenase 1 (HO1) which is involved in the enzymatic cascade reaction of phytochrome chromophore biosynthesis (Parks and Quail, 1991). Many previous studies have found that the genetic mutations of *hy1* can result in etiolated seedlings, as the mutations cause deficiency in both PΦB production and Pchlide production. Loss of haem oxygenase in *hy1* mutant causes a blocked PΦB biosynthesis pathway, resulting in increased accumulation of haem in plastids. The accumulation of haem acts as a tetrapyrrole signal repressing ALA biosynthesis by inhibiting glutamyl-tRNA reductase (*HEMA1/2*) expression in the nucleus through retrograde signalling, which in turn reduces the Pchlide production in the chlorophyll biosynthesis pathway (Terry and Kendrick, 1999). Phytochrome biosynthesis-defective plants respond poorly to light signals, so the chloroplasts cannot fully develop in *hy1* and the phenotype of *hy1* resembles in part that of dark-grown seedlings (Figure 1.7).

Previous researchers in my laboratory used mutagen ethyl methanesulfonate (EMS) to randomly generate point mutations on *hy1* mutants. A reporter, green fluorescent protein, which is driven by the promoter of *LHCB* gene was introduced into seedlings (López-Juez and Hills, 2011). Because *LHCB* is a light-induced gene, the expression level of this gene indicates the response level to the light signal (López-Juez and Pyke, 2005; Lopez-Juez, 2007). Thus, *LHCB* gene

expression is lower in *hy1* compared to WT. Surprisingly, they found that the reporter line, after mutagenesis, led to isolation of the *lyn1 hy1* mutant which seems to have a *LHCB* gene expression level intermediate between *hy1* and the WT (Hills, 2002). Thereby, they generated a novel mutant, *lyn1*, which can rescue the expression of nuclear-encoded plastid protein LHCB. They further studied the plant phenotypes. The double mutants showed a phenotype intermediate between *hy1* and WT. *lyn1* single mutant was obtained through backcross (Shindo, Didcock, López-Juez, unpublished).

1.4.2 LYN1 is not a light-inactivated repressor of photomorphogenesis

At the beginning, LYN1 was suspected to work as the central negative regulators of photomorphogenesis, COP/DET/FUS group proteins, which were identified due to their mutant seedling phenotype in darkness. When grown in the dark, these mutants present some light-grown seedling features, typically with a short hypocotyl, expanded cotyledons and enhanced levels of photosynthetic gene expression. They do not show complete chloroplast development in the dark because chlorophyll synthesis requires light, and photosystems cannot assemble without chlorophyll (Waters and Langdale, 2009). As mentioned above, DET1 is a nuclear-localized repressor of photomorphogenesis. *det1* was grown together with *lyn1* to compare their phenotypes in different light conditions. When grown in the light, *det1* early seedling phenotype is not very different compared with WT (Figure 1.7 A). Several transcriptional factors are activated to cause photomorphogenesis, because DET1 activity is repressed by photoreceptors in the WT. *det1* mutant is defective on DET1 protein production, hence it cannot repress transcription factors in the dark. Thus, *det1* seedlings display light-grown features even in the dark (Figure 1.7 B). 6-day light-grown *lyn1* has the same short hypocotyl length as *det1* (Figure 1.7 A), but 6-day dark-grown *lyn1* seedlings have longer hypocotyl length than *det1* and no different from the WT

(Figure 1.7 B). Therefore, LYN1 is not a light-inactivated repressor of photomorphogenesis.

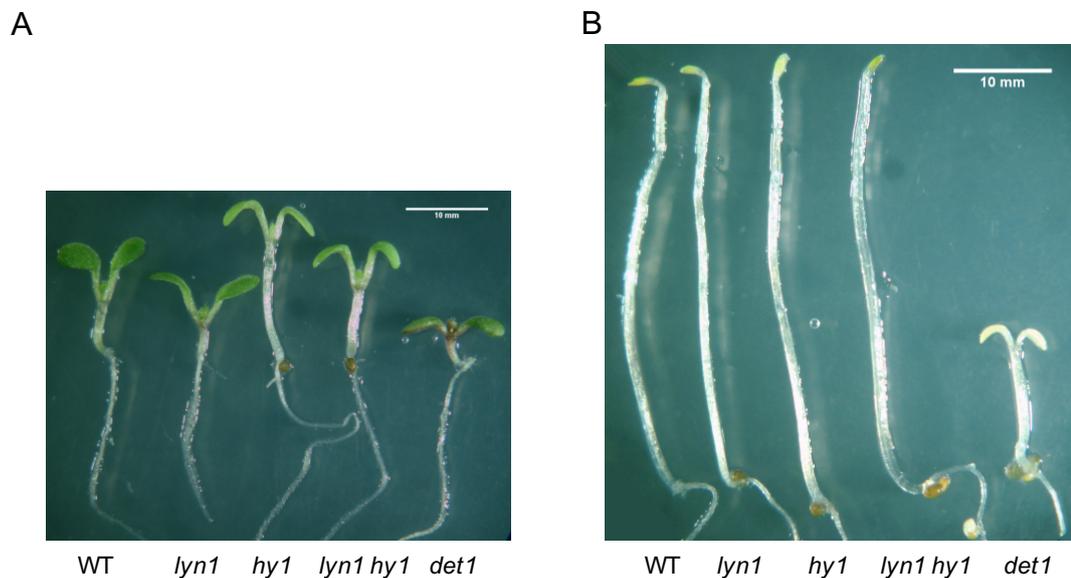


Figure 1.8 *lyn1* effect of phenotypes of *hy1* in light and dark

Seedlings of WT, *lyn1*, *hy1*, *lyn1 hy1* and *det1* grown in white light (100 μmol m⁻² s⁻¹) and dark for 6 days in 1% MS media. Scale bar = 10 mm. Figure from (He, 2013).

1.4.3 *lyn1* suppression of the growth and greening of seedlings

The dramatic differences between four phenotypes remain visible at different stages of development, on both early seedlings and later adult plants, to the naked eye (Figure 1.8). Therefore, the suppression of seedling phenotypes was first time studied on hypocotyl elongation and chlorophyll production at different light fluence rates on 6-day old light-grown seedlings by measuring the hypocotyl length and chlorophyll content by a previous MSc student (Tan, 2012). This experiment was also repeated on 5-day old seedlings during my MSc research. Both results had the same outcome which is that *lyn1* suppresses hypocotyl length and chlorophyll content on *hy1*, leading to a shorter hypocotyl length and a higher chlorophyll content of *lyn1 hy1*. *lyn1* single mutant causes an increased light response, leading to a shorter hypocotyl length compared to the WT. This

assay was repeated and shown as a part of a wider assay in this thesis (Figure 3.6 and 3.7).



Figure 1.9 *lyn1* suppression of the growth and greening of 20-day plants
Seedlings of WT, *lyn1*, *hy1* and *lyn1 hy1* grown in white light ($180 \mu\text{mol m}^{-2} \text{s}^{-1}$) for 20 days. Figure from (Shindo, 2005).

lyn1 suppression of Landsberg (Ler) ecotype *hy1-1* was also previously studied during my MSc research to address whether *lyn1* is allele-specific or also applied to other *hy1* alleles. *hy1-1*, generated by fast-neutron mutagenesis, carries a 13-bp deletion at the 3'boundary of the second exon (Davis et al., 1999). This deletion creates a frameshift during splicing and the translated protein becomes a truncated protein which misses 71 aa at the C-terminal end. According to the studies, *lyn1* showed a similar suppression of *hy1-1* (Figure 1.9) (He, 2013). Therefore, *lyn1* nonspecifically suppresses *hy1* mutations.

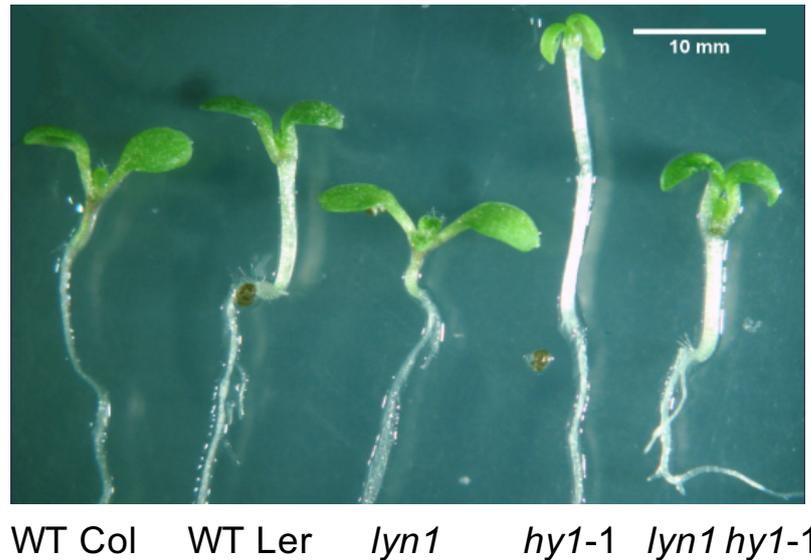


Figure 1.10 *lyn1* suppression of *hy1-1*

Seedlings of WT Col, WT Ler, *lyn1*, *hy1-1* and *lyn1 hy1-1* grown in white light ($100 \mu\text{mol m}^{-2} \text{s}^{-1}$) for 6 days in 1% MS media. Scale bar = 10 mm. Figure from (He, 2013).

1.4.4 *lyn1* eliminates the *gun* phenotype of *hy1*

As mentioned in the introduction, *hy1* is also defined as *gun2* mutant. Thus, it showed a *gun* phenotype which has high *PhANG* expression compared to the WT when the *hy1* seedlings were treated with NF. However, previous researchers found that *PhANG* expression was reduced in *lyn1 hy1* compared to that of *hy1* in the presence of NF (Tan, 2012) (Figure 1.10). It suggests that *lyn1* may reverse the *gun* phenotype of the *hy1* mutant by regulating the tetrapyrrole biosynthesis pathway.

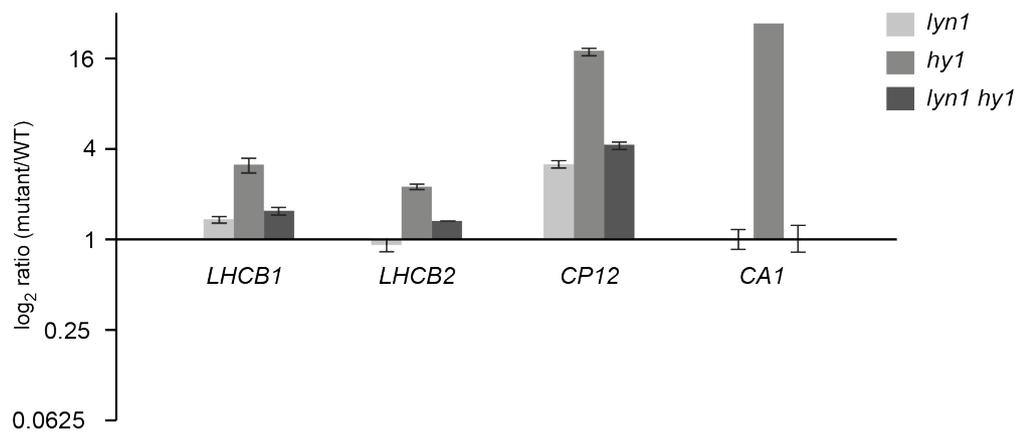


Figure 1.11 Quantification of the expression of light-induced photosynthetic genes in the presence of NF

Expression level in each mutant expressed as a ratio to the level in the WT grown under the white light ($200 \mu\text{mol m}^{-2} \text{s}^{-1}$) in the presence of NF. Figure adapted from (Tan, 2012).

1.4.5 *lyn1* does not rescue haem oxygenase

Because of the observation of *lyn1* suppressing the gun phenotype, a hypothesis was that *lyn1* decreases the haem accumulation by increasing the conversion from haem to biliverdin IXa, enhancing the production of phytochromobilin, and therefore of phytochromes. The reduced haem accumulation leads to the decreased retrograde signalling, in turn, reduce the *PhANG* expression. In this case, cytoplasmic located phytochrome apoproteins should decrease because they are used to synthesis phytochromes by binding to PΦB. Therefore, the previous researchers in my laboratory used the immuno-blotting to assess the phyA apoprotein decay (Shindo, 2005). They carried out this by exposing the 5-day dark grown seedling in R light and observing the presence of phyA apoprotein over time. Without the R light irradiation, the phyA apoproteins were present in all 4 genotypes of 5-day dark grown seedlings (Figure 1.11). Once the seedlings were exposed to R light, the phyA apoproteins immediately disappeared in the WT and *lyn1*. This is because the phyA apoproteins bind to PΦB to synthesise phyA (becoming holoprotein) in WT and *lyn1*. Once the seedlings are exposed to

the R light, phyA are rapidly degraded. In contrast, the phyA apoproteins were constantly present in *hy1* and *lyn1 hy1*. Because of the absence of PΦB in *hy1* and *lyn1 hy1*, phyA apoproteins remained in the cytoplasm without forming phyA holoprotein. This result indicates that *lyn1* does not suppress *hy1* by restoring PΦB, at least to a detectable extent.

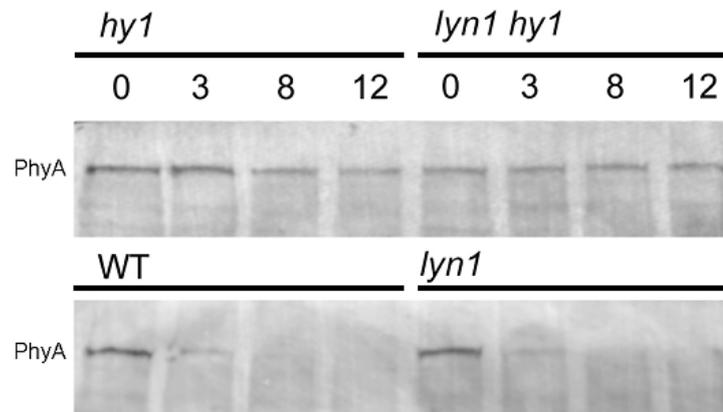


Figure 1.12 Immuno-blotting of phyA apoprotein decay assay

Total protein was extracted from 5-day old seedlings of WT, *hy1*, *lyn1* and *lyn1 hy1* grown in the dark or dark followed by 3h, 8h or 12h of red light. Figure from (Shindo, 2005)

1.4.6 Mapping and cloning of the *LYN1* gene

A previous investigator (Didcock and López-Juez, unpublished) mapped the mutation narrowly, using mutant plants from a recombinant (Col x Ler, i.e. *lyn1 hy1-100* x *hy1-1*) population. I continued the mapping as part of my MSc research. This region, identified to contain *lyn1*, was reassessed by looking at the critical recombinant mutant plants at a number of nearby polymorphic sequences. Localization of *LYN1* was carried out by high throughput sequencing between *lyn1 hy1* and the reference Col sequence (Aubry and Hibberd, personal communication). Among 200,000 polymorphisms identified, only one fell within a coding sequence within our narrowly-mapped region: a single nucleotide change was found at position 11,012,014 (as defined in the TAIR10 sequence) on *JMJ14* (At20400) gene (Figure 1.12). Further Sanger sequencing showed that the single

nucleotide changed from G to A. The change was consistent with the role of the mutagen ethyl methanesulfonate (EMS), which is responsible for changing G to A or C to T. A nonsense mutation is induced by the *lyn1* mutation to change tryptophan into a stop codon, resulting in early termination of protein translation (Figure 1.13).

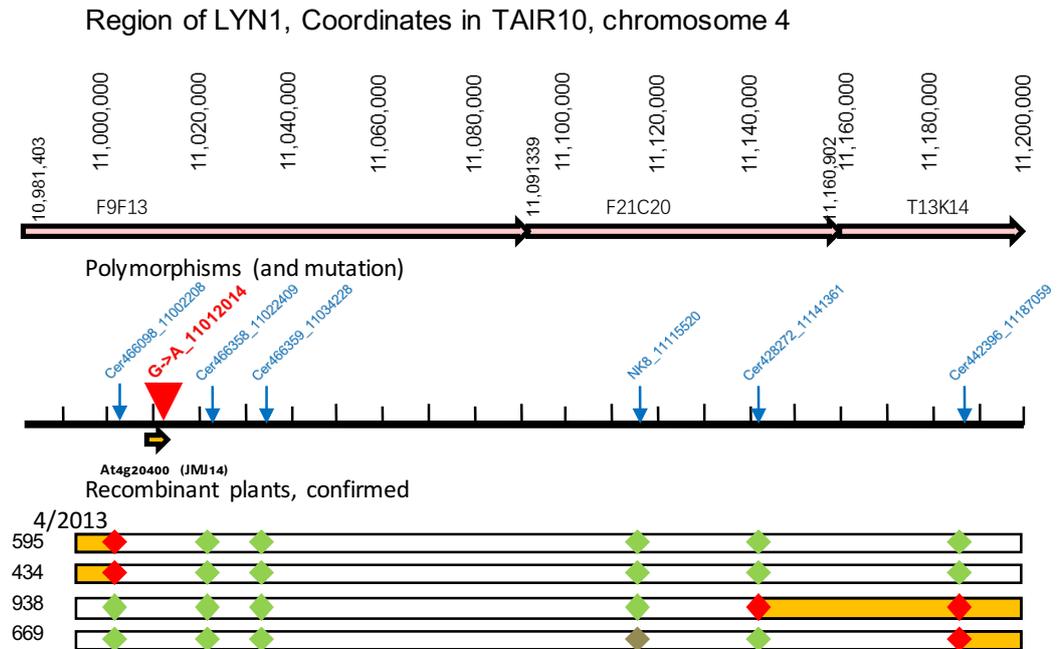


Figure 1.13 The region of *LYN1* on chromosome 4 (coordinates in TAIR10)

The red point signifies at least one *lyn1* recombinant plant is heterozygous (Col and Ler) at this polymorphism point. The green points show all recombinant plants carry the Col allele. Polymorphisms are shown in blue. *lyn1* mutation is shown within a red arrow. The marker NK8 was unconfirmed in plant 669, so it is shown in grey. The recombinant regions are shown in yellow. The Col regions are shown in white. For example, for recombinant plant 938 markers cer466098, cer466358, cer466359 and NK8 show a Col genotype and markers cer428272 and cer442396 are heterozygous, this means from this point to the right they are also expected to be heterozygous. The recombination point in one gamete which gave rise to this plant must have occurred between the markers NK8 and cer428272. *lyn1* mutation is found to change nucleotide from G to A at position 11,012,014 on protein JM14 (At4g20400). Figure from (He, 2013).

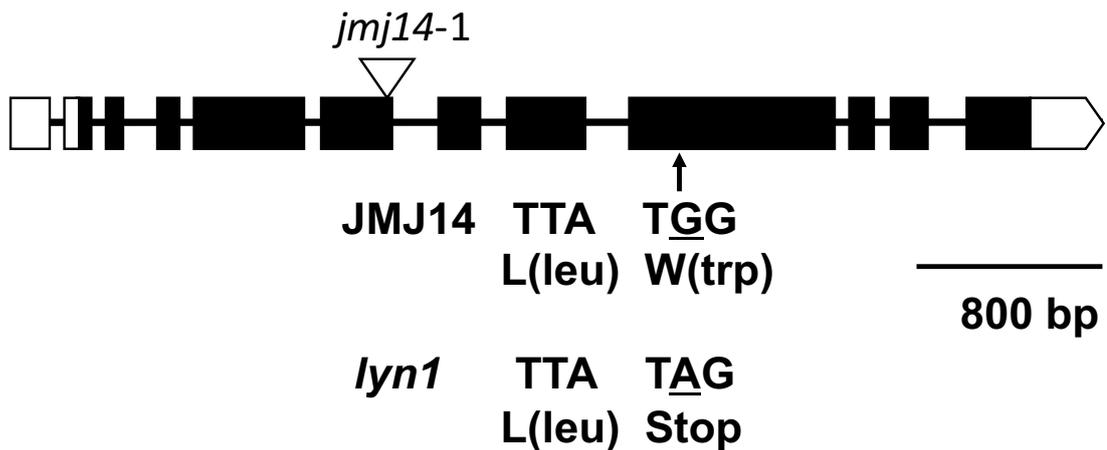


Figure 1.14 *JMJ14* gene structure and the location of the *lyn1* and *jmj14-1* mutations. *jmj14-1* is a Salk line mutation obtained from the stock centre and later used in *lyn1* identity experiments (see chapter 3). Figure from (He, 2013).

1.4.7 Current published studies about the role of JMJ14.

JMJ14 is a histone H3 lysine 4 (H3K4) demethylase, participating in the modulation of histone methylation as part of the control of epigenomic regulation (Jeong et al., 2009; Lu et al., 2010; Yang et al., 2010). Methyl groups, which are one type of epigenetic marks, can be added to histone protein lysine residues by histone lysine methyl-transferases (HKMTs) and removed by histone demethylases (HDMs) to remodel chromatin in order to maintain cell fate and genomic stability (He et al., 2011; Black et al., 2012). Epigenetic marks do not affect DNA-sequence, but are attached onto DNA or onto DNA-associated proteins. The reason to refer to this mode of regulation of gene expression, these marks, as “epigenetic” is that these marks are inherited, they can be passed down to the next generation of cells, through mitosis, marks acquired even during embryogenesis can continue to be present for the life of the organism. In some occasions this mode of regulation has been found to be inherited even by the progeny, the next generation of organisms (in a true genetic sense) (Trerotola et al., 2015). In eukaryotic cells, chromatin is formed by packing DNA around nuclear histone protein, forming nucleosomes. A nucleosome comprises four

core histones: H2A, H2B, H3 and H4. Each core histone consists of a H3-H4 tetramer and two H2A-H2B heterodimers that bind together to form a protein octamer (Eickbush and Moudrianakis, 1978). The N-terminal tails residues of the histones can be altered by post-translational covalent modifications (PTMs) in order to influence numerous biological processes. PTMs exist in various forms, including methylation, acetylation, phosphorylation, ubiquitination and others (Fisher and Franklin, 2011; He et al., 2011). Histone methylation modifies chromatin structure by adding methyl groups on lysine and arginine residues of the histone N-terminal tails. Lysine methylation on histone 3 has been found at least in five residues (H3K4, H3K9, H3K27, H3K36 and H3K79) and on histone 4 has been found at least in one single residue (H4K20) (Martin and Zhang, 2005). According to the position, the effects of histone methylation can be either gene activation or repression. The study of several histone lysine methylation states in the *Arabidopsis* genome has demonstrated that H3K4, H3K9 and H3K36 methylation show a strong positive correlation with gene expression levels in general, whereas H3K27 methylation shows a strong negative correlation with gene expression levels (Roudier et al., 2011). A lysine residue can be mono-, di- or trimethylated to cause differential biological effects of methylation (Martin and Zhang, 2005). For example, the position of histone 3 lysine 4 can be mono- (H3K4me1), di- (H3K4me2) or trimethylated (H3K4me3) (Zhang et al., 2009). Histone acetylation is another intensely studied epigenetic marker. It also occurs on lysine residues of the histone N-terminal tails (Verdone et al., 2005). Acetyl groups are established on the lysine residue by histone acetyltransferases (HATs) and removed by histone deacetylases (HDACs) (He et al., 2011). In *Arabidopsis*, 12 HAT and 18 HDAC genes have been identified (Pandey et al., 2002). Histone acetylation is generally associated with active gene expression. It can regulate transcription by recruiting transcriptional activators or neutralizing the positive charges of histone proteins to relax and open promoter regions for facilitating transcription (Graham and Karen, 2016). Histone phosphorylation can occur on

serine, threonine and tyrosine residues and it has been found to be involved in DNA damage repair, transcription regulation and chromatin compaction (Rossetto et al., 2012). Ubiquitination of histone is a modification not correlated with degradation. It occurs in the largely mono-ubiquitinated form on lysine residues. Monoubiquitination of histones H2A and H2B associates with transcription regulation also by regulating the assembly of transcription complexes on the promoter (Zhang, 2003). H2A monoubiquitination is generally associated with transcription repression while H2B monoubiquitination is involved in transcription activation (Feng and Shen, 2014).

Histone H3K4 methylation is mediated by Trithorax (Trx) group proteins. Trx group proteins have an evolutionarily conserved SET-domain, which is to establish histone methylation marks on lysine residues. In *Arabidopsis*, ARABIDOPSIS TRITHORAX 1 (ATX1) and ARABIDOPSIS TRITHORAX 2 (ATX2) are responsible for H3K4me3 and H3K4me2 methylation, respectively (Saleh et al., 2008). Additionally, other SET-domain genes (SDG) such as SDG2 and SDG4 also have function on mediating H3K4 methylation (Cartagena et al., 2008; Berr et al., 2010; Guo et al., 2010). Because of the crucial role of histone lysine methylation in gene transcription, organisms have to carefully and dynamically control and balance the methylation state by histone lysine demethylation (Shilatifard, 2008). Until recently, the discovered demethylase enzymes have been classified into two large groups. LSD1 (Lysine specific demethylase 1) is one class which requires action of a cofactor, Flavin Adenine Dinucleotide (FAD), to directly reverse histone H3K4 modification by an oxidative demethylation reaction (Klose et al., 2006; Forneris et al., 2007). It can only demethylate mono- and dimethylated H3K4 but not trimethylated H3K4 (Shi et al., 2004). Another class and also the largest of demethylase enzymes is Jumonji C (JmjC)-domain-containing histone demethylases (JHDMS). It catalyses lysine demethylation by oxidative reaction that requires two cofactors, α -ketoglutarate

(α KG) and iron Fe(II). In addition to the methyl group removing function of LSD1, JHDMS can also demethylate the trimethylated histone lysine-methylation states (Klose et al., 2006; Lu et al., 2008). JMJ14 has been found as one of the Jumonji C (JmjC) domain-containing demethylases which are considered as the major histone demethylases in eukaryotic cells (Jeong et al., 2009; He et al., 2011).

Almost two-thirds of *Arabidopsis* genes contain at least one type of H3K4me. The study of H3K4me distribution shows that three types of H3K4me are frequently present at the different positions of transcribed regions (Figure 1.14). H3K4me1 is present in relatively low levels at both 5' and 3' end of transcribed regions, and enriched in the internal region. The distribution of H3K4me2 and H3K4me3 are quite similar. They are both enriched in the 5' end and promoter (about 200-bp regions upstream of transcription start sites). H3K4me3 is present slightly closer to the 5' end and it occurs at a higher level than H3K4me2 (Zhang et al., 2009; He et al., 2011). The highly enriched H3K4me2/3 at the promoter and the beginning of the coding region suggests that they are very likely to play a functional role in initiation of transcription. Some studies have shown that trimethylation of H3K4 opens the binding region of DNA that helps to facilitate recruitment of RNA polymerase II (Martin and Zhang, 2005; Black et al., 2013). H3K4 methylation is not only associated with gene expression but also with gene splicing. As shown in Figure 1.15 A, H3K4me3 is enriched particularly in the 5' untranslated region and the beginning of the coding region, both of which are close to the promoter. It has also been found to influence splicing factor recruitment that determines transcript composition (Black et al., 2012). This observation supports that H3K4 is involved in regulation of gene expression by different methylation marks at different regions of genes. Especially, H3K4me2/3 is very likely to play a role in transcriptional regulation. As several currently published studies have suggested that JMJ14-regulated genes have an enriched H3K4me2/3 in the 5' end and promoter regions, and JMJ14 is more efficient in

carrying out H3K4 di- and tri-demethylation (Jeong et al., 2009; Lu et al., 2010), I suspected the *jmj14* mutation to rescue light response through regulating the H3K4me2/3 marking state in order to mediate gene transcription level. According to this model, some of these *lyn1*-regulated genes would be involved in chloroplast development.

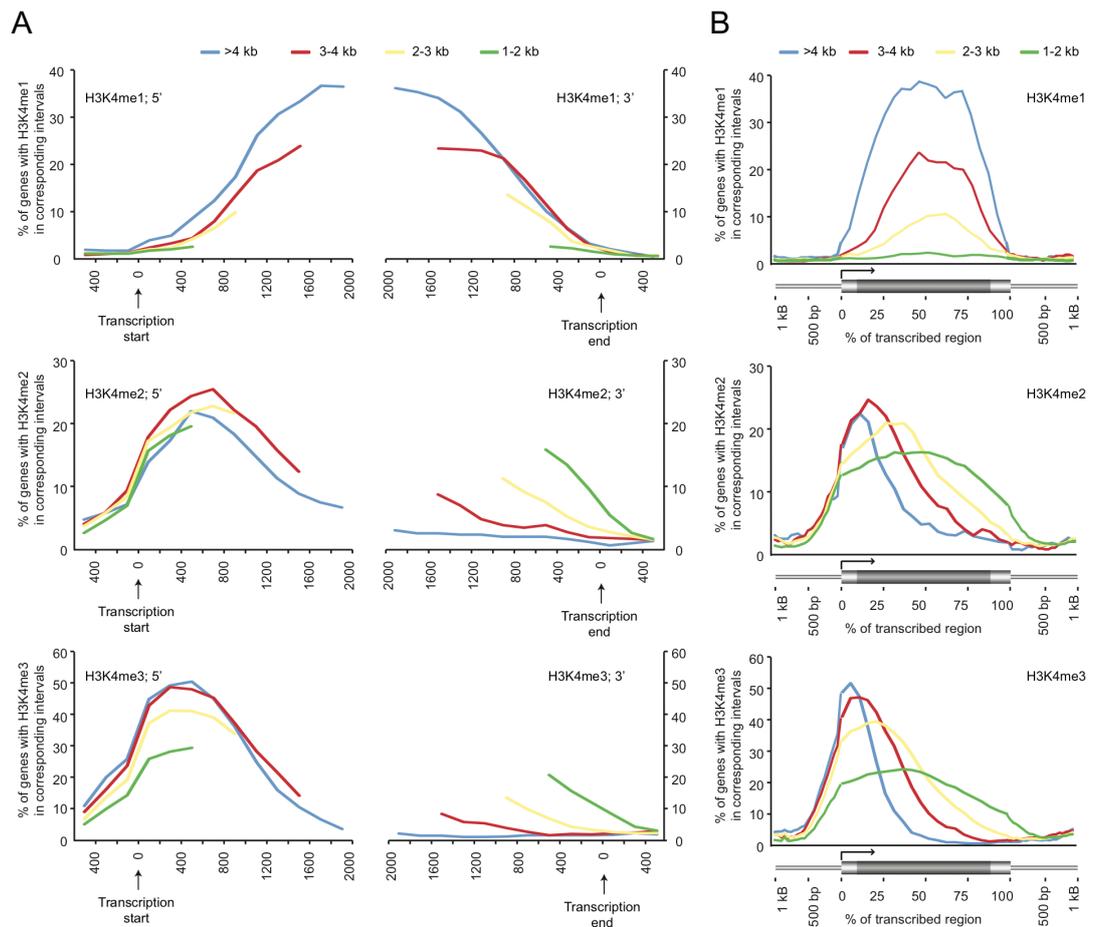


Figure 1.15 Distribution of H3K4me within genes

A. Distribution of H3K4me at the 5' and 3' ends of genes. Because the position of the mark in relation to the transcription start and termination sites is greatly dependent on the length of the transcript, genes were categorised into 4 groups according to their transcript length and shown in different colour as indicated at top. y-axis represents the percentage of genes containing H3K4me in their promoters or 5' ends and 3' ends. x-axis represents the distance between DNA

position and gene transcription start or end site.

B. Distribution of H3K4me across genes. Gene transcribed regions are represented in a thick horizontal bar along the x-axis in a percentage scale (to normalise for gene length). The 1- kb regions upstream and downstream of each gene are represented in a thin horizontal bar.

A and B demonstrate that the peaks of both di- and tri-methylation occur at the transcription start sites, both marks extend in both directions, being present at the end of the promoters, and within the coding region. Figure adapted from (Zhang et al., 2009)

1.4.7 Statement of authorship

As this project continues the work from my MSc research, a number of previous results are shown in this thesis. This paragraph clarifies all the assays obtained from my MSc work before the start of this thesis. *lyn1* suppression of the growth and greening assay was carried out during my MSc in 5 days dark and light grown seedlings before the *LYN1* gene location was identified. This experiment was repeated during my PhD study including also *jmj14* mutants, in order to confirm the *LYN1* gene identity (shown in section 3.2.1.2). *LYN1* was demonstrated not to be a light-inactivated repressor of photomorphogenesis by previous researchers in the lab. This assay was repeated during my MSc research and the result of this assay is shown in section 1.4.2. To confirm that *lyn1* does not only suppress *hy1-100* but also other phytochrome biosynthetic mutations (*hy1-1* and *hy2-1*), the double mutants were generated by previous researchers, but no further work had been carried out on them. The *lyn1* suppression of the growth and greening of other phytochrome biosynthetic mutations was carried out during my MSc research, and is shown in sections 1.4.3 and 4.1.3, respectively. The Pchlide level was quantified during my MSc by extracting Pchlide from 5 day-dark grown seedlings (Appendix Figure 7.1). The Pchlide accumulation was analysed through an alternative physiological strategy (Far-red light block greening) in this

thesis to support the results from my MSc. The *LYN1* gene mapping was partially carried out by previous MSc students. It was completed during my MSc research (shown in section 1.4.6). The chloroplast compartment quantitation assay was performed in mesophyll cells of rosette leaves and cauline leaves during my MSc research (Appendix Figure 7.2). The same assay was repeated in mesophyll cells and bundle sheath cells of cauline leaves (shown in section 4.2.1).

Unless otherwise indicated, all other assays carried out after my MSc and described in this thesis were carried out by myself. The exceptions are: the detection of presence and quantitation of phytychromobilin-containing photochemically active phytochrome, the RNA-Seq assay of global gene expression (the library preparation and Illumina sequencing itself), and the principal component analysis (PCA) of those data comparing the different RNA-Seq samples. Detecting of presence of phytychromobilin-containing photochemically active phytochrome was performed by the laboratory of Tim Kunkel, University of Freiburg. The sequencing was performed by the laboratory of Prof. Beemster's, University of Antwerp. PCA was performed by a bioinformatics researcher at the University of Antwerp.

1.5 Hypothesis

According to the studies which preceded this project, there are two preliminary hypotheses which could explain how the *lyn1* mutation increases light response. The first hypothesis is that the *lyn1* suppression is caused by direct amplification of the weak light response. An alternative hypothesis is that *lyn1* primarily elevates etioplast development and subsequently enhances chloroplast development, resulting in rescuing a small amount of PΦB to synthesise phytychromes, and as a result the light response. This compensated light response further improves the development of chloroplasts to, in turn, increase PΦB synthesis and allow perception of an increased light signal. Eventually, this

feedback generates a “virtuous cycle” that leads to a significantly rescued light response in the *lyn1* mutant (Figure 1.14). These two hypotheses would lead to two contrasting potential applications of the *lyn1* mutation, enhancing light responses or enhancing plastid development. How is the light response rescued by the *lyn1* mutation? This question was addressed experimentally in this thesis.

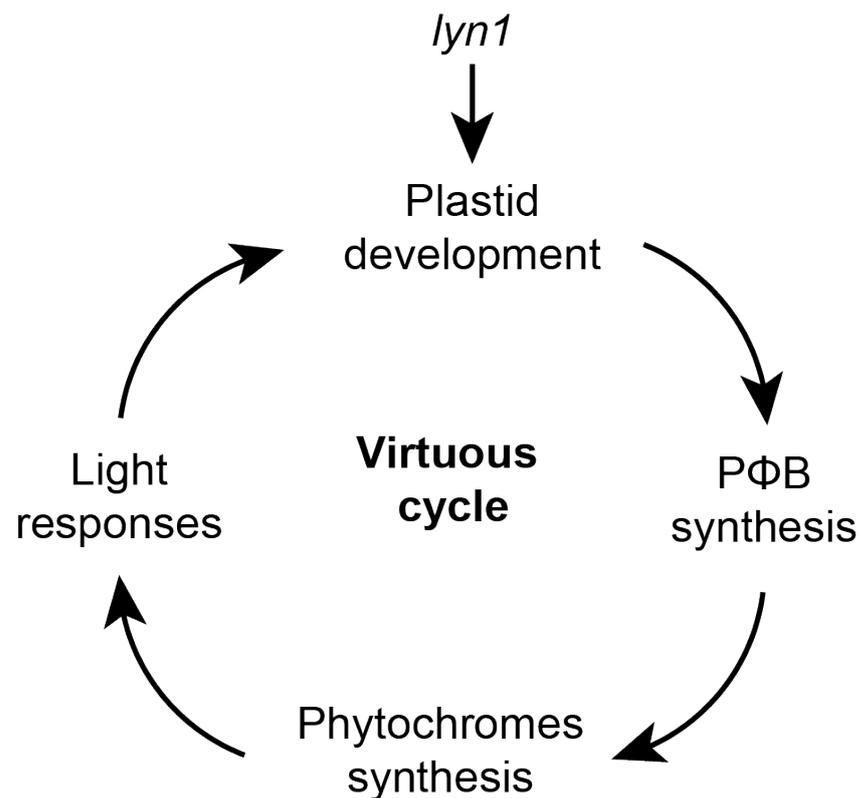


Figure 1. 16 A schematic diagram showing the hypothesis of *lyn1* rescuing light responses in a virtuous cycle.

1.6 Aims

1. Confirming the *LYN1* gene identity and analysing *JMJ14* evolution

After *lyn1* location was identified to occur on *JMJ14*, it was important to confirm the gene identity to prove that this was the causative mutation. The *lyn1* location was first to be confirmed by comparing the growth and greening between *lyn1* and *jmj14-1* (containing a T-DNA KO lesion on *JMJ14*). *lyn1* and *jmj14-1* mutants were crossed to see if they failed to complement each other. Finally, the

functional *JMJ14* gene was transferred to the *lyn1* mutant to see if *JMJ14* can complement the *LYN1* function. Once the *LYN1* gene was confirmed, I assessed *LYN1* presence and distribution in other species.

2. Analysing the *lyn1* impact of chloroplast development including chloroplast compartment in different leaf cells and the capacity of tetrapyrrole synthesis

The enhanced chloroplast development needed to be assessed by observing and comparing the chloroplasts between different genotypes in different leaf cells under the microscope. If *LYN1* and *JMJ14* are same gene, chloroplast development was also expected to be enhanced in *jmj14*. The data of this analysis, if confirmed, would later allow me to make use of the mutation in biotechnological applications, for example to boost chloroplasts in specific cells. By conducting a variety of experiments, I addressed whether enhanced chloroplast development is the primary effect of *jmj14*, or the secondary effect of an enhanced light response.

3. Analysing the *lyn1* impact on global gene expression

To ultimately find out how the *LYN1* gene acts (how the mutation rescues *hy1*), I lastly made use of global gene expression analysis to identify which biological processes, or which specific genes, are affected by *lyn1*.

Chapter 2 Materials and Methods

2.1 Plant materials

Arabidopsis thaliana wild type ecotypes Columbia (Col) and Landsberg erecta (Ler) and mutant *det1* (Col) were obtained from Joanne Chory (Salk institute, San Diego, California, USA). *phyA*-211, *phyB*-9 and *cry1*-304 (Col) were obtained from Christian Fankhauser (University of Geneva, Geneva, Switzerland). *hy1*-100 (Col), *hy1*-1 (Ler) and *jmj14*-1 (Col) were obtained from the *Arabidopsis* stock centre (Nottingham, UK). *hy2*-105 (Col) were obtained from Nobuyoshi Mochizuki (Kyoto University, Japan). Lines containing the *Lhcb::Ω::GFP* reporter transgene (LOS5) *lyn1 hy1* and *lyn1* mutants (Columbia background) were generated in our laboratory (López-Juez and Hills, 2011) and obtained from our stocks. The location of each gene on *Arabidopsis* genetic map is provided in the appendix (Figure 7.1).

2.2 *Arabidopsis thaliana*-related preparations

2.2.1 Seedling growth media preparation

Plant growth media consisted of 4.4 g Murashige-Skoog salt and Gamborg's vitamins, 500 mg MES, 8 g phyto-agar (Duchfa), with or without 1% sucrose, with sterile water added to make a final 1 L volume. The final solution pH was adjusted to 5.7 with KOH. Media was autoclaved at 121 °C and cooled down to 50 °C, then poured into Petri dishes in a laminar flow hood cabinet. Antibiotics or herbicide were added into media before pouring for transformed plant selections in the cloning experiments. Media did not contain sucrose for the FR block of greening experiment.

2.2.2 Seed sterilisation

The required number of seeds were placed in a 1.5 ml eppendorf tube. Seeds were washed in 95% ethanol for 1 min, the ethanol removed and seeds resuspended in 1 ml of diluted bleach (2 volume of water: 1 volume of bleach) for 10 mins and occasionally the tube was placed in a rotator. After removing bleach, seeds were washed 4 times with sterile water in the laminar flow hood. The last step was to resuspended seeds in 0.5 ml of 0.1% autoclaved agar and stored in a 4°C fridge.

2.2.3 Seed plating

Seeds were homogeneously distributed on the surface of the agar plates using a Gilson pipette in the laminar flow hood. Plates were left open for a short time in order to dry the excess water and then sealed with a micro-porous tape. The plates containing seeds were incubated at 4°C for cold treatment for 3 days.

2.2.4 Plant soil preparation

A mixture of 6 scoops of Levington M3, 6 scoops of John Innes No. 3, 1 scoop of Perlite and 20 g of Intercept or 10 g of Exemptor insecticide (total approximate volume 25 L) were prepared and used for growing *Arabidopsis*.

2.2.5 Seedling and plant growth conditions

Seedlings were transferred to a growth chamber with different light fluence rates ($0\sim 180 \mu\text{mol m}^{-2} \text{s}^{-1}$) and a constant temperature of 22 °C. Seedlings were transferred from agar plates to soil after 5 days and placed in growth room with temperature of 21 °C and $180 \mu\text{mol m}^{-2} \text{s}^{-1}$ light fluence rate. The fluence rates of $20 \mu\text{mol m}^{-2} \text{s}^{-1}$ FR light and $180 \mu\text{mol m}^{-2} \text{s}^{-1}$ white light were used for the FR light block of greening.

2.3 *Arabidopsis* genotyping assays

2.3.1 DNA extraction method

The Edwards DNA extraction protocol (Edwards et al., 1991) was used for genotyping assays of *Arabidopsis* tissue. DNA extraction buffer stock (50 ml) was prepared to the following final concentrations, by mixing the following volumes from stocks: 200 mM Tris-HCl, pH 7.5 (10 ml of 1M stock), 250 mM NaCl (2.5 ml of 5M stock), 25 mM EDTA (2.5 ml of 0.5M stock), 0.5% SDS (2.5 ml of 10% stock) and 32.5 ml sterile water. The plant tissue was collected and placed in an eppendorf tube, all tubes were then cooled down by placing in N_{2(l)} together with blue pestle. The plant tissue was ground with the cooled blue pestle. 500 µl of DNA extraction buffer was added to the sample and the mixture was vortexed for 10 s. The DNA extraction mixture was centrifuged at full speed (microcentrifuge, 13,200 rpm) for 5 mins. 450 µl of the supernatant was transferred without disturbing any pelleted cellular material into a new tube containing 450 µl Isopropyl alcohol (IPA). The tube containing sample was vortexed and left at room temperature for 10 mins then centrifuged at full speed for 7 mins. After discarding the supernatant, the pellet was washed in 500µl 70% cold ethanol by inverting the tube without attempting to resuspend the pellet, and then ethanol was removed. The pellet was left to dry by placing the tube briefly in a 65 °C heat block. 50/10 TE buffer stock was prepared containing 50 mM Tris-HCl (pH 7.5, 2.5 ml of 1M stock), 10 mM EDTA (1 ml of 0.5M stock) by mixing stocks and 46.5 ml sterile water. The DNA pellet was re-suspended in 50 µl TE DNA suspension buffer. The sample should be kept frozen at -20 if not immediately used

2.3.2 Primers design

Primers were designed using Primer3 (<http://frodo.wi.mit.edu/>). DNA sequence was copied and pasted into Primer3. Primer picking conditions and DNA product

size were set according to the type of experiment. Primers were designed by selecting 'pick primers' tab. CAPS primers were designed to surround differential restriction sites as determined by Webcutter (<http://rna.lundberg.gu.se/cutter2/>). And dCAPS primers were designed by dCAPS Finder 2.0 (<http://helix.wustl.edu/dcaps/dcaps.html>). qPCR primers were selected with the following criteria: towards 3', about 80-100 bp product, annealing temperature about 60 °C; any self-complementarity <5 (3 if possible), 3' self-complementarity <2 (0 if possible), no GC clamp and no more than 3 repeats of Gs or Cs. All primer details are shown in the appendix (Table 7.1).

2.3.3 Real-time polymerase chain reaction

Promega GoTaq Flexi DNA Polymerase was used for general genotyping of plant material. A master mixture was made before adding extracted DNA sample. Taq polymerase was the last solution to be added to the master mixture. The PCR sample was prepared using the parameters shown below.

Solution		Large Reaction	Small Reaction	
Master mixture	PCR water	31.8 µl	15.4 µl	
	5 x GoTaq Green PCR buffer	10 µl	5 µl	
	dNTPs 12.5 mM each mix, 200 µM each	0.8 µl	0.4 µl	
	MgCl ₂ 25 mM stock, 2.5 mM final	5 µl	2.5 µl	
	Forward primer	5 picomoles (small) from 20 µM	0.5 µl	0.25 µl
	Reverse primer		0.5 µl	0.25 µl
	Taq 1 unit (small reaction)		0.4 µl	0.2 µl
	DNA template		1 µl	1 µl
Total volume		50 µl	25 µl	

PCR program was set according to the condition of the primers and the size of amplicon. The annealing temperature was based on the lowest T_m of primer pair and the extension time was based on 1 min/Kb amplicon length. A general PCR program is shown below.

Initial denaturation	94 °C	2 mins	} 30 cycles
Denaturation	94 °C	30 s	
Annealing	48-62 °C	30 s	
Extension	72 °C	1 min	
Final extension	72 °C	5 mins	

2.3.4 Restriction enzyme digestion of dCAPS and CAPS assay

The dCAPS assay was designed to genotype these mutants. It is a modification of CAPs technique for detection of single nucleotide, by using a mismatched PCR primer to the template DNA to introduce or destroy a restriction enzyme recognition site. All the restriction enzymes were fast digestion enzymes ordered from NEB (NEW ENGLAND BioLabs Inc.). Enzyme digestion of PCR products was prepared using the parameters shown below. The enzyme and primers details for each assay are shown in the appendix (Table 7.2).

	Solution	dCAPS	CAPS
Master mixture	Water	12 µl	15 µl
	Buffer	2 µl	3 µl
	Enzyme	1 µl	2 µl
PCR products		15 µl	10 µl
Total volume		30 µl	30 µl

2.3.5 Gel electrophoresis

TBE stock (10x) was prepared by mixing 242 g Tris Base, 100 ml EDTA (0.5 M, pH 8.0), made up to 1 L with sterile water. 1% agarose gel was made by

dissolving agarose powder with 1×TBE for separating DNA PCR products which were larger than 3000bp, 1.2% agarose gel was used for DNA PCR products larger than 900bp and 2% agarose gel for smaller DNA PCR products separation in dCAPS and CAPS assays. Agarose was heated in a flask with TBE until the solution became crystal clear. Ethidium bromide (EtBr) was added from a 1 mg/ml (x20,000) stock in the agarose gel and mixed well by gently shaking the flask. The melted agarose gel was poured into a tray containing a comb and allowed to set. Upon cooling, the gel was placed in gel electrophoresis tank containing clean 1xTBE buffer and the comb was removed. 5 µl DNA ladder and 15 µl sample were loaded into wells. The gel was run at a constant voltage (80V for the small tank and 120 for the large tank) until DNA PCR products were well separated. The gel was exposed to UV light to observe the DNA size. GeneRuler 1kb DNA ladder (Thermo) or Hyperladder IV (Bioline) were used in this project (Figure 2.1).

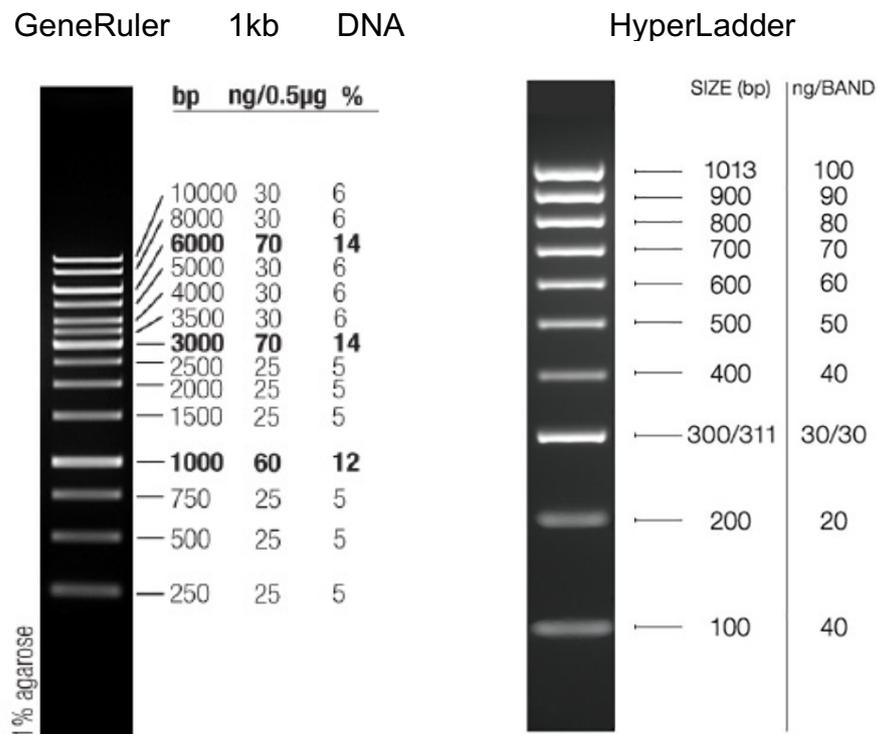


Figure 2. 1 Agarose gel images of molecular markers used during the project Figures from Thermo Fisher Scientific (GeneRuler) and BIOLINE (HyperLadder IV).

2.4 Reverse Transcription PCR

Up to 100 mg of plant material per sample was collected for RNA extraction. The RNeasy plant mini kit (Qiagen) was used for extraction. The procedure followed the manufacturer's instructions. RNA concentration was measured by Nanodrop ND-1000 spectrophotometer. A ratio of 260/280 around about 2 indicated the good quality of RNA (free from protein). The quality of extracted RNA was also confirmed on an agarose gel, confirming integrity through the presence of sharp bands of rRNA. 1 µg of RNA was reverse transcribed into cDNA by Qiagen QuantiTect[®] reverse transcription kit containing pre-mixed oligo-dT and random primers for all reverse transcription procedures in this study, followed by a general PCR programme described in chapter 2.3.3.

2.5 Plastid compartment size determination

Arabidopsis leaf fragments were collected and fixed in 4% glutaraldehyde for one hour in a 1.5 ml Eppendorf tube. They were transferred and submerged in PBM buffer after removing the glutaraldehyde. The samples were covered with foil and stored at 4 °C for no longer than a few days. To separate cells, the PBM Buffer was removed from the tubes, replaced with 500µl of EDTA solution, and the samples incubated for 1-2 h at 65 °C. A single leaf fragment was collected with forceps and placed on a microscope slide. A cover slip was placed on top of the fragment and gently squashed by tapping with the dissection needle's wooden handle. The microscope slide was observed using differential interference contrast (DIC) microscopy (Nikon Optiphot 2) with 20x and 40x objectives. Cell images were collected and chloroplasts were quantified using NIS elements R. Chloroplast and cell quantifications were carried out live, directly under the microscope by changing the focus to cover the entire cell, usually requiring three

focal planes (Pyke, 2011). The calculation of total chloroplast areas and cell index were shown below.

Total chloroplast area = Chloroplast number per cell x Average chloroplast size

Cell index = Total chloroplast area / Cell plan area

2.6 Hypocotyl length measurement

Arabidopsis seedlings were laid on agar plates along with a measuring ruler. Images were taken at a perpendicular angle. Image J (<https://imagej.nih.gov/ij/>) was used to analyse the hypocotyl length from the images (Schneider et al., 2012).

2.7 Chlorophyll and protochlorophyllide quantitation

20 seedlings were collected and placed into a 1.5 ml eppendorf tube containing 1 ml N,N-dimethylformamide (DMF). The tissue was kept overnight in DMF at 4 °C in dark conditions. Before quantitation, the sample was vortexed for 15 mins, then centrifuged for another 15 mins at 13200 rpm. The absorbance of the supernatant was measured using a Thermo Scientific Helios Beta Spectrophotometer at 664 nm and 647 nm. The chlorophyll content of each seedling was calculated using the formula below.

Total chlorophyll per seedling (nmol) = $(19.43 A_{646.8} + 5.08 A_{663.8}) / 20$ seedlings

LS50-B spectrofluorimeter (Perkin Elmer) and FL Winlab Scan v4.01 software were used for chlorophyll and protochlorophyllide quantitation of FR-grown *phyA* mutants. The seedlings were collected in a dark room under dim green light conditions. Protochlorophyllide and chlorophyll were extracted using the same protocol. They were quantified by spectrofluorimetry with 10 nm slit of 430 nm excitation and 635 nm emission light. The peak of fluorescence around 630 nm

and 672 nm corresponded to protochlorophyllide and chlorophyll, respectively. Fluorescence was expressed in relative units.

2.8 Photochemically-active phytochrome quantitation

This experiment was performed by the laboratory of Tim Kunkel, University of Freiburg. Dark-grown seedlings, in which phytochrome A accumulates, were packed in cuvettes, placed in a purpose-built dual wavelength spectrophotometer (Ratiospec) and exposed to consecutive red and FR light several times. The absorbances were measured at 730 nm and recorded after each light pulse. The absorbance changes were calculated from each set of light pulses and then averaged. The average of absorbance changes represented the relative amount of photochemically-active phytochromes.

2.9 Plastid genome to nuclear genome ratio

To analyze the DNA copy number ratio between plastid-encoded genome and nuclear-encoded genome, three plastid-encoded genes and two nuclear-encoded genes were selected. Three plastid-encoded genes were selected as being present in different regions of plastid genome: *rbcL*, *ndhG* and *ycf2* are in the large signal-copy, the small signal copy and the two inverted repeats region, respectively. Because plastid genome rearrangements can occur due to the presence of the inverted repeats and intra-molecule recombination, use of a single gene could have inaccurately over- or under-estimate the number of incomplete copies of the chloroplast genome. Two single-copy nuclear-encoded genes, *HO1* and *CHS*, were also selected to indicate the copy number of nuclear genome.

Up to 100 mg of seedlings were collected for DNA extraction. DNA extraction was performed using DNeasy plant Mini kit (Qiagen). The procedure followed the

manufacturer's instructions. The quality of extracted DNA was confirmed by NanoDrop. A 260/280nm ratio between 1.8 to 2.0 indicated high DNA purity (absence of protein). DNA sample was diluted to 1/10 and 1/100. 1/10 diluted sample was used for analysing nuclear-encoded genes (*HO1* and *CHS*) copy number and 1/100 diluted sample used for the plastid-encoded genes (*rbcl*, *ndhG* and *ycf2*). A PCR product standard was generated by endpoint PCR amplification. Endpoint PCR used the same programme as general real-time PCR, except it used at least 35 cycles. Purified amplicons (Qiagen) were quantified by Nanodrop, then used to generate a 10-fold dilution series. The dilution series started from 25 pg/μl to 0.00025 pg/μl and analysed with other samples by quantitative real-time genomic-PCR (qgPCR). 2x qPCRBIO SyGreen Mix Lo-ROX was used to quantify the copy number of genes. The qg-PCR sample was prepared using the parameters shown below.

		1 standard reaction
Master mixture	2 x qPCRBIO SyGreen Master Mix	10 μl
	RNase Water	6 μl
	Primers mixture (F+R) 10 μM	2 μl
cDNA sample		2 μl
Total volume		20 μl

q-PCR programme was set using the parameters shown below.

Hold	95 °C	2 mins	
Cycling	95 °C	5 s	} 35 cycles
Melt	60 °C	20 s	
	72 °C~95 °C	rise by 1 degree	
		pre-melt 90s for each step afterward 5 s	
		470 nm	
Channels	Green	510 nm	
	Source	10	
	Detector		
	Gain		

Rotor-Gene Q software (Qiagen) automatically set a threshold to calculate Ct values. A standard curve was generated using the Ct values and the concentrations of serial dilution that indicated the concentration of each sample according to the Ct value. According to the concentration, the copy numbers were calculated by the online DNA standard copy number calculator (<http://cels.uri.edu/gsc/cndna.html>) based to the length of cDNA template. The average copy numbers of two nuclear-encoded genes and three plastid-encoded genes (using half for *ycf2*, present as two copies per genome) were calculated respectively, before calculating the plastid genome to nuclear genome ratio.

2.10 Global gene expression analysis

2.10.1 RNA sample preparation

Up to 100 mg of plant material per sample was collected for RNA extraction. The RNeasy plant mini kit (Qiagen) was used for extraction. The procedure followed the manufacturer's instructions. RNA concentration was measured by Nanodrop spectrophotometer ND-1000. Ratio of 260/280 around about 2 indicated high RNA purity (absence of protein). The quality (intactness) of extracted RNA was also confirmed by gel electrophoresis.

The experimental design of the global gene expression analysis compared 4 genotypes (WT, *lyn1*, *hy1* and *lyn1 hy1*) and 2 light conditions ($100 \mu\text{mol m}^{-2} \text{s}^{-1}$

continuous white light and dark) of 5-day old seedlings. 8 RNA samples of different genotypes and light conditions were extracted. They are shown below.

Genotype \ Light condition	WT	<i>lyn1</i>	<i>hy1</i>	<i>lyn1 hy1</i>
Light (L)	WT_L	<i>lyn1_L</i>	<i>hy1_L</i>	<i>lyn1 hy1_L</i>
Dark (D)	WT_D	<i>lyn1_D</i>	<i>hy1_D</i>	<i>lyn1 hy1_D</i>

Each sample was obtained as 3 biological replicates. 4 µg RNA diluted in 50 µl water for each sample was submitted in dry ice to Prof. Beemster's laboratory at the University of Antwerp for RNA sequencing.

2.10.2 RNA sequencing data analysis

The Beemster laboratory made use of an Illumina 1500 DNA sequencing platform. All the genotypes in both light conditions were analysed as three biological replicates. However, not all the samples produced reasonable reads. A sample was indicated as successful if it generated more than 15,000,000 reads. The first biological replicate of *lyn1* dark sample and the third biological replicate of *hy1* dark sample only generated 1192 reads and 3858 reads, which are much lower than the expected number, therefore these two samples were rejected for use in the later RNA sequencing (RNA-Seq) analysis. A PCA following a preliminary gene expression analysis was carried out by a bioinformatics researcher at the University of Antwerp, and this indicated that the second biological replicate of the *lyn1* light sample showed very different expression levels compared to other two biological replicates. The sequence of the *LYN1* gene of this second biological replicate of the *lyn1* light sample was verified from the RNA-Seq file to establish whether this replicate was a correct *lyn1* mutant, but this was not the case and the sample was excluded. All the biological replicates of each sample which are included in the RNA-Seq data analysis are shown in table 2.1.

Genotype		WT	<i>lyn1</i>	<i>hy1</i>	<i>lyn1 hy1</i>
		Light condition			
Light (L)		1 2 3	1 2 3	1 2 3	1 2 3
Dark (D)		1 2 3	1 2 3	1 2 3	1 2 3

Table 2. 1 List of all the biological replicates of each genotype and each light condition included and excluded in the RNA-Seq data analysis

Each sample includes 3 biological replicates which are numbered 1, 2 and 3. The black colour indicates that this replicate is included in the RNA-Seq data analysis and red indicates that this replicate is excluded.

The RNA-Seq raw data were normalized by CLC Genomics Workbench. Gene expression values of all three biological replicates of each genotype were indicated and selected as “expressed” if they all had more than 4 reads. The false-discovery rate (FDR) based-pairwise multiple comparison was applied by CLC to pair up the samples and indicate whether there is any significantly different expression for each gene between the sample pair. Among all the comparisons, 16 relevant ones were selected for use in the RNA sequencing data analysis as a filter (Figure 2.2). A gene was selected as significantly expressed if the FDR p-value of any comparison of gene expression between two samples was lower than 0.05. The \log_2 value of normalized average gene expression ratio of each sample pair was calculated. The genes were selected as exhibiting more than 2-fold change if the \log_2 expression difference between any two genotypes in either light condition was larger than 1. Those selected genes were indicated as differentially expressed genes.

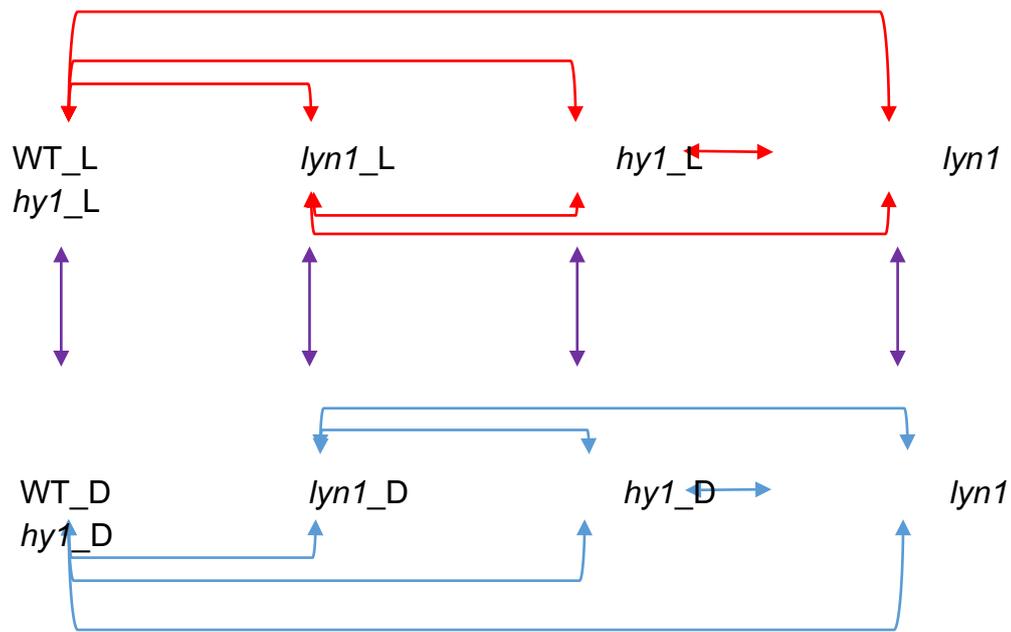


Figure 2. 2 Figure to show FDR-based pairwise multiple comparisons between different genotypes and light conditions

Red arrows indicate the comparisons between different genotypes in the light. Blue arrows indicate the comparisons between different genotypes in the dark. Purple arrows indicate the comparisons between different light conditions for each genotype.

GO (Gene Ontology) term frequencies of differentially expressed genes were analyzed by the Bingo plug-in of Cytoscape software (Shannon, 2003; Maere et al., 2005). According to the gene expression pattern, the differentially expressed genes were clustered into groups using the KMC (K-means/Medians clustering) algorithm of MeV (MultiExperiment Viewer) (Howe et al., 2011). The KMC was set to use Euclidean Distance, as a metric to calculate k-means distance, a limit of 5 clusters for clustering all differentially expressed genes and 10 clusters for subclustering cluster 1 of the initial 5, and a limit of 50 maximum iterations. The number of 5 clusters for all differentially expressed genes, and 10 subclusters for cluster 1 was chosen experimentally, as that which revealed what appeared to specifically be *lyn1*-dependent genes. The functional gene groups which were associated with biological processes in plants were obtained from MAPMAN and

then manually expanded in our laboratory (López-Juez et al., 2008). They were used for the functional gene classification of genes of interest. To obtain the frequency of functional genes among the genes of interest, the overlapping genes between the functional gene group and the genes of interest were identified using the AGI converter and Venn selector from the Bio-Array Resource (BAR, http://bar.utoronto.ca/ntools/cgi-bin/ntools_agi_converter.cgi) (Toufighi et al., 2005). The overrepresentation or underrepresentation of these functional genes among the genes of interest, i.e. the probability of the observed frequency, was computationally analysed with a hypergeometric test using R (R Core Team, 2014). The genes indicated as “expressed” at the first step of selection of differentially expressed genes from the RNA-Seq raw data were used as the total population of that test. If the $-\log_{10}$ of the hypergeometric p-value of any given biological function is greater than 1.3 ($p < 0.05$, $-\log_{10} 0.05 = 1.3$), that biological function is indicated as “overrepresented”.

2.10.3 Processing of publicly-available data

Data of the Arabidopsis Gene Atlas from Schmid et al. (2005) were obtained from ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>, accession number: E-TABM-17). GeneChip ID was converted to AGI codes using the _at to AGI converter of BAR. The overlap between the Schmid data and the differentially expressed genes of RNA-seq data were identified using Venn selector of BAR. The overlapping genes were used as the test population in hypergeometric tests for analysing the association of early differentiation and *lyn1*-dependent regulation. A list of gene expression change ratios was created between internal sets of Gene Atlas data which represented Early vs. Later stages of differentiation during development (6/5, 6/7), Early vs. Completed development (6/1), Ongoing development vs. late development (17/13), Ongoing development vs. Completed (17/12) and Photosynthetic vs. Non-Photosynthetic organs (7/3). The differentially expressed genes were selected from each ratio by setting up a

threshold. Any ratio above 2 was recognised as differentially expressed genes involved in cell entry into differentiation. Duplicates were removed by Duplicate Remover of BAR.

Data of Roudier et al. (2011), representing a genome-wide map of histone modifications present in young seedlings was obtained from Geo NCBI (<https://www.ncbi.nlm.nih.gov/geo/>, accession number - roots: GSM607834, seedlings: GSE24658). The H3K4me3-marked genes of green tissue only were defined as the H3K4me3 marked genes of whole seedling excluding the overlap between them and those marked in roots, identified using Venn selector tool. The overlaps of Roudier data and the differentially expressed genes of RNA-seq data were identified using Venn selector. The overlapping genes were used as the test population of hypergeometric tests for analysing the association of H3K4me3 marking state and *lyn1*-dependent regulation.

The Roudier data-based H3K4me3 chromatin marking state in whole seedlings' and in roots' DNA for all genes were observed as a histogram through a Generic Genome Browser at Epigara (<http://epigara.biologie.ens.fr/cgi-bin/gbrowse/a2e/>). Several representative genes (photosynthetic genes: *LHCB1.2*, *LHCB1.3* and *GLK1*; non-photosynthetic and light dependent gene: *CHS*; constitutive genes: *ACT2* and *UBQ10*) were selected. Their H3K4me3 chromatin marking state in whole seedlings' and roots' DNA were obtained from Epigara by selecting 'H3K4me3_WT-Col_roots (HD2.3X) – Roudier et al. (2001)' and 'H3K4me3_WT-Col_seedlings (AT7d1) – Roudier et al. (2001)' under the column of '2-A2E Genome-wide Tiling Array'.

2.11 Bioinformatics analysis

2.11.1 Multiple sequence alignments and phylogenetic tree construction

Information of *Arabidopsis* JmjC domain-containing proteins of JARID1 group was obtained from a published analysis (Lu et al., 2008). *Arabidopsis* Jumonji protein sequences were obtained from The *Arabidopsis* Information Resource (TAIR, <https://www.arabidopsis.org/>) (Berardini et al., 2015). *JMJ14 Arabidopsis* paralogues were identified using a BLASTP search (NCBI). The top 6 paralogues were selected. For other organisms, chosen to reflect a wide taxonomic range within and outside plants, the closest homologue was identified by BLASTP (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). The sequence of each homologue was obtained from Genbank using BLASTP (Benson et al., 2012). A multiple protein sequence alignment of full *JMJ14* protein sequence was performed by EMBL-EBI based Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). A phylogenetic tree was produced by SplitsTree (<http://www.splitstree.org/>) (Huson and Bryant, 2006) using the alignment file which was generated by Clustal Omega. Plaza (https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v4_monocots/genes/view/LOC_Os05g23670) was additionally used for *Oryza sativa* protein sequence BLAST (Van Bel et al., 2017). iTAK (http://itak.feilab.net/cgi-bin/itak/db_gene_seq.cgi?trans_ID=Glyma.09G083300.1) was additionally used for *Glycine max* protein sequence BLAST (Zheng et al., 2016). The protein domains were predicted using SMART (<http://smart.embl-heidelberg.de>) (Letunic et al., 2015).

2.11.2 Gene expression analysis

The *JMJ14* gene expression level and location across the plant developmental stages was analysed by the Arabidopsis eFP Browser at BAR (<http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi>) (Winter et al., 2007), visualising data from the *Arabidopsis* Gene Atlas (Schmid et al., 2005).

2.12 Complementation and overexpression

2.12.1 Bacterial growth media and conditions

All LB broth and LB agar were made up using LB broth buffered capsules (Duchefa, 1g per capsule, one capsule for 40 ml water). Additionally, LB agar contained 15 g/L agar (High gel strength). Both media were autoclaved at 121 °C for 20 mins. Before making the LB agar plates, antibiotics were added into the cooled media. All the antibiotics used in the bacterium selection are shown in Table 2.2. 50 µl LB broth containing transformed bacteria was streaked on an antibiotic-selective LB agar plate by a sterilized bacteria spreader, then the antibiotic plate was placed in a 37 °C incubator overnight for *E. coli* to grow. *Agrobacterium* required growing in a 28 °C incubator for 3 days. Plates was placed upside down in the incubator.

Chemical agent	Diluent	Storage temp.	Stock	Final	Bacterial and transformed plant selection
Ampicillin	H ₂ O	-20 °C	100 mg/ml	50-100 µg/ml	<i>E. coli</i>
Carbenicillin	H ₂ O	-20 °C	100 mg/ml	50-100 µg/ml	<i>E. coli</i>
Gentamicin	H ₂ O	-20 °C	25 mg/ml	25 µg/ml	<i>A. tumefaciens</i>
IPTG	H ₂ O	4 °C	0.1 M	0.5 mM	<i>E. coli</i>
Kanamycin	H ₂ O	-20 °C	50 mg/ml	25-50 µg/ml	<i>A. tumefaciens</i> <i>E. coli</i>
Rifampicin	MetOH	-20 °C	50 mg/ml	50 µg/ml	<i>A. tumefaciens</i>
Spectinomycin	H ₂ O	-20 °C	100 mg/ml	100 µg/ml	<i>A. tumefaciens</i> <i>E. coli</i>
X-Gal	N,N'-dimethyl-formamide	-20 °C	50 mg/ml	80 µg/ml	<i>E. coli</i>
Phosphinothricin	H ₂ O	-20 °C	10 mg/ml	5 µg/ml	<i>A. thaliana</i>
Cefotaxime	H ₂ O	-20 °C	100 mg/ml	100 µg/ml	<i>A. thaliana</i>

Table 2. 2 Chemical agents used for bacterial and transformed plant selection and their preparation details

2.12.2 Preparation of *JMJ14* template

RNA was extracted from a WT plant. The RNeasy plant mini kit (Qiagen) was used for extraction. The procedure followed the manufacturer's instructions. RNA concentration was measured by Nanodrop spectrophotometer ND-1000. Ratio of 260/280 around about 2 indicated high purity RNA. The quality of extracted RNA was also confirmed by gel electrophoresis. 1 µg of RNA was reverse transcribed into cDNA by Qiagen QuantiTect[®] reverse transcription kit using optimally-mixed oligo-dT and random primers provided by the reverse transcription kit. Thermo Scientific Phusion (*Pfu*) High-Fidelity PCR Master Mix kit was used with *JMJ14* full-length primers to amplify full length *JMJ14* cDNA with minimal probability of any amplification error during the PCR amplification. The PCR sample was prepared using the parameters shown below.

	Solution	1 Standard Reaction
Master Mixture	PCR water	18 μ l
	Phusion mix	25 μ l
	Forward Primer	2.5 μ l
	Reverse Primer	2.5 μ l
	DNA Template	2 μ l
Total volume		50 μ l

The PCR programme was set using the parameters shown below.

Initial denaturation	98 °C	30 s	} 30 cycles
Denaturation	98 °C	10 s	
Annealing	70 °C	45 s	
Extension	60 °C	1 min 30 s	
Final extension	72 °C	7 mins	

The PCR products were cleaned by PCR purification without gel extraction (avoiding the associated problems which will be mentioned in section 3.2.1.4.1) using wizard SV gel and PCR clean-up system kit. The procedure followed the manufacturer's instructions. Adenosines were added at both ends of the DNA fragments using the A-tailing procedure. A standard reaction for A-tailing was prepared using the parameters shown below.

	1 Standard Reaction
Purified PCR product (300 g~ 3000 ng)	4 μ l
5xGoTaq PCR Buffer	2 μ l
dATP (1 mM)	2 μ l
GoTaq polymerase	1 μ l
25 mM MgCl ₂	0.6 μ l
ddH ₂ O	0.4 μ l
Total volume	10 μ l

The mixture was incubated at 70°C for 15~30 mins. pGEM[®]-T Easy Vector System (Promega) was used to ligate *JMJ14* PCR products into a pGEM[®]-T Easy vector. A standard reaction for ligating was prepared using the parameters shown below.

	1 Standard Reaction	Positive control	Background control
2 x rapid ligation Buffer	5 µl	5 µl	5 µl
pGEM-T Easy Vector (50 ng)	1 µl	1 µl	1 µl
T4 DNA ligase	1 µl	1 µl	1 µl
PCR product	3 µl	n/a	n/a
Control Insert DNA	n/a	2 µl	n/a
H ₂ O	n/a	1 µl	3 µl
Total volume	10 µl	10 µl	10 µl

The mixture was incubated at 4 °C overnight.

The identification of successful insertion was indicated by the blue/white screening on indicator plates. pGEM[®]-T Easy Vectors contains *lacZ* gene of the *lac* operon which encodes β-galactosidase. The recombination between insert and vector interrupts the gene for, and therefore the activity of, β-galactosidase. The activity of β-galactosidase can be detected by X-gal (5-bromo-4-chloro-3-indoyl-β-d-galactopyranoside) which is a soluble colourless substrate. X-gal can be hydrolysed by β-galactosidase to form 5-bromo-4-chloro-indoxyl, which spontaneously dimerizes and oxidizes to form an insoluble and intensely blue product 5,5'-dibromo-4,4'-dichloro-indigo. Isopropyl β-D-1-thiogalactopyranoside (IPTG) acts as the inducer of the *lac* operon to enhance the expression of *lacZ*. Therefore, the vectors with insertions are present in white colonies, while empty vectors are present in blue colonies.

2.12.3 Generation *attB*-flanked *JMJ14* cDNA

A pair of primers shown in Figure 2.3 was used to amplify *attB*-flanked *JMJ14* fragments by PCR.

attB1+JMJ14F:

GGGGACAAGTTTGTACAAAAAAGCAGGCTATGGATCAGCTTGCATCTCTAGC

attB2+JMJ14R:

GGGGACCACTTTGTACAAGAAAGCTGGGTTTAAGGACTTATCTCCATCTTATCAACC

Figure 2. 3 Designing of *attB*-adapted *JMJ14* primers

The *attB* sites (29 bp) are added to the *JMJ14* full-length primers and are both underlined.

The PCR programme was set using the parameters shown below.

Initial denaturation	98 °C	30 s	} 30 cycles
Denaturation	98 °C	10 s	
Annealing	70 °C	45 s	
Extension	72 °C	1 min 30 s	
Final extension	72 °C	7 mins	

2.12.4 Transformation of competent cells

2.12.4.1 Heat-shock transformation procedure for *E coli*

The silver efficiency (BIO-85026, $\geq 10^8$ cfu/ μ g of pUC19) of α -Select Chemically Competent Cells (Bioline) were used to multiply pGEM-T Easy/*JMJ14* plasmids and pDONR201/*JMJ14* plasmids. 10 μ l of plasmid culture and 50 μ l *E coli* were mixed in a 1.5 ml eppendorf tube and incubated on ice for 30 mins. The mixture was transferred to a 42 °C water bath for 30 s, then immediately transferred to ice for 2 mins. 1 ml LB broth was added to the mixture and incubated on a shaker at 37 °C and 200 rpm for 1 hour. The cells were spun down for 1 min at 3000 rpm.

Most supernatant was removed and only 100 µl was left. The cells were re-suspended in the remaining supernatant for selection on antibiotic plate.

2.12.4.2 Preparation of competent *Agrobacterium tumefaciens* cells

1 ml overnight start culture was used to incubate with 240 ml of LB broth in a large autoclaved flask with antibiotics on a shaker at 28 °C and 120 rpm for 11 hours, until the OD 600 nm was between 0.5~0.8. The *Agrobacterium* culture was transferred from the flask into four 50 ml falcon tubes and cooled on ice. The cooled *Agrobacterium* culture was centrifuged at 4 °C and 4000 g for 15 mins. The supernatant was removed and the cells were re-suspended in ice-cold sterile 1xTE. 1xTE was prepared by mixing 5 ml Tris-HCl (1 M, pH 8.0), 1 ml mM EDTA (0.5 M) and 494 ml dH₂O. The mixture was centrifuged at 4 °C and 4000 g for 15 mins. The supernatant was removed and the cells were re-suspended in 10 ml cold LB. 250 µl of *Agrobacterium* stock was aliquoted into 1.5 ml Eppendorf tubes and stored in -80 °C after flash freezing by N_{2(l)} (Hofgen and Willmitzer, 1988).

2.12.4.3 Heat-shock transformation procedure for *Agrobacterium*

Agrobacterium tumefaciens strain GV3101 was used to multiply pB2GW7/JMJ14 plasmids. 3 µl of plasmid culture was added into 200 µl of frozen *Agrobacterium* stock and the mixture was thawed on ice for 10 mins. The mixture was transferred to a 28 °C water bath for 5 mins, then transferred to ice for 2 mins immediately. 1 ml LB broth was added to the mixture and incubated on a shaker at 28 °C and 120 rpm for 3 hours. The cells were spun down for 1 min at 8000 rpm. Most supernatant was removed until only 50 µl remained. The cells were re-suspended with the remaining supernatant for selecting on the antibiotic plate later.

2.12.5 Colony PCR

2.12.5.1 Colony PCR for *E coli*

A tip was used to pick a single colony from the antibiotic-selective LB agar plates and washed in 20 μ l H₂O by pipetting up and down several times. The cells were boiled within H₂O for 15 mins at 95 °C and then centrifuged for 5 mins at full speed. 2 μ l of the supernatant was used in colony PCR. Promega hot start master mix was used to amplify the insert, present within a plasmid. The sample for PCR was prepared using the parameters shown below.

	Solution	Small Reaction
Master mixture	PCR water	8 μ l
	GoTaq G2 Hot start master mix 2x	12.5 μ l
	Forward primer	5 picomoles
	Reverse primer	from 20 μ M
DNA template		2 μ l
Total volume		25 μ l

The PCR programme was set using the parameters shown below.

Initial denaturation	94 °C	2 mins	} 35 cycles
Denaturation	94 °C	30 s	
Annealing	48-62 °C	30 s	
Extension	72 °C	3 mins	
Final extension	72 °C	10 mins	

2.12.5.2 Colony PCR for *Agrobacterium*

Agrobacterium colony PCR lysis buffer stock was prepared by mixing 200 μ l 10% Triton X-100, 40 μ l Tris-HCl (1M, pH8), 80 μ l EDTA (0.5 M, pH8) and 1.75 ml H₂O. A tip was used to pick a single colony from the plate and washed in 20 μ l *Agrobacterium* colony PCR lysis buffer by pipetting up and down several times.

The cells were boiled within the buffer for 5 mins at 95 °C and then centrifuged for 2 mins at full speed. The preparation of PCR sample and PCR programme setting were described in 2.12.5.1

2.12.6 Bacterial DNA extraction

A single colony was picked by tip and re-streaked in a new antibiotic-selective LB agar plate and grown for new single colonies. A single colony was selected from the new antibiotic-selective LB agar plates and multiplied in 1 ml LB broth with antibiotics in a culture tube on a shaker. *E. coli* grew at 37 °C and 200 rpm overnight and *Agrobacterium* grew at 28 °C and 140 rpm for a few days until the culture turned turbid. Qiagen spin miniprep kit was used for extracting plasmids from competent cells. The procedure followed the manufacturer's instructions.

2.12.7 Gateway cloning technology

Gateway cloning technology was used to transfer gene fragments into a destination vector for 35S promoter complementation and overexpression. Gateway BP Clonase II Enzyme Mix was used to ligate an *attB*-flanked gene with donor vector by BP recombination. Gateway LR Clonase II Enzyme Mix was used to transfer the gene from the donor vector to the destination vector by LR recombination. Both reactions were prepared using the parameters shown below.

Solution	1 Standard Reaction
Donor vector/destination vector (150 ng/μl)	1 μl
PCR product (15 ng-150 ng)	7 μl
BP/LR Clonase II enzyme mix	2 μl
Total volume	10 μl

The mixture was incubated at 25 °C overnight.

2.12.8 *Agrobacterium*-mediated transformation

500 ml LB broth, 4 *Agrobacterium* glycerol stocks from independent strains and antibiotics were mixed in a large autoclaved flask to grow overnight. The *Agrobacterium* culture was centrifuged at 28 °C and 3000 g for 15 mins. The supernatant was removed and the cells were re-suspended in infiltration solution. The infiltration solution was prepared by mixing 25 g sucrose, 50 µl Silwet and 500 ml H₂O. Plant shoots were dipped into infiltration solution for 2 mins. Infiltrated plants were watered and covered with foil to keep in the dark for 24 hours. The foil was removed and the plants were watered again on the next day. The infiltration was repeated at least 3 times.

2.12.9 Plasmid DNA sequencing

Plasmid miniprep was quantified by nanodrop and diluted into 50~100 ng/µl. 15 µl of diluted plasmid miniprep with the separately-packed sequencing primers were sent to Eurofins for sequencing (<http://www.eurofins.co.uk/eurofins-online.aspx>). The sequencing results were downloaded and examined as chromatogram using Chromas (Goodstadt and Ponting, 2001).

2.13 Statistic multiple comparisons test

In chapter 3 and 4, most analyses contain more than two genotypes. This requires multiple comparisons to indicate whether there is any statistically significant difference between the reference genotype and the target genotype. WT is a reference group for all single mutants (*hy1*, *lyn1*, *jmj14-1*, *hy2-105*, *phyA-211*, *phyB-9* and *cry1-304*). *hy1* is a reference group for all double mutants (*lyn1 hy1*, *jmj14-1 hy1*, *hy2-105 lyn1*, *phyA-211 lyn1*, *phyB-9 lyn1* and *cry1-304 lyn1*). The statistically significant differences between genotypes were analysed by one way ANOVA, which is a multiple comparison test, and followed by Bonferroni and Holm multiple comparison T-test, which is a modification of a standard T-test

which controls for multiple comparison type-1 errors and only pairs relative to a reference group. When the p-value corresponding to the F-statistic of one-way ANOVA was lower than 0.05, this indicated that one or more treatments are significantly different. In contrast, when the p-value was higher than 0.05, this suggested that none of the treatments is significantly different. The degree of significant difference (* p-value <0.05, ** p-value <0.01, ***p-value <0.001) was represented as the number of asterisk on the top of bar charts.

Chapter 3 Result – Confirming the *JMJ14* Gene identity

3.1 Introduction and Approach

3.1.1 *LYN1* gene identification

3.1.1.1 T-DNA insertion knock-out lines

According to the mapping results from of previous MSc students' work and my MSc work, *lyn1* is a point mutation on *JMJ14* gene. It was essential to confirm the *LYN1* gene identity before any further research. One way of confirming the *JMJ14* identity involves the use of T-DNA insertion lines. T-DNA insertion generates a knock out (KO) mutation by introducing the T-DNA plasmid sequence into the target gene disrupting the gene transcription and preventing translation of a functional protein. *jmj14-1* (salk_135712), a T-DNA KO mutation on *JMJ14*, was obtained from the stock centre. If the *lyn1* location was identified correctly, then *lyn1* and *jmj14-1* would be allelic, two different mutations on the same gene. To confirm this, *lyn1* and *jmj14-1* were compared on their capacity to suppress hypocotyl elongation and enhance chlorophyll content. The same comparison was also performed between *lyn1 hy1* and *lyn1 jmj14-1* to determine if *jmj14-1* could independently suppress *hy1*. *jmj14-1* appeared to have the same function in suppressing the light response defect of *hy1* as *lyn1*, therefore the *LYN1* identity was confirmed in principle.

3.1.1.2 Gene complementation assay

Along with the phenotypical comparison between *jmj14-1* and *lyn1* and the comparison of the *jmj14-1* and *lyn1* suppression of *hy1*, whether or not *jmj14-1* and *lyn1* were two mutations on the same allele was also confirmed genetically

by complementation tests. *LYN1* identity was confirmed through two types of genetic complementation tests. The first one was the failure of complementation test which was carried out by crossing the *lyn1 hy1* with *jmj14-1 hy1* to obtain the F1 progenies. Whether *jmj14-1* failed to complement *lyn1* was studied on the F1 plants. Another complementation test was the transformation complementation test which inserted a full-length of *JMJ14* cDNA into the genomic sequence of *lyn1 hy1* mutants by gateway cloning technology. Whether the transferred functional *JMJ14* can complement the *lyn1* mutation in *lyn1 hy1* was studied on the transformed progenies.

3.1.1.3 Cloning strategy to generate a *JMJ14* overexpression construct

The Gateway Cloning Technology, commercialised by Invitrogen, is a very efficient method to transfer DNA fragments between vectors by 2 recombination reactions, BP reaction and LR reaction (Figure 3.1). The BP reaction happens between the *attB* sites flanking the DNA fragment and the *attP* sites of the donor vector (Figure 3.1 A). It can generate an entry clone which contains the gene fragment flanked by *attL* sites. The LR reaction happens between the *attL* sites of the entry clone and the *attR* sites of the destination vector (Figure 3.1 B). It can generate an expression clone which contains the *attB* sites flanking the gene fragment. All the vectors (donor vector and destination vector) carrying the insert were always multiplied within *E. coli* cells and selected by antibiotics according to their antibiotic resistances during the process of gateway cloning. Finally, the expression clones were transferred into *Agrobacterium tumefaciens* for multiplication, in order to transfer the gene fragment into *Arabidopsis* plants by the T-DNA binary system (Dandekar et al., 1989).

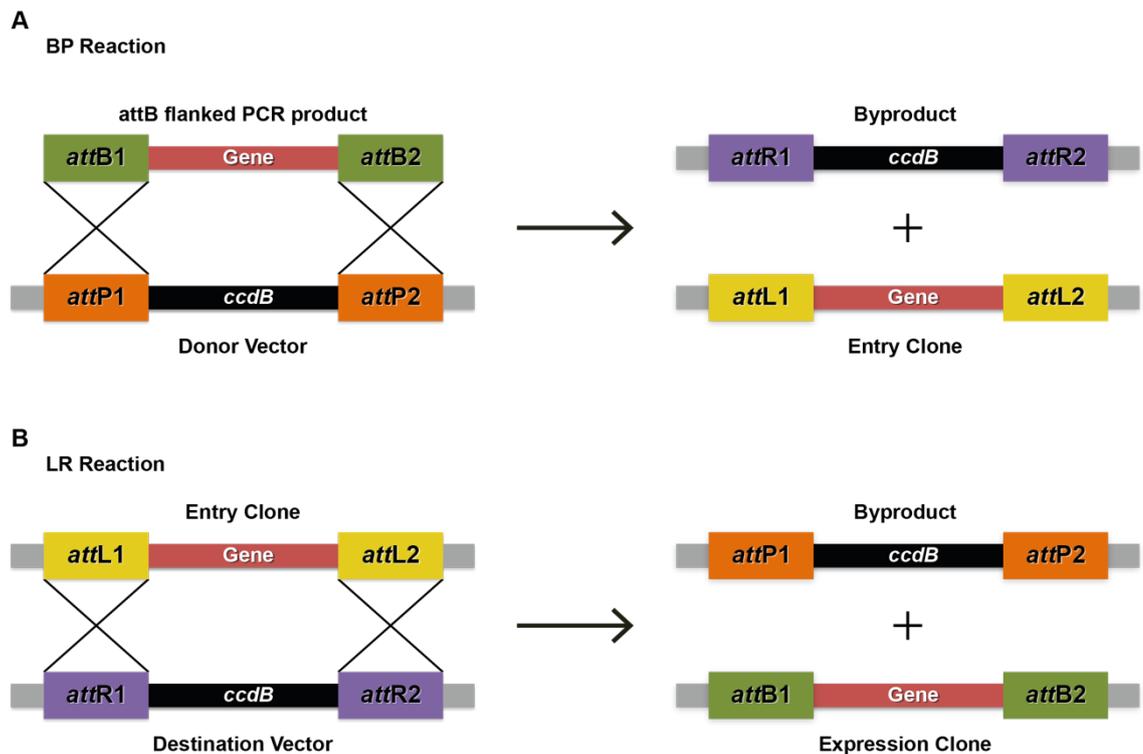


Figure 3.1 The principle of Gateway Cloning used to generate a *JMJ14* complementation and overexpression construct

A. BP recombination reaction. A gene is transferred into a donor vector to produce an entry clone and a by-product. The *attB* sites of a *JMJ14* cDNA fragment recombine with the *attP* sites of a donor vector to generate the gene flanked by *attL* sites in an entry clone.

B. LR recombination reaction. A gene is transferred from an entry clone into an expression clone while producing a by-product. The *attL* sites of the entry clone recombine with *attR* sites of a destination vector to generate an *attB* sites-flanked gene in a plant expression vector.

The initial cloning strategy was to directly generate an *attB*-flanked *JMJ14* PCR product from cDNA by PCR amplification. However, the BP recombination always failed. The possible reason of failure was summarised in the results (section 3.2.1.4.1). Given this failure, a new strategy was designed. The full process of transferring cloning is shown in Figure 3.2. Firstly, the *JMJ14* full length cDNA was inserted into a pGEM-T vector. Then the *JMJ14* was amplified and

transferred between 2 vectors. The pGEM-T/*JMJ14* plasmid was used as template to generate *attB*-flanked *JMJ14* cDNA by PCR. Using plasmid as a PCR template improved the efficiency of PCR amplification and reduced the production of primer dimers. Even without the purification, the *JMJ14* cDNA fragment was inserted into the donor vector and screening for the correct insert could also be carried out.

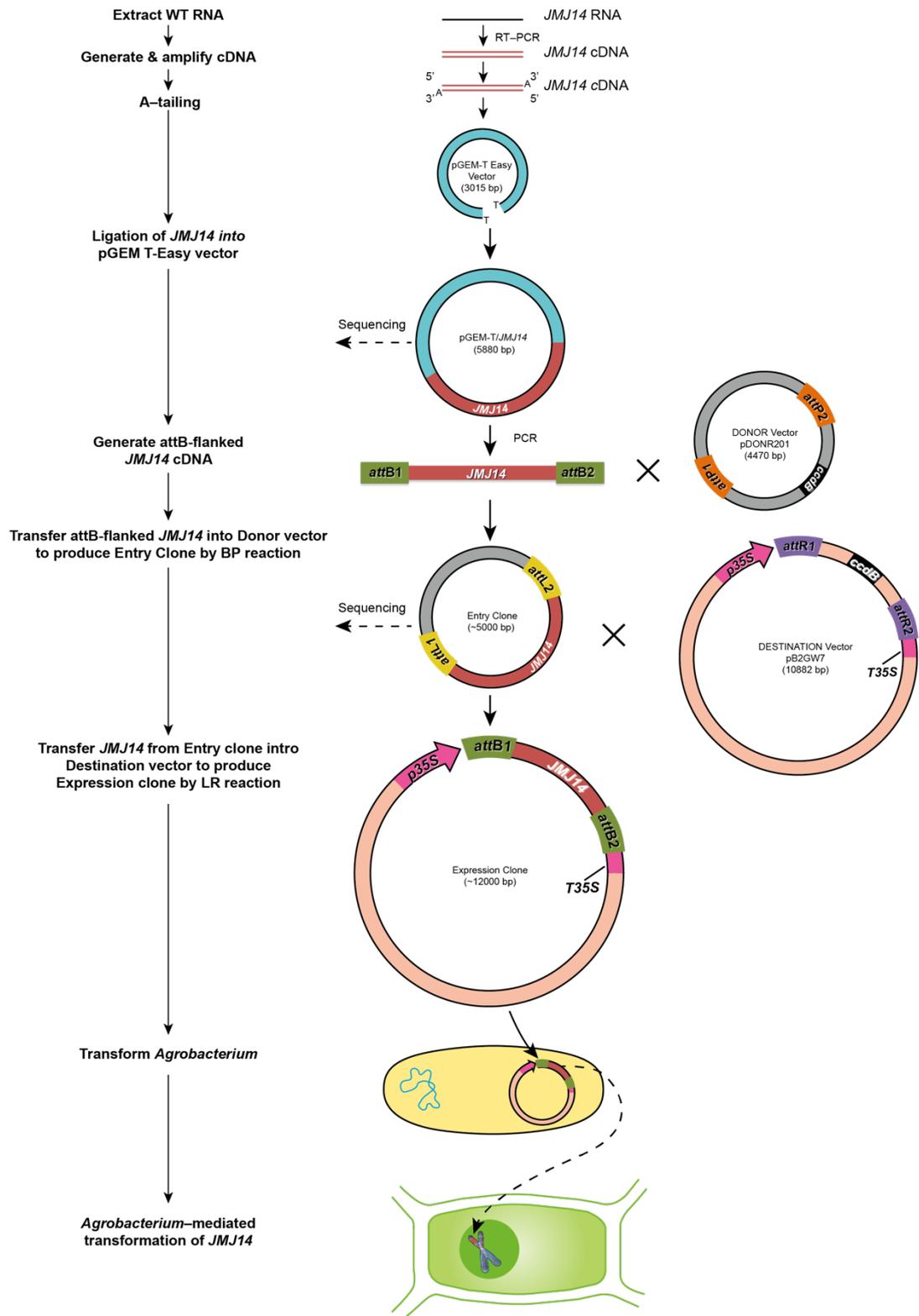


Figure 3.2 Schematic diagram of the cloning strategy for *JMJ14* complementation and overexpression in *Arabidopsis*

3.1.2 Homologues of *JMJ14* across the different species.

The first *jumonji* (*jmj*) gene was identified by a mouse gene trap strategy (Takeuchi et al., 1995). Subsequently, *jmj* genes that carried the evolutionarily conserved JmjC domain across various species were found (Yang et al., 2010). Klose and collaborators (2006) studied the phylogenetic relationship of JmjC-domain-containing proteins from several organisms by comparing the similarities of protein sequences of JHDMS (JmjC domain-containing histone demethylases). According to the domain architectures, JHDMS were categorised into 7 groups which included JHDM1, PHF/PHF8, JARID, JHDM3/JMJD2, UTX/UTY, JHDM2 and JmjC domain only. The JARID group is further subgrouped into JARID1 and JARID2 (Klose et al., 2006). These JHDM homologues may retain the same protein functions based on their similar domain architectures. In *Arabidopsis*, 21 JmjC domain-containing histone demethylases have been found and named JMJ11-JMJ32 (although no JMJ23 is annotated) (Lu et al., 2008). JMJ14-JMJ19 were identified as members of the JARID1 group by phylogenetic analysis. In *Arabidopsis*, all of them have extensive similarity to human JARID1 family (Lu et al., 2008). Studying the evolutionary relationships between *JMJ14* and its paralogues in *Arabidopsis* and also between *JMJ14* and its orthologues in other species, may contribute to clarifying JMJ14 functions.

3.2 Results

3.2.1 Confirmation of the identification of *lyn1*

3.2.1.1 Comparison between *lyn1* and *jmj14* plant phenotypes

WT, *lyn1* and *jmj14-1* plants were grown together in the same light condition (Figure 3.3). After four and half weeks growth, the differences in phenotype among the three genotypes were distinctly visible. *jmj14-1* has a very similar phenotype to *lyn1* but is significantly different to WT. Figure 3.3 clearly shows

that *jmj14-1* and *lyn1* stem growth began earlier, and as a result I can conclude they both have flowered earlier than WT.



Figure 3.3 Comparison of WT, *jmj14-1* and *lyn1* on the phenotypes. The plants grew in white light ($180 \mu\text{mol m}^{-2} \text{s}^{-1}$) for about four and a half weeks.

3.2.1.2 *jmj14* rescues the light response defect of the *hy1* mutant

3.2.1.2.1 Identification of *jmj14-1 hy1* double mutant

To identify whether *jmj14-1* acts as *lyn1* suppressing the *hy1* mutation, the *jmj14-1 hy1* double mutant was generated by crossing *jmj14-1* with *hy1* to obtain F1 seeds, then F1 plants (*jmj14-1^{+/+} hy1^{+/+}*), with a WT phenotype, self-fertilized to produce the F2 seeds. If the *lyn1* and *jmj14-1* are two different mutations on the *JMJ14* gene, *jmj14-1* should also suppress *hy1* independently. As expected based on Mendel's laws, segregation appeared among the F2 seedlings: some had long hypocotyls and pale cotyledons, like *hy1*, and some had shorter hypocotyls and better greening of the cotyledons compared to *hy1*, like *lyn1 hy1*. To obtain the *jmj14-1 hy1* double mutants, some F2 seedlings were carefully selected and genotyped. During the selection, some seedlings were observed to have a similar phenotype to *lyn1 hy1* (intermediate length of hypocotyls and

greening of cotyledons compared to WT and *hy1*). *jmj14-1 hy1* was expected to have a similar phenotype to *lyn1 hy1*, therefore these seedlings were genotyped. Figure 3.4 shows a sample of genotyping results. The *hy1* genotyping result shows that all the selected plants were *hy1* mutants (Figure 3.4 A) and the *jmj14-1* genotyping result shows that only plant no.1 is a *jmj14-1* mutant and all the other plants (no.2~4) are *jmj14-1* heterozygous (Figure 3.4 B). In summary, the genotype of plant no.1 is identified as *jmj14-1 hy1* double mutant. The genotypes of plants no.2, 3 and 4 are identified as *jmj14-1^{+/+} hy1*.

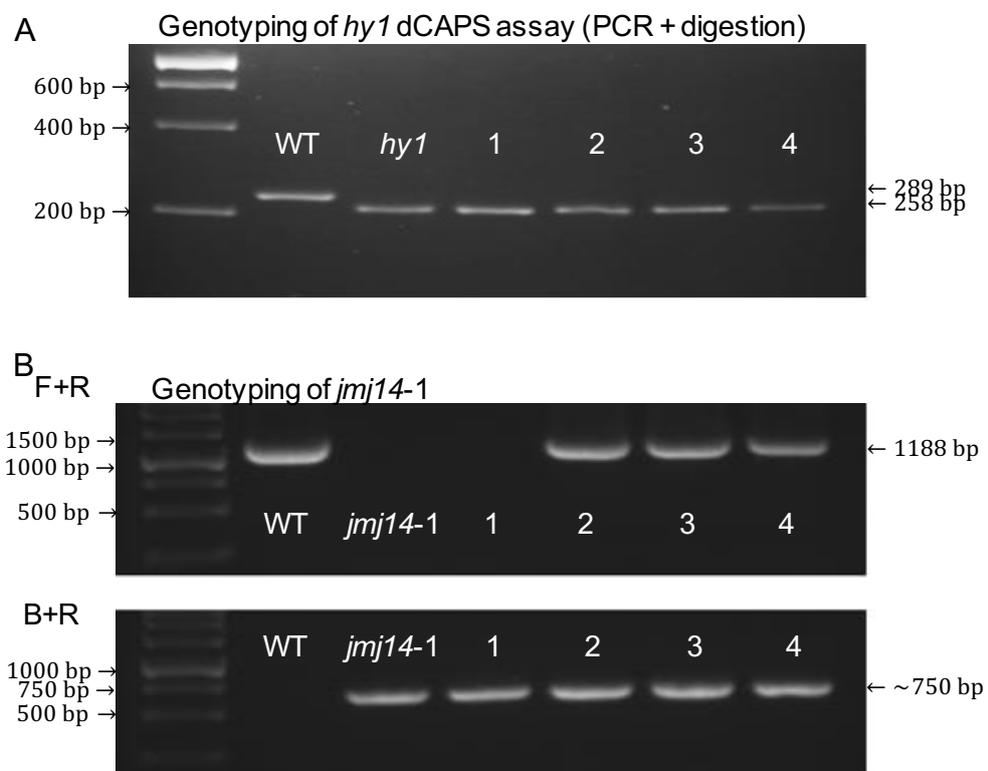


Figure 3. 4 Identification of *jmj14-1 hy1* by genotyping assays

A. dCAPS assay for the identification of *hy1* mutants from *hy1* x *jmj14-1* F2 plants. PCR products of WT, *hy1*, F2 plants no.1~4 were digested with Apol restriction enzyme and separated on 2% agarose gel are shown.

B. Genotyping assay of *jmj14-1* mutants from *hy1* x *jmj14-1* F2 plants. PCR products of WT, *hy1*, F2 plants no.1~4 which were amplified with the forward and reverse primers as well as the border and reverse primers, separated on 1.2% agarose gel are shown.

Plants no.1 and 2 were transferred to soil to grow and their phenotypes were observed after four and half weeks. Plants no.1 and 2 show a significant difference on their phenotype (Figure 3.5 A). Plant no.2 (*jmj14-1^{+/+} hy1*) whose genotype is *hy1* homozygous mutant and *jmj14-1* heterozygous, shows a light response-defective phenotype as *hy1*. Plant no.1 (*jmj14-1 hy1*) whose genotype is homozygous mutant on both *hy1* and *jmj14-1*, has a similar phenotype to *lyn1 hy1*: it exhibits earlier flowering, better growth and greening than plant no.2. *jmj14-1* does show a clearly independent suppression of *hy1* as I expected in the adult plants. The comparisons among six genotypes of 5-day light-grown seedlings, WT, *lyn1*, *jmj14-1*, *hy1*, *lyn1 hy1* and *jmj14-1 hy1*, clearly show that the cotyledons of *lyn1* and *jmj14-1* are larger and exhibit better greening compared to other mutants. *lyn1* and *jmj14-1* have shorter hypocotyls than WT. The cotyledons of double mutants, *jmj14-1 hy1* and *lyn1 hy1*, are larger and exhibit better greening than *hy1*. *jmj14-1 hy1* and *lyn1 hy1* have shorter hypocotyls than *hy1* and longer hypocotyls than WT (Figure 3.5 B). Therefore, the *jmj14-1* independent suppression of *hy1* is not only obvious in the adult plants but also in 5-day light-grown seedlings.

A



B



Figure 3. 5 *j mj14* suppression of seedling phenotypes

A. Differences between the phenotype of F2 *hy1* mutants which are also *j mj14-1* heterozygous (plant 2) and *j mj14-1* homozygous (plant 1) mutant. The plants grew in constant light ($180 \mu\text{mol m}^{-2} \text{s}^{-1}$) for about four and half weeks.

B. Comparison between 5-day light-grown seedlings' phenotypes of WT, *lyn1*, *j mj14-1*, *hy1*, *lyn1 hy1* and *j mj14-1 hy1*. Scale bar=10 mm.

3.2.1.2.2 Hypocotyl length and chlorophyll content

During my previous MSc research, it was demonstrated that *lyn1* rescues light response by suppressing hypocotyl elongation and rescuing chlorophyll content on 5-day light-grown seedlings. This assay was repeated during my PhD research, so that the same assays could also be carried out on *jmj14-1* to see if *jmj14-1* acts like *lyn1* suppressing the *hy1* mutation, in order to confirm the *lyn1* identity. The results are shown in Figure 3.6 and 3.7. WT, *lyn1*, *jmj14-1*, *hy1*, *lyn1 hy1* and *jmj14-1 hy1* were grown in either the dark or continuous white light ($10 \mu\text{mol m}^{-2} \text{s}^{-1}$ and $100 \mu\text{mol m}^{-2} \text{s}^{-1}$) for 5 days. The hypocotyl lengths are compared among 6 genotypes. The statistically significant differences between genotypes were indicated by one-way ANOVA, which is a multiple comparison test. The one-way ANOVA suggests that one or more comparisons are significantly different in medium and high fluence rates, but not in low fluence rate. This was followed by Bonferroni and Holm multiple comparisons T-test, which only pairs relative to a reference group. In the Bonferroni and Holm multiple comparisons, the single mutants, *lyn1*, *jmj14-1* and *hy1*, were compared to WT and the double mutants, *lyn1 hy1* and *jmj14-1 hy1*, were compared to *hy1*. All 6 genotypes showed very similar hypocotyl length in the dark (Figure 3.6 A), and the statistical test indicates no significant difference among 6 genotypes. The dark-grown seedlings of all 6 genotypes always had longer hypocotyl than their respective light-grown seedlings. The light-grown seedlings hypocotyl length decrease when light fluence rate increases. In both light fluence rates, *hy1* has a longer hypocotyl length than WT and the difference between *hy1* and WT is significant (Figure 3.6 B and C). Except between *hy1* and WT, no other difference is significant in low light fluence rate (Figure 3.6 B). In contrast, in high light fluence rate, *lyn1* and *jmj14-1* have shorter hypocotyl length than WT, while *lyn1 hy1* and *jmj14-1 hy1* have shorter hypocotyl length than *hy1* (Figure 3.6 C). All the differences are

significant in high fluence rate. In summary, rescue of hypocotyl length by *lyn1* was only observed and proven to be significant in high fluence rate.

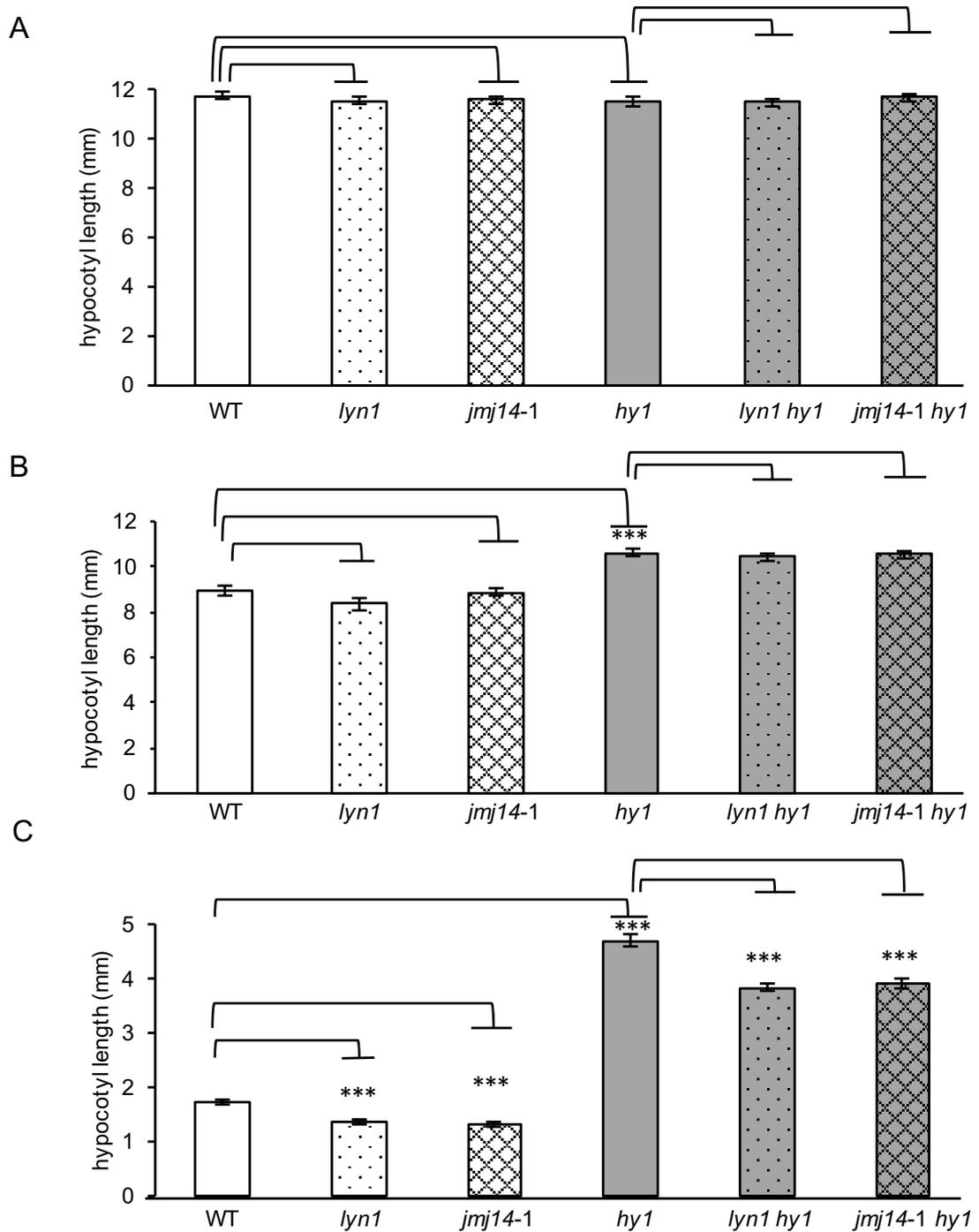


Figure 3. 6 *jmj14* effect of hypocotyl length of 5-day seedlings in different light fluence rate

A. Seedlings grew in the darkness, n=30;

B. Seedlings grew in 10 $\mu\text{mol m}^{-2} \text{s}^{-1}$ continuous white light, n=30;

C. Seedlings grew in 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$ continuous white light, n=30;

Whether *jmj14-1* acts like *lyn1* rescuing the greening response was also studied by measuring the chlorophyll content of seedlings. The same comparisons and statistical tests used in analysis of hypocotyl length suppression were also carried out in this assay. Because seedlings cannot produce chlorophyll in the dark, the chlorophyll content was only studied on 5-day light-grown seedlings. WT, *lyn1*, *jmj14-1*, *hy1*, *lyn1 hy1* and *jmj14-1 hy1* were grown in continuous white light (10 $\mu\text{mol m}^{-2} \text{s}^{-1}$ and 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$) for 5 days. Chlorophyll was extracted from 20 seedlings for each biological replicate. The total chlorophyll content per seedling of 6 genotypes is shown in Figure 3.7. Chlorophyll contents increase when the light fluence rate increases in all genotypes. In both light fluence rates, relative to the WT, chlorophyll content of *hy1* is lower while those of *lyn1* and *jmj14-1* are higher. *lyn1 hy1* and *jmj14-1 hy1* have higher chlorophyll content relative to *hy1*. The differences in all the comparisons are significant. In summary, *lyn1* rescue of chlorophyll content is significant in both light conditions.

The quantifications of hypocotyl length and chlorophyll content both indicate that *jmj14* can independently suppress growth and greening defect of *hy1*. However, the suppression is not obvious in all light fluence rates. In the higher fluence rate, the effects of *lyn1* suppressing hypocotyl length and chlorophyll content are always significant. Under the low fluence rate, *jmj14* rescue of the light response is more substantial in increasing chlorophyll content but no more so suppressing hypocotyl elongation. Because of this observation, in accordance with the expectation that *jmj14-1* would independently suppress *hy1*, I determine that *jmj14-1* and *lyn1* are very likely to be two mutations on the same gene.

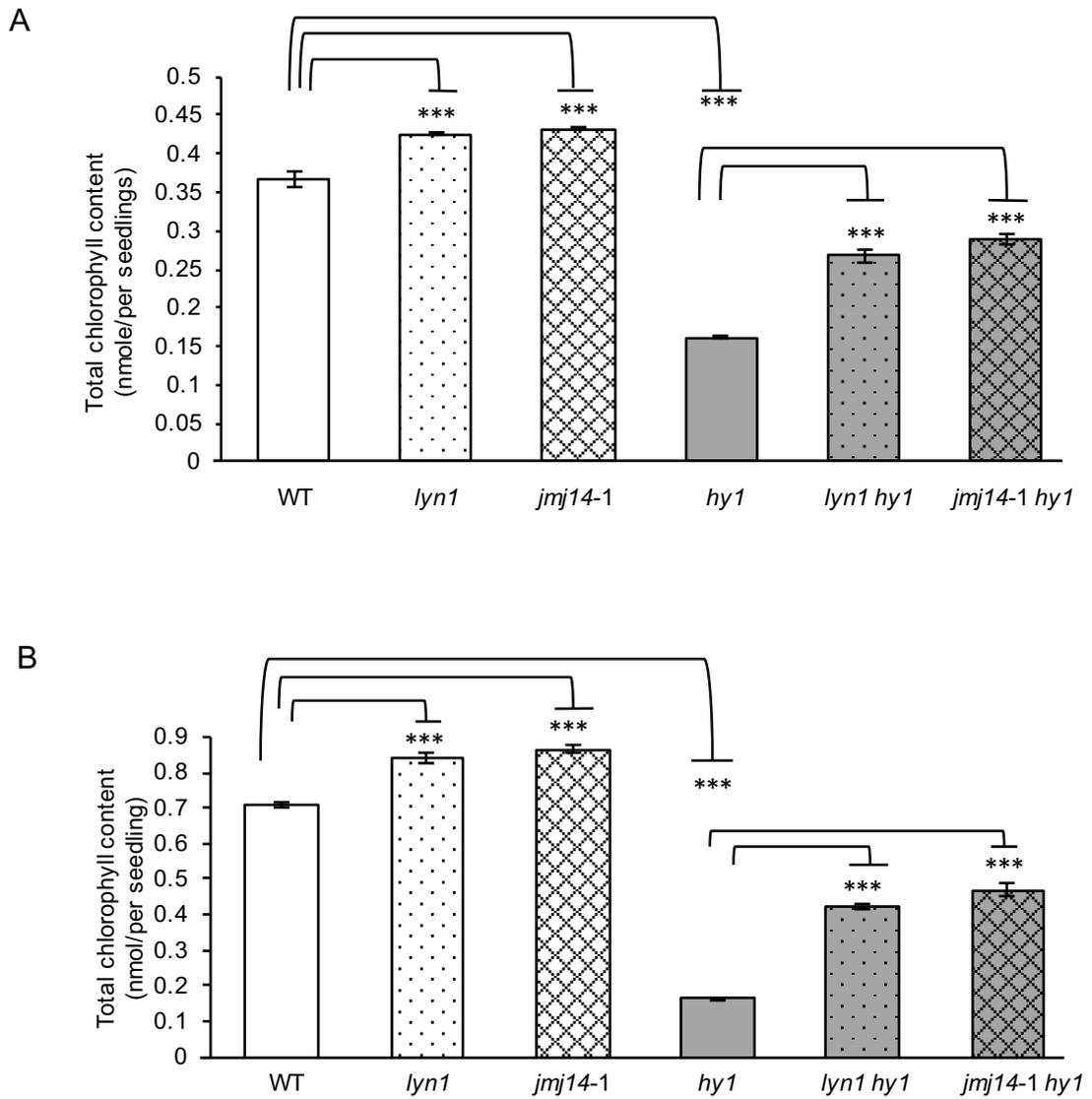


Figure 3.7 *jmj14* effect of chlorophyll content of 5-day seedlings in different light fluence rate

A. Seedlings grew in $10 \mu\text{mol m}^{-2} \text{s}^{-1}$ continuous white light, $n=3$;

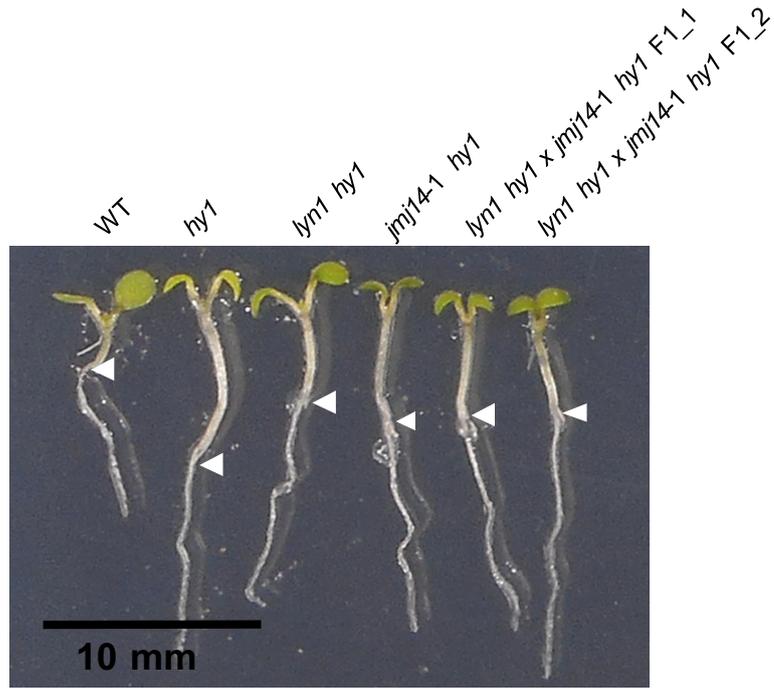
B. Seedlings grew in $100 \mu\text{mol m}^{-2} \text{s}^{-1}$ continuous white light, $n=3$.

3.2.1.3 Failure of complementation of *lyn1* by *jmj14-1*

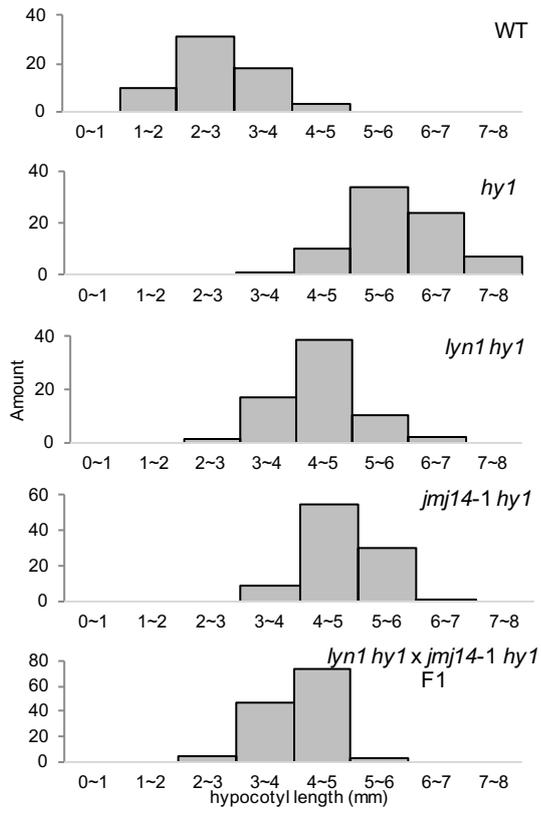
Although *jmj14-1* has been shown to act like *lyn1* independently suppressing *hy1*, just comparing the flowering time, hypocotyl length and chlorophyll content are not sufficient to confirm the *lyn1* identity. However unlikely, it remains possible that *lyn1* and *jmj14-1* are two independent mutations both capable of suppressing

hy1. On the other hand, given that *lyn1* and *jmj14-1* are recessive mutations, a mutant carrying one allele of *lyn1* and one of *jmj14-1* (a double heterozygous) would, if both are alleles of the same gene, carry no functional allele and therefore be a homozygous mutant. This would be genetic confirmation of gene identity. A failure of complementation test was carried out by crossing *lyn1 hy1* with *jmj14-1 hy1* and then the hypocotyl length of F1 seedlings was measured and compared to WT, *hy1*, *lyn1 hy1* and *jmj14-1 hy1*. If *jmj14-1* and *lyn1* were two mutations on the same gene, then *jmj14-1* should fail to complement *lyn1*. Thus, the F1 plants should have same phenotype as *lyn1 hy1* and *jmj14-1 hy1*. The hypocotyl length was quantified on 5-day light grown seedlings (Figure 3.8 A). The results are presented by the frequency distribution of hypocotyl length, showing that the most frequent hypocotyl length of WT and *hy1* seedlings is about 2~3 mm and 5~6 mm, respectively (Figure 3.8 B). The F1 seedlings have similar hypocotyl length to *lyn1 hy1* and *jmj14-1 hy1* which are about 4~5 mm. The comparison among young plants (Figure 3.8 C) also shows that F1 plants have a similar phenotype to *lyn1 hy1* and *jmj14-1 hy1*. Genotypes of 6 seedling pools were confirmed and are shown in Figure 3.8 D. This observation in which, as expected, F1 seedlings of a cross of *lyn1 hy1* and *jmj14-1 hy1* present similar hypocotyl length to *lyn1 hy1* and *jmj14-1 hy1*, this failure of complementation outcome, confirms that *jmj14-1* and *lyn1* are two mutant alleles of the same gene.

A



B



C



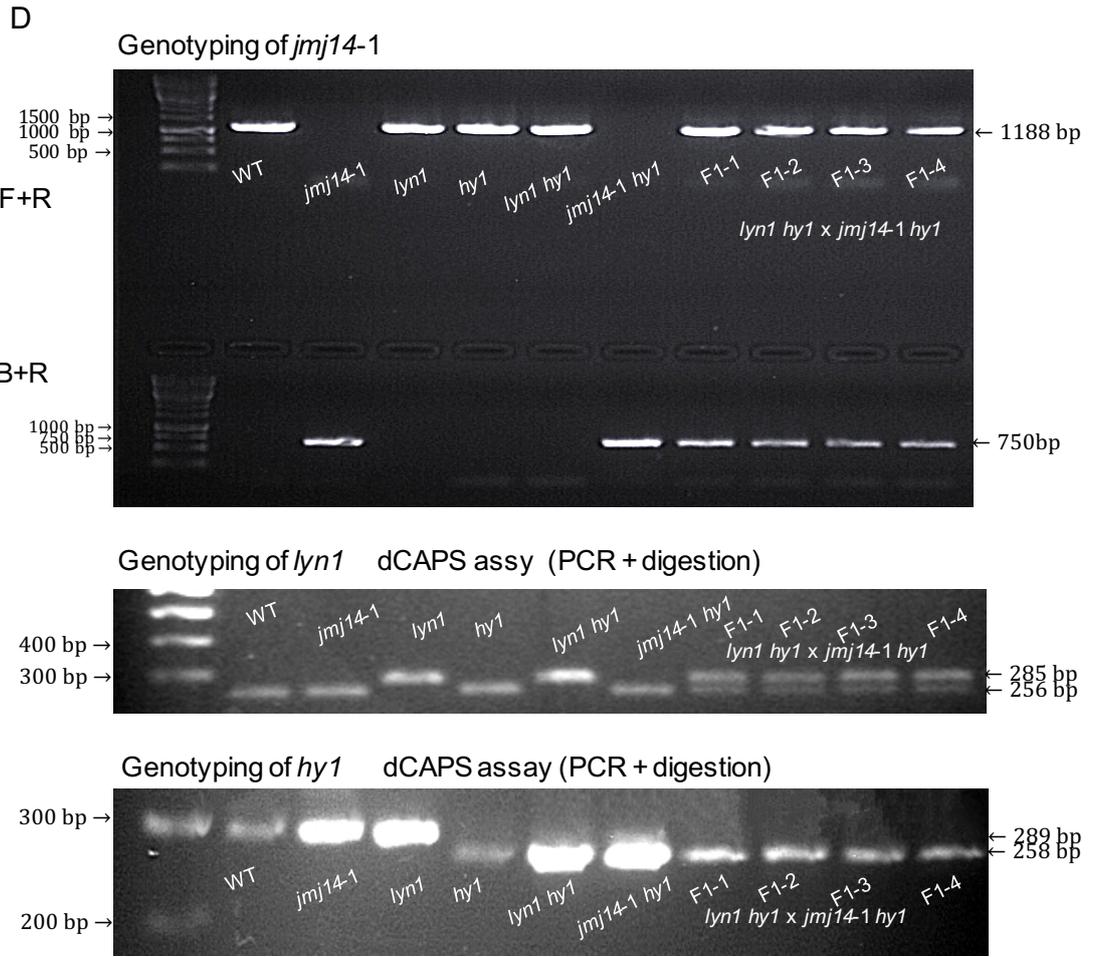


Figure 3. 8 Confirmation that *jmj14-1* fails to complement *lyn1* suppressing hypocotyl length

A. Comparison between 5-day light-grown seedlings' phenotypes of WT, *hy1*, *lyn1 hy1*, *jmj14-1 hy1* and F1 (*lyn1 hy1 x jmj14-1 hy1*). Scale bar=10 mm.

B. The figure shows the comparison of 5-day light-grown seedling hypocotyl lengths among WT, *hy1*, *lyn1 hy1*, *jmj14-1 hy1* and F1 plants. All the seedlings were grown in $100 \mu\text{mol m}^{-2} \text{s}^{-1}$ continuous white light. Hypocotyl lengths were categorised into 8 groups and shown in different bins on the x-axis. The Y-axis shows the frequencies of hypocotyl length. 31% of WT seedling hypocotyl lengths are between 2~3 mm. 34% of *hy1* seedling hypocotyl lengths are between 5~6 mm. 39% of *lyn1 hy1* seedling hypocotyl lengths are between 4~5 mm. 55% of *jmj14-1 hy1* seedling hypocotyl lengths are between 4~5 mm. 74% of F1 seedling hypocotyl lengths are between 4~5 mm.

C. The image shows 20-day light-grown WT, *hy1*, *lyn1 hy1*, *jmj14-1 hy1* and F1 plants in constant white light ($180 \mu\text{mol m}^{-2} \text{s}^{-1}$).

D. The top image shows the *jmj14-1* genotyping result. PCR products of WT, *jmj14-1*, *lyn1*, *hy1*, *lyn1 hy1*, *jmj14-1 hy1* and F1 seedlings were amplified by the forward and reverse primers and the border and reverse primers and separated on 1.2% agarose gel are shown. The middle image shows *lyn1* dCAPS genotyping assay result. PCR products of WT, *jmj14-1*, *lyn1*, *hy1*, *lyn1 hy1*, *jmj14-1 hy1* and F1 seedlings digested with HaeIII restriction enzyme and separated on 2% agarose gel are shown. The bottom image shows *hy1* dCAPS genotyping assay result. PCR products of WT, *jmj14-1*, *lyn1*, *hy1*, *lyn1 hy1*, *jmj14-1 hy1* and F1 seedlings digested with ApeI restriction enzyme and separated on 2% agarose gel are shown.

3.2.1.4 Complementation of *lyn1* mutants by transferring functional *JMJ14*

3.2.1.4.1 Failure of BP recombination

The initial cloning strategy was to directly generate an *attB*-flanked *JMJ14* PCR product using a pair of designed long primers (Forward primer: *attB1* + *JMJ14* full-length cDNA forward primer; Reverse primer: *attB2* + *JMJ14* full-length cDNA reverse primer). However, the recombination always failed, and no colony appeared on the antibiotic-selective agar plate. As the long primers are initially only semi-hybridised to DNA template, they cannot fully bind to DNA at the PCR annealing stage. The amplification becomes less efficient and a large amount of primer dimers are produced. One possible reason for failure of recombination is that the large amount of primer dimers competes with the small amount of PCR products, resulting in a very low probability of successful insertion of PCR product into the donor vector. DNA gel extraction was used to remove the primer dimers from PCR products before using the PCR product for gateway BP recombination reaction. The second possible reason of failure recombination is that the UV exposure caused DNA cross-linking during gel extraction which interfered with the recombination. Use of such gel purification was unavoidable, since most primer dimers cannot be simply removed if the PCR products only go through end-point PCR amplicon purification. In summary, no recombinant product was produced according to the initial cloning strategy. Given this failure, a new strategy was designed (shown in section 3.1.1.3).

3.2.1.4.2 Ligation of *JMJ14* cDNA into pGEM-T Easy vector

The *JMJ14* cDNA PCR products which are 2865 bp were inserted into pGEM-T Easy vector first. pGEM[®]-T Easy vector is 3015 bp and contains an Ampicillin resistance gene (Amp) (Appendix Figure 7.4). Because pGEM[®]-T Easy vectors are linearized vectors with a single 3'-terminal thymidine at both ends, the DNA fragments must have 3'-terminal adenosine at both ends in order to ligate.

However, not all thermostable polymerases generate 3'A-tailed PCR products. *Pfu* DNA Polymerase generates blunt-ended fragments and pGEM-T can only ligate with a fragment with 3'A-overhangs (Knoche and Kephart, 1999). Thus, adenosines were added at both ends of the *JMJ14* cDNA fragments using the A-tailing procedure. The A-tailed DNA fragments were ligated with vectors and colonies were selected on antibiotic selective LB plates. The identification of successful insertion was indicated by the blue/white screening on indicator plates. Sixteen white colonies were selected from the blue/white screening plate and then genotyped to find the correct insertion by colony PCR. The inserts were genotyped with *JMJ14* full-length primers. The result shows that colony no.5 carries the full-length *JMJ14* cDNA (Figure 3.9).

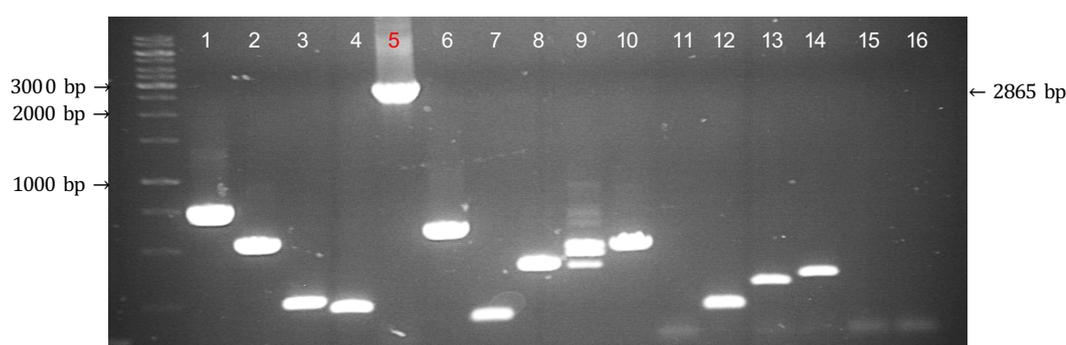


Figure 3. 9 Colony PCR for identification of full-length *JMJ14* cDNA insertion in the pGEM-T Easy vector

Colony PCR among 16 colonies shows colony no. 5 carries a fragment which has a size matching the expected one (2865 bp) of full-length *JMJ14* cDNA.

Although mutations are unlikely to be introduced during PCR amplification when using *Pfu* polymerase, the insert was sequenced to confirm that it was a full-length *JMJ14* without any sequence errors. Colony 5 was grown as a larger *E. coli* culture and then extracted by Qiagen spin miniprep kit to generate the plasmid minipreps for sequencing. Before sending for commercial Sanger sequencing, these plasmid minipreps were confirmed to be carrying full-length *JMJ14* by PCR (Figure 3.10). 15 μ l of 70 μ g/ml plasmid miniprep was sent to

Eurofins for sequencing. The insert was sequenced using 4 sequencing primers. Two primers which were provided by Eurofins annealed against the T7 promoter and the SP6 inverted promoter sequences, respectively. The other two primers (*JMJ14* sequencing primer 2 and 3) were designed in the middle region of *JMJ14*. The sequencing result confirmed the insert is *JMJ14*. However, the result shows that the first nucleotide (Adenosine) of the start codon (ATG) of the complete *JMJ14* cDNA is missing. Fortunately, this error was easily corrected in the next step when the gateway *attB* sites are added to the *JMJ14* cDNA during the PCR amplification (the codon is part of the primer).

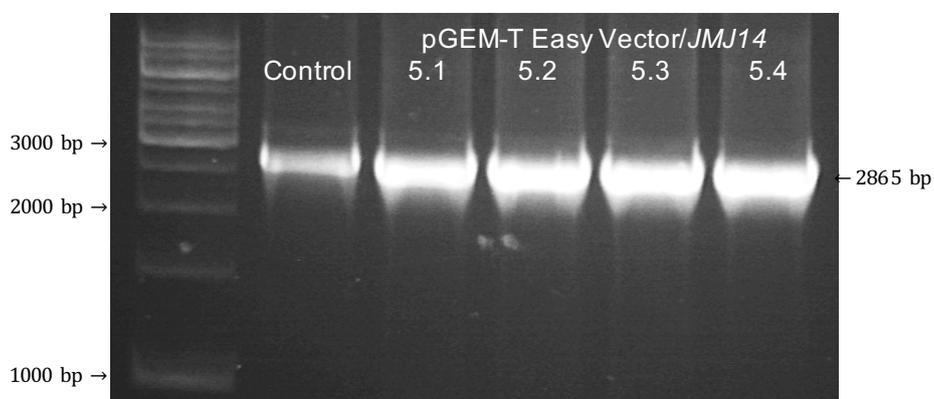


Figure 3. 10 Confirmation of correct insertion of *JMJ14* fragment in pGEM-T Easy vector by PCR

All the PCR products have a similar size (2865 bp) as full-length *JMJ14* cDNA.

3.2.1.4.3 Generation of *JMJ14* cDNA fragments with Gateway *attB* sites

The pGEM-T Easy Vector/*JMJ14* construct was used as a DNA template to generate a large quantity of *JMJ14* amplicons. To add the *attB* sites to both ends of the *JMJ14* cDNA sequence, a pair of long primers containing the *attB* sites and around 20 nucleotides of either the 5' or 3' ends of the *JMJ14* cDNA sequence were used. These *attB*-flanked PCR fragments I generated allow the recombination between the *attB*-flanked *JMJ14* PCR products and the donor vectors during the BP reaction. Clearly, the gel image shows that PCR products

have a stronger band of *JMJ14* and less obvious primer dimers (Figure 3.11). The *attB*-flanked full-length *JMJ14* cDNA PCR products are slightly longer than *JMJ14* cDNA.

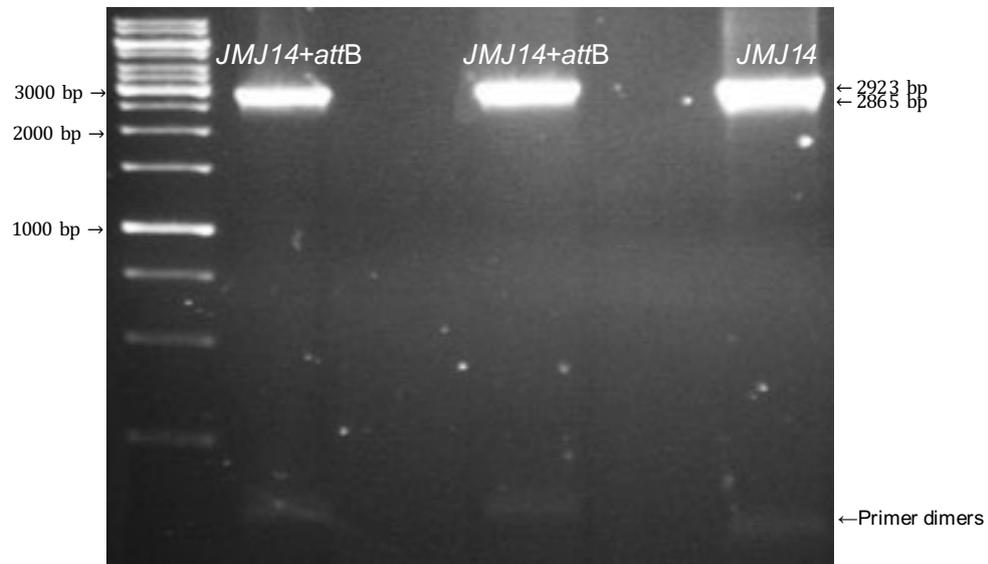


Figure 3. 11 Gel image of PCR products of *attB*-flanked *JMJ14* and *JMJ14* amplified from the pGEM-T Easy vector/*JMJ14* plasmid template

The *attB*-flanked *JMJ14* fragments were generated by amplifying *JMJ14* with *attB*-adapted *JMJ14* primers. The *attB*-flanked *JMJ14* PCR products were separated on the 1% agarose gel. After adding the *attB* sites, the *attB*-flanked *JMJ14* PCR products are 2923 bp which is 58 bp larger than the control. Few primer dimers were produced.

3.2.1.4.4 Gateway™ vector-related works

3.2.1.4.4.1 Transformation into Gateway™ Donor vector p201

Once the *attB*-flanked *JMJ14* PCR products were generated, they needed to be inserted into a donor vector. Donor vector pDONR™201 was used for transferring *JMJ14* to a destination vector. pDONR™201 is 4470 bp and has *attP* recombination sites with Kanamycin (Kan) resistance (Appendix Figure 7.5). *JMJ14* fragments were ligated with pDONR vector by Gateway BP reaction. The recombination of *attB* sites of *JMJ14* fragment and *attP* sites of pDONR vector

generates *attL* sites that transfers *JMJ14* into an entry clone. The recombinant products were multiplied by transformation into *E. coli*. Sixteen colonies were selected from Kan LB plates and genotyped to identify the colonies with the correct insert by colony PCR. The inserts were genotyped with *JMJ14* full-length cDNA primers. The result shows that the recombinants of colony 2 and 5 probably carry the correct inserts which have a similar size to *JMJ14* (Figure 3.12).

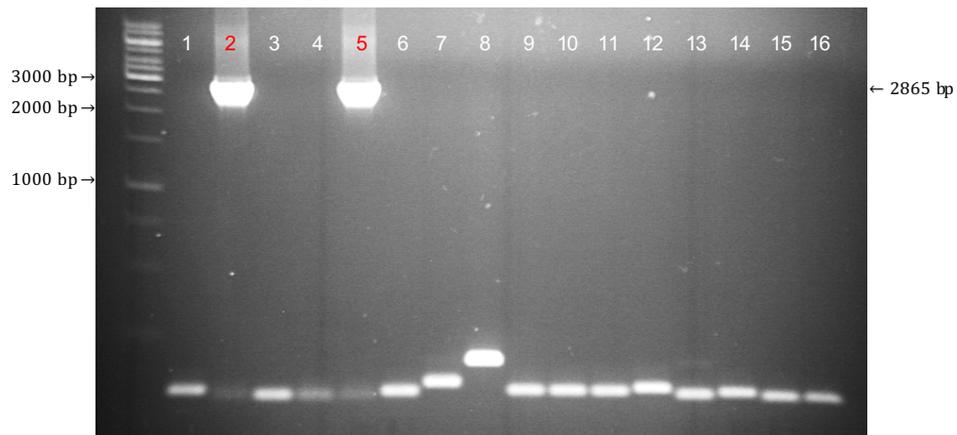


Figure 3. 12 Colony PCR for identification of full-length *JMJ14* cDNA insertion in the pDONR201 vector

Colony PCR among 16 colonies shows colony no.2 and 5 carry fragments which have a similar size (2865 bp) to full-length *JMJ14* cDNA.

The colonies 2 and 5 were both cultured in Kan-containing LB to obtain larger *E. coli* cultures for plasmid DNA isolation (miniprep). The isolated pDONR201/*JMJ14* plasmid constructs from colonies 2 and 5 were extracted from the cultures to be sequenced and to perform the LR reaction in the later step. The plasmid minipreps were run with empty pDONR vector on 1% agarose gel to compare the plasmid size (Figure 3.13 A). The figure indicates that a fragment is inserted into the pDONR™201 vector of both colonies 2 and 5, because the super coiled plasmid 2 and 5 show a larger size compared to the super coiled empty pDONR™201 vectors. Figure 3.13 B indicates that an insert which is about 2865 bp is present in the plasmid of clones 2 and 5. Plasmid no.5 was selected and

sent for sequencing, targeting the insert using 4 designed sequencing primers. Two primers are on pDONR™201 vector close to the *attL* recombination sites and the other two primers (*JMJ14* sequencing primer 2 and 3) are in the middle region of *JMJ14*. Sequencing results showed that no errors had been introduced and that the recombination reaction had inserted the fragment correctly and in the correct orientation. The error by which the first nucleotide of the start codon was missing had been rectified.

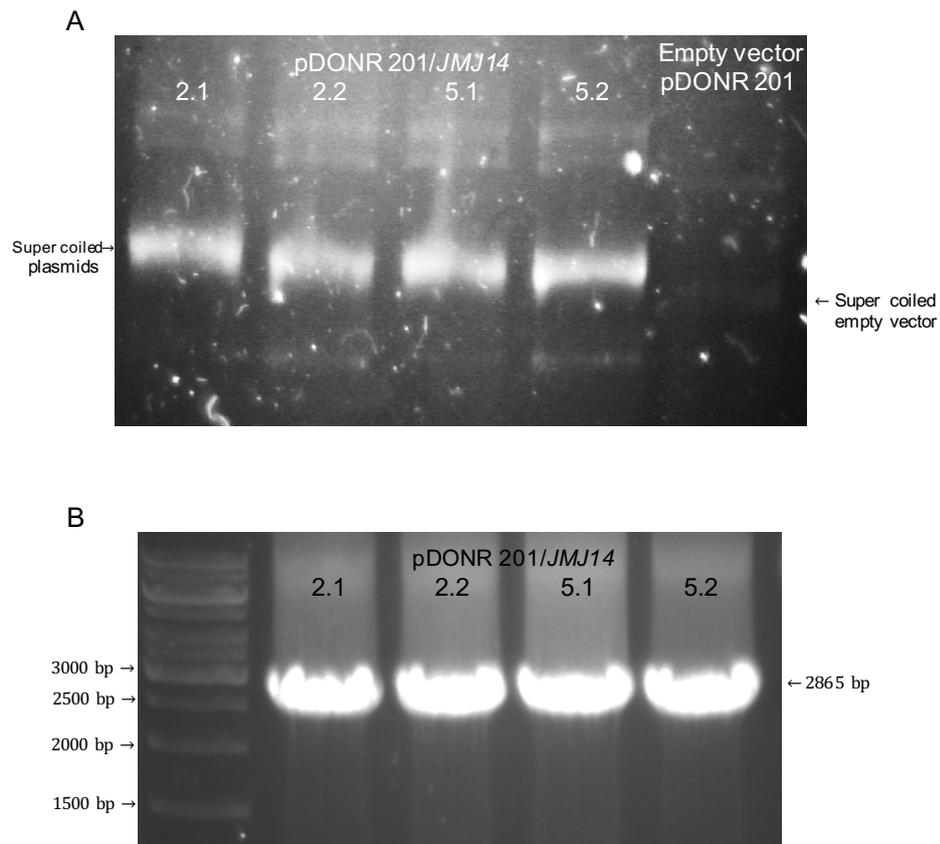


Figure 3. 13 Confirmation of correct insertion of *JMJ14* fragments in the entry clones

A. Comparison of the size between plasmids no.2 and 5 and empty pDONR™201 vectors on 1% agarose gel.

B. Confirmation of correct insertion of *JMJ14* fragments in the pDONR201 vector by PCR. All the PCR products have a similar size (2865 bp) to full-length *JMJ14* cDNA.

3.2.1.4.4.2 Transformation into Gateway™ Destination vector pB2GW7

Once the BP recombination happened, a pair of *attL* sites were generated in the entry clone, which could, in turn, recombine with the *attR* sites of the destination vector pB2GW7 used in this research. Destination vector pB2GW7 is 10882 bp long, contains the cauliflower mosaic virus 35S coat protein promoter (CaMV 35S, p35S) upstream of the *attR* recombination sites, as well as 35S terminator (T35S) and a Spectinomycin (Sp) resistance gene (Appendix Figure 7.6). *JMJ14* fragments were exchanged from entry clones into destination vectors by Gateway LR recombination reaction, resulting in the generation of a plant expression construct. The *JMJ14* cDNA was cloned under the control of the 35S promoter for constitutive expression in plant cells after transformation. The recombinant vector was cloned and multiplied by introducing into *E. coli*. 8 colonies were selected from the Sp LB plates and genotyped to confirm the correct insertion by colony PCR. All the colonies were genotyped using *JMJ14* full-length primers. The result shows that all recombinants carried the correct inserts which were about 2865 bp (Figure 3.14 left). All the colonies were also genotyped using pB2GW7 sequencing primers (one is on the pB2GW7 35S promoter and the other one is on the 35S terminator). These PCR products are slightly larger than full-length *JMJ14* (Figure 3.14 right). The *JMJ14* cDNA in pB2GW7 was not sequenced at this step, because this step is to transfer the *JMJ14* cDNA by recombination reaction between the two vectors without PCR amplification. Mutations are most likely to occur during PCR amplification. The *JMJ14* sequence was already confirmed when *JMJ14* cDNA was present within pDONR201 before transferring into pB2GW7. Therefore, the *JMJ14* inserts were at this point trusted to have the correct sequence.

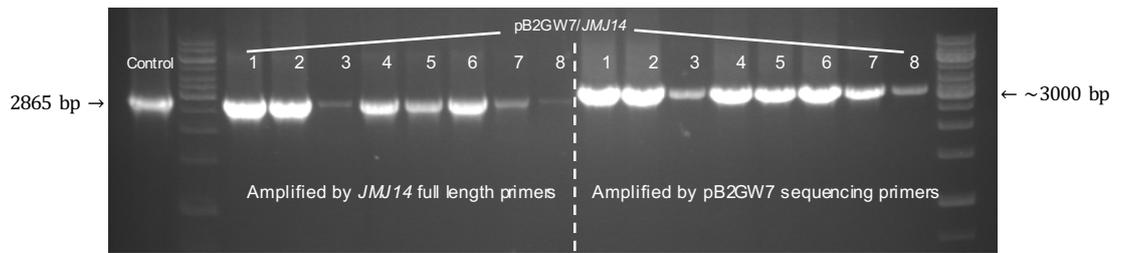


Figure 3. 14 Colony PCR for identification of full-length *JMJ14* cDNA insertion in the pB2GW7 vector

8 colonies were amplified by *JMJ14* full-length primers (left) and pB2GW7 sequencing primers (right). The PCR products which are amplified by *JMJ14* full-length primers have a similar size (2865 bp) to full-length *JMJ14*. The PCR products which are amplified by pB2GW7 sequencing primers have a slightly larger size (~3000 bp) than full-length *JMJ14*.

The colonies 1, 4 and 5 were cultured in Sp LB to obtain larger amounts of *E. coli* cultures. They were extracted to generate the pB2GW7/*JMJ14* plasmid minipreps for transferring into *Agrobacterium*. The insertions of pB2GW7/*JMJ14* in the plasmid minipreps were double-confirmed using *JMJ14* full-length primers (Figure 3.15 A left) and pB2GW7 sequencing primers (Figure 3.15 A right). pB2GW7 empty vector was used as control of the pB2GW7/*JMJ14* plasmid in 1% agarose gel electrophoretic separations, to compare the plasmid size (Figure 3.15 B).

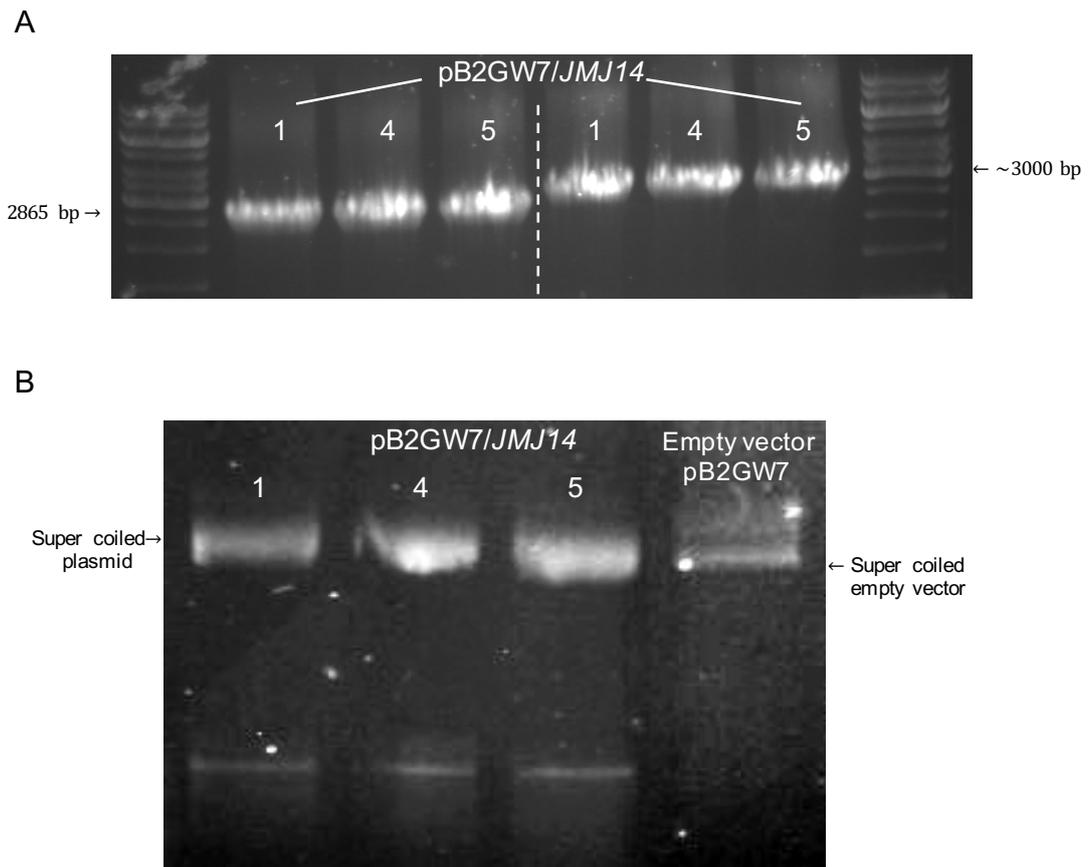


Figure 3. 15 Confirmation of correct insertion of *JMJ14* in the plant expression vector

A. PCR amplification of an inserted full-length *JMJ14* fragment in the pB2GW7 vector by *JMJ14* full-length primers (left) and pB2GW7 sequencing primers (right). The PCR products amplified using *JMJ14* full-length primers have a similar size (2865 bp) to full-length *JMJ14* cDNA. The PCR products amplified using pB2GW7 sequencing primers have a slightly larger size (~3000 bp) than full-length *JMJ14* cDNA.

B. Comparison of the size between expression clone and empty pB2GW7 vector on 1% agarose gel.

3.2.1.4.5 Transforming and screening of plants

3.2.1.4.5.1 *Agrobacterium*-mediated transformation into plants

The pB7WG2/*JMJ14* plasmid miniprep of colony 1 was selected to transform *Agrobacterium tumefaciens* (GV3101). Eight colonies were selected and genotyped to confirm the correct insertion by colony PCR. All the colonies were genotyped by *JMJ14* full-length primers. The result shows that all *Agrobacterium* colonies carried the correct inserts (Figure 3.16). They were all individually cultured in a larger volume of LB, then mixed with glycerol to make stocks and store at -80°C until the plants were ready for infiltration.

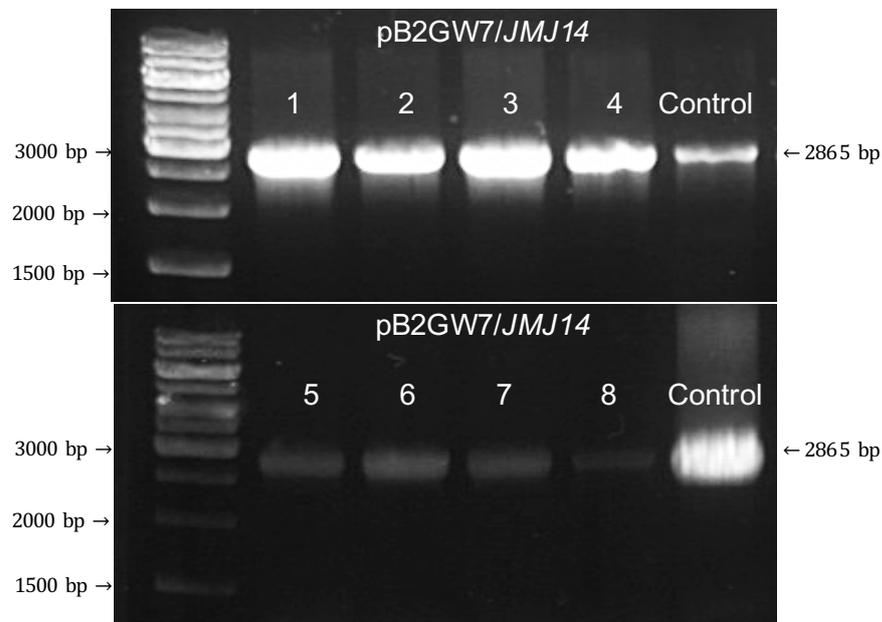


Figure 3. 16 Colony PCR confirming correctly inserted *JMJ14* in the expression vector

All the colonies were amplified by *JMJ14* full-length primers. They all can amplify full-length *JMJ14*.

The *Agrobacterium*-infiltration solution was prepared by mixing 4 different strains of *Agrobacterium* colony stocks. This is to avoid unsuccessful transformation caused by potential mutations carried in any single colony strain. The infiltrations were performed on *lyn1 hy1* plants at early stages of flowering (when their

flowering stems were about 10-15 cm tall). They were repeated four times on the same plants (T0). The T1 seeds were collected from these T0 plants and then grown on antibiotic-containing plates for screening.

3.2.1.4.5.2 Screening and verification of complemented plants

The T1 plants were selected by growing T1 seeds on phosphinothricin and cefotaxime-containing agar plates. Phosphinothricin (PPT) is a herbicide which kills the untransformed plants. Cefotaxime is an antibiotic which eliminates any remaining *Agrobacterium* (Okkels and Pedersen, 1988). Eight T1 plants survived on the selective agar plates. These eight plants were transferred into soil and grown with *hy1* and *lyn1 hy1* plants side by side. An individual T1 plant is shown in Figure 3.17 as a sample, with *hy1* and *lyn1 hy1* plants as controls. If *JMJ14* can complement *lyn1* defect in T1 plants, then T1 plants should appear *hy1* phenotype. As expected, all (8/8) T1 plants showed a *hy1* phenotype instead of a *lyn1 hy1* phenotype and no significant difference of phenotype among T1 plants was observed.

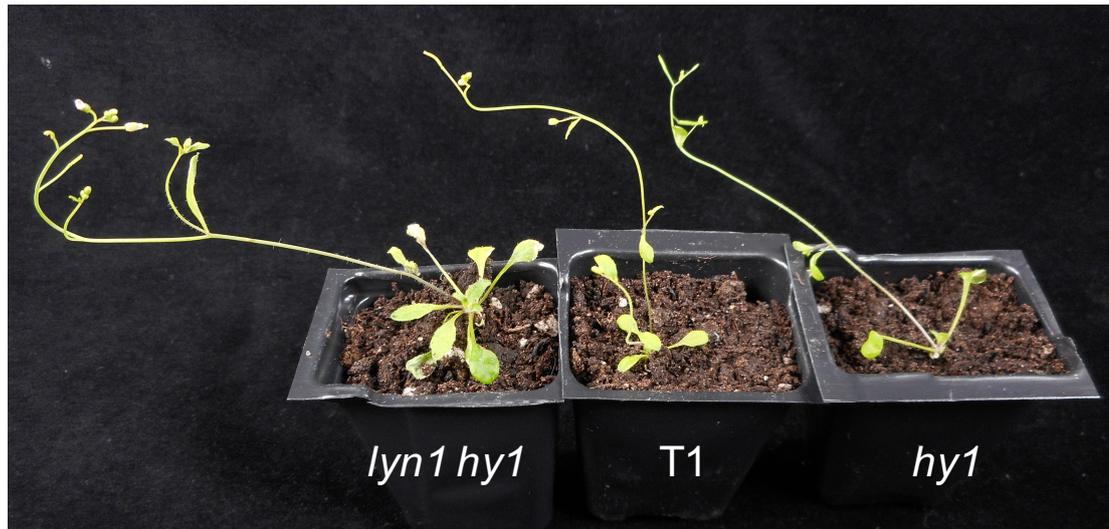


Figure 3. 17 Comparison of phenotype between *lyn1 hy1*, T1 and *hy1*

The plants were grown in ($100 \mu\text{mol m}^{-2} \text{s}^{-1}$) constant light for four and half weeks.

To further confirm the presence of inserts, pB2GW7 sequencing primers were used to amplify the inserts in the genome of T1 plants. The PCR result shows that all the plants carried the correct inserts (Figure 3.18 A). To double check that the genotype of these transformed T1 plants was *LYN1 overexpressor* in a *lyn1 hy1* background (*LYN1ox lyn1 hy1*), PCR genotyping was done using the *lyn1* dCAPS primers. If these T1 plants are *lyn1 hy1* while also carrying one or more copies of inserted functional *JMJ14* cDNA, then they should show that they are *lyn1* heterozygous mutants. As expected, all T1 transformed plants showed that they are *lyn1* heterozygous mutants and *hy1* homozygous mutants (Figure 3.18 B). Transferring functional *JMJ14* gene into *lyn1 hy1* mutants can complement the *lyn1* suppression, so this transgenic complementation assay confirms that *lyn1* is a mutation on the *JMJ14* gene.

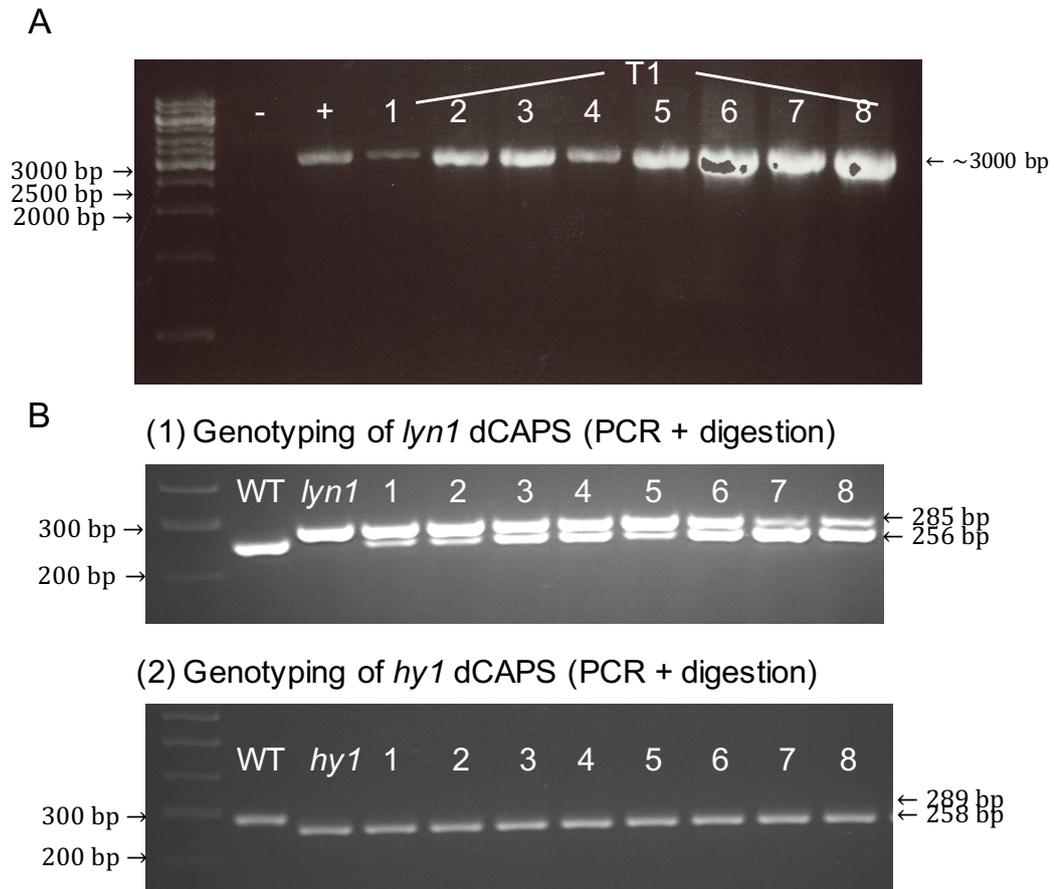


Figure 3. 18 Confirmation of the present of *JMJ14* inserts in the T1 plants

A. PCR amplification of inserts by pB2GW7 sequencing primers. All the T1 plants contain the inserts which are about 3000 bp. The negative control does not contain the insert.

B. Confirmation of the genotype of T1 plants. (Top) *lyn1* dCAPS genotyping assay. PCR products of T1 plants were digested with *HaeIII* restriction enzymes and separated on 2% agarose gel are shown. (Bottom) *hy1* dCAPS genotyping assay. PCR products of T1 plants were digested with *ApoI* restriction enzymes and separated on 2% agarose gel are shown.

3.2.2 Detecting *JMJ14* mRNA in *jmj14-1* and *lyn1* mutants

Although *lyn1* and *jmj14-1* are located on the same gene, they are two different mutant alleles that lead to different mRNA and protein being synthesised. mRNA and proteins may or may not be synthesised according to different types of mutation. If *JMJ14* protein is partly translated and retains some of its original functions in mutants, then it can lead to a reduced suppression of *hy1* than a null mutant would, and as a consequence retention of a decreased light response. To analyse whether *lyn1* and *jmj14-1* retain the *JMJ14* proteins, RT-PCR was carried out. Instead of detecting the protein by Western blotting, RT-PCR detects the presence of mRNA in the mutants. *jmj14* mRNA was reverse transcribed to *jmj14* cDNA and then amplified by PCR. The RNA and genomic DNA were extracted from WT, *lyn1*, *jmj14-1*, *lyn1 hy1* and *jmj14-1 hy1*. Figure 3.19 A shows that all the plants were of the correct, expected genotypes. The RNA samples were reverse transcribed into cDNA, on which *JMJ14* full-length primers were used to amplify *JMJ14* by PCR (Figure 3.19 B). If any *JMJ14* cDNA was present, then the PCR products should show a band on the gel. Figure 3.19 B shows that *jmj14-1* does not contain any full-length *jmj14* cDNA, confirming its knockout character. Unexpectedly, *lyn1* does contain full-length *jmj14* cDNA, in spite of the similarity of the phenotype between *lyn1* and *jmj14-1*.

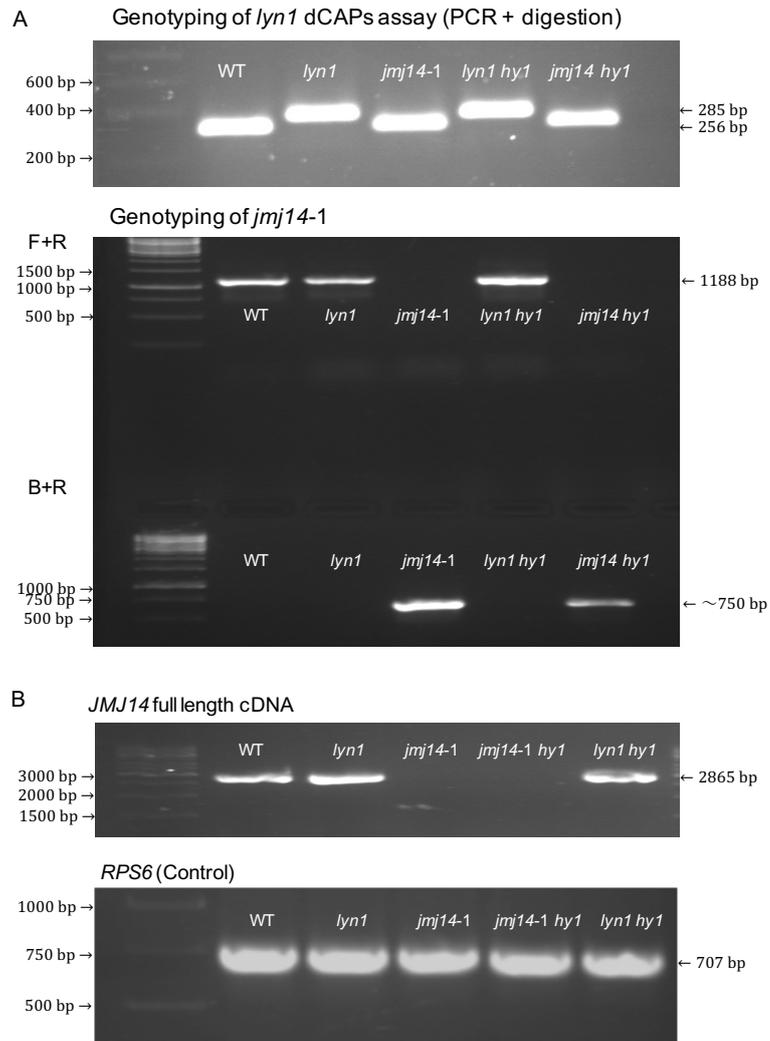


Figure 3. 19 RT-PCR of full-length *JMJ14* revealed that transcription of *JMJ14* was eliminated in *jmj14-1* mutants but not in *lyn1* mutants

A. Confirmation of the genotype of the plants which were used for RT-PCR of full-length *JMJ14* cDNA. The top image shows the *lyn1* dCAPS assay result. PCR products of WT, *lyn1*, *jmj14-1*, *lyn1 hy1* and *jmj14-1 hy1* plants were digested with *HaeIII* restriction enzymes and separated on 2% agarose gel are shown. The bottom image shows the *jmj14-1* genotyping result. PCR products of WT, *lyn1*, *jmj14-1*, *lyn1 hy1* and *jmj14-1 hy1* plants were amplified by the forward and reverse primers and the border and reverse primers and separated on 1.2% agarose gel are shown.

B. RT-PCR with *JMJ14* full-length primers. *RPS6* was used as an internal control.

3.2.3 Structure, expression and evolutionary relationships of JMJ14

Understanding the function of JMJ14 is the purpose of this research. Although the function of JMJ14 has been determined by several studies to be a histone demethylase in *Arabidopsis*, the biological process which JMJ14 is involved in, in order to rescue the light response, is still unclear. Analysing the evolutionary relationships of JMJ14, and the presence of homologues in other species, may uncover the range of organisms in which this function is essential. Identifying also other homologous genes in *Arabidopsis* (“paralogues”) is necessary before one can conclude on “orthologous” relationships, because orthologues are homologs from different species which have similar or identical functions. Based on the study of these JMJ14 homologues, the function of JMJ14 in *Arabidopsis* may become more intelligible. Therefore I carried out a bioinformatic analysis of JMJ14 structure and conserved domains, a search for homologues within *Arabidopsis* and other species, and lastly a phylogenetic analysis.

According to TAIR10 annotation and the TAIR10 gene model At4g20400.1, the *Arabidopsis JMJ14* gene (At4g20400) contains 11 exons (shown in Figure 3.20 A). It encodes a 954-amino-acid polypeptide. Protein domains were analysed by Simple Modular Architecture Research Tool (SMART). SMART is an online tool which allows the identification and annotation of genetically mobile domains and the analysis of domain architectures. After inserting the JMJ14 amino acid sequence into SMART, it automatically provided its domain architecture. The *JMJ14* encoded protein contains 5 distinct domains which are the JmjN domain, JmjC domain, C5HC2 zinc finger (zf-C5HC2), F/Y-rich N-terminal (FYRN) and F/Y-rich C-terminal (FYRC) domains (Figure 3.20 B) (Lu et al., 2008).

After a wide search for homologues across phyla, and based on published information (Lu et al. 2008), I selected a list of organisms representing eukaryotic

phyla to assess the distribution of JMJ14-like proteins. For each of those organisms, a BLAST search (NCBI) was used to identify the closest homologue. The top 6 *JMJ14* paralogues were selected, as they were all also annotated as JARID1 family proteins. The list of organisms and homologues is provided in Table 3.1. A multiple protein sequence alignment based on the full length of JMJ14 protein sequence across that list of homologues in different species evidenced the evolutionary relationships of the JMJ14 protein. The amino acid sequences of JMJ14 homologues were obtained from NCBI and aligned by Clustal Omega. A multiple protein sequence alignment of ALN file (shown in the Appendix Table 7.3) was generated by Clustal Omega, then opened by SplitStree software. A phylogenetic tree was constructed using SplitStree to reveal the evolutionary relationships between *JMJ14* homologues. Organisms were chosen to represent dicot, monocot and non-seed vascular plants, land plants, green algae, yeast and animals. The result is shown in Figure 3.20 C. In *Arabidopsis*, *JMJ15* and *JMJ18* are the closest paralogues of *JMJ14*. Although all JMJ14-19 proteins share a common origin across eukaryotes, including fungi (represented by yeast), those of animals and plants are much more closely related, and they have diversified within plants. Although JMJ17 is also identified among the *Arabidopsis* JARID1 family proteins, it is the *Arabidopsis* representative of a protein which had already likely diverged in the last common ancestor of land plants. The yeast homologue of *JMJ14* diverged in the very early precursors.

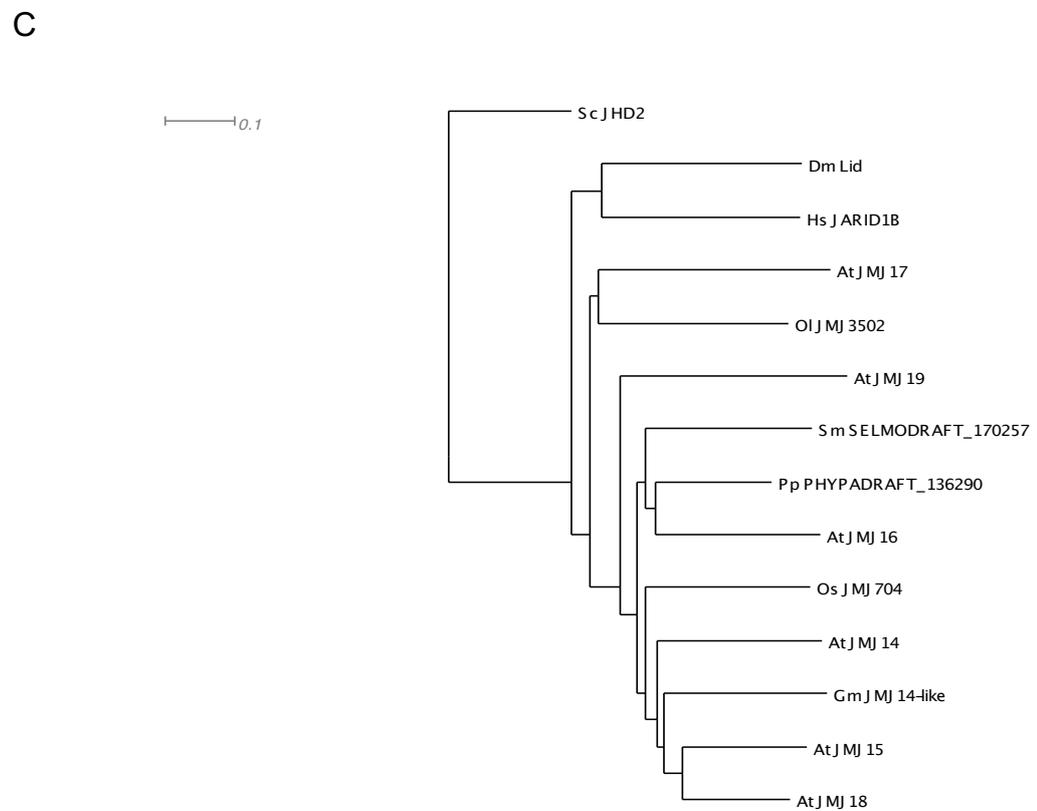
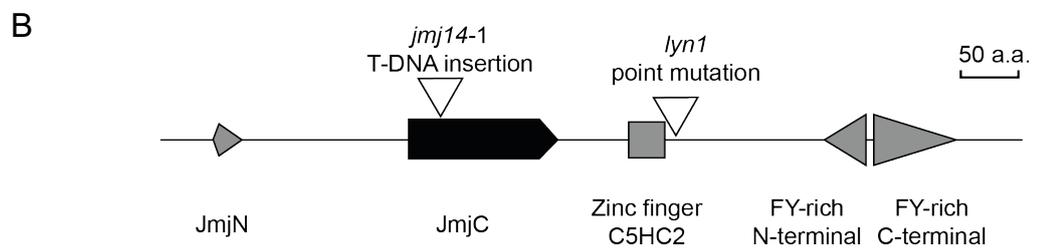
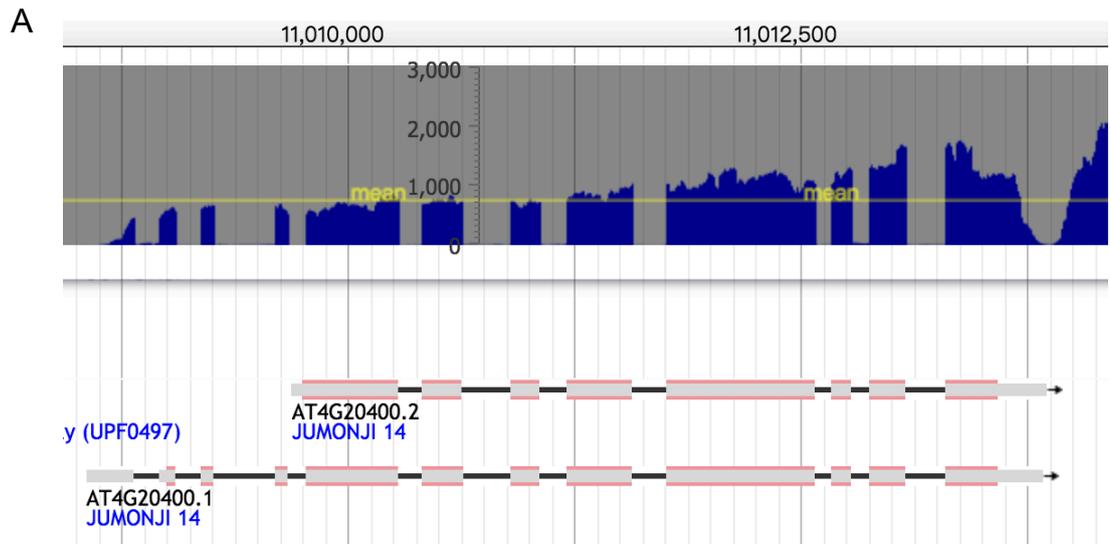


Figure 3. 20 Phylogenetic relationship of JmjC-domain-containing proteins from some selected species

A. Image showing the two JMJ14 protein coding gene models, together with evidence of mRNA expression in light-grown seedlings. Two models of the *JMJ14* gene are predicted to result in different mRNAs caused by alternative splicing. RNA-Seq data were used to detect whether the exon/intron prediction is correct, and whether any alternative splicing events are likely to occur in significant levels. In the RNA-Seq analysis, only exon transcribed mRNA can be detected and read, intron cannot be detected. The frequencies of RNA-Seq reads mapping at each position along the genomic sequence are shown in blue at the top grey area. The frequencies of reads show a bias towards the 3' end, but distribute across the differential region between the two gene models. If model AT4G20400.2 is a real alternatively spliced mRNA, then at least in light-grown seedlings model AT4G20400.1 is predominant. Figure from the light grown seedling of mapping coverage of RNA-Seq-based evidence at Araport 11 (Cheng et al., 2017).

B. Schematic domain structure of JMJ14. Domains were predicted by SMART. Lines indicate inter-domain regions. *jmj14-1* T-DNA insertion site and *lyn1* point mutation site on the genomic sequence of *JMJ14* are marked on the corresponding positions of *JMJ14* translated protein.

C. Phylogenetic tree showing the evolutionary relationships of JMJ14-related proteins across species. The tree was constructed by aligning the whole amino acid protein sequences in the SplitsTree software. Phylogenetic relationship of jmjC domain-containing proteins from *Arabidopsis thaliana*, *Glycine max*, *Oryza sativa*, *Selaginella moellendorffii*, *Physcomitrella patens*, *Ostreococcus lucimarinus*, *Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Homo sapiens* are indicated with At, Gm, Os, Sm, Pp, Ol, Sc, Dm and Hs prefix, respectively.

	Accession	Gene name	Gene description	Ident
1	NP_193773.2	<i>JMJ14</i>	JUMONJI 14 [<i>Arabidopsis thaliana</i>]	100%
2	NP_174367	<i>JMJ18</i>	Transcription factor jumonji family protein/ zinc finger (C5HC2) family protein [<i>Arabidopsis thaliana</i>]	61%
3	NP_181034.1	<i>JMJ15/MEE27</i>	Transcription factor jumonji family protein/ zinc finger (C5HC2) family protein [<i>Arabidopsis thaliana</i>]	58%
4	NP_001184940.1	<i>JMJ16/PKDM7D</i>	Transcription factor jumonji family protein/ zinc finger (C5HC2) family protein [<i>Arabidopsis thaliana</i>]	56%
5	NP_181429.2	<i>JMJ19</i>	Transcription factor jumonji family protein/ zinc finger (C5HC2) family protein [<i>Arabidopsis thaliana</i>]	41%
6	NP_001319304.1	<i>JMJ17</i>	Transcription factor jumonji (jmiC) domain- containing protein [<i>Arabidopsis thaliana</i>]	43%
7	XP_003535005.1	<i>JMJ14</i> -like	PREDICTED: lysine-specific demethylase JMJ18-like [<i>Glycine max</i>]	46%
8	XP_015640426.1	<i>JMJ704</i>	PREDICTED: putative lysine-specific demethylase JMJ16 [<i>Oryza sativa Japonica</i> <i>Group</i>]	88%
9	XP_002968749.1	<i>SELMODRAFT_170257</i>	hypothetical protein SELMODRAFT_170257 [<i>Selaginella</i> <i>moellendorffii</i>]	50%
10	XP_001770596.1	<i>PHYPADRAFT_136290</i>	predicted protein [<i>Physcomitrella patens</i>]	51%
11	XP_001416696.1	<i>JMJ3520</i>	predicted protein [<i>Ostreococcus lucimarinus</i> <i>CCE9901</i>]	40%
12	XP_011507394.1	<i>JARID1B</i>	PREDICTED: lysine-specific demethylase 5B isoform X3 [<i>Homo sapiens</i>]	42%
13	NP_523486.1	<i>Lid</i>	little imaginal discs, isoform A [<i>Drosophila</i> <i>melanogaster</i>]	41%
14	NP_012653.1	<i>JHD2</i>	histone demethylase [<i>Saccharomyces</i> <i>cerevisiae</i> S288C]	39%

Table 3. 1 List of homologues in different species evidenced by the evolutionary relationships of the JMJ14 protein

3.2.4 *JMJ14* gene expression pattern during *Arabidopsis* development

The *JMJ14* gene expression was examined on publicly-available data from *Arabidopsis*. The *JMJ14* expression pattern was analysed with the *Arabidopsis* eFP Browser, which presents data collected by the *Arabidopsis* Gene Atlas (Schmid et al. 2005). In *Arabidopsis*, *JMJ14* is widely expressed, but appears particularly active in the shoot apex (collected at the inflorescence stage) and during embryo development within the future seeds (Figure 3.21).

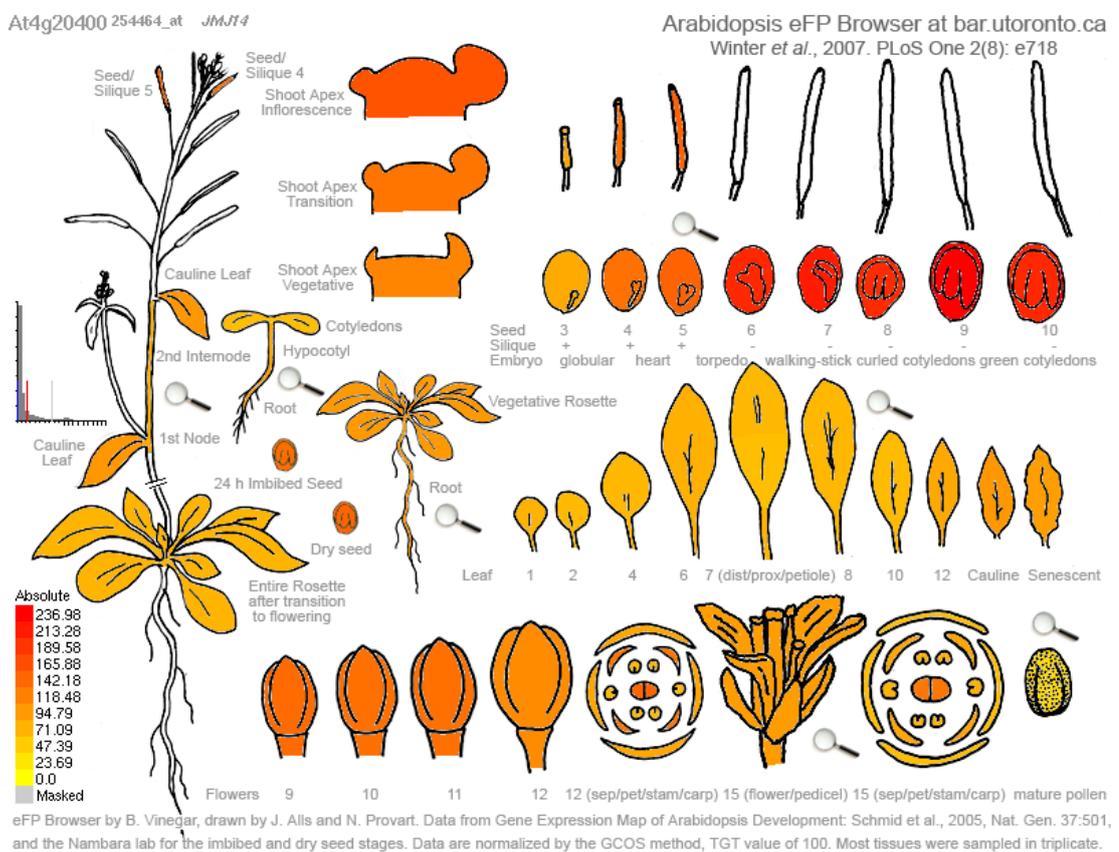


Figure 3. 21 Image shown the *JMJ14* gene expression map of *Arabidopsis*. Colours represent gene expression level, where the higher the number the higher the gene expression. Figure from the *Arabidopsis* eFP Browser at BAR (Winter et al., 2007), visualising data from the *Arabidopsis* Gene Atlas (Schmid et al., 2005).

3.3 Discussion

3.3.1 Confirmation of the identification of *lyn1*

The identity of *lyn1* was confirmed phenotypically and genetically by several analyses. The observation of earlier flowering on *jmj14-1* and *lyn1* preliminary supports the identification of *lyn1*. However, this argument is only correlative. In the hypocotyl elongation and greening response of double mutants, *jmj14-1* shows that it can independently suppress *hy1*. It shows a similar extent of suppression of hypocotyl elongation and chlorophyll content to that of *lyn1*. Moreover, the two types of complementation analysis reveal the most compelling evidence to support the *lyn1* identification result. The *jmj14-1* failure of complementation of *lyn1* resulting in the *lyn1 hy1* phenotype of F1 plants indicates that *jmj14-1* and *lyn1* are the two allelic mutations within the *JMJ14* gene. Once *JMJ14* is transferred into *lyn1 hy1* mutants, *JMJ14* can express functional *JMJ14* protein to complement the missing functional *LYN1* protein in the *lyn1 hy1* mutants. The *hy1* defect is no longer suppressed, so the complemented *lyn1 hy1* plants show the *hy1* phenotypes. According to all the analyses, *lyn1* has been confirmed to locate at the *JMJ14* gene.

So far, by numerous studies, *JMJ14* has been found to regulate flowering time. It represses the expression of FLOWERING LOCUS T (*FT*), which is a key floral activator, and TWIN SISTER OF *FT* (*TSF*), which is an *FT* homologue, to inhibit the floral transition in *Arabidopsis* (Yang et al., 2010). *jmj14* mutation leads to a reduced gene expression level of *FT* and *TSF* and also an earlier flowering time (Jeong et al., 2009; Lu et al., 2010). However, the dramatically different phenotype of *jmj14* compared to WT was noticed when the adult plants show different flowering time. All the studies missed the more subtle phenotype difference between WT and *jmj14* in early developing seedlings. This study observed this phenotype indirectly by screening *lyn1* on *hy1*. *lyn1* suppression of

hypocotyl length and chlorophyll content was easier to observe on *hy1*. Subsequently, the subtle suppression of hypocotyl length and chlorophyll content in *lyn1* compared to WT was detected. *JMJ14* was also identified in another study by screening for a suppressor of a mutant which has a defect in transgene-induced posttranscriptional gene silencing (PTGS). The *jmj14* mutation was found to suppress PTGS by increasing promoter methylation and reducing transgene transcription (Le Masson et al., 2012).

3.3.2 **JMJ14 protein domains**

Although *lyn1* shows a suppression of hypocotyl elongation and chlorophyll content, a slightly different extent of suppression between *lyn1* and *jmj14-1* was observed. This difference was also observed in another assay in the next chapter. Whether any mRNA is still produced in *lyn1* became the obvious question. The RT-PCR of *jmj14* cDNA revealed that *jmj14* mRNA is still present in *lyn1* but is not detectable in *jmj14-1*. The RT-PCR only determines the presence of mRNA but not the quantity. The *lyn1* mRNA generates premature termination codons (PTCs) which are targeted for degradation by nonsense-mediated decay (NMD) (Nolte and Staiger, 2015). The absence of *JMJ14* mRNA in *jmj14-1* may be due to the fact that the T-DNA insertion caused mRNA degradation during the post-transcriptional processing. Another possible reason for the non-detection of mRNA in *jmj14-1* may be because of the insufficient reverse transcription time and PCR extension time applied in the RT-PCR. The Salk line T-DNA insert is 4501 bp (O'Malley and Ecker, 2010). The T-DNA inserted *jmj14* becomes a very large DNA fragment which is 7366 bp. Even if the *jmj14-1* DNA is transcribed into a long mRNA, the RT-PCR time used for *lyn1* is not long enough for the *jmj14-1*. The aberrant mRNA is not the target mRNA in this assay. In both cases, *jmj14-1* does not reveal a detectable *jmj14* mRNA.

Immediately, the next question is whether the JMJ14 protein can be transcribed. If it is fully or partially transcribed whether the protein function is completely lost in *lyn1* and in *jmj14-1*. As mentioned in the general introduction (chapter 1.4.5), JHDMS catalyses lysine demethylation by an oxidative reaction that requires two cofactors, α -ketoglutarate (α KG) and iron Fe(II). The JmjC domain folds into eight β -sheets to form an enzymatically active pocket for coordinating with Fe(II) and α KG (Klose et al., 2006). Within the JmjC domain, three amino acid residues bind to the Fe(II) cofactor and other two residues bind to α KG. As the JmjC domain is the highly conserved domain shared by all JHDMSs, the JmjC domain is predicted as the most crucial domain for demethylation activity of the protein (Takeuchi et al., 2006). However, there is no evidence to show that the JmjC domain is the only domain to supply the JMJ14 protein biochemical function. Like most JHDMS family proteins, JMJ14 contains a JmjN domain and a C5HC2 zinc-finger domain. The domain organization in which JmjC accompanies JmjN and C5HC2 zinc-finger domains is also called the Jmj family of transcription factors and it has been found that it may be implicated in gene transcription or chromatin remodelling (Balciunas and Ronne, 2000). JHD2, a protein of the JHDMS group, was identified in yeast. Its JmjN domain maintains the protein stability by interacting with JmjC domain to prevent the proteasomal degradation (Huang et al., 2010). The C5HC2 zinc-finger domain has eight potential zinc ligand-binding residues, it is a putative DNA binding domain because it has been demonstrated to have a specific DNA binding activity (Lu et al., 2008; Searle et al., 2010). Unlike the JHDMS domain structure of animal and yeast proteins in the JARID1 group, instead of having AT-rich interaction domain (ARID) and plant homeodomain (PHD), all JmjC domain-containing protein (except JMJ17) in *Arabidopsis* additionally contain an F/Y-rich N terminus (FYRN) domain and an F/Y-rich C terminus (FYRC) domain, both (paradoxically) in the C terminus (Zhang et al., 2015). FYRN and FYRC need to be both present to function. These two domains contain chromatin-binding activity that helps the protein to bind to some specific chromatin regions (Lu et

al., 2008). They have been found to interact with NAC050/052 proteins, which are a pair of NAC (NAM, ATAF, CUC) domain-containing transcription factors, to play the role of histone demethylase recruitment (Zhang et al., 2015). According to the previous sequencing result of my MSc research, the point mutation in *lyn1* occurs about two thirds into the total *JMJ14* sequence (Figure 1.11) (He, 2013). In the corresponding positions of *JMJ14* translated protein, *lyn1* mutation point is after the JmjC domain towards to the C-terminus (Figure 3.20 B). Compared to *lyn1*, the *jmj14-1* T-DNA insertion point occurs at the beginning of the JmjC domain (Berardini et al., 2015). Although the *lyn1* point mutation caused a stop codon leading to an earlier termination of *JMJ14* translation, a partial protein containing three of the five conserved domains, and including the JmjC, catalytic domain, may still be translated and retained in the cytoplasm of *lyn1* mutant. This partial translated *JMJ14* is proposed to retain some *JMJ* protein function because of the presence of entire JmjC domain. However, this does not happen in *jmj14-1* because of, presumably, mRNA degradation. Even if the long aberrant mRNA is transcribed in *jmj14-1*, it cannot be translated to a functional *JMJ14* protein because the T-DNA insertion completely disrupts the *JMJ14* protein translation. Especially, JmjC domain is no longer present. Therefore, I propose that the *jmj14-1* completely losses *JMJ14* function and *lyn1* may retain some.

3.3.3 *JMJ14* evolution

3.3.3.1 Current studies of *JMJ14* paralogues

The *JMJ14* homologues show nearly universal distribution in eukaryotes. *AtJMJ15* and *AtJMJ18* are the two closest paralogues of *AtJMJ14*. The ancestor of *JMJ15* and *JMJ18* share a common ancestor with *JMJ14*, which suggests they probably have a very similar cellular function. Although they are all from the JARID1 family, they still retain some unique functions. Indeed, *JMJ15* and *JMJ18* have been found to play the opposite role of *JMJ14* in controlling flowering time

(Yang et al., 2012a,b). In *Arabidopsis*, several distinct flowering-promoting pathways, such as photoperiod and vernalisation pathways, are controlled by the environmental signal (day length, temperature and light, etc.) through a complex regulatory network. In this network, there are two central flowering regulators, FT and FLOWERING LOCUS C (FLC). FT is the floral activator regulating the expression of floral identity genes and involved in the photoperiod pathway, which responds to day length. Moreover, it is also directly repressed by the flowering time negative regulator, FLC. A subsequent analysis found that JMJ18 is a repressor of *FLC* gene expression by H3K4me2/3 demethylation (Yang et al., 2012a). This was carried out by studying *FLC* expression in 2 Salk lines carrying T-DNA KO mutations and 1 Salk line carrying T-DNA knock down mutation of *JMJ18*, all of which show a weak late-flowering phenotype (Yang et al., 2012a).

Another study from the same research group found that JMJ15 has a similar function to JMJ18. While the mutation of *JMJ15* did not cause any obvious phenotype, overexpressing *JMJ15* repressed *FLC* expression, which subsequently increased *FT* expression resulting in an obvious early flowering phenotype (Yang et al., 2012b). This suggests that although JMJ15 has the same function as JMJ18, JMJ18 dominantly regulates *FLC*. In the first study, the JMJ18 binding of *FLC* chromatin was studied by ChIP analysis in the transgenic plants. JMJ18 was found to directly bind to the *FLC* and mediate levels of H3K4 methylation (Yang et al., 2012a). However, no evidence exists that JMJ15 directly binds to *FLC*. Probably, JMJ15 indirectly regulates *FLC*. According to these studies, JMJ15, JMJ18 and JMJ14 are all involved in regulating flowering time. This supports an evolutionary relationship in which paralogues may be involved in similar biological functions, even with contrasting phenotypic effects.

3.3.3.2 Current studies of *JMJ14* orthologues

The first *jmj* gene was identified in mice and it was found to involve in neural tube formation (Takeuchi et al., 1995). Further studies found that *jmj* genes are expressed in a wide range of adult tissues in mouse and humans. The *JMJ14* orthologue from the human genome was selected for sequence analysis in this thesis. Nevertheless, it was found that the protein-coding regions of the mouse and human genomes are about 84% identical. In contrast, only about 30% of the protein-coding regions are identical on average between human and yeast (Pearson, 2013). As a clue to function, many studies have found that the *jmj* genes are strongly expressed in embryonic stem cells (Bergé-Lefranc et al., 1996; Takeuchi et al., 2006). This is in accordance with the *Arabidopsis* Gene Atlas gene expression analysis result of *JMJ14* in which the strongest expression is in *Arabidopsis* embryos, at all stages of embryo development before arresting into future seeds, with additional elevated expression in the shoot apex.

Human JARID1B is the closest homologue of *Arabidopsis* JMJ14 in the JARID1 subgroup of the JHDMS family. In mammals, the JARID1 subfamily encompasses four protein members: JARID1A (RBP2); JARID2B (PLU1); JARID1C (CMCX) and JARID1D (SMCY) (Blair et al., 2011). JARID1A and JARID1B are highly associated with cell proliferation, cell differentiation and several cancer types. (Johansson et al., 2014; Rasmussen and Staller, 2014). Importantly, they were found to play opposing roles in embryonic stem cells. JARID1A promotes differentiation (Benevolenskaya et al., 2005) and JARID1B promotes proliferation (Yamane et al., 2007). Therefore, the pluripotency of embryonic stem cells is to some extent controlled by JARID1A and JARID1B. Additionally, it is critical for differentiation of neural stem cells from embryonic stem cells before the established neural stem cell stage. (Schmitz et al., 2011). Moreover, JARID1B is not only involved in repressing cells entering into differentiation in embryonic stem cells but also in haematopoietic stem cells by H3K4 demethylation (Dey et

al., 2008; Cellot et al., 2013). Interestingly, it was found that JARID1B localises near promoters of active genes to force H3K4 methylation at promoter region, avoiding spreading H3K4 methylation into gene bodies (Kidder et al., 2014). This seems to contradict the essential function of H3K4me3 demethylase. Nevertheless, the average H3K4me3 levels are relatively low in JARID1B target genes compared to those of non-productive genes (Rahl et al., 2010). In fact, this suggests that rather than completely removing the H3K4me3 markers, JARID1B maintains a low level of H3K4me3 marking state. Moreover, JARID1B has been found to play an important role in tumorigenesis. It acts as a deregulated factor in several cancer types. For example, it causes proliferation of breast cancer cells by directly repressing the transcription of tumour suppressor genes (Yamane et al., 2007), and it expresses in melanoma specifically (Kuźbicki et al., 2013).

It has been widely supported by global evolution that fungi (including yeast) and animals share a common ancestor more recently than the last common ancestor of theirs and plants. The direct ancestor of *JMJ14* may have been lost in yeast, even though it is retained in animals. Yeast *JHD2* retains a precursor of *JMJ14* and *JMJ17*. *JHD2* function was specifically investigated during yeast sporulation, because several studies did not detect any impact of *JHD2* on transcription in vegetative cells (Xu et al., 2012). It was found that *JHD2* extended the duration of the spore differentiation by sustaining the postmeiotic gene expression to prevent the sporulating cells from terminating differentiation prior to entering transcriptional quiescence (Xu et al., 2012).

Normally, orthologues retain the same functions in the course of evolution. Identification of orthologues is critical for reliable prediction of gene function. By studying the *JMJ14* orthologues, *JARID1B* and *JHD2*, I conclude that *JMJ14* cannot have had an ancestral function specific to chloroplast development. Instead, it was inferred that *JMJ14* may play a role in regulating cell differentiation by histone demethylation, as its homologues do in animal cells. This may help to

clarify the biological processes that JMJ14 is involved in to rescue light response in *Arabidopsis*.

3.4 Conclusion

lyn1 has been confirmed to contain a point mutation in the *JMJ14* gene. *lyn1* retains *jmj14* mRNA which may still produce a partial JMJ14 protein. If *JMJ14* protein is produced to some degree, it may retain some demethylation function, as the conserved catalytic JmjC domain is present in it. Proteins with a JmjC domain are universally distributed in eukaryotes. The high conservation of the JmjC domain predicts it to play a crucial role in the demethylation function. The analysis of known functions of JARID1 proteins in other species hints that JMJ14 may play a conserved role in decisions of cell fate, by regulating cell differentiation. The wide tissue distribution of JMJ14 in *Arabidopsis*, with no relation to the “green” nature of the tissue, but with elevated expression in shoot apices and during embryogenesis, supports such a broad role. This prediction may help interpret outcomes of further research into the function of JMJ14 rescuing light responses.

Chapter 4 Result – Characterisation of the *JMJ14* Gene

4.1 Introduction and Methodology

4.1.1 *jmj14* impact on plastid development

As the *HY1* locus encodes the haem oxygenase which is involved in biosynthesis of the phytochrome, the phytochrome defective mutant, *hy1*, has shown a significant defect in light responses. Meanwhile, *lyn1* dramatically suppressed growth and greening defects of *hy1* can be observed. The chloroplast is the organelle where photosynthesis takes place. Its development is fundamental for plant growth and greening. *jmj14* was suspected to have an impact on chloroplast development. Quantifying the chloroplast compartment of a single cell can determine to what extent cells are filled with chloroplasts. Chloroplast compartment is dependent on chloroplast size and number. If *jmj14* causes an increased chloroplast size or number, the chloroplast compartment should also increase. During my previous MSc research, chloroplasts were observed in mesophyll cells of rosette and cauline leaves (Appendix Figure 7.2 A). This is because mesophyll cells are the place where chloroplasts fully develop and most photosynthesis occurs in C₃ plant. However, the chloroplasts amount measurements from the captured images was difficult and of limited confidence in its accuracy because chloroplasts overlapped with each other in one single focal plane on mesophyll cells of fully developed rosette leaf and cauline leaves of adult plants of the all genotypes. Therefore, I decided on the quantitation being done live, directly under the microscope by changing the focus and capturing all focal planes within a single cell. This allowed us to measure with certain every single chloroplast in the cell under observation. As the chloroplast compartment analysis showed a similar result between rosette and cauline leaves among 4

genotypes (WT, *lyn1*, *hy1* and *lyn1 hy1*) (Appendix Figure 7.2 B), this assay was repeated in mesophyll cells of cauline leaves only in the WT, *lyn1*, *jmj14-1*, *hy1*, *lyn1 hy1* and *jmj14-1 hy1*. Additionally, the chloroplast compartment of bundle sheath cells of cauline leaves was also studied to determine whether *lyn1* affects a particular type of cell or specific tissue. Another reason of study in bundle sheath cells is these cells are less fully occupied by chloroplasts than mesophyll cells (Kinsman and Pyke, 1998). The chloroplast compartment change can be easily observed compared to mesophyll cells, if *lyn1* also affects chloroplasts in bundle sheath cells. Although *Arabidopsis* is a C₃ plant whose chloroplasts in the bundle sheath cells were only assumed to contribute a small quantitative role in photosynthesis, it has been found some C₄ chloroplast functions are retained in the bundle sheath of C₃ plants. The study of *lyn1* impact on chloroplasts in bundle sheath cells may also be useful for a C₄ rice project later.

Another way of detecting the *lyn1* effect of plastid development is by looking at the difference of plastome (chloroplast genome) copy number among the 6 genotypes. Plastid size and DNA quantity have been found to show a positive correlation (Rauwolf et al., 2010). Because better developed plastids contain more copies of plastomic DNA than relatively smaller plastids, I speculated that if *jmj14* enhanced chloroplast size, it could increase plastome copy number. Technically, a “chloroplast compartment” analysis, using DIC microscopy, can only indicate the *jmj14* effect on the plastid compartment in the light. Measuring the relative copy number ratios between plastome and nuclear genome in dark-grown seedlings is an alternative way to indicate whether *jmj14* can enhance the plastid compartment in the dark. This experiment was carried out by using quantitative real-time genomic-PCR (qgPCR) to quantify gene copy number of plastomic DNA and genomic DNA and then calculating the relative copy number ratios. For greater accuracy, absolute qgPCR quantitation was carried out, a standard curve was also run together with samples in order to convert the

sample's qPCR Ct values into the number of copies. In an initial experiment, the gene copy number was only studied on one plastid-encoded gene (*rbcL*) and one nuclear-encoded gene (*HO1*). The results were not always very robust after several repeats as the error bars were too large to distinguish if there was any statistically significant difference among the 6 genotypes. Many factors can cause the variability of results. Firstly, the standard curves were very easy to shift resulting in a different ratio. This can be caused by many reason, such as cycles of freezing and thawing, which may reduce solubility of the DNA molecules. Secondly, good quality of DNA samples was also required from DNA extraction. Several different methods of extracting DNA were tried. CTAB method can extract large amounts of DNA but the purity was relatively low. The contamination, probably of carbohydrates, in DNA samples gave a peak between absorbance of 220-230 nm. This may introduce variability in the qPCR result. The extracted DNA samples using a commercial kit, Qiagen DNeasy Plant Mini Kit, were finally selected to use for this experiment. Thirdly, the ratio differences between the 6 genotypes in 5-day dark-grown seedling are relatively small, so any small impact such as pipetting error can increase the variability and resulting error bars. Fourthly, the chloroplast genome contains 2 inverted repeats (IRA and IRB) in its map, separating large single-copy (LSC) and small single-copy (SSC) regions, and there are circumstances in which rearrangements of the genome can occur (Figure 4.1) (Shaver et al., 2006; Rowan and Bendich, 2011). To overcome these issues, this experiment was repeated by using three plastomic genes (*rbcL*, *ycf2* and *ndhG*) and two genomic genes (*HO1* and *CHS*) to obtain a more robust result. *HO1* and *CHS* are two nuclear genes selected for calculating the average copy number of nuclear genome. Three genes were selected from each individual region of chloroplast genome. *rbcL* was selected from LSC. *ndhG* was selected from SSC. *ycf2* is present in the both inverted repeat regions, so each plastid genome always contain two copies of *ycf2*. The average copy number of plastid genome was also calculated by these three genes.

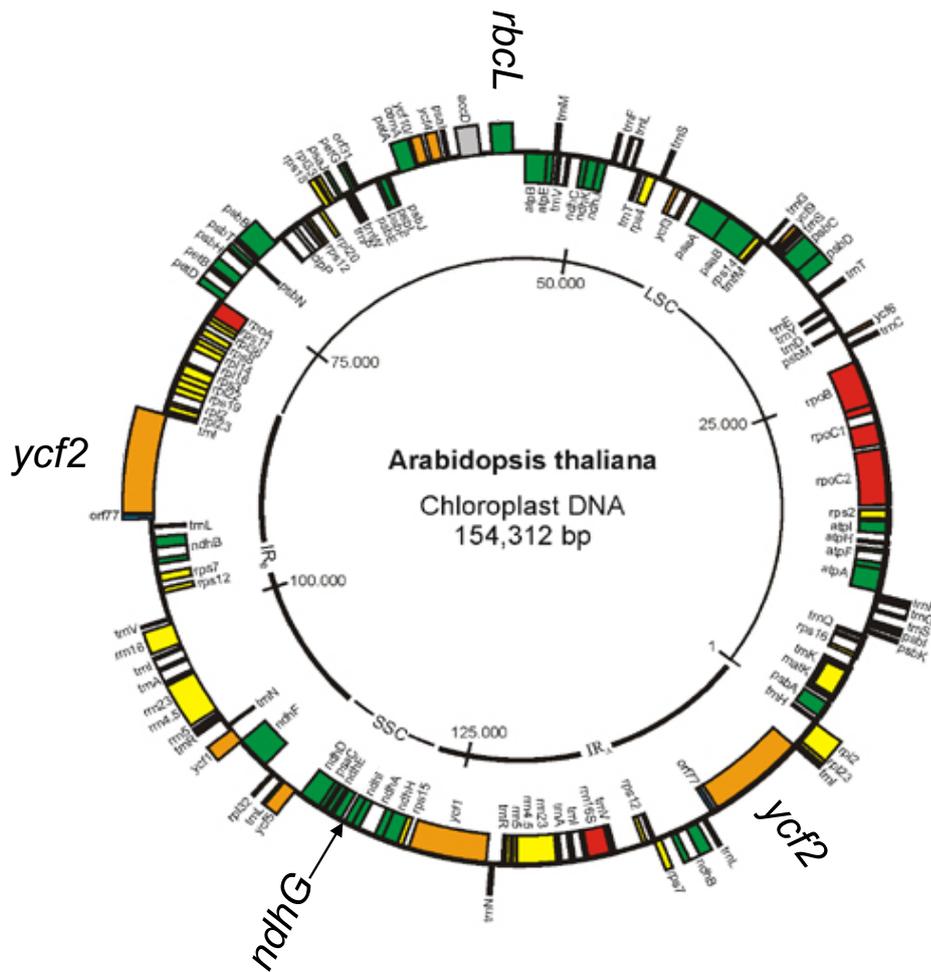


Figure 4. 1 *Arabidopsis* chloroplast genome map

Figure adapted from (Sato et al., 1999)

4.1.2 *jmj14* impact on tetrapyrroles synthesis

During my MSc research, a Pchlide measurement experiment suggested that Pchlide level is increased by the *lyn1* mutation in the dark (Appendix Figure 7.1). An alternative way of testing the change of Pchlide level is by the FR light block of greening experiment. The principle of the FR light block of greening experiment has been explained in chapter 1.3.2. The accumulation of excess Pchlide leads to cell death of etiolated seedlings. According to our previous Pchlide measurement experiment, I predicted that *lyn1* and *jmj14* seedlings would have an increased mortality rate compared to WT in the FR-block of greening experiment. As *hy1* is defective in light detection, the mutants with recovered light

sensitivity, *lyn1 hy1* and *jmj14 hy1*, would be expected to have significantly higher mortality rates than *hy1*.

Is the rescued light response of *lyn1 hy1* caused by an elevated phytochrome level produced in the light? In other words, it is possible that *lyn1* resulted in some recovery of photochemically-active phytochrome, which was absent in *hy1*. *hy1* causes the loss of PΦB, one of tetrapyrrole synthesis pathway end products, which is the phytochrome chromophore and therefore a necessary component for phytochrome synthesis. This was investigated by quantifying the level of photochemically-active phytochrome.

4.1.3 Assessment of whether *lyn1* rescues a PΦB synthase defective mutant

PΦB synthase (HY2) is another enzyme, which catalyses biliverdin IX α to PΦB conversion, and is involved in the phytochrome biosynthesis pathway (Rockwell et al., 2006). According to the hypothesis, if *lyn1* rescues *hy1* by enhancing the capacity to synthesize PΦB in plastids, through enhanced plastid metabolic activity, then *lyn1* is also likely to rescue *hy2*. Therefore, *lyn1* suppression of *hy2* was also studied to test the hypothesis. Two *hy2* mutations, *hy2* T-DNA insertion Salk line (SALK_104923) and *hy2-1*, were studied during my MSc research. *hy2* inserted Salk line (SALK_104923) is in the Col ecotype. It was expected to be a *hy2* KO mutant as it carries a T-DNA insertion within the *HY2* locus. Unfortunately, it does not show a phytochrome-deficient phenotype (Didcock and López-Juez, unpublished observations). It had been confirmed that *HY2* mRNA was still present by reverse transcription PCR, so it cannot be used for *lyn1* suppression research (He, 2013). *hy2-1* is in the Ler ecotype, and I observed a small degree of suppression by *lyn1* (Figure 4.2). In order to study the phenotype of *lyn1 hy2* double mutants in a single genetic background, I made use of *hy2-105*, which carries a splicing mutation in the *HY2* locus in the Col ecotype with a 25 bp

deletion in the second intron (Kohchi et al., 2001). Using the same ecotype, therefore a single genetic background, avoids the heterogeneity of phenotypes which can be observed in the progeny of crosses between different backgrounds.

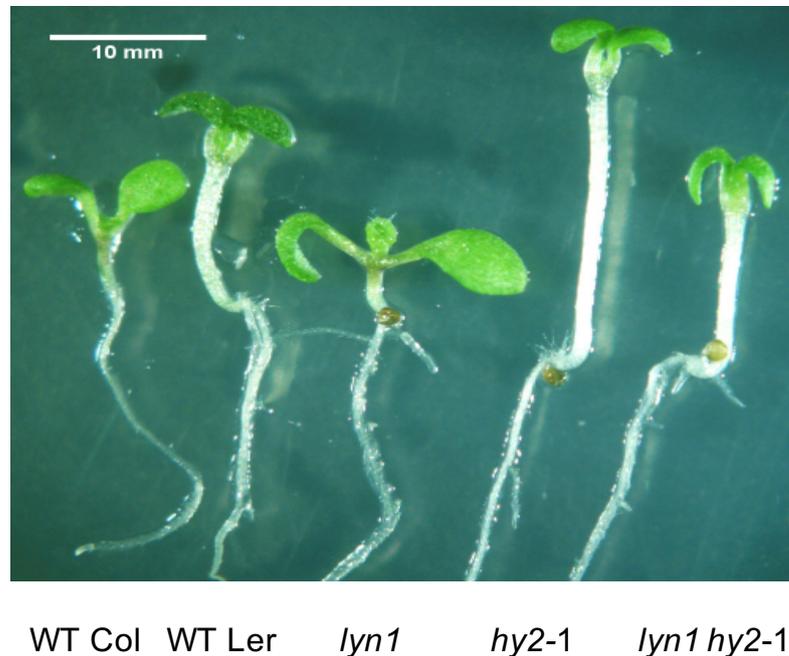


Figure 4.2 *hy2-1* mutation effect of phytochrome and chlorophyll biosynthesis pathway

Seedlings of WT Col, WT Ler, *lyn1*, *hy2-1* and *lyn1 hy2-1* were grown in white light ($100 \mu\text{mol m}^{-2} \text{s}^{-1}$) for 6 days in 1% MS media. Scale bar = 10mm.

4.1.4 Assessing whether *lyn1* suppresses photoreceptor apoprotein mutants, *phyA-211*, *phyB-9* and *cry1-304*

So far, all the observations shown in this chapter determined the impacts of *lyn1* mutation on plastids only. However, phytochrome synthesis takes place only partially within plastid. The synthesis of phytochromes is completed in the cytoplasm by binding P Φ B covalently to phytochrome apoproteins (Rockwell et al., 2006). Whether *lyn1* has any impact on phytochrome apoprotein mutants is also important to know in order to confirm the hypothesis by which *lyn1* rescues the light response by increasing the plastid-synthesised P Φ B, but not altering the

phytochrome apoproteins. In this case, *lyn1* effects on hypocotyl elongation and chlorophyll content of *phyA* (defective in the FR light photoreceptor) and *phyB* (defective in a main R light photoreceptor) were studied. Additionally, *lyn1* effect of *cry1* (defective in a B light photoreceptor) was also studied at the same time.

phyA-211 mutants were identified among M2 progeny of gamma irradiation-treated *Arabidopsis*, then the long hypocotyl seedlings isolated after growing under FR light (Reed et al., 1994). *phyA-211* has a large deletion in the *PHYA* gene and it is a presumed null mutation (Reed et al., 1994). The *phyB-9* mutation has a stop codon and is a presumed null mutation (Reed et al., 1993). *cry1-304* were isolated from fast-neutron mutagenized populations (Bruggemann et al., 1996). It has a large deletion and it is a presumed null mutation (Mockler et al., 1999; Kang and Ni, 2006)

One thing to be noted is that the quantitation of chlorophyll content of the FR treated seedlings was not as simple as that of R light and B light treated seedlings. Theoretically, seedlings should only contain Pchl_a in the FR light. Pchl_a, synthesised in the dark, is converted to Chl_a in the presence of POR and R light. Chl_a then converts to chlorophyll, independently of light. In theory, spectrally pure FR light should not be capable of converting Pchl_a to Chl_a. However, it is technically difficult to generate a pure FR light treatment. There is always some mixture of R light within the FR light treatment. The FR grown seedlings always produce Pchl_a as well as some chlorophyll, resulting in two absorbance peaks at 630nm (Pchl_a) and 664nm (chlorophyll), respectively. Chlorophyll can be detected and quantified by spectrophotometer directly when seedlings detect R or B lights. Pchl_a, the precursor of chlorophyll, is produced all the time even in the dark and it can be detected and measured by spectrofluorimeter. Pchl_a converts into chlorophyll only once seedlings are exposed to R-containing light.

The aim of this chapter is to investigate whether the *lyn1* mutation boosted plastid development, and whether enhanced plastid tetrapyrrole metabolic activity could lead to a rescued light response in phytochrome chromophore mutants, but not in other photoreceptor mutants.

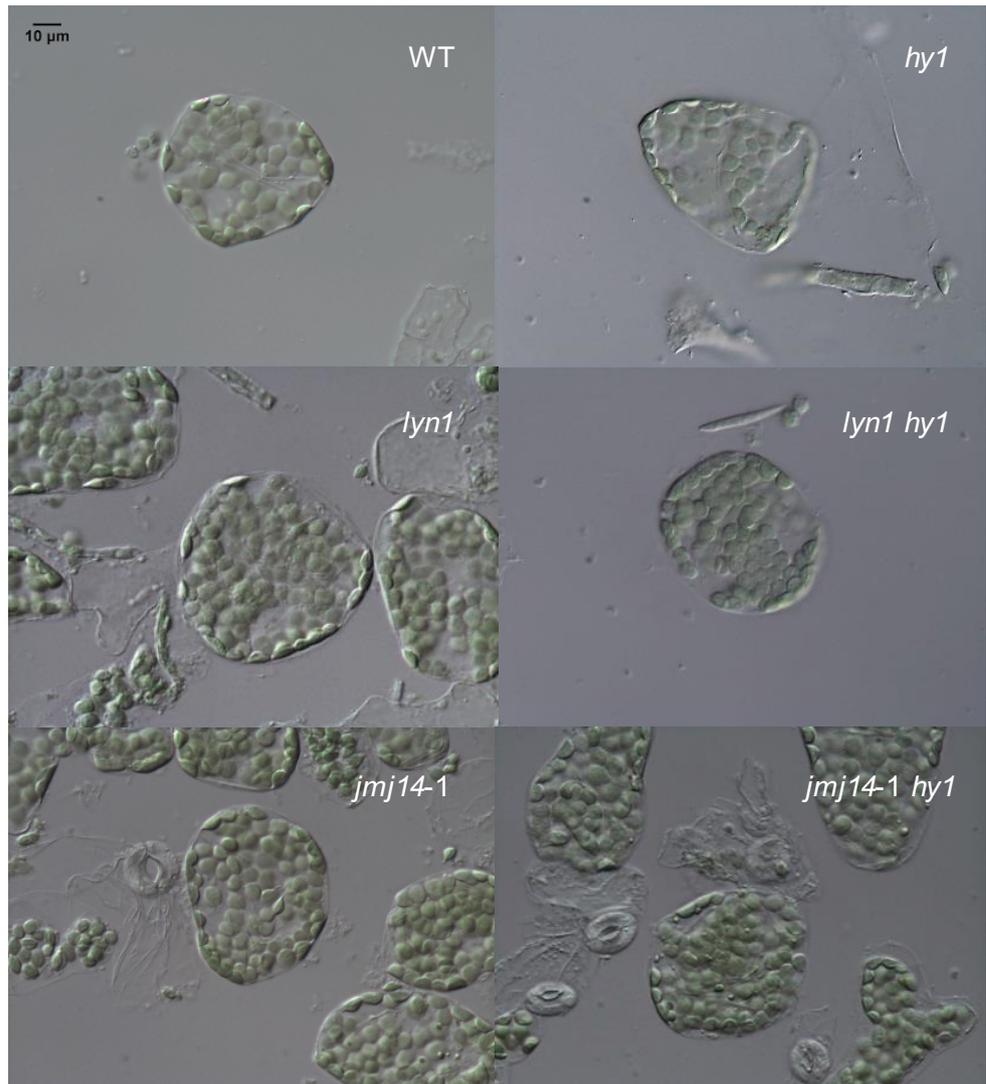
4.2 Results

4.2.1 Chloroplasts compartment analysis

Leaf tissue samples were collected from the fully-developed cauline leaves of WT, *lyn1*, *jmj14-1*, *hy1*, *lyn1 hy1* and *jmj14-1 hy1* plants. These tissue fragments were fixed treated in order to separate individual intact cells. Mesophyll cells and bundle sheath cells were particularly selected to observe under the microscope. Mesophyll cells are much larger than bundle sheath cells (Kinsman and Pyke, 1998). These two types of cells, and the chloroplasts they contained, were observed using DIC microscopy (Figure 4.3). I use “chloroplast compartment” to refer to the total amount of chloroplast in a cell (the product of the number of chloroplasts by the average chloroplast plan area), and “cell index”, which represents the occupancy of cell plan area by chloroplasts, to express the relative chloroplast compartment, the ratio between the chloroplast compartment and the plan area of the cell it was derived from.

According to the observation of chloroplasts of 6 genotypes in both types of cells, *hy1* contains less developed chloroplasts compared to the WT, while *lyn1 hy1* and *jmj14-1 hy1* are more occupied by chloroplasts than *hy1*. It is hard at first glance to distinguish whether *lyn1* and *jmj14-1* are more occupied by chloroplasts than WT. In accordance with the published data, bundle sheath cell chloroplasts are always smaller than mesophyll cell chloroplasts in every genotype.

A



B

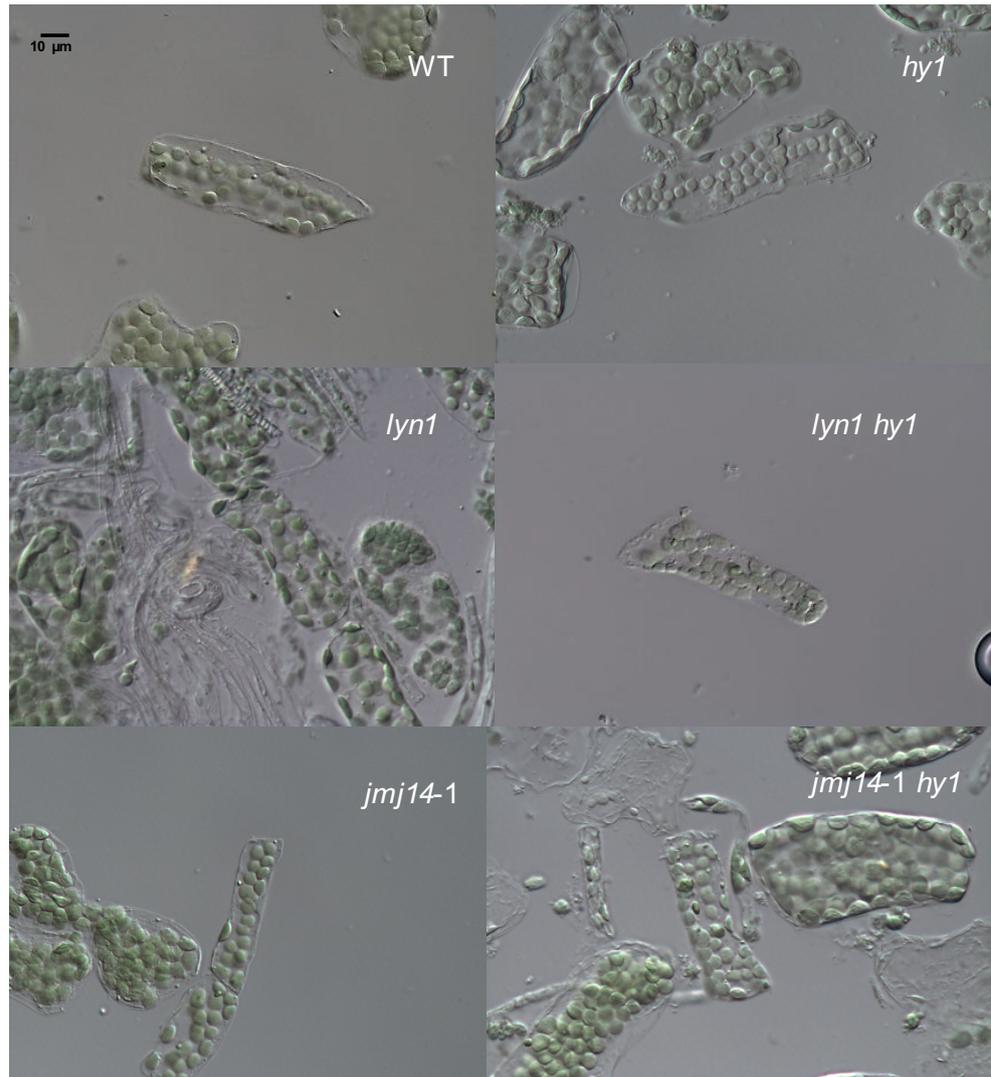


Figure 4. 3 Observation of chloroplast development under DIC microscopy

The fully developed mesophyll (A) and bundle sheath (B) cells and their chloroplasts from cauline leaves of adult plants for WT, *lyn1*, *jmj14-1*, *hy1*, *lyn1 hy1* and *jmj14-1 hy1* were selected and observed under the microscope. Plants grown on soil under 16 h photoperiodic white light of $180 \mu\text{mol m}^{-2} \text{s}^{-1}$. Scale bar=10 μm .

Chloroplast amount (total number), average chloroplast size for each cell and the cell size were measured to calculate total chloroplast area, cell index and chloroplast density. The cell indices were compared among 6 genotypes. The statistical significance of differences between genotypes were tested by one-way

ANOVA, the positive result suggesting that one or more comparisons are significantly different in both types of cells. This was followed by Bonferroni and Holm multiple comparisons T-testing, which only pairs groups relative to a reference group. In the Bonferroni and Holm multiple comparisons, the single mutants, *lyn1*, *jmj14-1* and *hy1*, are compared to WT, and the double mutants, *lyn1 hy1* and *jmj14-1 hy1*, are compared to *hy1*. In accordance to the observation under the microscope, *lyn1* and *jmj14-1* do not have significantly different cell indices relative to WT, while *hy1* has a lower cell index than WT in mesophyll cells. *lyn1 hy1* and *jmj14-1 hy1* both have higher cell indices than *hy1*, a result in accordance with the visual observations of *lyn1*-increased chloroplast occupancy of cells (Figure 4.4 A). The differences between double mutants and *hy1* are significant. In the bundle sheath cells, similar to the observation of chloroplast indices of mesophyll cells, the WT cell index is higher than that of *hy1*. *lyn1 hy1* and *jmj14-1 hy1* both have a higher cell index than *hy1* (Figure 4.4 B). Statistical testing indicates that the differences are significant. In contrast to mesophyll cells, in which the cell index of *lyn1* and *jmj14-1* was no different from that of the WT, in bundle sheath cells it was elevated in the mutants. This was not immediately apparent under the microscope, demonstrating the value of obtaining precise quantitative data. Statistical testing indicates that the differences between the two single mutants (*lyn1* and *jmj14-1*) and the WT are both significant. Kinsman and Pyke (1998) found that bundle sheath cells are always less occupied by chloroplasts than mesophyll cells. Accordingly, the result showed that bundle sheath cells have generally lower cell index than mesophyll cells in all genotypes.

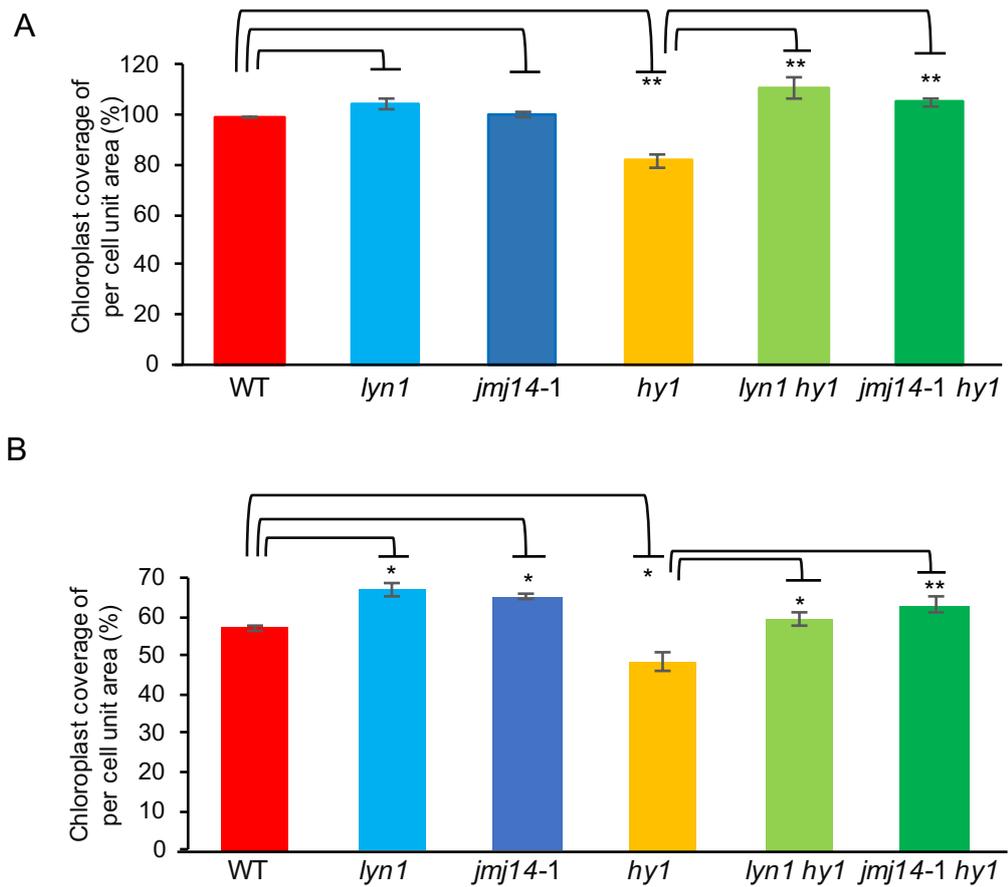


Figure 4.4 Quantitative analysis of cell index

A. Mesophyll cell, n=20;

B. Bundle sheath cell, n=20.

The correlations between cell size and total chloroplast area of each genotype were also plotted. This represented how total chloroplast area changes when cell size changes, given the tendency of chloroplasts to distribute in one layer (under the cell membrane) within the available cellular space. Clearly, when cell size increases, the total chloroplast area also increases in all genotypes of both types of cell (Figure 4.5). However, the rates of increase are different between genotypes. In mesophyll cells, the trendlines of WT, *lyn1* and *jmj14-1* almost overlap (Figure 4.5 A). The proportions of total chloroplast area of these three genotypes are always similar regardless of the size of the cell. In Figure 4.5 B, the trendlines of *lyn1 hy1* and *jmj14-1 hy1* almost overlap. The slopes of their trendline are steeper than that of *hy1*. The total chloroplast area of *hy1* increases

much slower than that of *lyn1 hy1* and *jmj14-1 hy1* when cell size increases. In bundle sheath cells, the trendline of WT is almost parallel with that of *lyn1* and *jmj14-1*. The slope of WT trendline is slightly lower than those of *lyn1* and *jmj14-1* (Figure 4.5 C). The slopes of total chloroplast area against total cell area among the 3 genotypes do not increase as fast for bundle sheath cells as those in mesophyll cells do when cell size increases. In *hy1* mutants, the result for bundle sheath cells is very similar to that for mesophyll cells between the 3 genotypes (Figure 4.5 D). By studying both types of cells, our results suggest that *lyn1 hy1* and *jmj14-1 hy1* have an increased ability to generate larger chloroplasts than *hy1* does. In bundle sheath cells, *lyn1* and *jmj14-1*, the single mutations, also have a slightly increased ability to generate larger chloroplasts than the WT does. In any genotype, the slope of the trendline in mesophyll cells is always greater than that in bundle sheath cells. The chloroplast coverage of mesophyll cells increases faster than that in bundle sheath cells, when cell size increases. The total chloroplast area of mesophyll cells is always much larger than that of bundle sheath cells for any genotype. This is in accordance to previously published observations that mesophyll cells have a stronger ability to generate better-developed chloroplasts than bundle sheath cells do (Kinsman and Pyke, 1998).

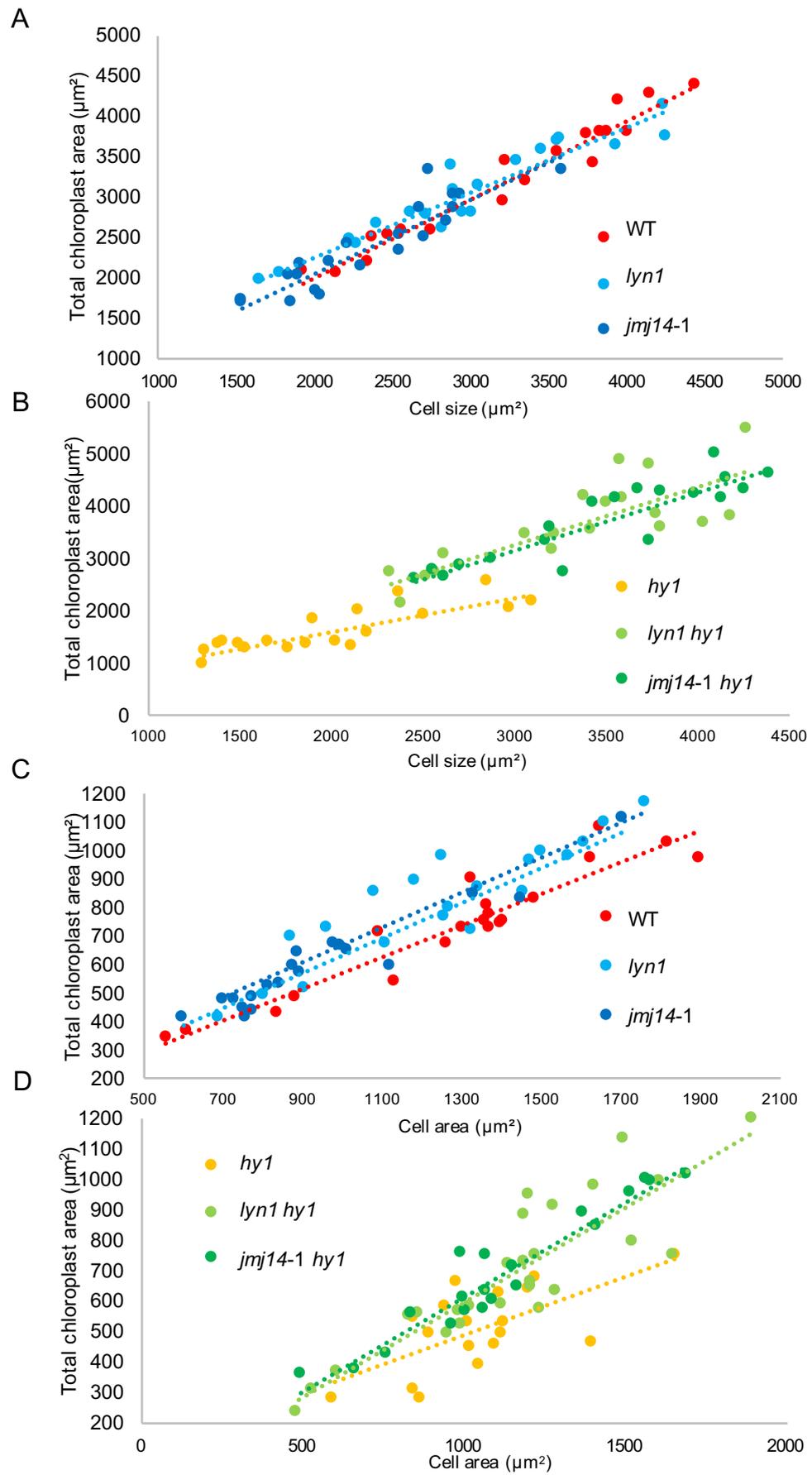


Figure 4. 5 Correlation between total chloroplast area and cell size of all genotypes

X-axis represents the total chloroplast area (μm^2) which is calculated using the total number of chloroplast times the average of chloroplast size in a single cell. y-axis represents the cell area (μm^2). The total chloroplast areas and cell areas reveal a linear relationship by drawing a trendline. The trendlines (linear regression lines) were calculated by Excel.

A. Scatterplot of WT, *lyn1* and *jmj14-1* mesophyll cells, n=20 (20 individual cells were measured);

B. Scatterplot of *hy1*, *lyn1 hy1* and *jmj14-1 hy1* mesophyll cells, n=20;

C. Scatterplot of WT, *lyn1* and *jmj14-1* bundle sheath cells, n=20;

D. Scatterplot of *hy1*, *lyn1 hy1* and *jmj14-1 hy1* bundle sheath cells, n=20.

The cell index depends not only on chloroplast size but also on chloroplast number. Therefore, chloroplast density was also calculated using the total amount of chloroplast per cell, divided by the total cell plan area. The results show that the chloroplast densities remain constant among all genotypes in both types of cell (Figure 4.6). In accordance to Kinsman's results, chloroplast density of mesophyll cells is very similar to that of bundle sheath cells in all genotypes. The one-way ANOVA test indicated that there are no significant differences, neither between the two types of cell nor between the six genotypes. This suggests that *lyn1* does not elevate the cell index by increasing the number of chloroplasts in the cell. Cell index is related to the amount (number) and the size of chloroplasts. If the amount of chloroplasts does not change, then *lyn1* can only increase the chloroplast size to achieve elevated cell index. Because the chloroplast densities of mesophyll cell and bundle sheath cell are very similar, mesophyll cells have relatively larger chloroplasts than bundle sheath cells.

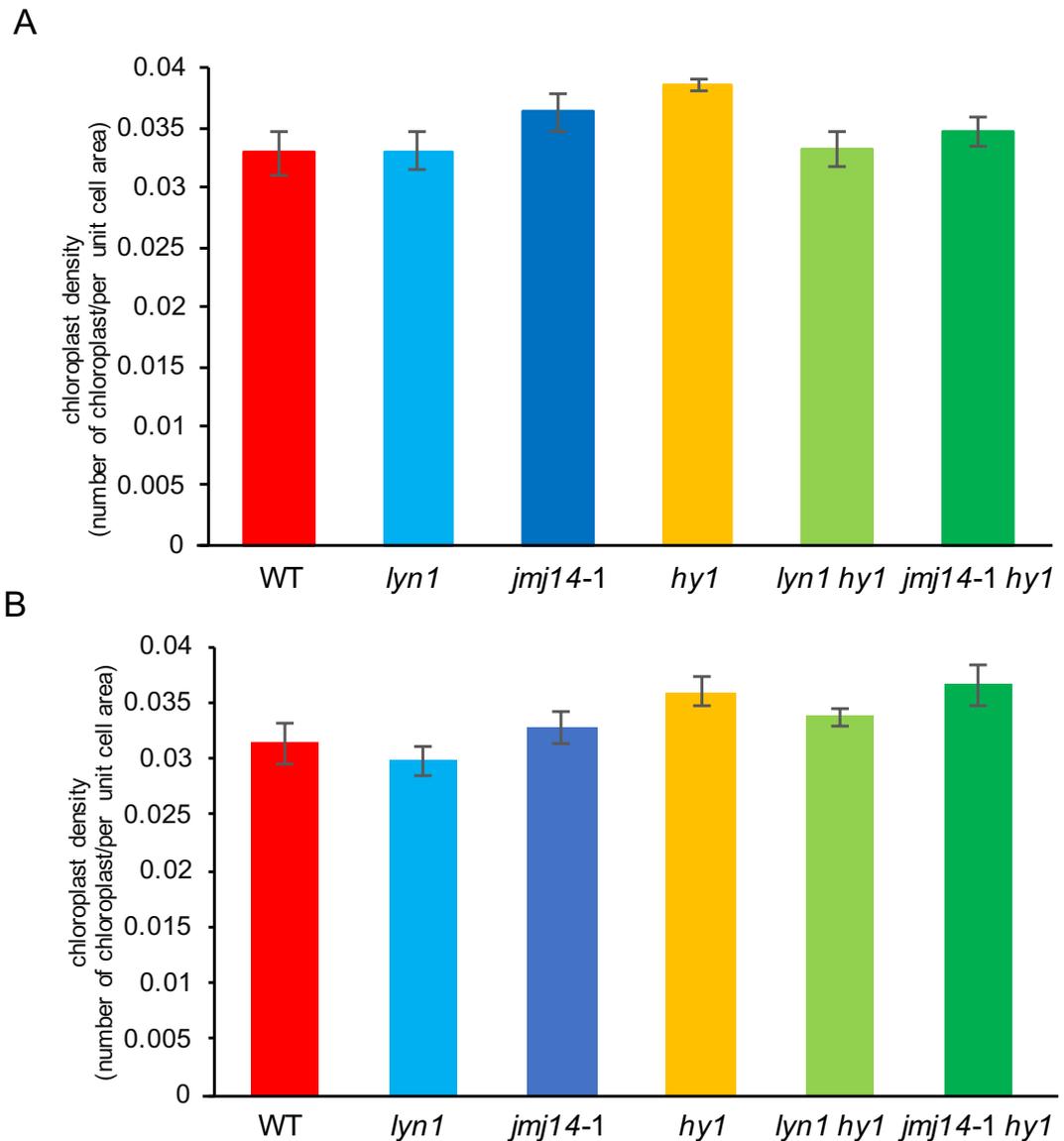


Figure 4. 6 Quantitative analysis of chloroplast density

A. Mesophyll cell, n=20;

B. Bundle sheath cell, n=20.

4.2.2 Analysis of plastid to nuclear genome copy number ratio

From the plastid compartment assay, the increased chloroplast compartment size occurring in the *lyn1* mutants can be detected only in the light condition. Given that the enhanced chloroplast compartment size can lead to increased plastid to nuclear genome copy number ratio (Enfissi et al., 2010), I speculated that an

increased plastid to nuclear genome copy number ratio can also be detected in the other types of plastid, for example in the dark, if *jmj14* impacts the global plastid development regardless of light. The genomic copy number ratio of the six genotypes was analysed in 5-day light-grown seedlings and 5-day dark-grown seedlings. The genomic copy ratios were calculated using the average copy number of plastid genes divided by copy number of nuclear genes. All genomic copy ratios of dark-grown seedlings are generally lower than those of their corresponding light-grown seedlings (Figure 4.7). The statistical analyses applied for this assay between genotypes in each light condition were the same as those used for the quantitation of chloroplast compartment. The results show that *hy1* has a lower ratio than *lyn1 hy1* and *jmj14-1 hy1*, while WT is higher than *hy1* in both light conditions. Statistical testing indicates that the differences are all significant. However, the genomic copy ratios of *lyn1* and *jmj14-1* are both significantly higher than that of WT only in the dark-grown condition (Figure 4.7 A). They do not have significant difference with WT in the light-grown condition (Figure 4.7 B). These results were concordant with those of the chloroplast compartment experiment as I will address in the discussion. The *jmj14* mutation was suspected to elevate the genomic copy ratio not only on *hy1* in both light and dark conditions, but also in the single mutants, *lyn1* and *jmj14-1*, in the 5-day dark-grown seedlings. That hypothesis has been confirmed. It is noteworthy that *jmj14* significantly elevates the genomic copy ratio in the 5-day dark-grown seedlings in the single mutants, in contrast to the 5-day light-grown seedlings.

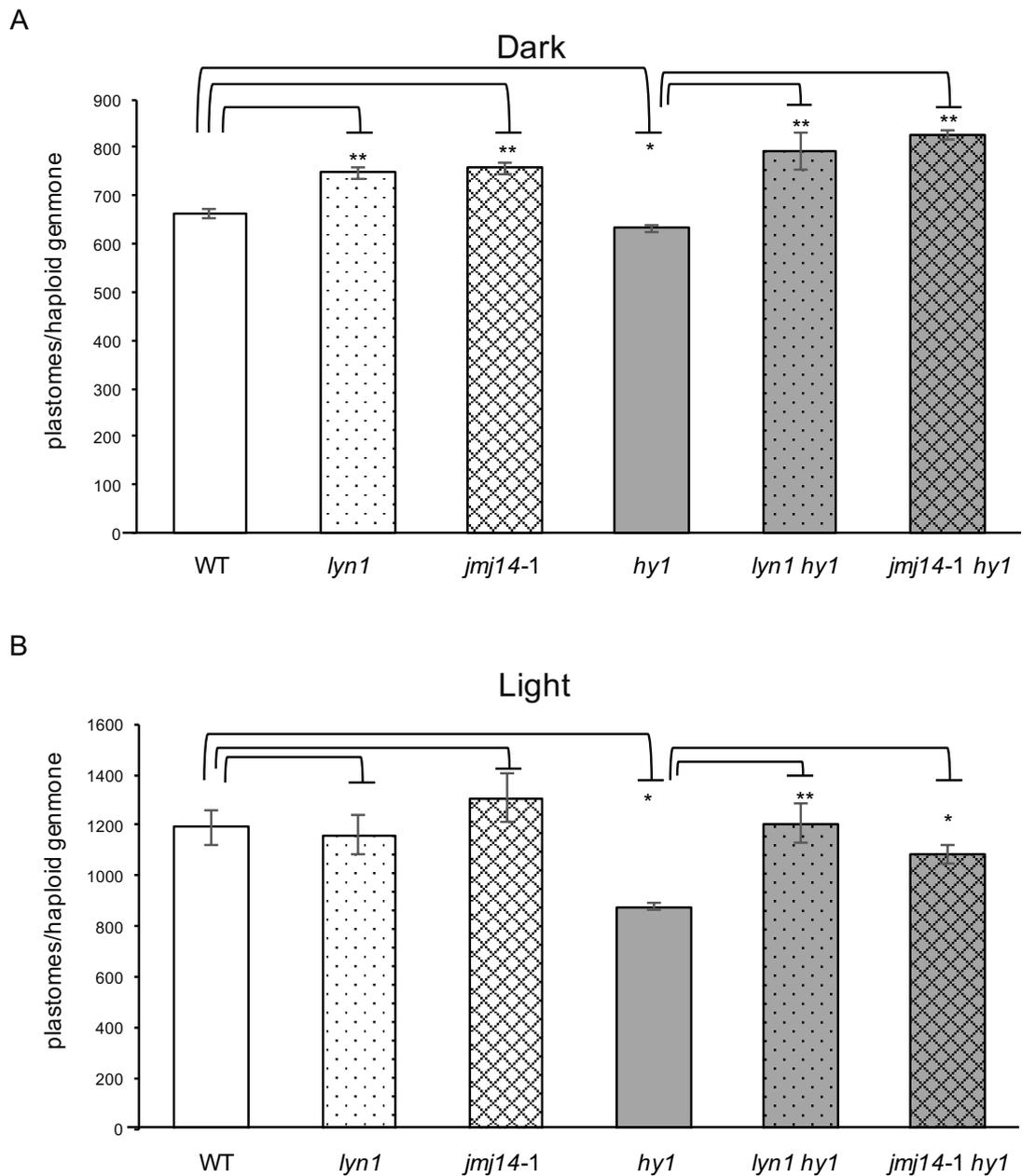


Figure 4. 7 Quantitation of plastome to nuclear genome copy number ratio
 Number of copies of the plastid genome per haploid nuclear genome, determined by qPCR, for 5-day dark-grown seedlings (A) and 5-day light-grown seedlings (B) of WT, *lyn1*, *jmj14-1*, *hy1*, *lyn1 hy1* and *jmj14-1 hy1*, n=4.

4.2.3 *lyn1* elevates Pchlide level

I set out to quantify the accumulation of excess Pchlide through its lethality-causing effects through the FR-light block of greening experimental approach.

Other seedlings were grown in the dark for 1, 2 and 3 day as control for FR light-grown seedlings. As shown in Figure 4.8, 4.9 and 4.10, all seedlings present could green upon transfer to white light directly, without FR or dark treatment. Figure 4.11 A illustrates the FR block of greening responses of WT, *lyn1* and *jmj14-1*. All 3 genotypes of deetiolated seedlings show a slightly decreased greening ability with increased dark growth periods. Even in the absence of a FR light high irradiance response during the dark growth followed by white light transfer, the impact on the seedlings of the transition from etiolation to deetiolation causes a decrease in greening ability. Longer etiolation causes lower greening ability of seedlings. 3d etiolated seedlings led to about 83% of green seedlings after 3d deetiolation. FR light-grown seedlings already appear photo-bleached even after 1 day FR exposure (Figure 4.9 A). Only about 18 % of seedlings were green after 3d FR exposure (Figure 4.9 C and 4.11 A). Although the greening ability of all three genotypes rapidly decreases and their greening percentages are very similar after 2d FR or 3d FR exposure (Figure 4.9 B, C and 4.11 A), the results show a clear difference among the three genotypes after 1 day FR exposure: under those conditions, *lyn1* and *jmj14-1* have a lower percentage of green seedlings than WT (Figure 4.9 A and 4.11 A). About 93% of WT seedlings were green while 78% of *lyn1* and 72% *jmj14-1* were so (Figure 4.11 A). The differences of greening ability are statistically significant. Figure 4.11 B illustrates the FR block of greening responses of *hy1*, *lyn1 hy1* and *jmj14-1 hy1*. Just like the dark-grown seedlings of WT and *jmj14* single mutants, all 3 *hy1* genotypes of deetiolated seedlings show a slight decrease in greening ability with increased dark growth periods (Figure 4.10). 95% *hy1* seedlings are green and 90% *hy1* double mutant seedlings are green after 3d deetiolation (Figure 4.10 C and 4.11 B). Because *hy1* is a light response defect mutant, the impact on the seedlings from etiolation to deetiolation is not as strong in *hy1* double mutants. When comparing the FR and dark pre-treatment results for *hy1*, it is apparent that FR irradiation has a very small effect of *hy1*, with almost 92% of seedlings

retaining their ability to green after 3 days FR treatment (Figure 4.10 C and 4.11 B). However, the FR treatment has dramatic impacts on *hy1* double mutants. Their greening ability gradually decreased with increased FR treatment periods, with about 52% of *lyn1 hy1* and 38% of *jmj14-1 hy1* remaining green after 3d FR exposure. Therefore, this experiment shows that the *jmj14* mutations restore the FR-HIR on the double mutants but also increase it in the single mutants, and also indirectly supports the conclusions of the previous experiment and the hypothesis in which *jmj14* can increase the Pchl_a level in the dark.

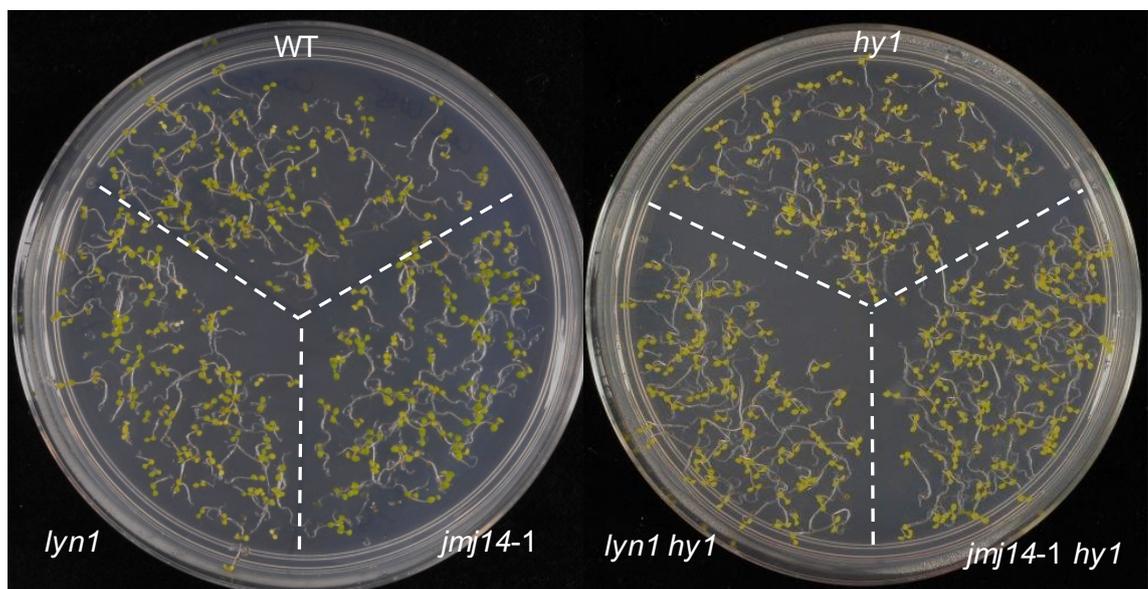


Figure 4. 8 Cotyledon greening of (control) plates without light treatment
Phenotypes of WT, *lyn1*, *jmj14-1*, *hy1*, *lyn1 hy1* and *jmj14-1 hy1* seedlings after exposure to 3-day continuous dark and 5-day continuous white light.

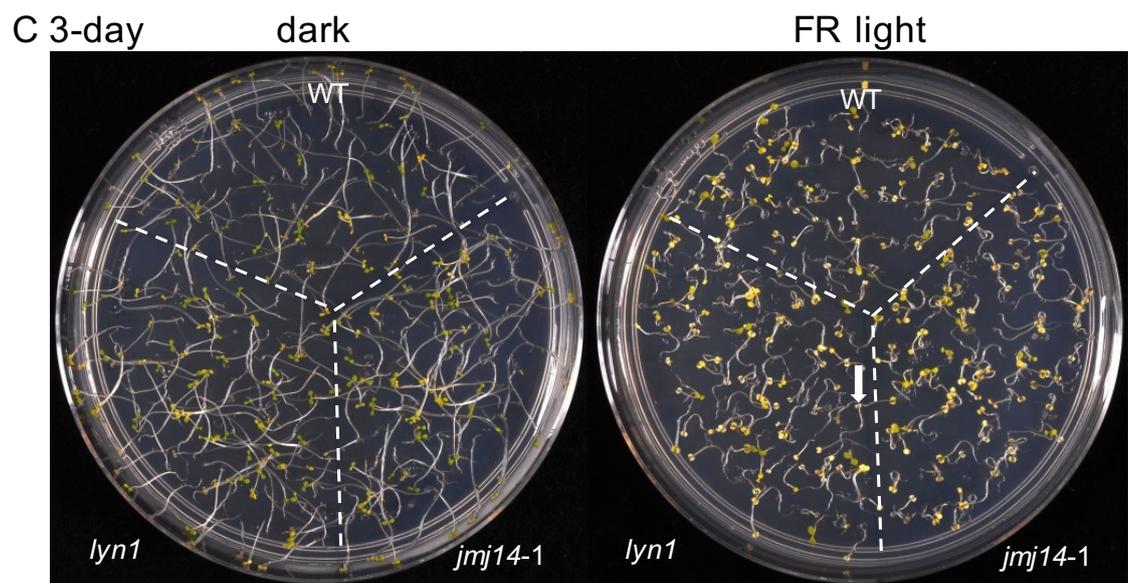
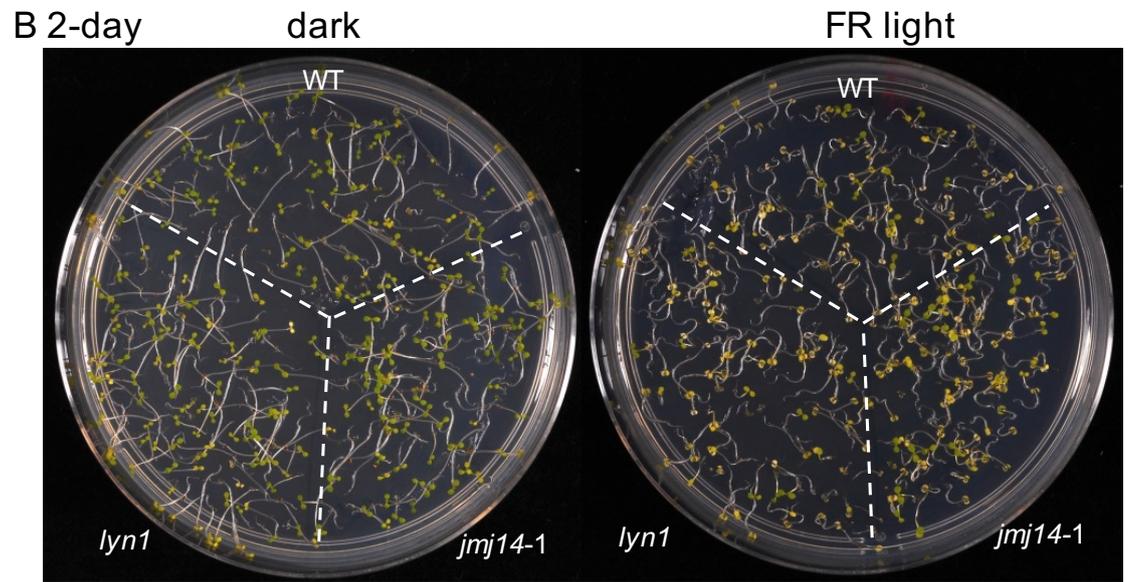
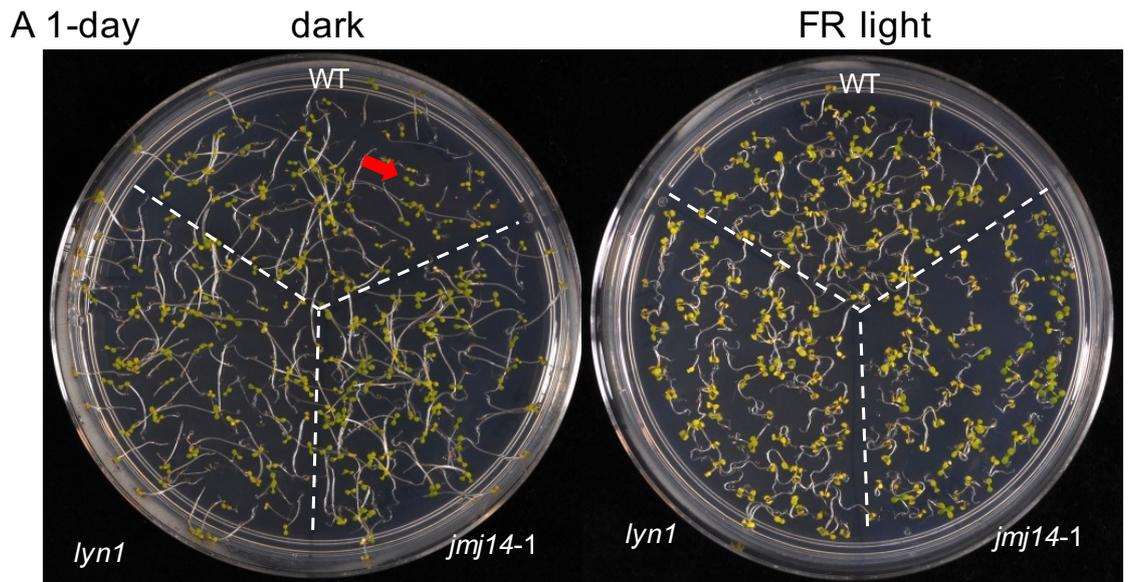




Figure 4. 9 FR-HIR of cotyledon greening of WT and *jmj14* single mutants by treating with different periods of FR light exposure

Phenotypes of WT, *lyn1* and *jmj14-1* seedlings after exposure to 3-day dark, then the following light pre-treatments: (A) 1-day of either continuous dark (left) or continuous FR light (right) followed by 5-day continuous white light; (B) 2-day of either continuous dark or continuous FR light followed by 5-day continuous white light; (C) 3-day of either continuous dark or continuous FR light followed by 5-day continuous white light. (D) The red arrow within the image displays an example of a seedling which was able to green well upon exposure to white light. The white arrow points to an example seedling which has undergone photo-bleaching.

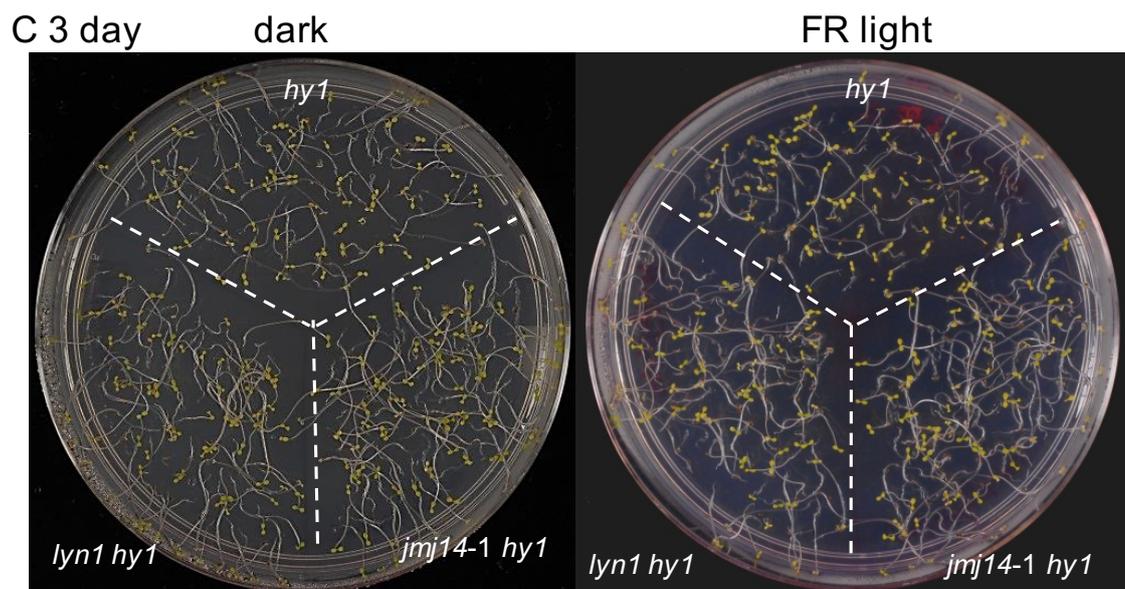
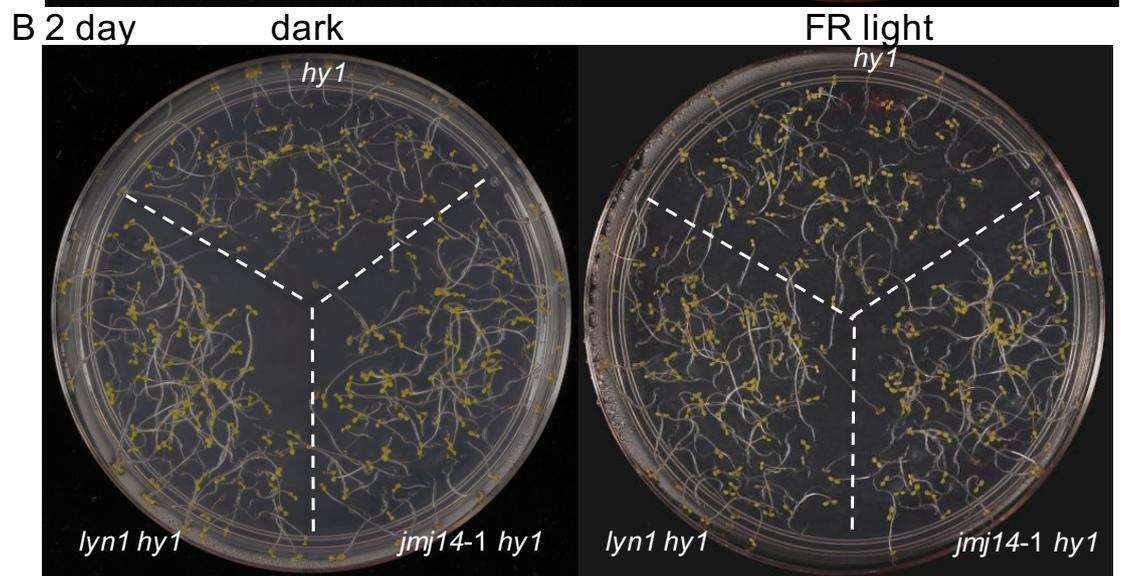
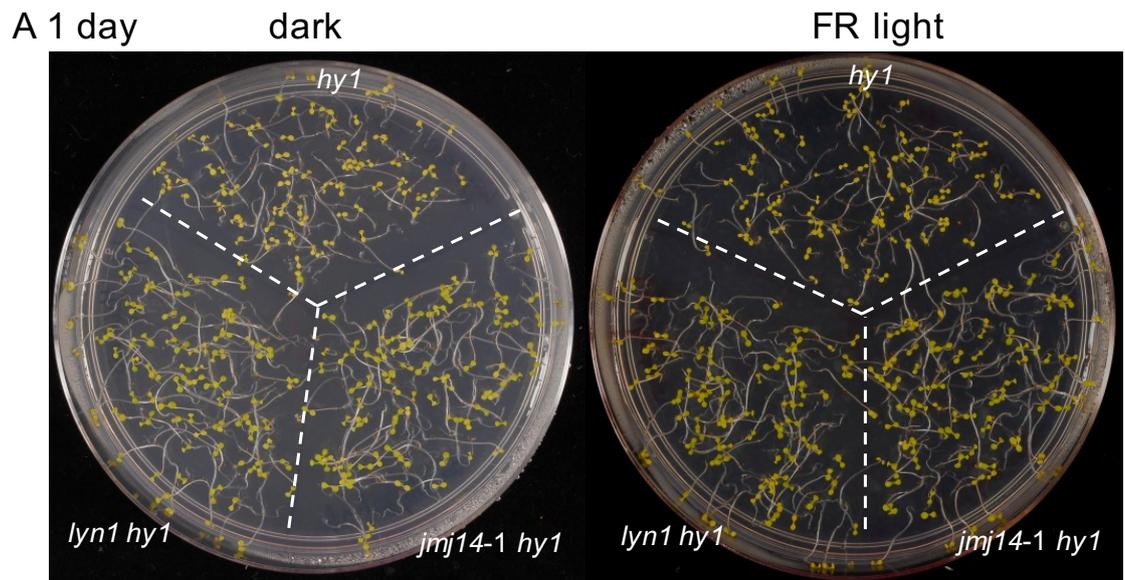


Figure 4. 10 FR-HIR of cotyledon greening of *hy1* mutants by treating with different periods of FR light exposure

Phenotypes of *hy1*, *lyn1 hy1* and *jmj14-1 hy1* seedlings after exposure to 3-day dark, then the following light pre-treatments: (A) 1-day of either continuous dark (left) or continuous FR light (right) followed by 5-day continuous white light; (B) 2-day of either continuous dark or continuous FR light followed by 5-day continuous white light; (C) 3-day of either continuous dark or continuous FR light followed by 5-day continuous white light.

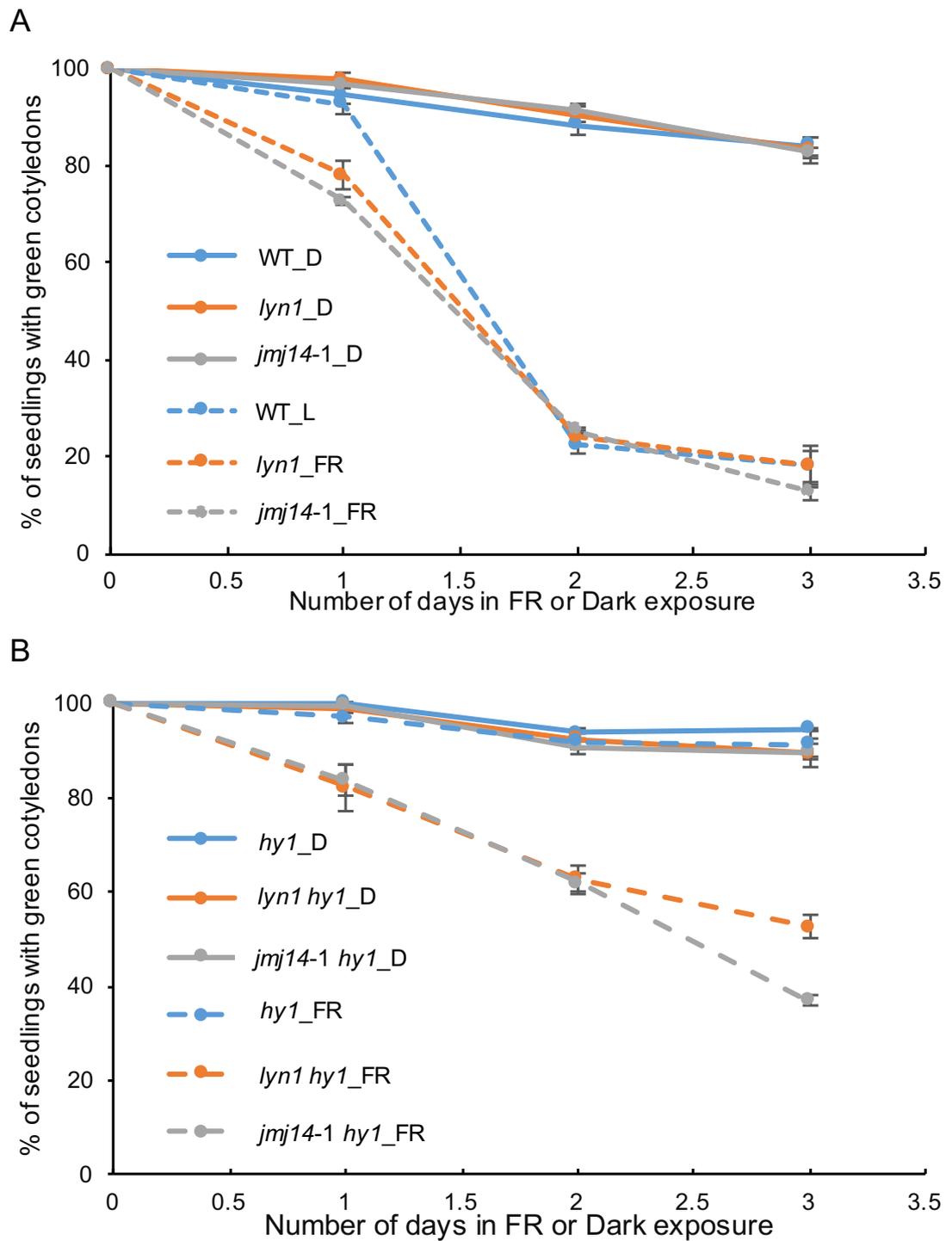


Figure 4. 11 FR-HIR of cotyledon greening rate of the different genotypes

A. WT, *lyn1* and *jmj14-1* seedlings greening rate; After 1 day FR exposure, the WT has a higher greening ability than *jmj14* mutants. The differences are significant, n=100;

B. *hy1*, *lyn1 hy1* and *jmj14-1 hy1* seedlings greening rate, n=100.

4.2.4 *lyn1* causes a small degree of suppression of *hy2-105*

4.2.4.1 Identification of the *lyn1 hy2-105* double mutant

Before doing any further experimental work with it, I assessed whether *hy2-105* is a KO mutant for the *HY2* gene. *hy2-105* phenotype shows only partially light-grown seedling features (Figure 4.14). RT-PCR confirmed that the full-length *HY2* mRNA is not present in *hy2-105* (Figure 4.12). Therefore, I refer to this genotype as *hy2-105* KO.

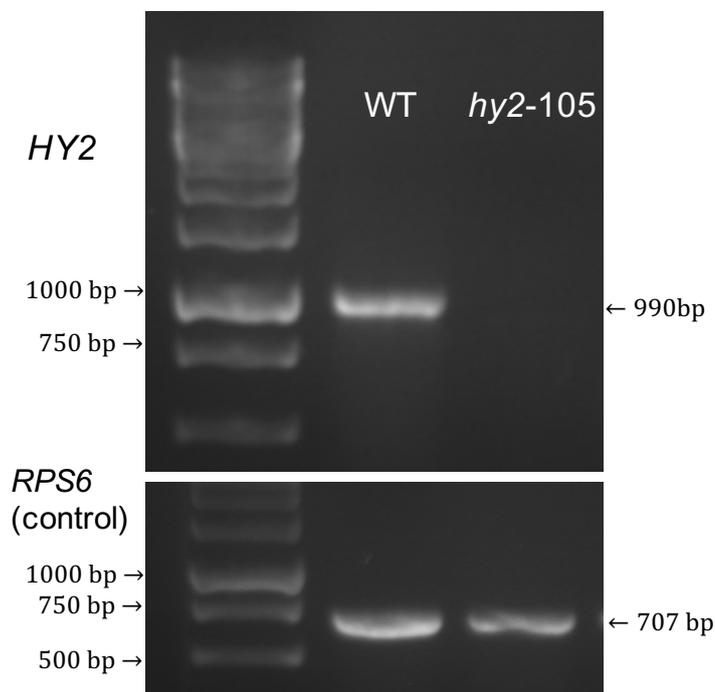


Figure 4. 12 *HY2* mRNA was confirmed as absent in the *hy2-105* KO plant (Col background) by reverse transcription

(Top) RT-PCR with *HY2* full-length primers. A 990 bp amplicon was amplified in the WT. (Bottom) *RPS6* was used as an internal control.

Confirmed *hy2-105* plants were crossed with *lyn1* to obtain F2 plants. Four *lyn1 hy2-105* double mutant plants were identified and selected from F2 plants by *lyn1* dCAPs assay and *hy2-105* CAPs assay. *lyn1* dCAPs assay shows that F2 (*lyn1*

x *hy2-105*) plants no. 3, 9, 10, 11, 21 and 26 are all *lyn1* homozygous mutants. Plant no. 2 is *lyn1* homozygous WT and plant no. 23 is *lyn1* heterozygous (Figure 4.13 A). Plants no. 3, 9, 10, 11, 21 and 26 were selected to test their *hy2-105* genotype. *hy2-105* CAPs assay shows that plants no. 3, 9, 10 and 11 are *hy2-105* homozygous mutants. Plant no. 21 is *hy2-105* heterozygous and plant no. 26 is *hy2-105* homozygous WT. Therefore, plants no. 3, 9, 10 and 11 are identified as *lyn1 hy2-105* double mutant plants (Figure 4.13 B).

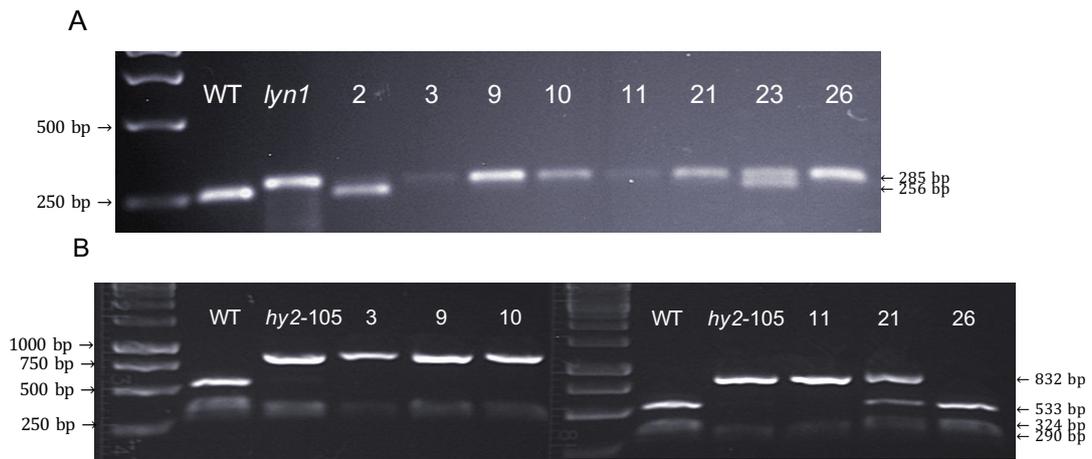


Figure 4. 13 Identification *lyn1 hy2-105* double mutants

A. *lyn1* dCAPs assay for the identification of *lyn1* mutants. PCR products of dCAPs assay of WT, *lyn1* and *lyn1* x *hy2-105* F2 plants, digested with *HaeIII* restriction enzyme and separated on 2% agarose gel are shown.

B. *hy2-105* CAPs assay for the identification of *hy2-105* mutants. PCR products of CAPs assay of WT, *hy2-105* and *lyn1* x *hy2-105* F2 plants, digested with *MfeI*-HF restriction enzyme and separated on 2% agarose gel are shown.

After *lyn1 hy2-105* double KO mutants were identified, four pools of F2 *lyn1 hy2-105* seeds were pooled together and grown on plates. As a preliminary result, the *lyn1* suppression of *hy2-105* was visible to the naked eye in 5 day-light grown seedlings. *lyn1 hy2-105* seedlings had shorter hypocotyl length and larger cotyledons (Figure 4.14). Hypocotyl length and chlorophyll content were further quantified to determine the *lyn1* suppression of *hy2-105*. The significance of differences among the different genotypes was assessed.



Figure 4.14 *lyn1* suppression of phenotype of *hy2-105*

Image of individual 5-day light grown WT, *lyn1*, *hy2-105* and *lyn1 hy2-105* seedlings. Seedlings were grown in white light ($100 \mu\text{mol m}^{-2} \text{s}^{-1}$) for 5 days in 1% MS media. Scale bar=10 mm.

4.2.4.2 Hypocotyl length

lyn1 suppression of hypocotyl elongation was observed under 3 light fluence rates (10, 50 and $100 \mu\text{mol m}^{-2} \text{s}^{-1}$). Statistically significant differences between genotypes were tested by one-way ANOVA, and revealed that one or more comparisons are significantly different in all three light conditions. This was followed by Bonferroni and Holm multiple comparisons testing, which only pairs sample groups relative to a reference group. In the Bonferroni and Holm multiple comparisons, the single mutants, *hy2-105* and *hy1*, are compared to WT, and the double mutant, *lyn1 hy2-105*, is compared to *hy1*. In all 3 light conditions, *hy2-105* hypocotyl length has a significant difference with that of the WT, its hypocotyl length being longer than WT (Figure 4.15). *lyn1 hy2-105* hypocotyl lengths were always shorter than that of *hy2-105* in any light conditions. However, the differences were only significant in intermediate and high fluence rates. Although *lyn1* suppression of *hy2-105* was seen in intermediate and high fluence rates, the extent of the suppression was different under the different fluence rates. The significance of the difference was highest at the intermediate fluence rate, so the *lyn1* suppression of hypocotyl elongation is the greatest at this light condition.

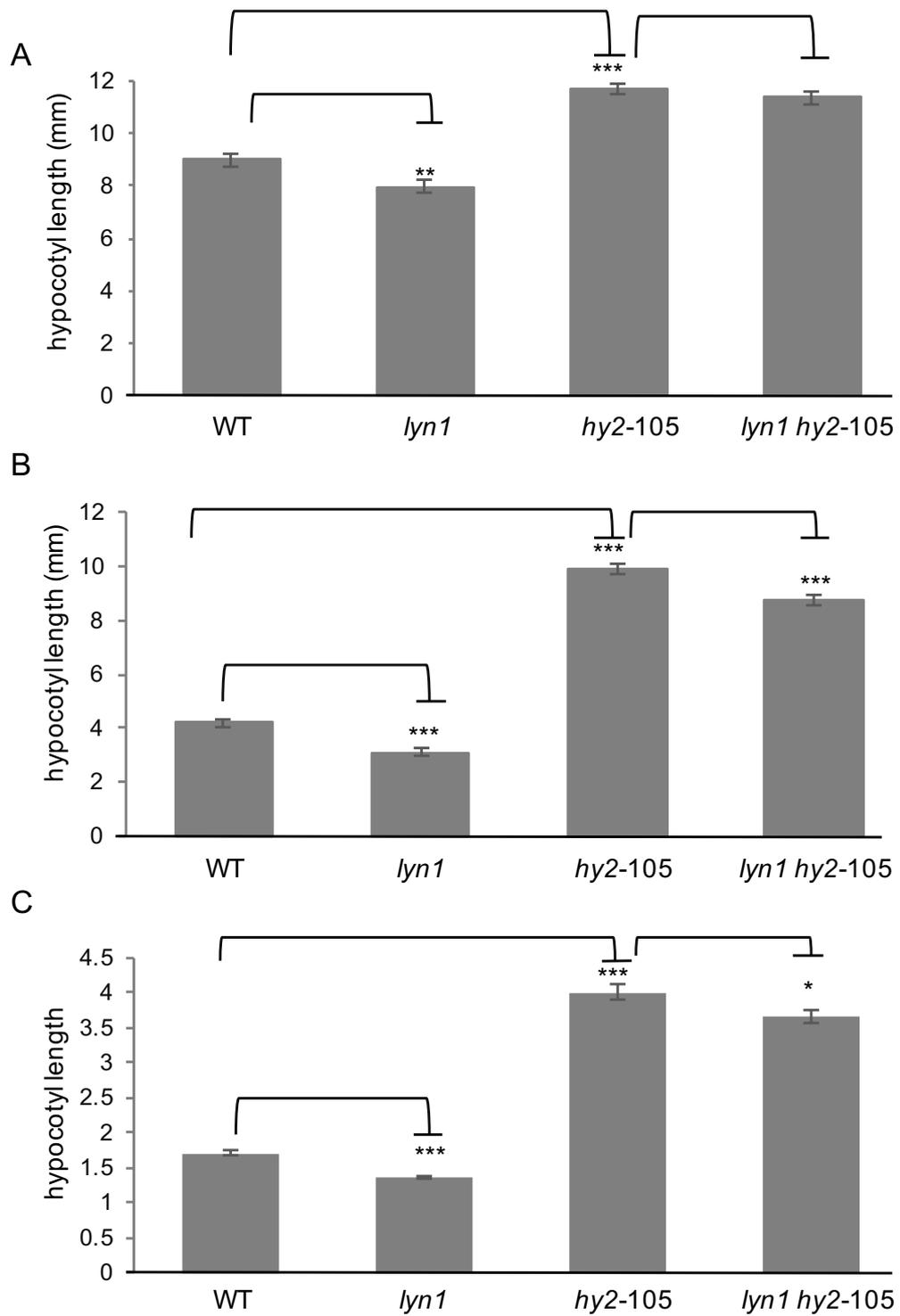


Figure 4.15 *lyn1* effect of hypocotyl length of *hy2-105*

Seedlings of WT, *lyn1*, *hy2-105* and *lyn1 hy2-105* were grown in continued white light (10, 50 and 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$) for 5 days on 1% MS media plate.

A. Low fluence rate light (10 $\mu\text{mol m}^{-2} \text{s}^{-1}$), n=30;

B. Medium fluence rate light (50 $\mu\text{mol m}^{-2} \text{s}^{-1}$), n=20;

C. High fluence rate light (100 $\mu\text{mol m}^{-2} \text{s}^{-1}$), n=20.

4.2.4.3 Chlorophyll content

lyn1 suppression as manifested by chlorophyll content of *hy2-105* was also observed under the 3 light conditions (Figure 4.16). The same statistical test used to indicate the difference of hypocotyl length was also carried out in this assay. In all 3 light conditions, *hy2-105* has lower chlorophyll content compared to WT. The differences are significant under intermediate and high fluence rates. *lyn1 hy2-105* chlorophyll content was higher than that of *hy2-105* in all three light conditions. However, the difference was only significant in intermediate and high fluence rates, but not in lower fluence rate. *lyn1* chlorophyll content was always higher than that of WT, but the extent of the suppression was different under the different fluence rates. The significance of the difference was highest at the intermediate and high fluence rates, so the *lyn1* suppression of chlorophyll content is the greatest at this two light conditions (Figure 4.16 B and C).

The purpose of testing the *lyn1* suppression of *hy2-105* was to determine whether *lyn1* suppresses the *hy1* mutation specifically or it does also any other PΦB synthesis defective mutant, like *hy2-105*. My hypothesis was that *lyn1* can rescue any PΦB synthesis defective mutant, so I expected that *lyn1* should also suppress growth and greening of *hy2-105*. Although *lyn1* did not always show the suppression of *hy2-105*, it showed suppression in certain conditions, which is sufficient to support the hypothesis. The best light condition to observe *lyn1* suppression of *hy2-105* turned out to be under the intermediate fluence rate light.

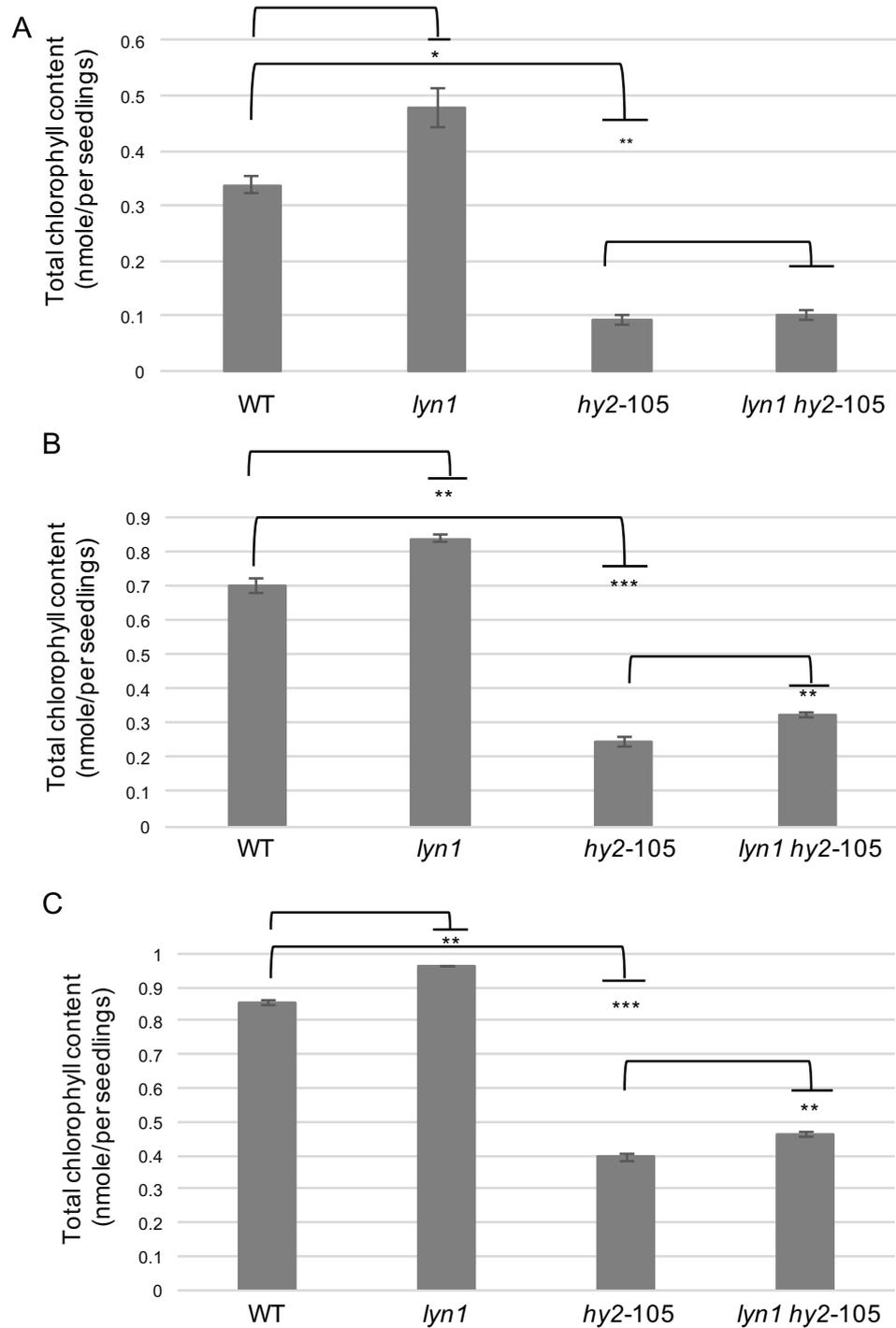


Figure 4. 16 *lyn1* effect of chlorophyll content of *hy2-105*

Seedlings of WT, *lyn1*, *hy2-105* and *lyn1 hy2-105* grown in continuous white light (10, 50 and 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$) for 5 days in 1% MS media.

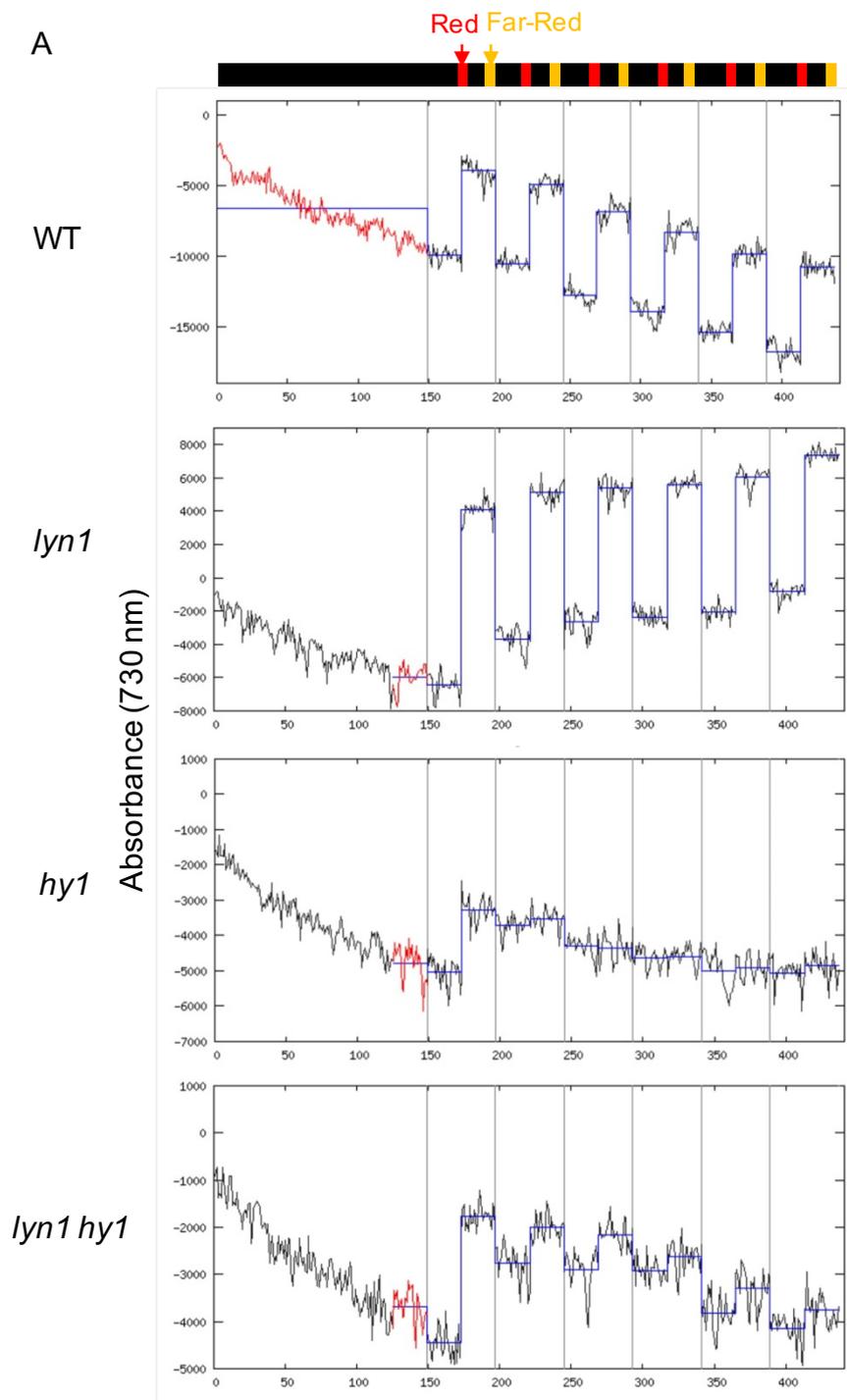
A. low fluence light (10 $\mu\text{mol m}^{-2} \text{s}^{-1}$), n=3;

B. medium fluence rate (50 $\mu\text{mol m}^{-2} \text{s}^{-1}$), n=3;

C. high fluence rate (100 $\mu\text{mol m}^{-2} \text{s}^{-1}$), n=4.

4.2.5 *lyn1* causes a detectable presence of phytochromobilin-containing photochemically active phytochrome in *hy1* and elevates its level in the WT

This experiment was performed by the laboratory of Tim Kunkel, University of Freiburg. The intrinsic photochemical activity of the chromophore prosthetic group allows phytochromes to convert between the two forms, either the R light-absorbing (Pr) form or the FR light-absorbing (Pfr) form. The Pr form absorbs maximally at 660 nm (and therefore less at 730 nm), while the Pfr form absorbs maximally at 730 nm (Li et al., 2011). The phytochromes kept converting between two photo-interconvertible forms and the level of photochemically active phytochromes can be detected as the average of absorbance change at 730nm (Figure 4.17 A). From the absorbance change diagrams received from three independently-grown seedling batches in Tim Kunkel's laboratory, the averages of absorbance change were measured and plotted (Figure 4.17 B). The WT showed exceptionally clear photochemically active phytochrome signals. Such signals were virtually absent from *hy1*. Crucially, *lyn1* caused a detectable photochemically active phytochrome level in *lyn1 hy1*. This can lead to a better light response in the double mutant. The statistical test indicates the difference of photochemically active phytochrome level between *hy1* and *lyn1 hy1* is significant. Unexpectedly, *lyn1* had a significantly higher photochemically active phytochrome level than the WT. Therefore, the *lyn1* mutation caused the appearance of a small amount of photochemically active phytochrome.



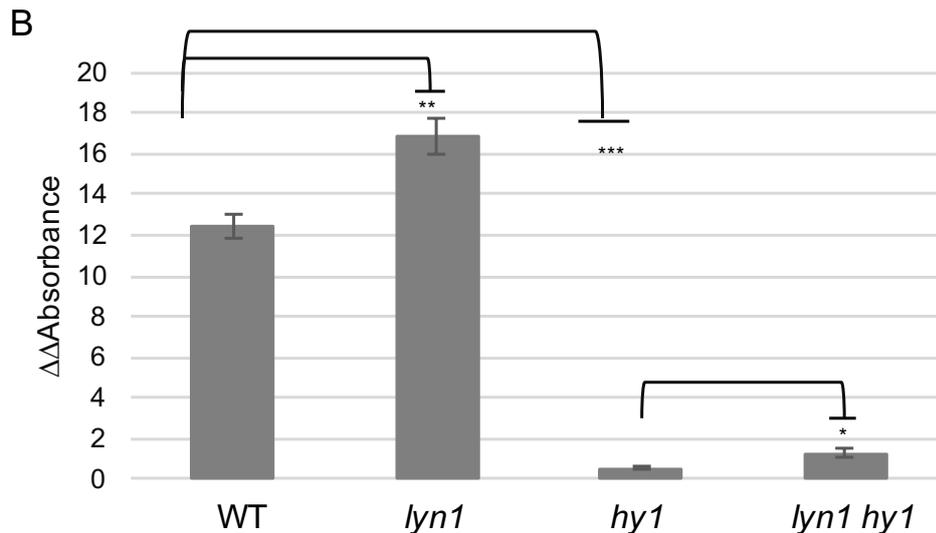


Figure 4. 17 Relative different amount of photochemically active phytochrome in 4 genotypes

A. Absorbance trace measurements of photochemical activity of phytochromes, representative individual experiments. Dark-grown seedlings were exposed to R light followed by FR light. R and FR light are consecutively applied several times (see chapter 2.8). Photochemical activity of phytochromes can be detected as the absorbance change at 730 nm by a ratio spectrophotometer. (Katrin Biergman and Tim Kunkel, University of Freiburg).

B. Quantitation of spectrophotometrically active phytochromes for dark-grown seedlings (n=3).

4.2.6 Confirming whether *lyn1* suppresses photoreceptor mutants *phyA-211*, *phyB-9* and *cry1-304*

Phytochromes are synthesized when phytochrome apoproteins covalently attach to PΦB in the cytoplasm. The *hy1* mutation causes a PΦB biosynthesis defect which results in the light responses defect. According to the phytychromobilin-containing photochemically active phytochrome level quantification assay, the hypothesis according to which *lyn1* rescues the light response defect by increasing the PΦB biosynthesis in plastids was supported. Alternative support

for this hypothesis would be provided by evidence showing that *lyn1* is unable to suppress a phytochrome apoprotein mutant. A series of experiments was carried out by studying the effect of *lyn1* on phytochrome apoprotein defective plants, *phyA* and *phyB*, and cryptochrome apoprotein defective plants, *cry1*. If the hypothesis is correct, then *lyn1* is expected to not rescue *phyA*, *phyB* and *cry1* under their diagnostic light conditions. This does not deny the possibility of some rescue of defects under generic conditions. Possibly, this could be due to the fact that photoreceptors cooperate and interact (see later).

4.2.6.1 Identification of *lyn1 cry1-304*, *lyn1 phyB-9* and *lyn1 phyA-211*

To assay the effect of *lyn1* on phytochrome apoprotein defective mutants, *phyA-211 lyn1*, *phyB-9 lyn1* and *cry1-304 lyn1* were generated by crossing *phyA-211*, *phyB-9* and *cry1-304* with *lyn1* to obtain double mutants from F2 plants. Because cryptochromes respond to B light, so *cry1* causes response defects during B light treatment. Although phytochromes also respond to B light, they do not respond as efficiently as to R and FR light. Therefore, *phyB* causes response defects during R light treatment and *phyA* causes response defects during FR light treatment (Fankhauser and Casal, 2004). All three mutants are etiolated-like when they grow under their corresponding diagnostic light conditions. This characteristic was used to quickly identify *cry1-304*, *phyB-9* and *phyA-211* mutants from F2 plants without the need for genotyping.

cry1-304 mutant was selected by growing the F2 (*lyn1* x *cry1-304*) seedlings under constant B light ($50 \mu\text{mol m}^{-2} \text{s}^{-1}$). F2 plants no. 1, 3, 4, 5 and 6 were selected as they had long hypocotyl lengths when they grew under constant B light. *lyn1* dCAPs assay shows that plant no. 5 is *lyn1* homozygous mutant. Plants no. 1 and 4 are *lyn1* heterozygous mutants. Plants no.3 and 6 are *LYN1* homozygous WT (Figure 4.18). Therefore, plant no. 5 is identified as *lyn1 cry1-304* double mutant.

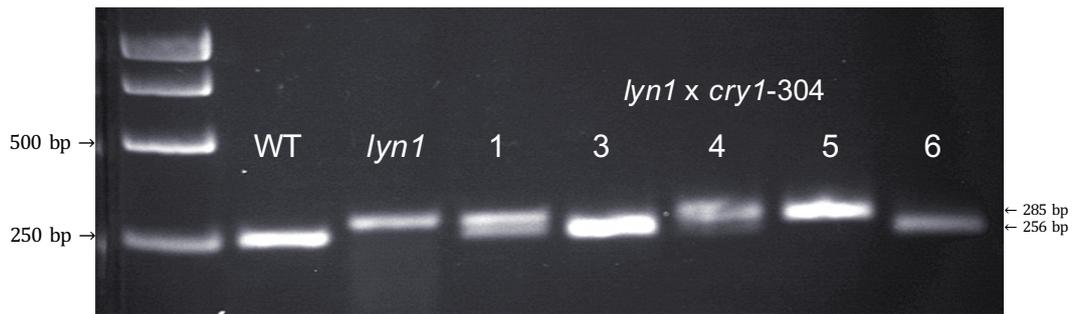


Figure 4. 18 Identification *lyn1 cry1-304* double mutants

lyn1 mutants were identified from *cry1-304* mutants by *lyn1* dCAPs assay. PCR products of dCAPs WT, *lyn1*, F2 (*lyn1* x *cry1-304*) plants no. 1, 3, 4, 5 and 6, digested with HaeIII restriction enzyme and separated on 2% agarose gel are shown.

PhyB-9 mutants were selected by growing the F2 (*lyn1* x *phyB-9*) plants under constant R light ($50 \mu\text{mol m}^{-2} \text{s}^{-1}$). F2 plants no. 9-3, 9-4, 9-5, 9-6, 9-7, 9-8, 9-9, 11-1 and 11-2 were selected as they had long hypocotyl lengths when they grew under constant R light. *lyn1* dCAPs assay shows that plants no. 9-4, 9-5, 9-6, 9-7, 9-9 and 11-1 are *lyn1* homozygous mutants. Plant no. 9-8 plant is *LYN1* homozygous WT. Plants no. 9-3 and 11-2 are *lyn1* heterozygous (Figure 4.19). Therefore, plants no. 9-4, 9-5, 9-7, 9-9 and 11-1 are identified as *lyn1 phyB-9* double KO mutants.

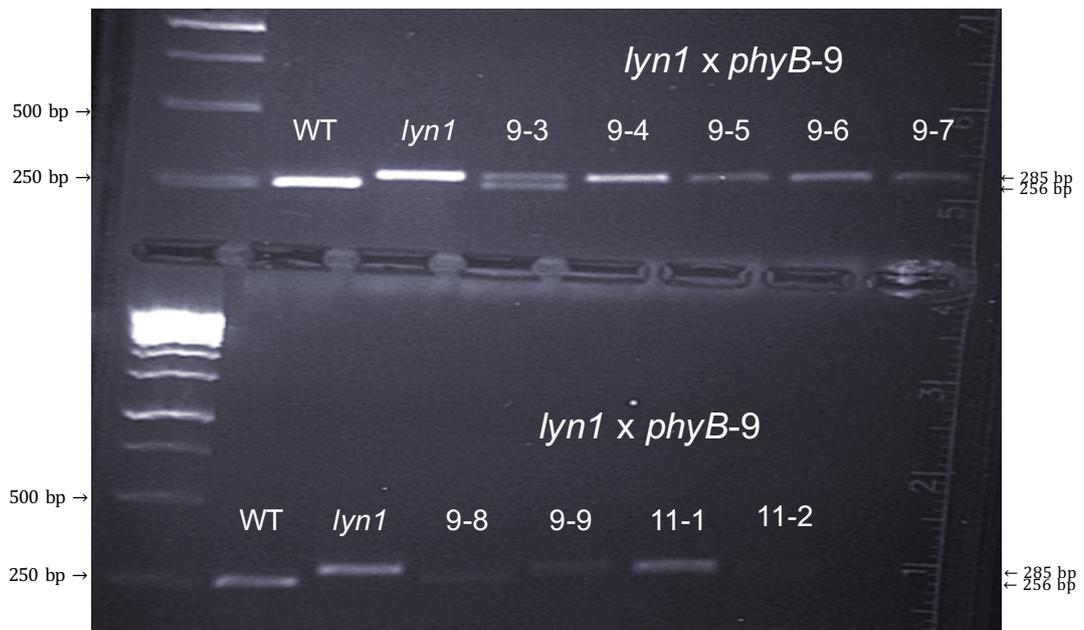


Figure 4. 19 Identification *lyn1 phyB-9* double mutants

lyn1 mutants were identified from *phyB-9* mutants by *lyn1* dCAPs assay. PCR products of dCAPs WT, *lyn1*, F2 (*lyn1 x phyB-9*) plants no. 9-3, 9-4, 9-5, 9-6, 9-7, 9-8, 9-9, 11-1 and 11-2 of digested with HaeIII restriction enzyme and separated on 2% agarose gel are shown.

phyA-211 mutants were selected by growing the F2 (*lyn1 x phyA-211*) plants under constant FR light ($30 \mu\text{mol m}^{-2} \text{s}^{-1}$). Using similar reasoning, I identified plant no. 8 plant is *lyn1* homozygous mutants (Figure 4.20). Plants no. 1, 4 and 5 are *LYN1* homozygous WT. Plants no. 2, 3, 6 and 9 are *lyn1* heterozygous. Therefore, plant no. 8 is identified as *lyn1 phyA-211* double mutant.

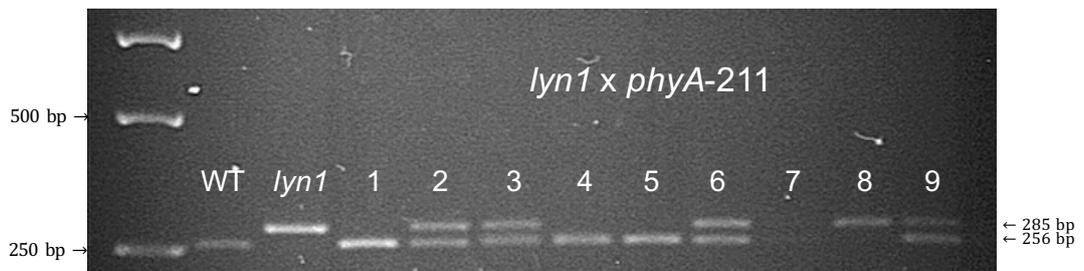


Figure 4. 20 Identification *lyn1 phyA-211* double mutants

lyn1 mutants were identified from *phyA-211* mutants by *lyn1* dCAPs assay. PCR products of dCAPs WT, *lyn1*, F2 (*lyn1 x phyA-211*) plants no.1~9 of digested with HaeIII restriction enzyme and separated on 2% agarose gel are shown.

Unexpectedly, I encountered two problems. The first was the old *phyA-211* seeds had very poor germination rate. The second was that seedlings from new *phyA-211* and *lyn1 phyA-211* showed segregation of phenotypes, namely some seedlings had long hypocotyl with small and unopened cotyledons and some had short hypocotyl with large and opened cotyledons, when they grew under FR light (Figure 4.21). I speculated that this was a result of a mixed *phyA* genotype within my seed pool. I addressed this by harvesting longer hypocotyl (arrowed in Figure 4.21) and shorter hypocotyl seedlings (circled in Figure 4.21) separately, and then genotyping the *phyA-211* mutation on both types of seedlings.

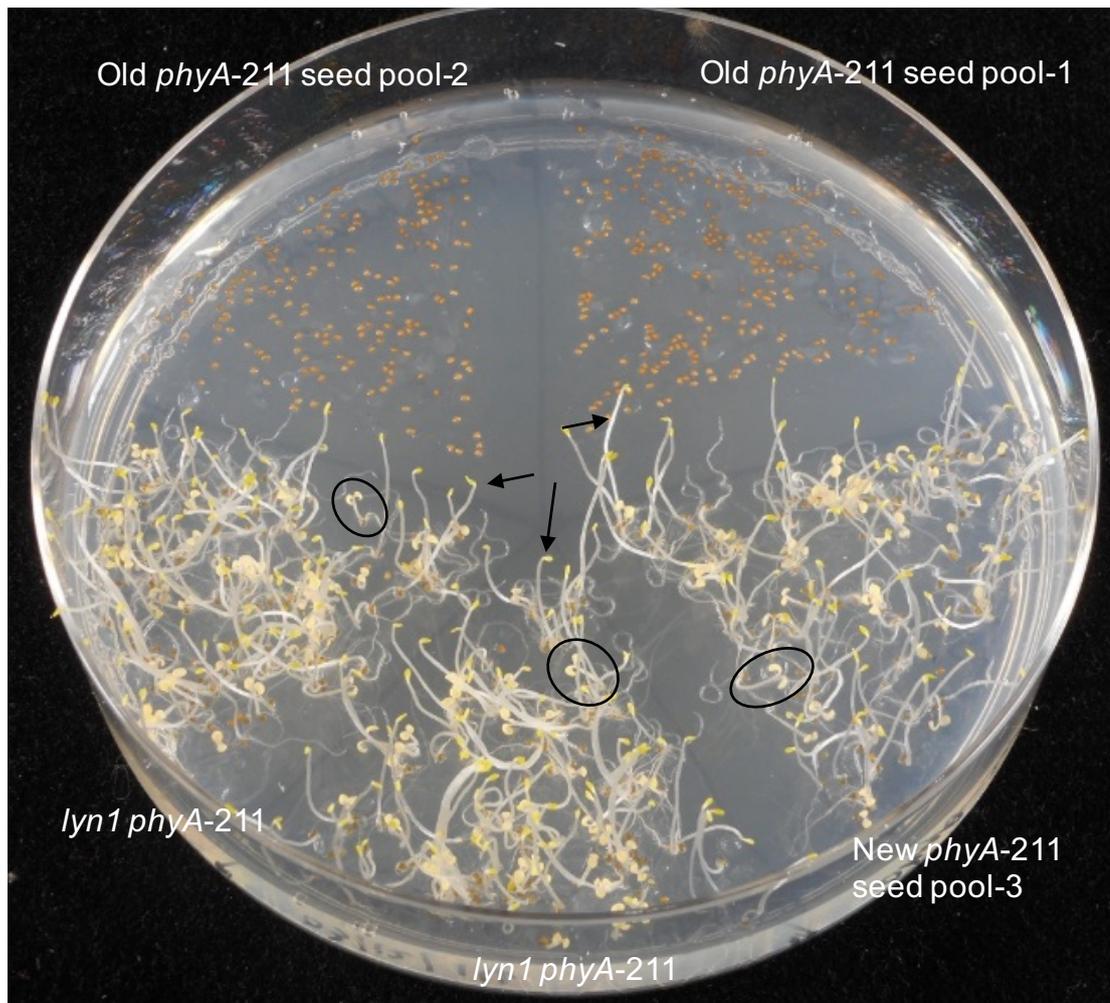


Figure 4. 21 Germination and segregation of *phyA-211* and *lyn1 phyA-211*

Two very old pools (1 and 2) of *phyA-211* seeds cannot germinate at all. New *phyA-211* seed pool was propagated from the *phyA* plants which germinated from the old *phyA-211* pools and used to cross with *lyn1*. Unexpected segregation occurred on both *phyA-211* and *lyn1 phyA-211*. Short hypocotyl seedlings (circled) and long hypocotyl seedlings (arrowed) appeared in both *phyA-211* and *lyn1 phyA-211*.

The *lyn1* dCAPS assay result confirmed that neither short nor long hypocotyl seedlings of *phyA-211* were *lyn1* mutants. Both short and long hypocotyl seedlings of *lyn1 phyA-211* were *lyn1* mutants (Figure 4.22)

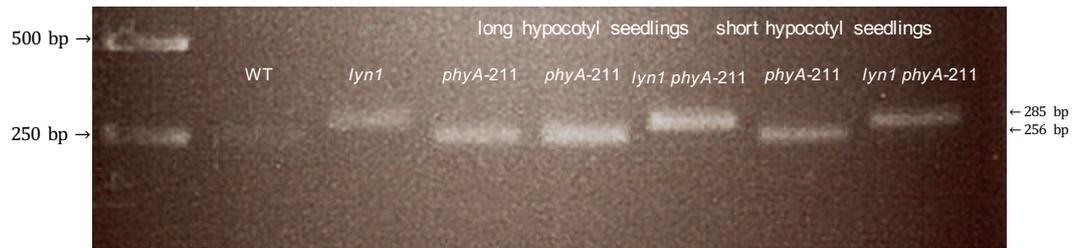


Figure 4. 22 *lyn1* dCAPs assay of re-identification *phyA-211* and *lyn1 phyA-211* mutants

PCR products of dCAPs WT, *lyn1*, long hypocotyl seedlings of *phyA-211*, long hypocotyl seedlings of *lyn1 phyA-211*, short hypocotyl seedlings of *phyA-211* and short hypocotyl seedlings of *lyn1 phyA-211* digested with *HaeIII* restriction enzyme and separated on 2% agarose gel are shown.

After *phyA-211* and *lyn1 phyA-211* had been confirmed as having the correct genotype at the *LYN1* locus in both long hypocotyl and short hypocotyl seedlings, only the long hypocotyl seedlings of *phyA-211* and *lyn1 phyA-211* were subjected to *phyA-211* genotyping assay. Primer sequences of *phyA-211* genotyping assay were provided by C. Fankhauser at the University of Lausanne, who provided *phyA-211* seeds. The *phyA-211* genotyping primers can only give a 1237pb amplicon in the WT. They cannot amplify the homozygous mutant *phyA-211* (Christian Fankhauser lab, personal communication). *phyA-211* genotyping proved that *phyA-211* and *lyn1 phyA-211* long hypocotyl seedlings were both *phyA-211* mutants (Figure 4.23). Their seeds were subsequently used for the further research.

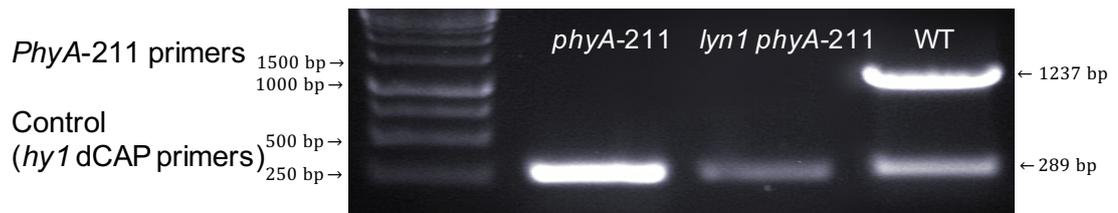


Figure 4. 23 Identification *phyA-211* mutants

PCR master mixture contained the forward and reverse primers for both *phyA-211* genotyping assay and *hy1* dCAP assay. (top band) *phyA-211* long seedlings and *lyn1 phyA-211* long seedlings produce no long amplicon. A 1237 bp amplicon was amplified in the WT. (Bottom band) *hy1* dCAP primers were used as a control.

4.2.6.2 Can *lyn1* suppress photoreceptor mutants in white light?

After *lyn1 phyA-211*, *lyn1 phyB-9* and *lyn1 cry1-304* were identified, they were grown together with WT, *lyn1* and their respective single mutants in the intermediate white light ($50 \mu\text{mol m}^{-2} \text{s}^{-1}$). As in the previous experiments, the *lyn1* effects on photoreceptor mutants were also observed through hypocotyl length and chlorophyll content quantification assay. If my hypothesis, according which *lyn1* suppresses the *hy1* light response defect by increasing PΦB biosynthesis but not affecting *phyA-211*, *phyB-9* and *cry1-304* photoreceptor apoprotein defects, was true, I expected that *lyn1* would not suppress hypocotyl length and chlorophyll content of these three photoreceptor apoprotein mutants. The suppression was tested by comparing *lyn1*, *phyA-211*, *phyB-9* and *cry1-304* to WT. *lyn1 phyA-211*, *lyn1 phyB-9* and *lyn1 cry1-304* were compared to their respective single mutants, *phyA-211*, *phyB-9* and *cry1-304*. The statistically significant differences between genotypes were tested by one-way ANOVA, then followed by Bonferroni and Holm multiple comparisons, which only pairs relative to a reference group.

4.2.6.2.1 Hypocotyl length

According to the results, *lyn1* seedlings had shorter hypocotyl length than the WT (Figure 4.24), a result similar to that shown in the last chapter. The difference between WT and *lyn1* was significant. *phyB-9* and *cry1-304* showed very similar results. Both of them were longer than the WT (Figure 4.24 A and B). The double mutants, *lyn1 phyB-9* and *lyn1 cry1-304*, were shorter than their respective single mutant control (*phyB-9* and *cry1-304*). The differences were significant in both cases. In contrast, *phyA-211* did not show a significant difference with the WT (Figure 4.24 C). *lyn1 phyA-211* was shorter than *phyA-211* and the difference was significant. Unexpectedly, *lyn1* was found to suppress hypocotyl elongation under white light on all three photoreceptor mutants.

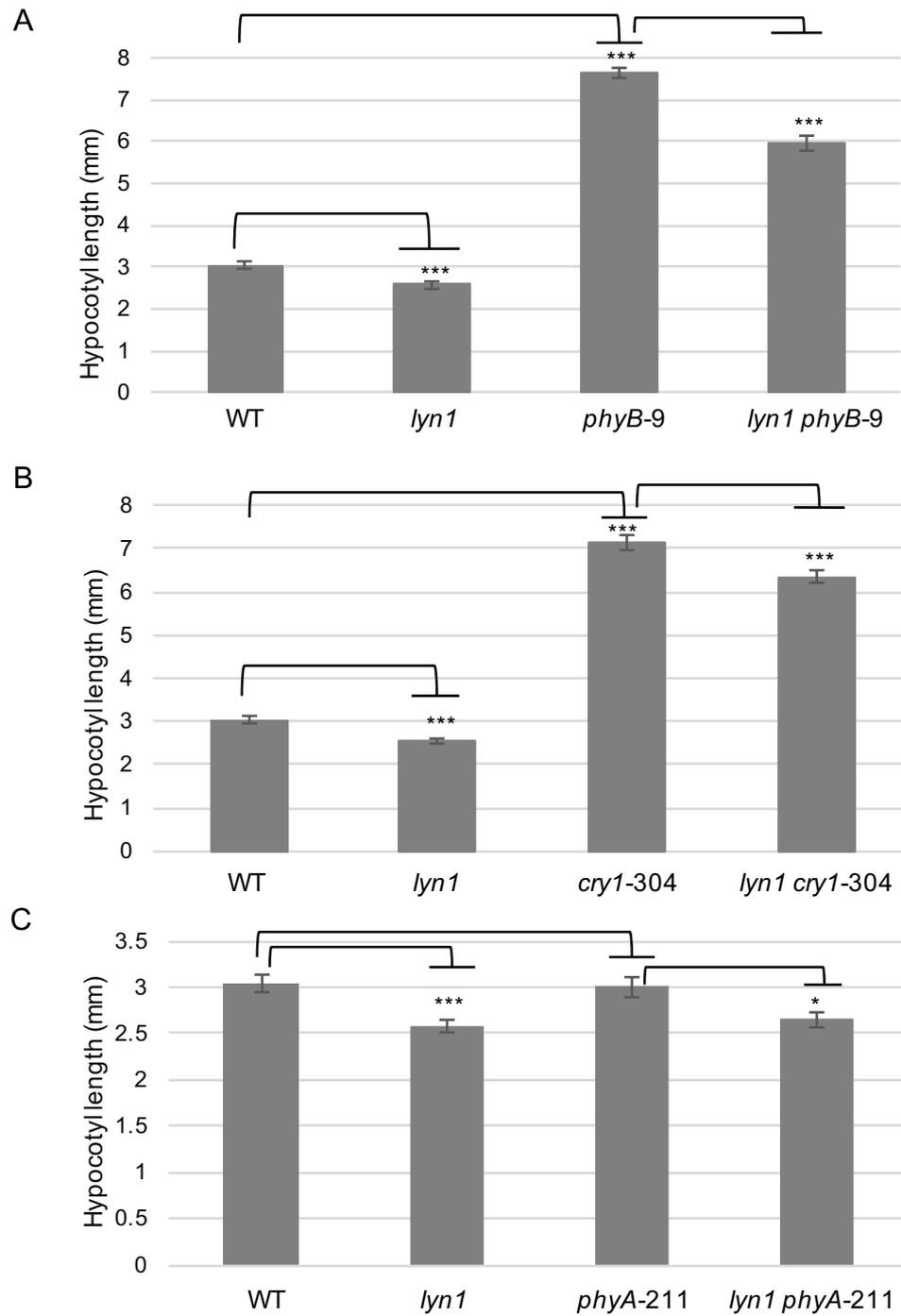


Figure 4. 24 *lyn1* suppression of hypocotyl length of photoreceptor apoprotein mutants in $50 \mu\text{mol m}^{-2} \text{s}^{-1}$ white light

A. *lyn1* suppression of *phyB-9*, n=20;

B. *lyn1* suppression of *cry1-304*, n=20;

C. *lyn1* suppression of *phyA-211*, n=20.

4.2.6.2.2 Chlorophyll content

Consistent with the result from the last chapter, the chlorophyll content of *lyn1* was higher than that of the WT, while in the WT it was higher than that of *cry1-304* and *phyB-9* and lower than that of *phyA-201*, all the differences being significant (Figure 4.25). *lyn1 phyB-9* and *lyn1 cry1-304* both had a higher chlorophyll content than their respective single mutant control (*phyB-9* and *cry1-304*) (Figure 4.25 A and B). The differences were significant. Although *lyn1 phyA-211* was lower than *phyA-211*, the difference was not significant (Figure 4.25 C). Therefore, unexpectedly, *lyn1* was capable of rescuing the chlorophyll content of *phyB-9* and *cry1-304*. Because *phyA-211* already had very high chlorophyll content, there was nothing for *lyn1* to rescue.

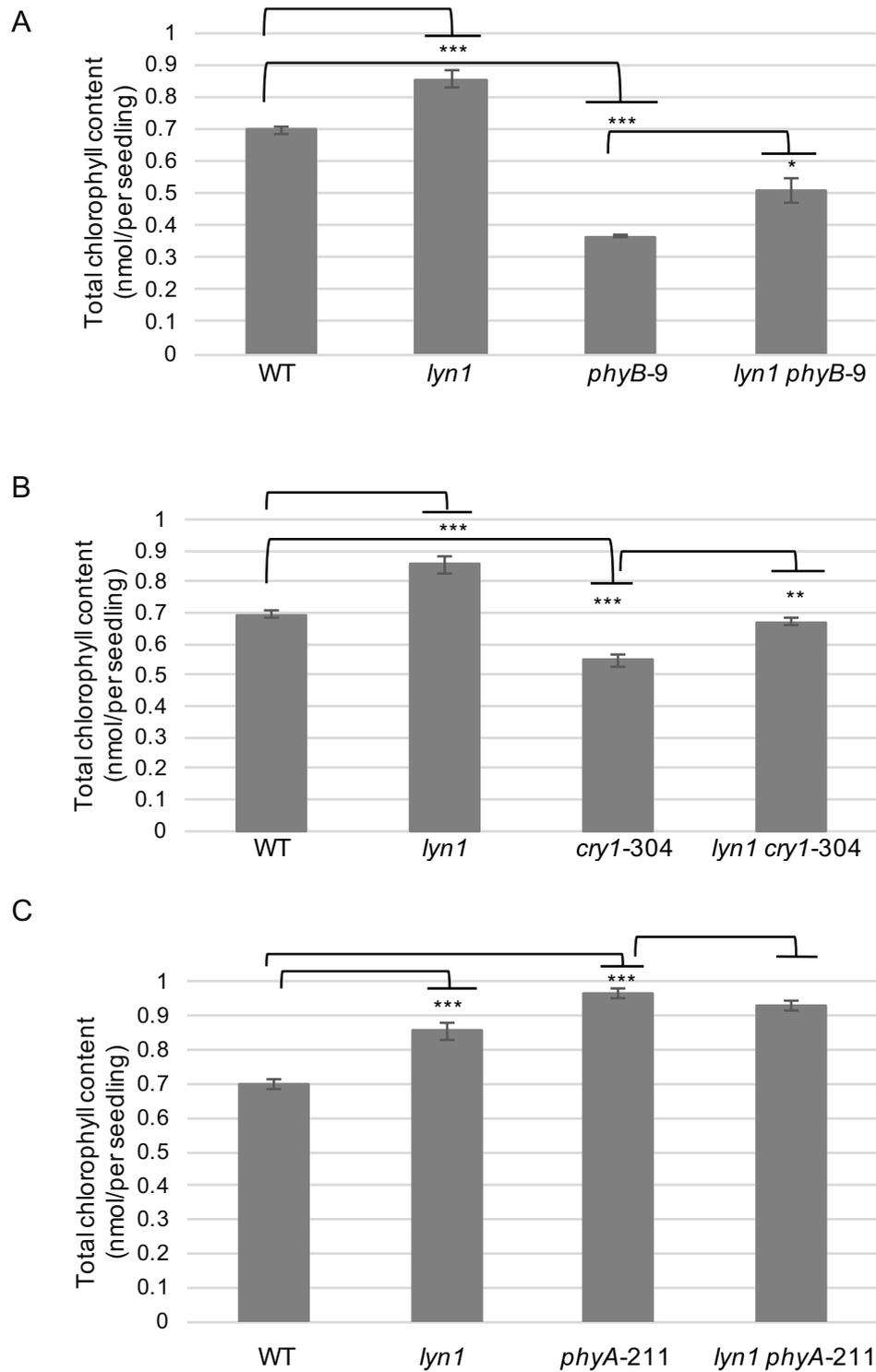


Figure 4. 25 *lyn1* suppression of chlorophyll content of photoreceptor apoprotein mutants in $50 \mu\text{mol m}^{-2} \text{s}^{-1}$ white light

A. *lyn1* suppression of *phyB-9*, n=3;

B. *lyn1* suppression of *cry1-304*, n=3;

C. *lyn1* suppression of *phyA-211*, n=3.

4.2.6.3 Can *lyn1* suppress photoreceptor mutants in their diagnostic light wavelength?

In its hypocotyl elongation and greening response, *lyn1* causes suppression of photoreceptor mutants in white light. This, at face value, does not support my hypothesis of a rescue of *hy1* based on enhanced plastid development leading to detectable chromophore synthesis. *phyA*, *phyB* and *cry1*, as active photoreceptors, mainly respond to FR, R and B lights, respectively. However, analysing the *lyn1* suppression of their respective photoreceptor apoprotein mutants in white light, which contains all 3 types of monochromatic light may introduce unwanted additional factors affecting the mutant phenotypes. To gain clarity, another set of experiments was carried out to further determine the *lyn1* suppression of the three photoreceptor mutants by measuring the hypocotyl length and chlorophyll content over a range of fluence rates of the light wavelength that the photoreceptor *specifically* responded to. In other words, the *lyn1* effect of hypocotyl elongation and chlorophyll content of *phyA-211* in FR light, *phyB-9* in R light and *cry1-304* in B light were measured to generate a fluence rate response curve. This assay was set in 4 light conditions. In white light treatment, WT, *lyn1*, *hy1* and *lyn1 hy1* were grown together on a plate over a fluence rate range (0~100 $\mu\text{mol m}^{-2} \text{s}^{-1}$) and used as the control. WT, *lyn1*, *phyA-211* and *lyn1 phyA-211* were grown over a fluence rate range (0~15 $\mu\text{mol m}^{-2} \text{s}^{-1}$) of FR light. WT, *lyn1*, *phyB-9* and *lyn1 phyB-9* were grown over a fluence rate range (0~40 $\mu\text{mol m}^{-2} \text{s}^{-1}$) of R light. WT, *lyn1*, *cry1-304* and *lyn1 cry1-304* were grown over a fluence rate range (0~50 $\mu\text{mol m}^{-2} \text{s}^{-1}$) of B light. The 5-day old seedlings were harvested after exposure to different light fluence rates in each treatment. Their hypocotyl length and chlorophyll content were quantified as in the previous assays.

4.2.6.3.1 Hypocotyl length

Hypocotyl measurement results were displayed as actual length and also as a ratio relative to the length of respective dark-grown seedlings of all genotypes, because dark-grown seedling hypocotyl length is found to differ among genotypes. I speculated this could be caused by different size of seeds, reflecting different size of embryos. The weights of seeds of each genotype were measured. As suspected, *phyA-211*, whose dark grown seedlings have a longer hypocotyl length, has larger seeds (about 0.031 mg/seed) than WT (about 0.024 mg/seed). *phyB-9*, whose dark grown seedlings have a shorter hypocotyl length, has smaller seeds (about 0.016 mg/seed) than WT. *cry1-304* has similar hypocotyl length to the WT, because it has very similar size of seeds (about 0.023 mg/seed). Although double mutants of *lyn1* combined with photoreceptor apoprotein mutations have similar size of seeds to the WT (between 0.021 mg/per seed and 0.025 mg/per seed), the double mutants still exhibit shorter hypocotyl length of dark-grown seedlings than the WT. No differences in the length of dark-grown seedling are revealed by statistical testing among the WT, *lyn1*, *phyB-9* and *lyn1 phyB-9* and among the WT, *lyn1*, *cry1-304* and *lyn1 cry1-304*. *lyn1* mutation causes a slightly reduced hypocotyl length. The dark-grown *lyn1* single mutant is shorter than WT even when it has similar size of seeds (0.022 mg/per seed) as WT.

Consistent with previous results, *lyn1* has shorter hypocotyl length than WT and *lyn1 hy1* has a shorter hypocotyl length than *hy1* under any fluence rate of white light. *lyn1* clearly suppresses *hy1* under any fluence rate of white light (Figure 4.26). The suppression is the most significant at 50 $\mu\text{mol m}^{-2} \text{s}^{-1}$ fluence rate.

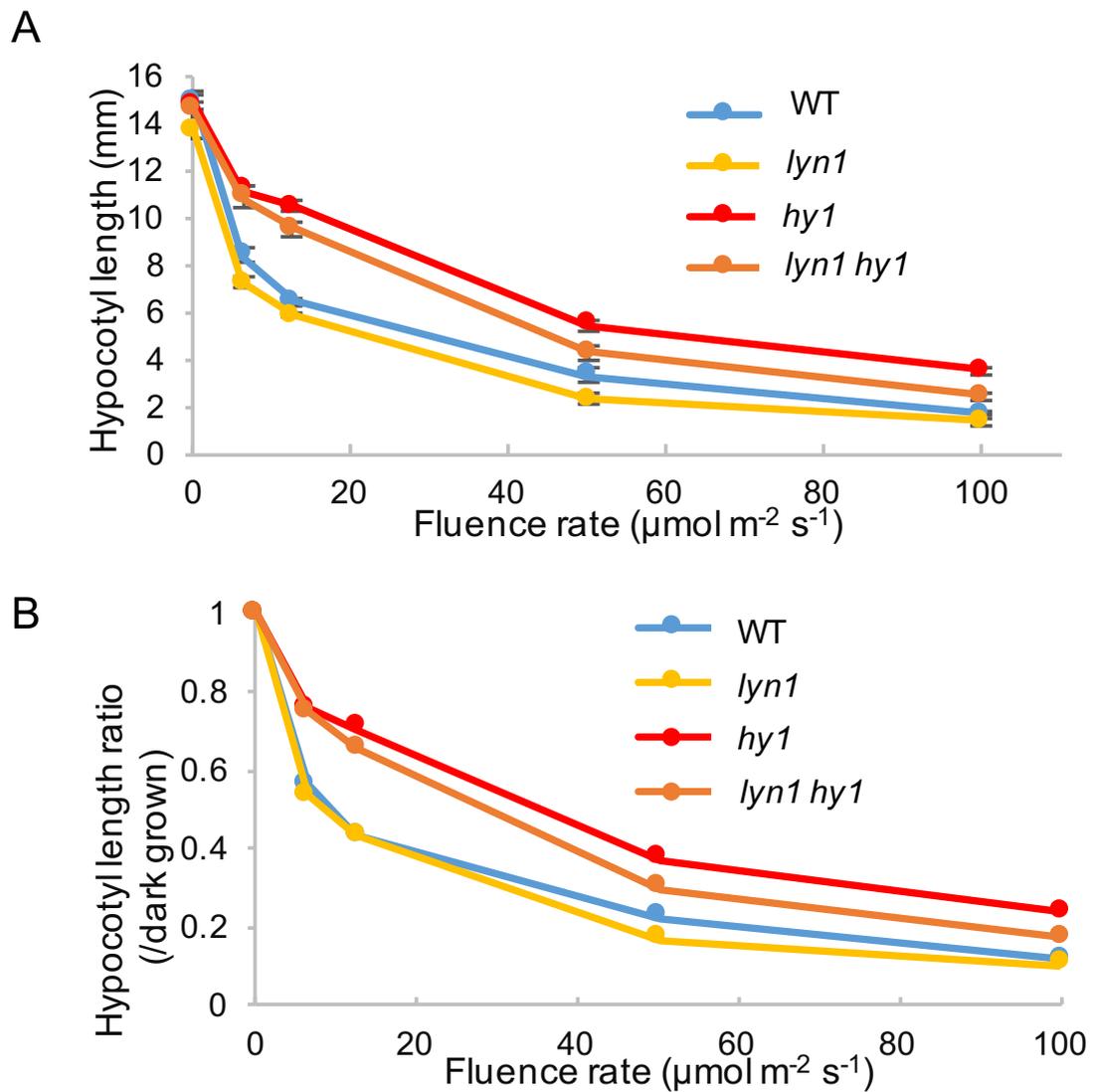


Figure 4. 26 Assessment of *lyn1* suppression of hypocotyl length of *hy1* over a fluence rate range of white light

A. WT, *lyn1*, *hy1* and *lyn1 hy1* were grown in 0, 5, 10, 50 and 100 $\mu\text{mol m}^{-2} \text{s}^{-1}$ fluence rate of white light. *lyn1* suppression of *hy1* in white light was used as an experimental control for the *lyn1* suppression of photoreceptor apoprotein mutants in monochromatic light. Y axis represents absolute hypocotyl length, in mm. Error bars represent SEM (n=20).

B. As (A) but hypocotyl length expressed relative to the value in the dark for each genotype.

In R light, *lyn1 phyB-9* does not have shorter hypocotyl length than *phyB-9* in any fluence rate. Hence *lyn1* does not show any suppression of *phyB-9* (Figure 4.27). *lyn1* single mutant has a shorter hypocotyl length than WT only at the highest fluence rate.

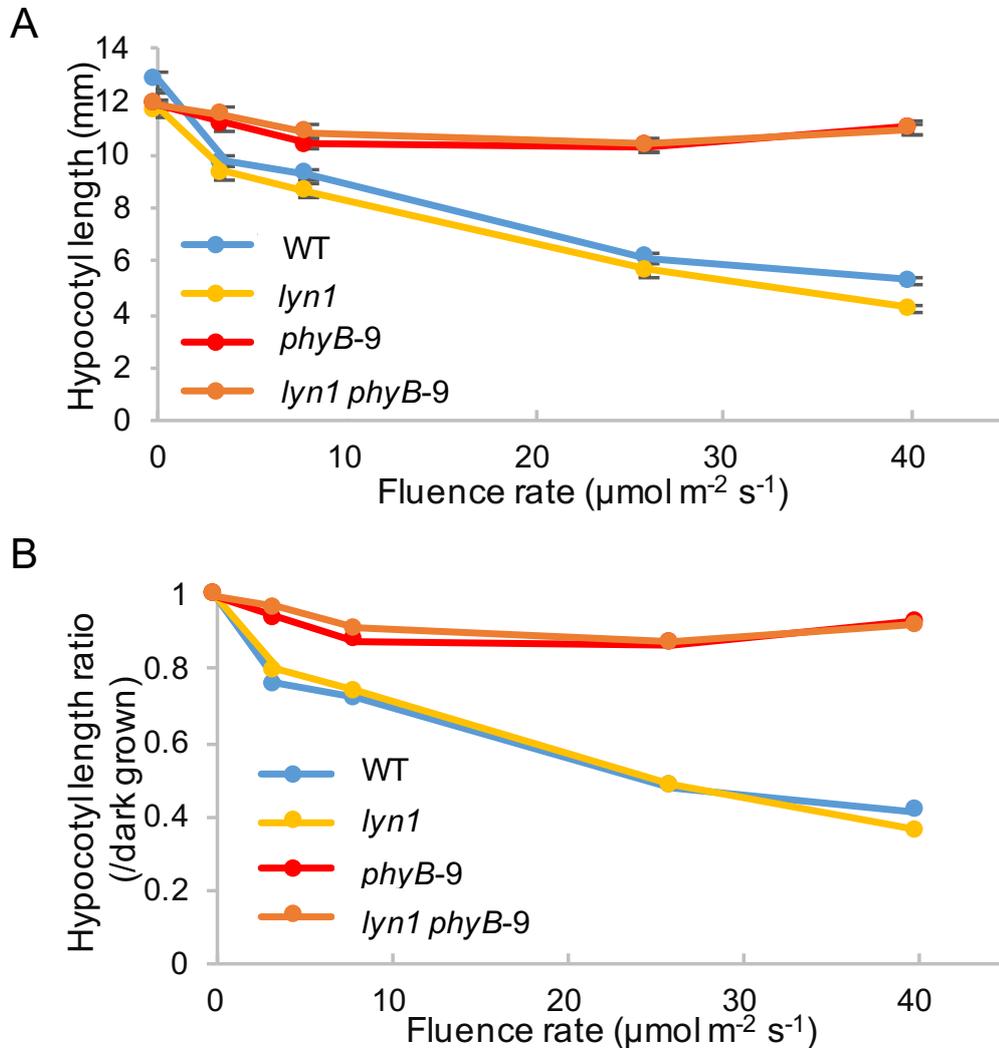


Figure 4. 27 Assessment of *lyn1* suppression of hypocotyl length of *phyB-9* over a fluence rate range of R light

A. WT, *lyn1*, *phyB-9* and *lyn1 phyB-9* were grown in 0, 2.5, 8, 26 and 40 $\mu\text{mol m}^{-2} \text{s}^{-1}$ fluence rate of R light. Y axis represents absolute hypocotyl length in mm. n=20.

B. As (A) but hypocotyl length expressed relative to the value in the dark for each genotype.

Under B light, *lyn1* does not show any suppression of *cry1-304* at any fluence rate (Figure 4.28). *lyn1* single mutant has shorter hypocotyl length than WT only at the lower fluence rate.

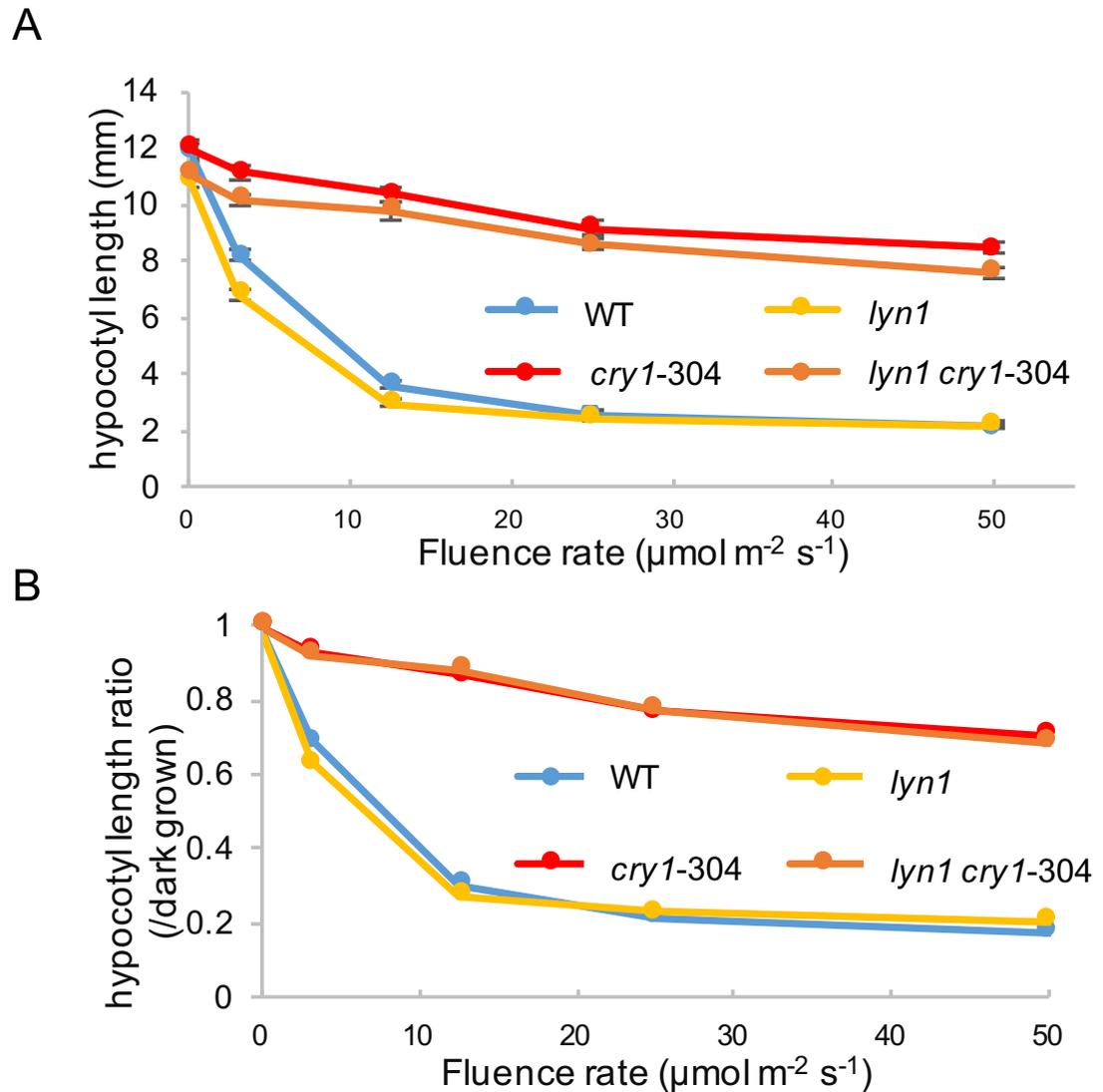


Figure 4. 28 Assessment of *lyn1* suppression of hypocotyl length of *phyB-9* over a fluence rate range of B light

A. WT, *lyn1*, *cry1-304* and *lyn1 cry1-304* were grown in 0, 2.5, 12, 25 and 50 $\mu\text{mol m}^{-2} \text{s}^{-1}$ fluence rate of R light. Y axis represents absolute hypocotyl length in mm. n=20.

B. As (A) but hypocotyl length expressed relative to the value in the dark for each genotype.

Under FR light, although the statistical test indicates that *phyA-211* has a significantly longer hypocotyl length than the other three genotypes, the ratios of four genotypes to their dark-grown seedling hypocotyl lengths indicate that *lyn1* does not suppress *phyA-211* at any fluence rate (Figure 4.29). *lyn1* single mutant has shorter hypocotyl length than WT only at the lower fluence rate.

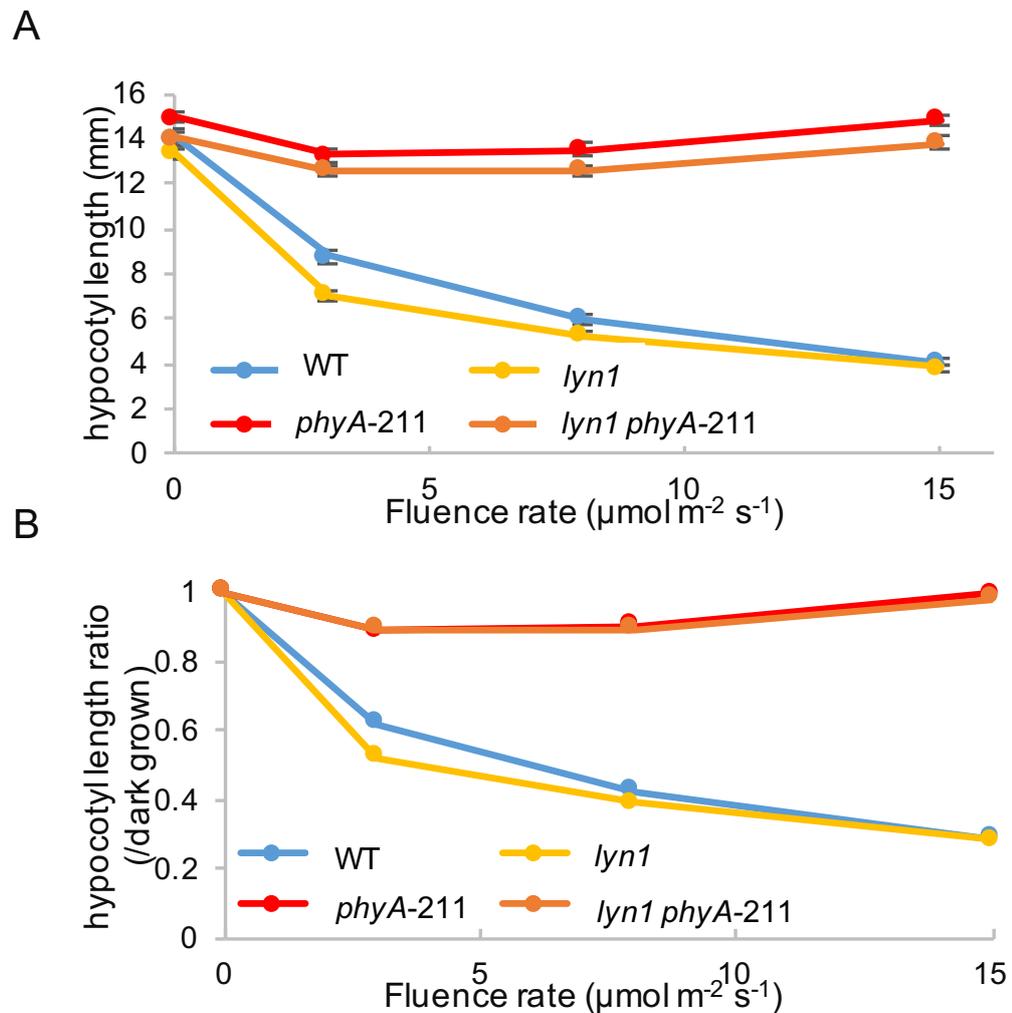


Figure 4. 29 Assessment of *lyn1* suppression of hypocotyl length of *phyA-211* over a fluence rate range of FR light

A. WT, *lyn1*, *phyA-211* and *lyn1 phyA-211* were grown in 0, 3, 8 and 15 $\mu\text{mol m}^{-2} \text{s}^{-1}$ fluence rate of R light. Y axis represents absolute hypocotyl length in mm. n=20.

B. As (A) but hypocotyl length expressed relative to the value in the dark for each genotype.

Overall the experiment asking whether *lyn1* suppresses hypocotyl elongation on *hy1* in white light, *phyB-9* in R light, *cry1-304* in B light and *phyA-211* in FR light reveals that *lyn1* does not substantially suppress photoreceptor apoprotein mutants under their diagnostic light wavelength. This result supports the hypothesis by which *lyn1* rescues the light response by increasing the PΦB biosynthesis and not by amplifying light signalling overall.

4.2.6.3.2 Chlorophyll content

In white light, *lyn1 hy1* has an increased chlorophyll level compared to *hy1* and *lyn1* also has an increased chlorophyll level compared to the WT across the entire fluence rate range (Figure 4.30 A). This is consistent with the previous results shown in the last chapter. The difference of chlorophyll level between *hy1* and *lyn1 hy1* and between *lyn1* and WT becomes increasingly significant as the fluence rate increases. Therefore, *lyn1* suppression strength increases with increased light fluence rate. In R light, *lyn1 phyB-9* has a higher chlorophyll level than *phyB-9* and *lyn1* has a higher chlorophyll level than WT across the entire fluence rate range. *lyn1* suppression of *lyn1 phyB-9* increases with increased light fluence rate as the differences of chlorophyll level are more significant. The *lyn1* single mutant also shows an increasingly significant difference of chlorophyll level, compared to WT, when the fluence rate increases. This indicates that *lyn1* can suppress *phyB-9* on chlorophyll level in R light (Figure 4.30 B). In B light, *lyn1* suppression of *cry1-304* has a similar outcome that on *phyB-9* in R light. *lyn1 cry1-304* is slightly higher than *cry1-304*. *lyn1* still exhibits a significant increase compared to WT (Figure 4.30 C).

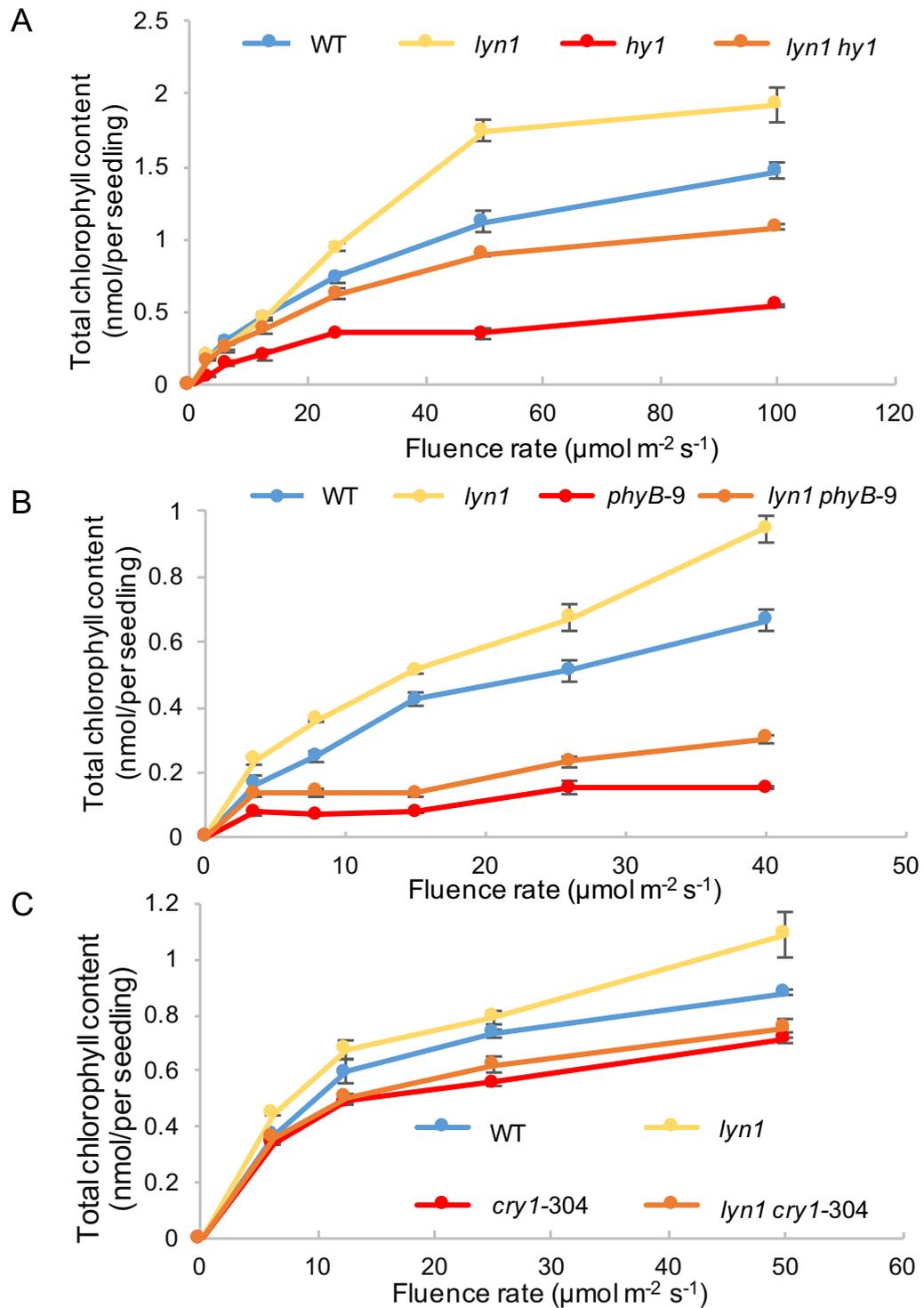


Figure 4. 30 *lyn1* suppression of chlorophyll level of *phyB-9* and *cry1-304* over a fluence rate range

A. *lyn1* suppression of *hy1* in white light, used as a control treatment, n=20;

B. *lyn1* suppression of *phyB-9* in R light, n=20;

C. *lyn1* suppression of *cry1-304* in B light, n=20.

The measurement of chlorophyll on FR-grown seedling has a complex result, because FR-grown seedlings produce both Pchl_a and chlorophyll in the FR light treatment. The proportion of Pchl_a and chlorophyll content is not always consistent. The ratio between the fluorescence peak of Pchl_a and chlorophyll is variable following different fluence rates of FR light. The chlorophyll measurement shows that all genotypes contain no chlorophyll in dark-grown seedlings and they start to accumulate chlorophyll with increased fluence rate (Figure 4.31 A). The chlorophyll levels of all 4 genotypes rapidly rise following the increased fluence rate, up to 2 $\mu\text{mol m}^{-2} \text{s}^{-1}$ FR light. Then chlorophyll levels of *phyA-211* and *lyn1 phyA-211* continue to rise steadily following the increased fluence rate, even above 2 $\mu\text{mol m}^{-2} \text{s}^{-1}$. Chlorophyll levels of WT and *lyn1* are not increased as fast as those of *phyA-211* and *lyn1 phyA-211*. The chlorophyll levels of WT and *lyn1* almost overlap as do *phyA-211* and *lyn1 phyA-211* at any fluence rate, so *lyn1* does not show a suppression of chlorophyll content. Meanwhile, Pchl_a levels of all four genotypes are also enhanced following the increased FR fluence rate until 2 $\mu\text{mol m}^{-2} \text{s}^{-1}$, then gradually decreased (Figure 4.31 B). The Pchl_a level of *lyn1* and WT are both enhanced much faster than those of *phyA-211* and *lyn1 phyA-211* in low fluence rate of FR light (< 2 $\mu\text{mol m}^{-2} \text{s}^{-1}$). Pchl_a measurement shows that *lyn1* has a degree of unexpected interaction with *phyA-211*, because *lyn1 phyA-211* had lower, rather than higher, Pchl_a level than *phyA-211* at any fluence rate, while *lyn1* single mutant seemed to have an enhanced light response in the FR treatment, because *lyn1* always has a higher Pchl_a level than WT. Because the amounts of conversion from Pchl_a to chlorophyll are different at different fluence rates, the conversion ratios between Pchl_a and chlorophyll were also studied (Figure 4.31 C). The relative ratio between Pchl_a and chlorophyll indicated that WT and *lyn1* have higher proportion of Pchl_a in the low fluence rate than *phyA-211* and *lyn1 phyA-211*. As the fluence rate increased, WT and *lyn1* convert Pchl_a to chlorophyll much

faster than *phyA-211* and *lyn1 phyA-211* do. *lyn1* does not show any effect of the conversion between Pchl_a and chlorophyll under FR light treatment.

Overall, for the three different light treatments, *lyn1* always shows some suppression of chlorophyll level on the *cry1-304* and *phyB-9* single photoreceptor apoprotein mutants, under their diagnostic light wavelength. The suppression is significant on *phyB-9* and very small on *cry1-304*. It does not show any suppression of *phyA-211*. This result appears to contradict my hypothesis, because I did not expect that *lyn1* suppresses the chlorophyll level of photoreceptor apoprotein mutants. The implications will be addressed in the discussion later.

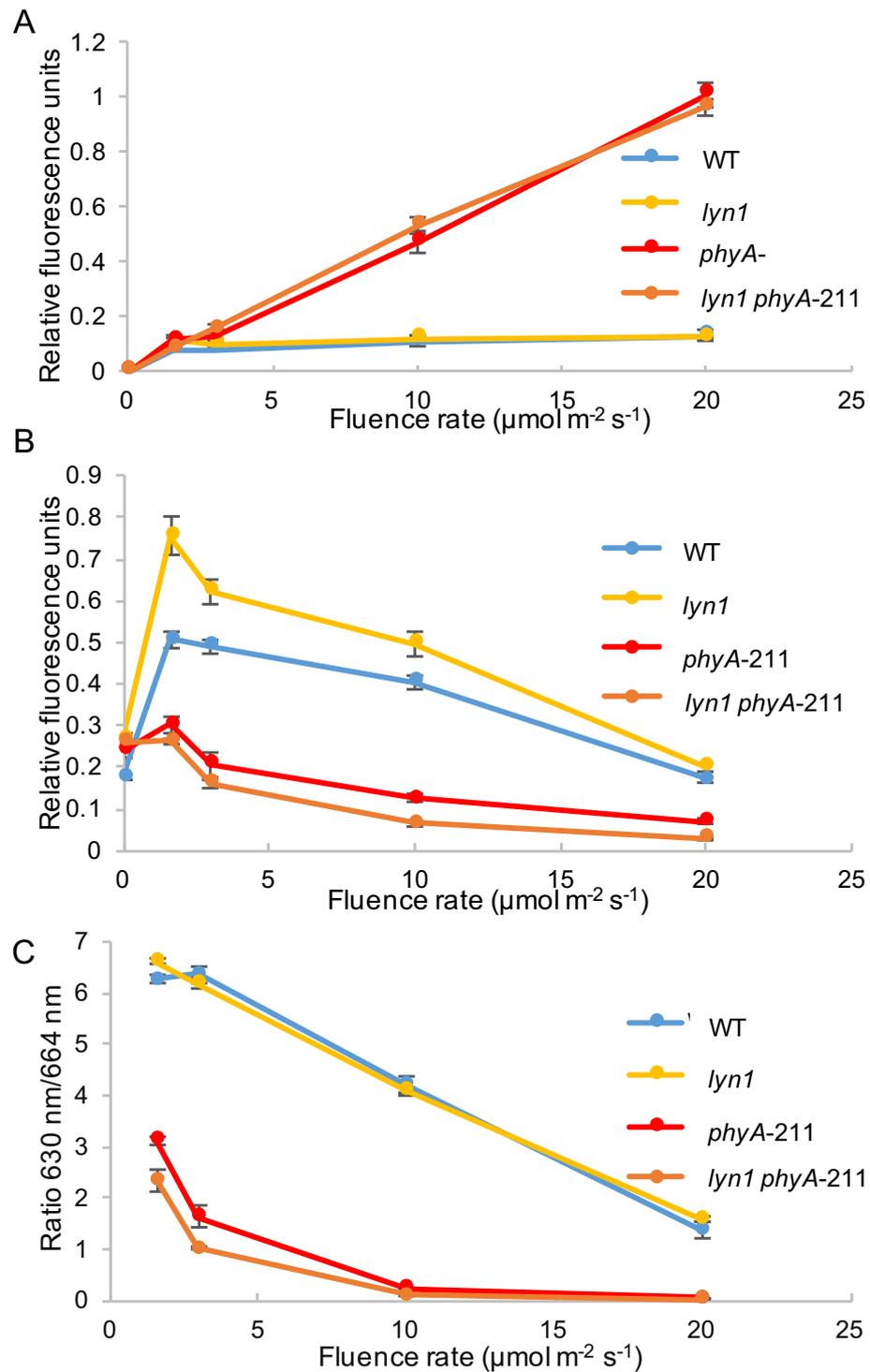


Figure 4. 31 *lyn1* suppression of chlorophyll level of *phyA-211*, over a range of FR fluence rate light

A. Chlorophyll level change at 664 nm

B. Pchl level change at 630 nm

C. Ratio between Pchl level (630 nm) and chlorophyll level (664 nm).

4.2.7 Overexpression of *JMJ14*

JMJ14 was overexpressed in the WT in order to further understand *JMJ14* function. Due to time constraints, this analysis is preliminary, as only T1 plants were examined. In all the overexpressed plants the presence of the insertion was confirmed using pB2GW7 sequencing primers (Figure 4.32). Nearly one-third of the T1 overexpressed plants had larger rosette leaves than WT (Figure 4.33 A). The rest of the plants had similar size of leaves and longer shoots, possibly reflecting slightly earlier flowering, compared to WT (Figure 4.33 B). None of the T1 plants had a *hy1* phenotype.

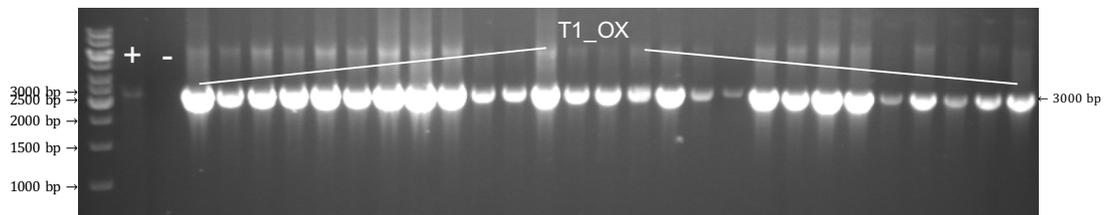


Figure 4. 32 Confirming the presence of *JMJ14* transgene in the T1 *JMJ14* overexpressed plants

PCR amplicons amplified using pB2GW7 sequencing primers are about 3000 bp long. All the T1 plants contain the transgene. The negative control does not contain the transgene.

A



B



Figure 4. 33 WT plant overexpressed with *JMJ14*

Comparison of phenotype between WT, T1 *JMJ14* overexpressed WT and *hy1*.

The plants were grown in the ($180 \mu\text{mol m}^{-2} \text{s}^{-1}$) constant light for about four and half weeks.

A. A T1_OX plant shows larger leaves than WT.

B. A T1_OX plant showing longer shoots (earlier flowering) than WT.

4.3 Discussion

4.3.1 *jmj14* enhances the plastid compartment and plastome to nuclear genome ratio

My hypothesis presumes that *jmj14* rescues the light response by increasing chloroplast development. This increase was supported by both the chloroplast compartment quantitation and the plastome to nuclear genome ratio assays. *jmj14* enhanced chloroplast compartment by increasing the size of individual chloroplasts rather than their amount. I also examined whether *jmj14* enhanced the relative average size of chloroplast to cell plan area in both types of cell of cauline leaves. However, *jmj14* only showed an increased chloroplast compartment in both cell types in the double mutants compared to *hy1*, but it did not in the single mutants compared to WT in mesophyll cells. This was supported by the result of the plastome to nuclear genome ratio assays. *jmj14* always showed an increased plastome to nuclear genome ratio on the double mutants, *lyn1 hy1* and *jmj14-1 hy1* (relative to their *hy1* control) in both light conditions, and in single mutants, *lyn1* and *jmj14-1*, in dark conditions, with one exception, for *lyn1* and *jmj14-1* the plastome to nuclear genome ratio does not differ significantly with that of WT in 5 day-light grown seedlings. According to the observation of chloroplast compartment quantitation assay, this may be because chloroplasts have fully developed in mesophyll cells of adult plant leaves or the majority of cotyledon cells of light-grown seedlings. The WT cells were fully occupied by chloroplasts, reaching a ceiling, therefore the mesophyll cells of *lyn1* and *jmj14-1* had no cellular space available for chloroplasts to further occupy. Although bundle sheath cells are also present in the seedlings, the plastome to nuclear genome ratios from the presumably much larger amount of mesophyll cells can overwhelm the smaller amounts of bundle sheath cells. Therefore, overall ratios would exhibit no difference between WT and the single mutants.

However, the chloroplast compartment and the plastomic to nuclear genome copy number ratio measurements in 5 day-light grown seedlings do both show that *jmj14* has an impact on chloroplast size.

Previously published data indicate that, as far as observed, there is a positive correlation between the plastid compartment and the plastome to nuclear genome ratio (Enfissi et al., 2010). This is in accordance with the results of simultaneously-increased chloroplast compartment and plastome to nuclear genome copy number ratio of this research. This suggests that *jmj14* does not only cause an impact on the plastid compartment size, but also other aspects of chloroplast development, such as plastome to nuclear genome ratio. Nuclear chloroplast-related transcriptome responses reveal the likelihood of a regulatory master switch of chloroplast development (Richly et al., 2003). This was identified by studying the majority of nuclear genes for chloroplast-targeted proteins, involved in different biological functions, including amino acid metabolism, carbohydrate metabolism, secondary metabolism, other metabolic functions, photosynthesis (dark and light reactions), protein modification and fate, protein phosphorylation, sensing, signalling and cellular communication, stress response, transport, transcription, protein synthesis and other unknown functions (Richly et al., 2003; Biehl et al., 2005). These genes were classified into the clusters of co-expressed genes. The expression of these genes was detected in many genetic and environmental contexts. The results showed that the expression of most genes was always upregulated or downregulated simultaneously in any conditions, i.e. most genes were coexpressed (Richly et al., 2003; Biehl et al., 2005). This strongly suggests that any genetic mutation or environmental signalling affects chloroplast development in general.

Whether the increased chloroplast development is the secondary effect of the light response, or a primary effect, remained unclear even after my observations of an enlarged chloroplast compartment. If it was a primary effect, then plastid

compartment size and the plastome to nuclear genome ratio should have also increased regardless of light. Although plastid compartment of dark-grown seedling is impossible to observe under DIC microscopy, plastome to nuclear genome ratio is still possible to quantify in 5-day dark-grown seedlings. Clearly, the light signal induced plastid development, so the plastids of dark-grown seedlings, etioplasts, have a generally lower plastome to nuclear genome ratio compared to the chloroplasts of light-grown seedlings. Even without the light, *jmj14* still causes an increased the ratio in the double mutants compared with *hy1*, as well as in the single mutants compared with WT. This is in accordance to my expectation by which plastid development does not only increase in the light but also in the dark in *lyn1* mutants. The increased plastome to nuclear genome ratio in 5-day dark-grown *jmj14* seedlings strongly suggests that *jmj14* effect of plastid compartment is a *primary* effect regardless of light. It also implies that general plastid development probably is also elevated by *jmj14* in dark.

4.3.2 *jmj14* enhances tetrapyrroles synthesis

jmj14 was found to increase the plastid genome copy number in the both dark and light, so plastid development was predicted to be elevated by *jmj14* regardless of light. If *jmj14* increases the plastid development in general, then all the metabolic compound biosynthesis should be also increased. To further support this hypothesis, the tetrapyrrole compounds, Pchlide (directly) and PΦB (indirectly), were quantified. The fact that the *lyn1* mutation increased Pchlide level in the dark had been supported by Pchlide level assays in my earlier work. FR block of greening response is an alternative way to assess the capacity to accumulate Pchlide, and its consequences, in *jmj14* mutants. Both physiological strategies to quantify Pchlide level showed the predicted result. *hy1* contains no active phytochrome and, as a result, FR treatment does not lead to the loss of PORA. There no free Pchlide accumulates, contrary to what happens in the WT, and very few of the FR-treated *hy1* seedlings are bleached compared to WT. As

jmj14 increased the Pchl_a synthesis, *jmj14* mutants increased accumulation of Pchl_a in dark-grown seedlings resulting in a higher bleaching rate on FR-treated seedlings upon transfer to white light. Because *lyn1 hy1* and *jmj14-1 hy1* present a rescue of the light responses absent in *hy1*, loss of PORA, absent in *hy1*, will occur in the double mutants, leading to a FR-block of greening. The accumulation of Pchl_a and *jmj14*-rescued loss of PORA in FR-treated *lyn1 hy1* and *jmj14-1 hy1* seedlings leads to an obvious photobleaching. Although Pchl_a quantification assay showed an increased level triggered by *jmj14* mutation during the experiments in my MSc, the increased photobleaching frequency in *lyn1* and *jmj14-1* only appeared in 1-day FR-treated seedlings. This is perhaps not surprising given that after 2 day FR-treatment, the bleaching was almost complete also in the WT. A remarkable phenomenon of a difference between the two allelic mutants, by which *jmj14-1* caused a higher bleaching rate than *lyn1* was noticed. The greening rate of *jmj14-1* was 5% lower than that of *lyn1* when seedlings had just 1 day FR treatment, and the greening rate of *jmj14-1 hy1* was nearly 15% less than *lyn1 hy1* when seedlings had 3 days FR treatment. This may be explained by the result of the RT-PCR of *JMJ14* cDNA, which failed to detect any cDNA in *jmj14-1* but observed a very small amount in *lyn1*. Any cDNA present in *lyn1* would still be from an mRNA which encoded a protein with an early stop codon. No full-length protein can possibly be produced. However, any potential truncated *JMJ14* protein present in *lyn1*, containing the conserved JmjC domain, as discussed in the previous chapter, might cause the retention of some function of *JMJ14*, resulting in a significantly increased greening ability of the *lyn1* mutants compared with the *jmj14-1* mutants.

Whether the *lyn1* increased photochemically-active phytochrome level in *lyn1 hy1* compared to *hy1* in order to cause a small amount of increased light response is a key to know. Based on photochemically-active phytochrome quantification, I realized that the *lyn1* mutation does cause an increased small but detectable

amount of photochemically-active phytochrome in *lyn1 hy1*. Additionally, an unexpected outcome is that phytochromobilin-containing photochemically active phytochrome level is even increased in the *lyn1* single mutant. Photochemically active phytochrome is synthesized in the cytoplasm by binding the PΦB to the phytochrome apoprotein. There is an evidence shown that both WT and *hy1* dark-grown seedlings contain an excess amount of phytochrome A apoprotein. This is evidenced by the fact that artificial chromophores are incorporated into holoprotein when exogenously supplied to dark-grown seedlings (Murphy and Lagarias, 1997). Because seedlings always have sufficient phytochrome A apoprotein accumulated in the cytoplasm, any plastid synthesised PΦB can immediately bind to phytochrome A apoprotein in the cytoplasm to produce phytochrome A by a single light pulse. The increased photochemically active phytochrome level in *lyn1* mutants is probably because of the increased the capacity of PΦB production along with a sufficient phytochrome A apoprotein. The small amount of photochemically active phytochrome detected reveals a small amount of PΦB being produced, and this, in turn, rescues light response in the double mutants. The increased amount of PΦB in both *lyn1* and *lyn1 hy1* can be explained by the *lyn1*-increased plastid compartment. The increased plastid development leads to enhance both plastid size and tetrapyrroles' synthesis. Both the FR light block of greening and the measurement of phytochromobilin-containing photochemically active phytochrome provide evidence to support the notion that *jmj14* enhances the synthesis of tetrapyrroles, regardless of light. This indicates that *jmj14* enhances the plastid development in general regardless of light.

4.3.3 *lyn1* partially rescues a PΦB synthase defect mutation, *hy2-105*

To confirm that *lyn1* does not rescue the light response defect by rescuing *hy1* mutation specifically, looking at whether *lyn1* also suppresses the effect of loss

of another enzyme in the PΦB biosynthesis pathway is an indirect way. However, the *lyn1* suppression of *hy2* was not obvious as that on *hy1*, because *hy2* has shown a less light response defective phenotypes. Although the less *lyn1* suppression of *hy2* is currently unclear, the less dramatic light response defect of *hy2* possibly can be explained by two reasons. Existing *hy2* mutants in Col and Ler ecotypes have found that still contain some mRNA which may afford a partially active enzyme with a lower affinity for converting biliverdin IXα to PΦB (Kohchi et al., 2001). Kohchi and collaborators indicated that the most *hy2* mutants, including *hy2-1* and *hy2-105*, retain some *hy2* mRNA. In this regard, the results of Kohchi were supported by the outcome of the monitoring of cDNA using RT-PCR on *hy2* T-DNA insertion Salk line (SALK_104923). Kohchi and collaborators also found that *hy2-105* contains relatively less *hy2* mRNA than *hy2-1*, as well as other *hy2* mutants of Col ecotype. Although both *hy2* T-DNA insertion Salk line (SALK_104923) and *hy2-105* contain some *hy2* mRNA, the much lower amount of mRNA in *hy2-105* was not sufficient to convert into cDNA template and then amplify a visible band by RT-PCR, in contrast to a clear cDNA amplicon of *hy2-104*. However, the phenotype of *hy2-105* is still not as dramatic as that of *hy1*. It is possible that *hy2* is a more physiologically leaky mutant and it contains some redundant enzymes to convert biliverdin IXα to PΦB. In *Arabidopsis*, PΦB synthase might be not the only enzyme to convert biliverdin IXα to PΦB. *hy1* and *hy2* mutants had been identified in other plant species (tomato and pea) (Terry, 1997). In both cases, the corresponding haem oxygenase mutants in these plant species showed less severe phenotypes (vanTuinen et al., 1996; Weiler et al., 1997). Kohchi suggested that the phenotypes might be determined by the relative allele strength of the two loci. Nevertheless, *lyn1* still exhibited some suppressions of *hy2-105* that was sufficient to support my hypothesis. This experiment determined that *lyn1* does not suppress the *hy1* mutation specifically. *lyn1* also can suppress other PΦB synthesis defective mutants. This suggests that *lyn1* rescues the light response

by compensating the haem biosynthetic pathway end product, PΦB. The elevated PΦB level, combined with the observation of an increased plastid genome copy number in the dark (consistent with an enlarged plastid compartment), suggests that an increased capacity of the PΦB biosynthesis pathway occurred in the mutant. Although the PΦB biosynthesis pathway is still largely blocked in *hy1* and *hy2*, a small amount of PΦB biosynthesis may therefore now occur. *jmj14* increases plastid development, and the small amount of PΦB synthesized in *hy1* is elevated to become detectable. Consequently, more phytochrome molecules are synthesised in the cytoplasm to perceive light better. Therefore, *jmj14* is predicted to be able to rescue any defect that cannot synthesise PΦB.

4.3.4 *lyn1* does not suppress photoreceptor apoprotein mutants

The interpretation of these experiments was not trivial, because of some apparent contradictions. I expected that *lyn1* would not suppress the growth and greening of photoreceptor apoprotein mutants. The assessment of hypocotyl phenotypes provided a seemingly clear-cut, expected result, by which *lyn1* did not suppress hypocotyl length phenotypes of all 3 types of photoreceptor apoprotein mutant under their diagnostic light wavelength. The assessment of pigment phenotypes was less clear. *lyn1* caused a phenotype suppression (an elevation) of chlorophyll content of *phyB-9* and *cry1-304* in R light and B light, respectively. These were unexpected results that did not meet my expectation. Nevertheless, *lyn1* did not suppress *phyA-211* in either white light or FR light conditions. However, the enhanced chlorophyll levels cannot simply be interpreted as *lyn1* suppression of *phyB* and *cry1*. Because the *lyn1* single mutants had an enhanced level of chlorophyll in R light and B light, the enhanced chlorophyll level probably is the result of increased tetrapyrroles synthetic capacity in the dark. In other words, the effect of the *lyn1* and photoreceptor mutations are simply additive. According to the Pchlde quantitation, the increased chlorophyll level in *lyn1* mutants is

probably due to the fact Pchl_a level of *lyn1* mutants is already increased in the dark. Because more Pchl_a was produced in the dark, more chlorophyll was synthesized after exposing to the light. Moreover, increased chlorophyll level is a secondary effect of the rescued light response. Concerning double mutants, the most dramatic contrast between the outcomes of hypocotyl elongation (where there was no suppression) and greening assays (where there apparently was) concerns *phyB*. *phyB* dominates in regulating hypocotyl cell elongation in response to R light (Tepperman et al., 2004). Tepperman and collaborators found that *phyA* and *phyB* work cooperatively to account for 96% of the R light induced early response genes expression (Tepperman et al., 2006). Unexpectedly, those authors found that the majority of R light responsive genes remain relatively strongly responsive to R light signal in the *phyB* null mutant. In red light, the *phyA phyB* double mutant, but neither single mutant, had reduced *LHCB* gene induction and chlorophyll accumulation (Reed et al., 1994). This finding, combined with the fact of *lyn1* increasing *phyA* holoprotein level, can contribute to explaining the enhanced chlorophyll level in *lyn1 phyB*, in addition to the enhanced plastid development capacity. *jmj14*-increased *phyA* can rescue the light response defect of *lyn1 phyB*, resulting in an enhanced chlorophyll level in the double mutant. Although *phyA* and *cry1* were the major photoreceptors involved in hypocotyl growth regulation in blue light, *phyA* acts as pre-programmed amplifier of certain *cry1*-regulated responses rather than interacting with *cry1* (Whitelam et al., 1993; Casal and Boccalandro, 1995; Neff and Chory, 1998). *cry1* is the major photoreceptor involved in chlorophyll accumulation in blue light (Neff and Chory, 1998). In contrast, *phyA* is defective in the only photoreceptor responsive to FR light. Indeed, *lyn1* did not suppress *phyA* in any light condition, which strongly supports my hypothesis.

The measurement of greening after the FR light treatment is more complex as Pchl_a and chlorophyll were both produced. The total amount of Pchl_a or

chlorophyll depended on expression of the *HEMA1* gene, whose product acts upstream in the Pchlde synthesis pathway. *HEMA1*, which encodes glutamyl-tRNA reductase, is regulated by light, including an FR-HIR (McCormac and Terry, 2002). *HEMA1* expression is not induced in a very low-fluence rate response, but low fluence rate used here is already higher than the very low-fluence rate (McCormac et al., 2001). In this case, the effect of R light on *HEMA1* expression was too small to be considered. The proportion between Pchlde and chlorophyll in the Pchlde synthesis pathway is dependent on the PORA level, which is regulated by the light signal. The way of FR-HIR degrades PORA was explained in the introduction of this chapter. The PORA catalysed reaction is a NADPH-dependent trans-reduction of the double bond in ring D of Pchlde to Childe. This reaction requires light. During catalysis, POR first binds NADPH, then Pchlde to form a ternary POR-Pchlde-NADPH complex which is able to absorb light (Reinbothe et al., 1996). FR and R, which act as brake and inducer, both control the conversion rate between Pchlde and chlorophyll, as follows: in the low fluence rate, the Pchlde production was increased because FR-HIR induced *HEMA1* expression in order to synthesise more Pchlde. Meanwhile, the majority of PORA was degraded by phytochrome action, activated by FR, while the fluence rate of R (present in the FR) was low and could only induce a small amount of conversion of Pchlde to chlorophyll. Therefore, Pchlde increased, as it kept accumulating in the low fluence rate. In the higher fluence rate, more Pchlde was synthesised given that *HEMA1* expression is mediated by a FR-HIR (McCormac and Terry, 2002) and that HIRs increase under high irradiance. Meanwhile, PORA began to be degraded in FR but only partially and R started to induce the conversion from Pchlde to chlorophyll. Although increased FR caused more degradation of PORA, the simultaneous increased R also induced more conversion. The overall conversion rate is still increased. More and more Pchlde converted into chlorophyll resulting in a decreased Pchlde level. *phyA* mutant is blind to FR, so it cannot detect FR. *HEMA1* expression and PORA

degradation were no longer controlled by FR. As a result, Pchl_{ide} level in *phyA* mutant could not suddenly increase. All synthesised Pchl_{ide} could rapidly and completely convert into chlorophyll, so Pchl_{ide} accumulation was not apparent in *phyA*.

It is worth noticing that the proportion of R to FR light is a fixed property of the light utilised as FR source, and therefore remains constant in any fluence rate. Given that I also observed the ratio of Pchl_{ide} to chlorophyll remained identical at all fluence rates of FR between *lyn1* and the WT, I can conclude that *lyn1* does not simply elevate the tetrapyrrole synthesis capacity prior to Pchl_{ide} without elevating the conversion to chlorophyll. Both were elevated to the same extent by the *lyn1* mutation. I predict that *lyn1* enhances the plastid metabolic capacities overall.

4.3.5 Future work

So far, the absence of *JMJ14* has shown various impacts on plastid development. It is interesting to know what impacts can be caused if *JMJ14* is overexpressed, especially on plastid development. This can help to discover more about *JMJ14* protein function, and its cell biological impact. The initial results in the current study have shown that the overexpressed T1 plants have a range of different leaf sizes. The different phenotypes of overexpressed T1 plants could be caused by the different expression level of inserted *JMJ14* transgene. Alternatively, these could be effects of the transformation event, and not be transmissible. The results are therefore very preliminary. These T1 plants should be propagated to T2 and, to obtain the homozygous *JMJ14* overexpressed plants, the T3 generation. All T3 progeny of a homozygous *JMJ14*-overexpressed T2 plant should survive on the antibiotic selective plates and not show segregation. The expression level of *JMJ14* gene can be detected from these T3 overexpressed plants by qPCR. A

study on plastid development of *JMJ14* overexpressed plants can then be carried out.

4.4 Conclusion

In summary, *jmj14* enhances the plastid development regardless of light, because the chloroplast compartment analysis and plastid to nuclear genome copy number ratio assays suggest that *lyn1* enhances plastid development not only in the light but also in the dark. *jmj14* was confirmed to not only suppress *hy1*, because it can also rescue another PΦB synthesis defective mutation, *hy2*. The enhanced plastid development leads an increased Pchl_a synthesis and the appearance of a detectable level of the previously missing PΦB. *lyn1*-enhanced PΦB level causes the appearance of a small but detectable amount of photochemically active phytochrome in *hy1* (as well as an increased amount in the wild type), in turn rescuing the light response defect in *hy1*. I suspected that the increased tetrapyrrole levels resulting from the enhanced plastid development of *jmj14* were caused by boosting the plastid metabolic capacities, resulting in a slightly rescued light response. The rescued light response further increases plastid development resulting in a “virtuous cycle”. This eventually results in clearly-increased light response.

Chapter 5 Result – Global Gene Expression Analysis

5.1 Introduction

5.1.1 Global gene expression analysis

JMJ14 is identified as a histone lysine demethylase which can demethylate H3K4me2 and H3K4me3 by many studies (Jeong et al., 2009; Lu et al., 2010; Yang et al., 2010). Jeong and collaborators (2009) determined that JMJ14 directly repress FT gene expression by H3K4me3 and H3K4me2 demethylation. According to the several studies of epigenetic marks among the whole genome, H3K4me2 and H3K4me3 methylation have been found to associate with gene expression (Zhang et al., 2009; Roudier et al., 2011). The data of Roudier identified that, among the whole genome, 23244 genes are H3K4me2 marked, 17861 genes are H3K4me3-marked and 17666 genes are marked by both H3K4me2 and H3K4me3. Because JMJ14 can regulate the expression of H3K4me2 and H3K4me3-marked gene by histone demethylation, which genes are JMJ14-regulated by removing the di- and trimethyl groups from H3K4 to mediate light response is a key to assess in this study. Although the majority of genes are H3K4me2/3 marked, the expression of these genes is not all correlated with H3K4 marking level. Only H3K4me3 shows a positive correlation with transcript abundance but H3K4me2 does not (Zhang et al., 2009; Roudier et al., 2011). Therefore, I am mainly forced to study H3K4me3-marked genes in global gene expression analysis. My hypothesis is that the JMJ14 H3K4me3-suppressed genes remain in the active state in *jmj14*, resulting in an enhanced light response.

To find *jmj14*-regulated chromatin modified genes or biological process is key to understanding the way of *jmj14* rescues the light response of *hy1*. During my

previous MSc research, several genes were selected to study the *lyn1* regulated gene expression. According to observations showing *lyn1* enhanced plastid development, *lyn1* was suspected to be a regulator of a nuclear-encoded, chloroplast-related gene TF. Therefore, genes such as *HY5*, *GLK1*, *GLK2*, *APRR2*, *GNC*, *CGA1* and *CRF2* (positive) and *PIF1*, *PIF3* (negative regulators) were selected and their expression in the WT, *lyn1*, *hy1* and *lyn1 hy1* were analysed by qPCR. Meanwhile, the expression of some chloroplast-encoded genes such as *rpoC1*, *psaA*, *ndhA* and *rbcL* were also analysed. The expression of many chloroplast-encoded genes was up-regulated in *lyn1 hy1* and *lyn1* compared to *hy1* and WT, respectively. This supports the observation that *lyn1* enhances chloroplast development. However, none of the tested nuclear-encoded chloroplast gene TFs showed a consistently increased expression in *lyn1 hy1* and *lyn1* compared to *hy1* and WT, respectively. Only *GLK1* showed just a slightly elevated expression (He, 2013). It is difficult to determine the significantly elevated expression of overexpressed nuclear-encoded genes which are induced by *lyn1* through qPCR. Instead of doing analysis on the individual genes, the gene expression analysis was carried out on the global genes by RNA sequencing (RNA-Seq). RNA-Seq is a more effective and efficient way to determine *lyn1*-regulated global gene expression changes.

RNA samples of four genotypes (WT, *lyn1*, *hy1* and *lyn1 hy1*) were extracted from 5-day light- and dark-grown seedlings. The RNA-Seq was performed by the laboratory of Prof. G. Beemster's, University of Antwerp. The expression of a total of 33602 genes from TAIR10 Gene Annotation Data (https://www.arabidopsis.org/portals/genAnnotation/gene_structural_annotation/annotation_data.jsp) were tested by revealing the presence and quantity of RNA in the samples. RNA-Seq data was analysed by myself to select a group of significantly differentially expressed genes through a series of processes. The whole strategy is illustrated in Figure 5.1. The RNA-Seq raw data was normalized

by CLC Genomics Workbench first, then the differentially expressed genes were selected by setting up several filters. If the expression values of all three biological replicates of each genotype was above 4 reads, this genotype was indicated as “expressed”. The normalized gene expression values of sample were compared between different samples through FDR. FDR-based pairwise multiple comparisons are illustrated in Figure 2.2. The gene expression values of contrasts between the different genotypes under the same light condition and between different light conditions of the same genotype were indicated as significantly differentially expressed if the FDR p-value of any contrast was lower than 0.05. Finally, a filter for a minimum two-fold expression change was applied. This was carried out by calculating the average of the \log_2 value of gene expression of each sample. The \log_2 expression difference was calculated using the highest expression value minus the lowest expression value of one gene among the 4 genotypes within the same light condition. If the \log_2 expression difference within any one light condition was greater than 1, the gene was indicated as significantly differentially expressed gene. The significantly differentially expressed genes were clustered into several bins according to their gene expression pattern. Their biological functions were also assessed by gene ontology (GO) term overrepresentation analysis.

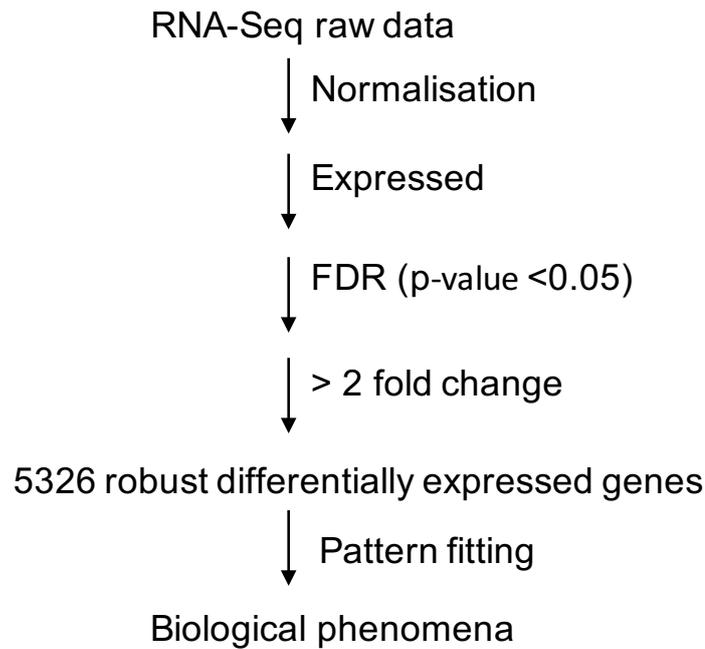


Figure 5. 1 Strategy of RNA-Seq analysis

For details of the methods and the pairwise sample comparisons see section 2.10 and Figure 2.2

A previous researcher from my laboratory has summarized many categories of functional genes (metabolism, photosynthesis, cell wall, cell cycle, development, etc.) which were associated with processes in plants obtained from MAPMAN and manually supplemented with additional genes (López-Juez et al., 2008). These groups of functional genes were used to analyse the overrepresentation or underrepresentation in the gene clusters of interest on any biological process. The statistical tool to assess over/underrepresentation was the hypergeometric test, which assesses the significance of a contingency table (observed/expected). In particular, in this research, a hypergeometric test was performed to test the significance of the probability of finding the observed (or greater) number of genes for a particular function or property in the sample (the particular cluster of differentially expressed genes) relative to the frequency with which such function or property occurs in the entire genome.

5.1.2 Assessment of global developmental roles through comparative analysis with published transcriptome datasets.

The main objective of this study is to identify which biological pathways are affected by the loss of *LYN1*. RNA-Seq analysis was carried out in order to determine the impact of the *lyn1* mutation (see later in this chapter sections 5.2.2 and 5.2.3). However, *lyn1* may not affect a specific biological pathway. This may occur if, for example, *jmj14* upregulates or modulates a very large number of genes, involved in multiple biological processes, at early developmental stages. Because *jmj* genes have been identified across multiple species and the evolutionary relationship analysis suggests that the *Arabidopsis* *JMJ14* homologues are nearly universally distributed in eukaryotes, the JMJ protein functions identified in other species may also be conserved in the *Arabidopsis* JMJ proteins. Many previous studies have found that JMJ14 homologues are tends to be involved in cell differentiation of several types of stem cell in wide range of mammal tissues (Takeuchi et al., 2006; Yamane et al., 2007; Dey et al., 2008; Schmitz et al., 2011; Cellot et al., 2013; Johansson et al., 2014; Rasmussen and Staller, 2014). If the JMJ14 protein function, which is involved in cell differentiation, is conserved globally, then overrepresentation may not appear on any particular biological function in the RNA-Seq analysis. To understand whether *lyn1* is involved in regulating gene expression at early developmental stages, I associated the existing observations and data from the *Arabidopsis* Gene Atlas (Schmid et al., 2005), which recorded global gene expression values in multiple plant tissues, at different developmental stages, with my own data from *lyn1* RNA-Seq analysis.

JMJ14 is a protein with clear homology to H3K4me3 demethylases of other organisms, so the *jmj14* up-regulated genes are expected to be H3K4me3-marked genes (or indirectly-dependent on such genes). A global analysis involving chromatin immunoprecipitation (ChIP) published by Roudier et al. (2011)

reported the H3K4me3 mark level of the entire genome in *Arabidopsis* whole seedlings and, separately, dissected roots. Determining the H3K4me3 marking state of differentially expressed genes identified in my RNA-Seq data using the data from Roudier and collaborators can confirm or rule out whether the *jmj14* up-regulated genes tend to be H3K4me3-marked genes. If the hypothesis that *jmj14*-regulated genes are not only involved in cell differentiation in early developmental stages, but are also H3K4me3-marked genes is confirmed, then the mechanism through which *jmj14* rescues the light response defect of *hy1* would be convincingly explained.

5.1.2.1 Gene expression analysis at early stages of differentiation

Schmid and collaborators developed a gene expression map of *Arabidopsis thaliana* development across different tissues and organs of the plant for a total 23257 genes (Schmid et al., 2005). This analysis was carried out by extracting RNA samples from different organs of 7-day old seedlings and 17-day old young plants of *Arabidopsis*, then using Affymetrix GeneChips which consist of hundreds of probes (a sequence of 25 bases oligonucleotide strands) that can hybridize the target cDNA sequence exactly. The gene expression estimates were normalised by gcRMA (GeneChip Robust Multiarray Averaging). The relevant samples for my study, which were extracted from different compartments of seedlings, are shown in Figure 5.2 and the compartment details are shown Table 5.1. I selected ratios between internal sets of data which represented the gene expression differences of Early vs. Later Development (6/5, 6/7), Ongoing vs. Late Development (17/13), Early vs. Completed Development (6/1), Ongoing vs. Completed Development (17/12) and Photosynthetic vs. Non-Photosynthetic organs (7/3). The genes whose expression are increased (above an arbitrary cutoff) from these lists of ratios can be used as a functional gene class which represents the early development genetic programme in plant growth, to assess whether these genes are overrepresented in the gene clusters of interest.



Figure 5. 2 Sample images to show the RNA samples collected from *Arabidopsis* at different developmental stages

Seedlings were grown under conditions equivalent to those reported by Smith et al. (2005). Samples had been collected from 7-day seedlings (left) and 17-day young plants (right). The plants were dissected and the tissue samples of different plant compartments were labelled with numbers. The sample details are shown in Table 5.1

Sample No.	Tissue	Age
17	rosette leaf 12 (newest leaf, early development)	17 days
13	rosette leaf 4 (late development)	17 days
12	leaf 2 (oldest leaf, complete development)	17 days
7	seedling, green parts	7 days
6	shoot apex, vegetative (early development)	7 days
5	leaves 1+2 (late development)	7 days
1	cotyledons	7 days
3	root (none photosynthesis tissue)	7 days

Table 5. 1 Table to show the details of RNA samples for the Affymetrix microarray carried out as the *Arabidopsis* Gene Atlas

RNA samples were extracted from different organs of *Arabidopsis* at different developmental stages.

5.1.2.2 Analysis of H3K4me3-marked level of the whole genome

Roudier and collaborators generated epigenomic maps of multiple chromatin modifications from 7-day whole seedlings and roots of *Arabidopsis* (Roudier et al., 2011). This analysis was carried out by ChIP-Chip assay. Importantly, one of the chromatin modifications which Roudier reported on was H3K4me3. The array data and genome annotation of this study is displayed using a Generic Genome Browser at Epigara (Roudier et al., 2011). By cross-comparing Roudier's data with *lyn1* differentially expressed genes data, the *lyn1* up-regulated genes can be assessed as to whether they tend to be H3K4me3-marked genes or not, based on the list of marked genes (statistically above background) reported in that study.

5.2 Results

5.2.1 Selection of differentially expressed genes and variability of RNA-Seq samples

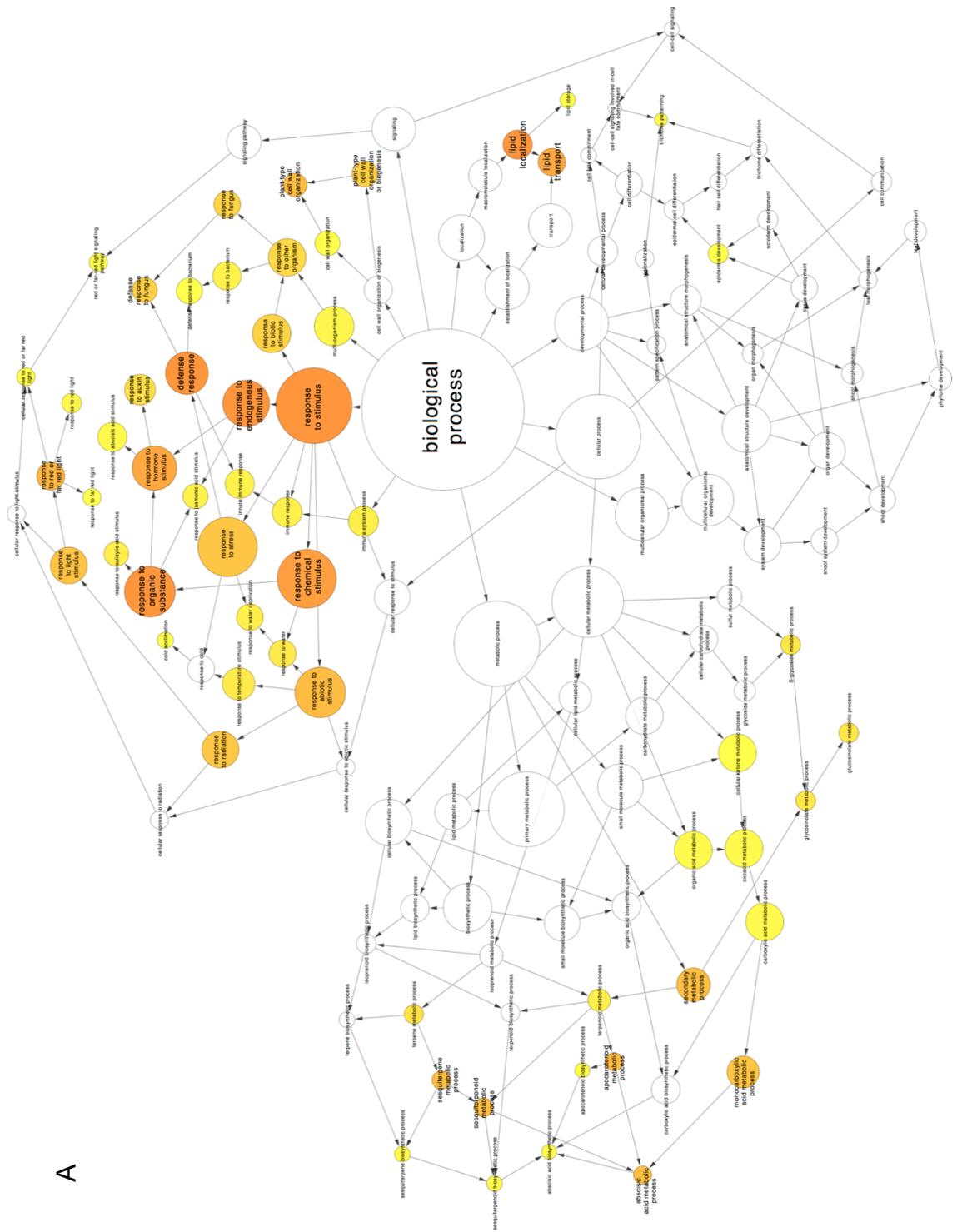
By following the RNA-Seq data analysis strategy, A total of 21613 genes were identified as expressed. The expressed genes were used as "population" in the hypergeometric test for analysing the overrepresentation of biological function later. 19692 genes were selected as significantly differentially expressed genes through FDR estimation. A total of 5326 genes were finally selected as differentially expressed genes across genotypes (WT, *lyn1*, *hy1* and *lyn1 hy1*) and not just between dark and light.

The samples sent for RNA-Seq analysis included 3 biological replicates of each genotype and each light condition. This should give a robust RNA-Seq data analysis result. However, *lyn1*_dark, *hy1*_dark and *lyn1*_light samples did not show a relatively similar gene expression level among all their biological replicates. In both cases one biological replicate for a genotype and condition gave very different expression results to the others of the same material, in the case of both *lyn1*_dark and *hy1*_dark, even after I manually rejected the failed biological replicate of both *lyn1*_dark and *hy1*_dark. The two biological replicates showed a substantial expression difference for some genes. Nevertheless, all the samples were subjected to PCA, which was performed by a bioinformatics researcher at the University of Antwerp, and showed that there is sufficient similarity between samples from the same genotype, even though this is not the case for all genes. Additionally, results of preliminary qPCR analyses of the samples show very little variation for each genotype and condition (results not shown). Therefore, the large variability of RNA-Seq results between replicates of *lyn1*_dark and *hy1*_dark samples could have been caused by RNA degradation or occur during the sequencing process. Although the large variations of

*lyn1*_dark and *hy1*_dark samples did cause difficulties during the analysis, as they reduced the significance of the statistical test, they were both used in the later RNA-Seq data analysis. In other words, although one can fully trust the differentially expressed genes identified (there are no false positives), there may have been other unnoticed changes (there will be false negatives). The gene expression of biological replicate of *lyn1*_dark and *hy1*_dark samples were checked individually in the RNA-Seq data after normalization. One biological replicate of the *lyn1*_light sample was suspected of some mix of dark and light sample, which must have occurred during RNA-Seq processing given highly consistent results of preliminary qPCR analyses in my laboratory of the samples submitted (results not shown). We were unable to identify how this occurred. Fortunately, the other two biological replicates of *lyn1*_light samples showed the robustly consistent gene expression, so the odd biological replicate of the *lyn1* light sample was rejected directly.

5.2.2 GO term analysis of total 5326 significantly differentially expressed genes

As a first step, the gene ontology (GO) term of total 5326 significantly differentially expressed genes were analysed by Cytoscape in order to understand the cellular component, molecular function and biological process of these genes (Shannon et al., 2003; Maere et al., 2005). In the biological process term, the total 5326 significantly differentially expressed genes have significant overrepresentation on response to stimulus, lipid localization and some metabolic processes (Figure 5.3 A). In the cellular component term, they have overrepresentation on the endomembrane system and light-harvesting complex (Figure 5.3 B). In the molecular function term, they have significant overrepresentation on several binding functions (tetrapyrrole binding, haem binding, iron ion binding and oxygen binding), structural constituent of cell wall and some catalytic activity (monooxygenase activity and oxidoreductase activity) (Figure 5.3 C).



A

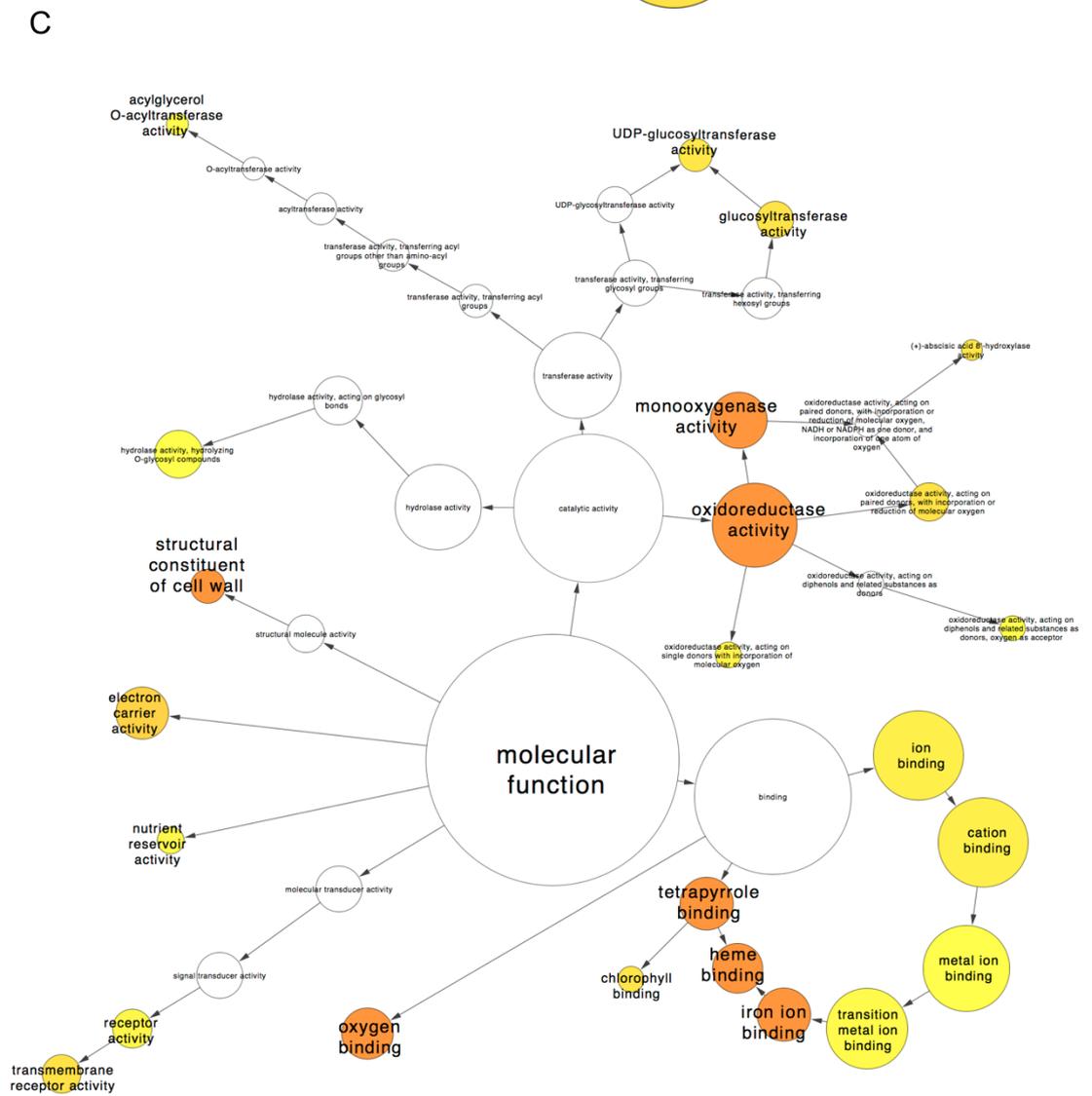
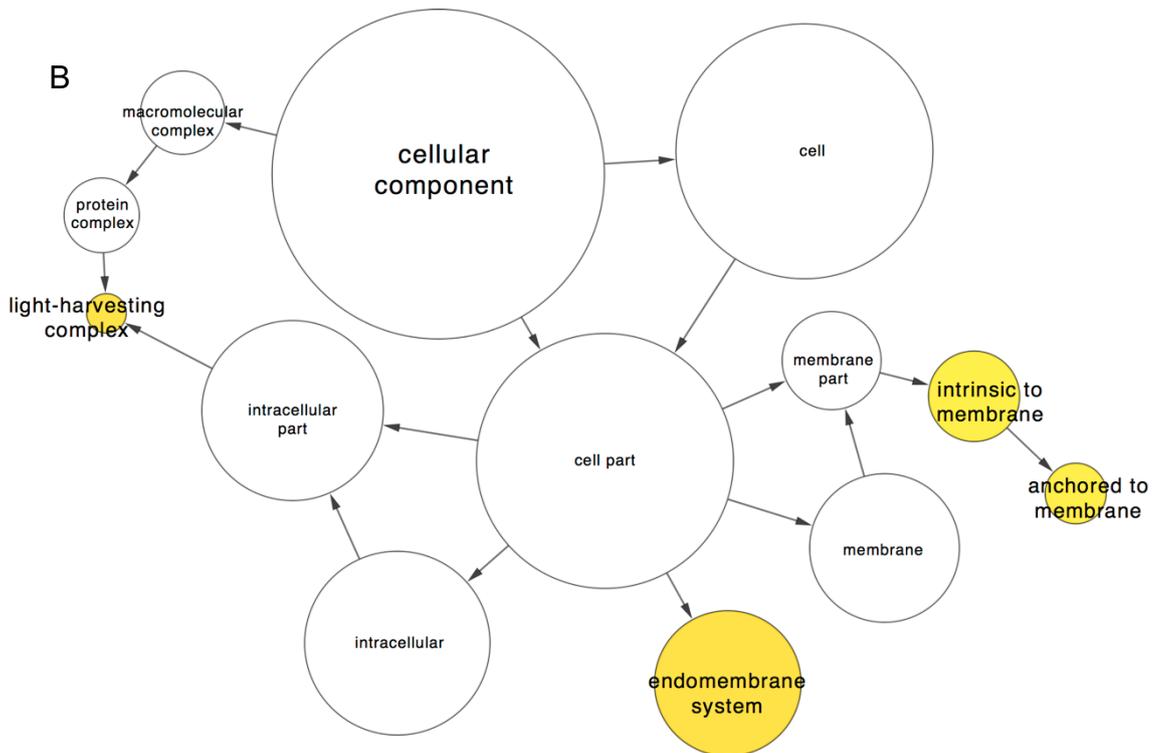


Figure 5. 3 GO term of total 5326 significantly differentially expressed genes

The images were produced by Cytoscape. GO categories were shown in biological process (A), cellular component (B) and molecular function (C) (figure best viewed on the electronic version). Each GO category is represented as a node. The map was produced according to the p-values and number of genes in the cluster belonging to a certain GO category. The node size represents the number of genes in this GO category. Yellow nodes represent GO categories that are overrepresented at the significance level. The node colour represents the significance of the p-value, darker orange colour representing lower p-value and greater significance.

5.2.3 Functional gene classification analysis

The total significantly differentially expressed genes were manually analysed to see if any of my functional gene categories representing an individual process were overrepresented, as measured by the hypergeometric test. The result shows that many functions were overrepresented in the total list of significantly differentially expressed genes (Figure 5.4). The significantly differentially expressed genes include a disproportionate number of typical photosynthetic, hormone and defence and cell wall-related genes. This primary functional gene classification analysis cannot give a clue as to which biological processes or group of functional genes are *lyn1*-regulated. These significantly different expression levels may be the effect of *lyn1* regulation or be caused by light. To identify the primary *lyn1*-regulated genes, these differentially expressed genes were clustered into different groups according to their expression pattern. A group of *lyn1* up-regulated genes was expected to appear by clustering, then be used for further functional gene analysis.

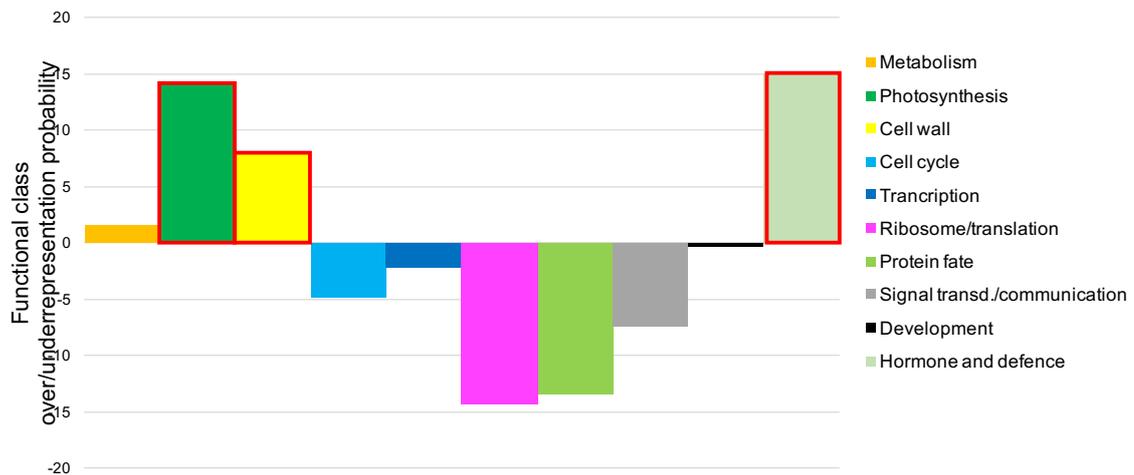


Figure 5. 4 Functional gene classification analysis of total 5326 significantly differentially expressed genes

Genes representing different gene ontology classes (functional categories) were selected and the frequency of their presence among the global differentially expressed genes (or clusters thereof) was quantified. Functional categories are represented in different colours. y-axis represents the probability of over/underrepresentation of each selected functional category among all the significantly, differentially expressed genes. Photosynthesis, cell wall and hormone and defence functional categories are overrepresented (hypergeometric test p value > 1.3), a fact highlighted by a red outline.

The total 5326 significantly differentially expressed genes were subsequently clustered into different numbers of bins (5, 10, 15 and 20) according to their gene expression patterns. The clearest group of *lyn1*-dependent genes was obtained by clustering into 5 bins (see Appendix Figure 7.8). Clustering into a different number of bins did not yield as clear *lyn1*-dependent genes (results not shown). According to the gene expression pattern of each cluster, genes of cluster 5 showed a perfectly *lyn1* up-regulated and light-independent behaviour (Figure 5.5 A). This cluster then underwent the functional gene classification analysis. However, the hypergeometric test showed that only the cell cycle and protein fate-related classes of genes were overrepresented (Figure 5.5 B). Additionally,

neither a single gene from the functional gene list of chloroplast developmental (global) genes nor chloroplast translation-related genes was present in cluster 5. The rest of functional gene classes were all underrepresented. The overrepresented functions of cluster 5 do not directly explain the *lyn1*-induced biological phenomena.

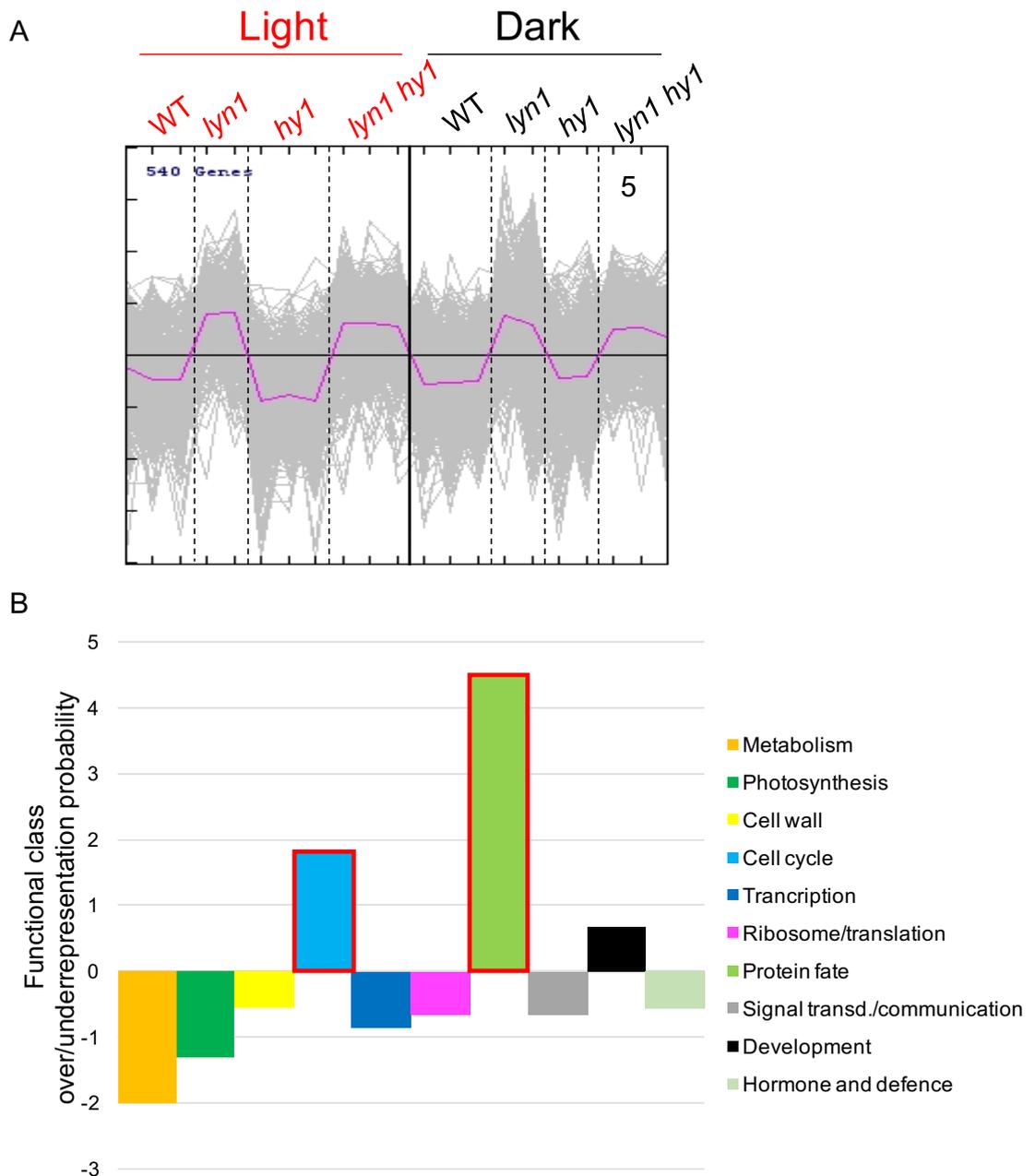


Figure 5. 5 Cluster 5 of differentially expressed genes reveals enrichment of biological process categories

A. Cluster 5, which is *lyn1* up-regulated and light-independent, was obtained by K-means clustering of differentially expressed genes using MeV. x-axis represents genotypes (WT, *lyn1*, *hy1* and *lyn1 hy1*) and light conditions (light and dark). Light and Dark samples are separated by solid lines. Four genotypes are separated by dotted lines. y-axis represents expression values (\log_2 scale) centred around the median for each gene. The pink line represents the average gene expression values of this cluster in \log_2 scale after normalization.

B. Probability of over/underrepresentation of selected functional categories for cluster 5. Functional categories are represented in different colours. Cell wall and protein fate functional categories are overrepresented (hypergeometric test p value > 1.3), a fact highlighted by a red outline.

The RNA-Seq was analysed globally without taking attention on individual genes at the beginning. The fold expression changes of genes present within *lyn1* up-regulated clusters was ranked from large to small and visually inspected. Notably, *HO3* and *HEMA3*, which are two enzymes involved in tetrapyrrole synthetic pathway, are present in cluster 5. Remarkably, *HO3* gene expression (\log_2 scale) ratio of *lyn1 hy1* / *hy1* in both light is about 8-fold, while the ratio *lyn1* / WT in the dark and in the light is about 9-fold and 4-fold respectively. The ratio is only 2-fold in *lyn1 hy1* / *hy1* in the dark, probably, because of the variability within RNA-Seq samples mentioned above (see Supplementary Table online). *HEMA3* gene expression ratio of *lyn1 hy1* / *hy1* in the light is about 4-fold and whereas the ratios of *lyn1* / WT in the dark, *lyn1 hy1* / *hy1* in the dark and *lyn1* / WT in the light are all about 2-fold (see Supplementary Table online).

More revealing results were obtained from other clusters. The cluster 1 of significantly differentially expressed genes shows a light up-regulated pattern of expression because all the genes were consistently highly expressed in the light and low expressed in the dark (Figure 5.6 A). This cluster also underwent the functional gene classification analysis. As expected, photosynthetic genes are

highly overrepresented as they are light regulated (Figure 5.6 B). The metabolic process-related genes are also overrepresented in cluster 1.

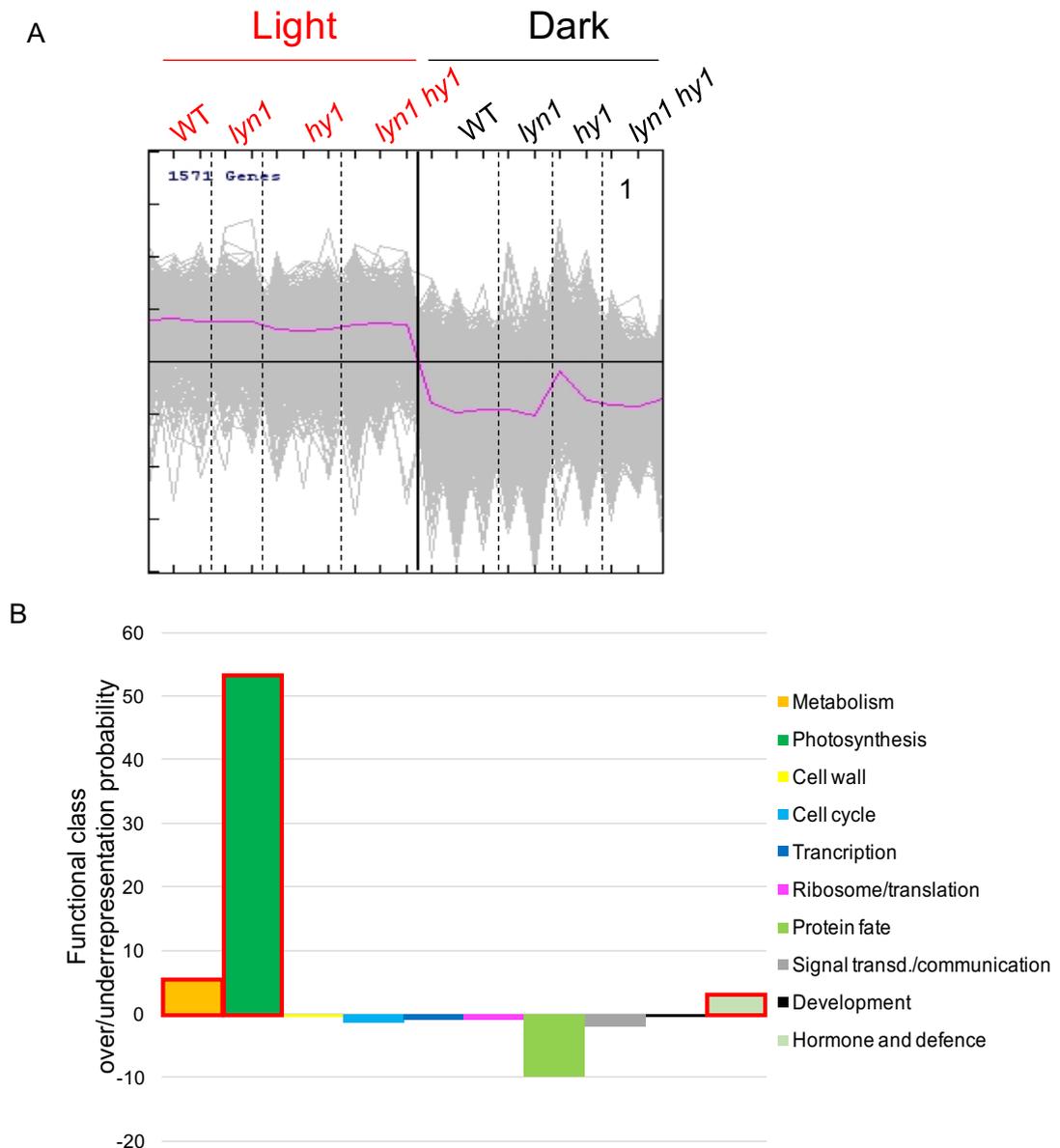


Figure 5. 6 Cluster 1 of differentially expressed genes reveals enrichments of biological function categories

A. Cluster 1, which is light dependent, was obtained as per Figure 5.5. x and y axes are also labelled as in Figure 5.5.

B. Probability of over/underrepresentation of selected functional categories for cluster 1, calculated as for Figure 5.5. Metabolism, photosynthesis and hormone and defence functional categories are overrepresented (hypergeometric test p value > 1.3), a fact highlighted by a red outline.

Cluster 1 is a large cluster which contains a total of 1571 genes. I suspected that some genes of cluster 1 were not only light up-regulated but also *lyn1*-regulated. Therefore, genes of cluster 1 were further classified into 10 subclusters (Appendix Figure 7.9). From the gene expression patterns of these, it is apparent that subclusters 1.1 and 1.10 are light-dependent and also *lyn1* up-regulated genes (Figure 5.7). Because the light-induced gene expression change is much stronger than *lyn1* regulation, the *lyn1*-regulated and light-dependent genes did not appear as an individual group other than as a strongly light-dependent genes when the significantly differentially expressed genes of cluster 1 were sub-clustered into fewer groups.

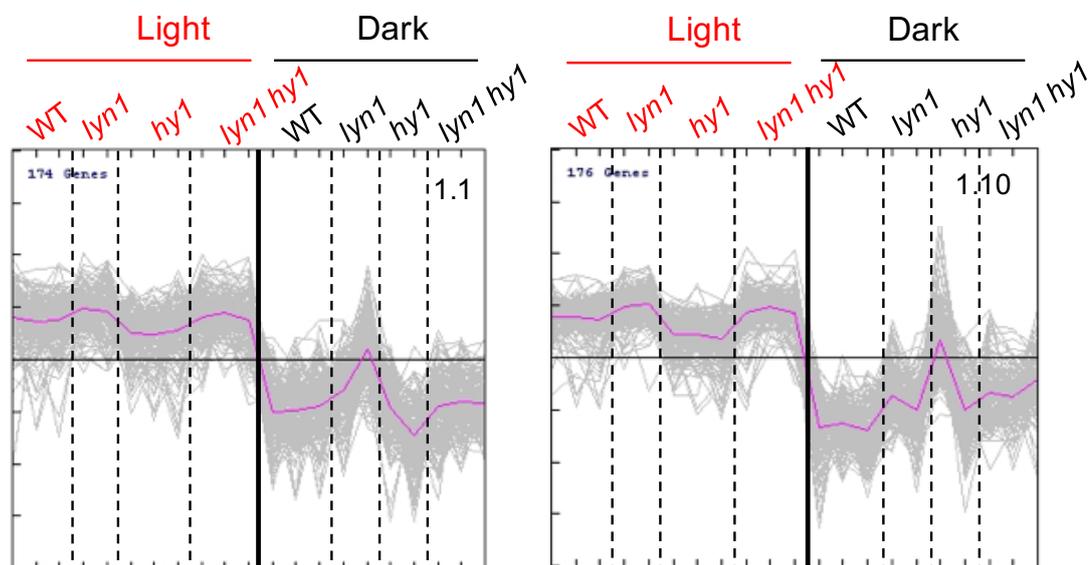


Figure 5.7 *lyn1* up-regulated subclusters selected from Cluster 1

Cluster 1, which contains light-dependent genes, was further classified into 10 subclusters by K-means clustering using MeV.

Subcluster 1.1 (left) and 1.10 (right) were the two subclusters which contained *lyn1* up-regulated genes. They were obtained as per Figure 5.5. x and y axes are also labelled as in Figure 5.5.

Subclusters 1.1 and 1.10 underwent a functional gene classification analysis. Subcluster 1.1 contains a slightly higher than expected number of

ribosome/translation-related genes and hormone-related genes. (Figure 5.8 A). Subcluster 1.10 exhibits over-representation of metabolic process and hormone-related genes (Figure 5.8 B). Surprisingly, photosynthetic functional genes are not overrepresented in either of these two subclusters, 1.1 and 1.10. Additionally, neither a single gene from the functional gene list of chloroplast developmental (global) genes nor chloroplast translation-related genes is present in these two subclusters. This suggests that *lyn1* may not regulate nuclear-encoded chloroplast genes directly.

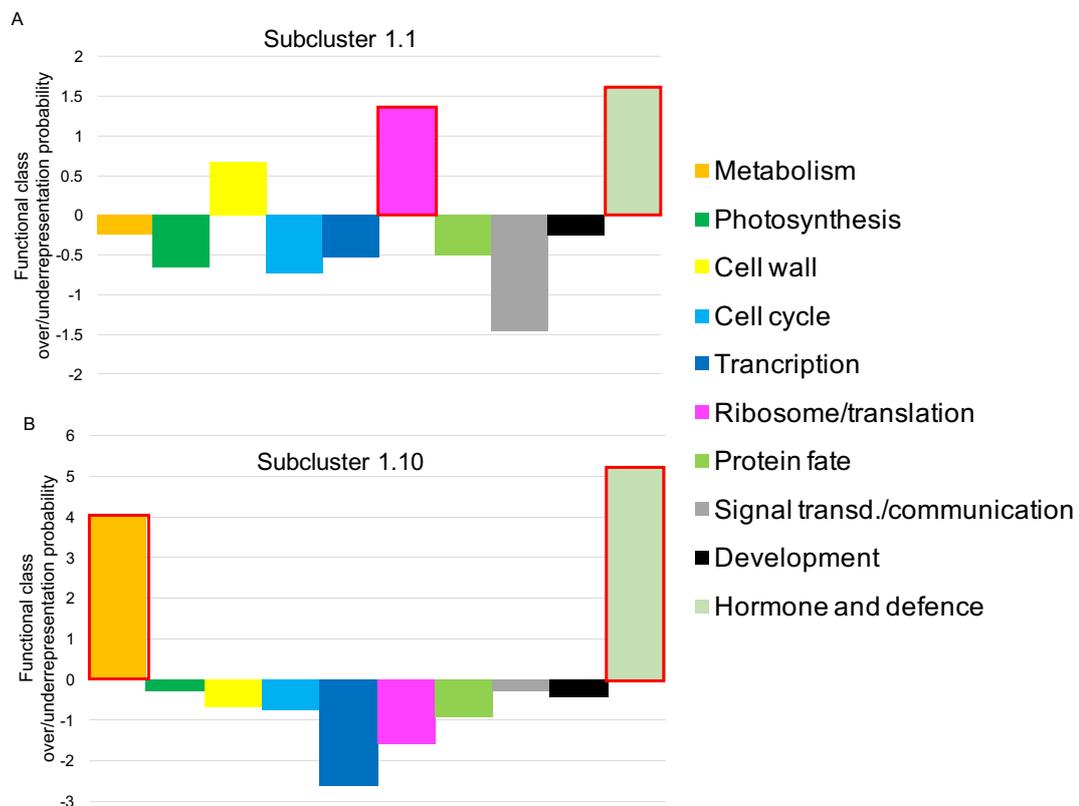


Figure 5. 8 *lyn1* up-regulated and light-dependent subclusters 1.1 and 1.10 reveal enrichments of biological function categories

y-axis represents the probability of over/underrepresentation of selected functional categories for *lyn1* up-regulated and light-dependent subcluster 1.1 (A) and 1.10 (B). Function categories of ribosome/translation and hormone and defence are overrepresented in subcluster 1.1. Function categories of metabolism and hormone and defence are overrepresented in subcluster 1.10. This fact is highlighted by a red outline.

Through this analysis of functional gene classification of *lyn1*-regulated genes, the explanation of how *lyn1* rescues the light response defect remains unclear. There are two possible reasons why a *lyn1*-regulated biological process is not clearly detected. One possible reason is that *lyn1*-regulated expression of a distinctive biological process which is not included in the functional gene classes. The other one is *lyn1* regulates multiple biological processes, so that none of the biological processes show clear overrepresentation. Therefore, publicly available data from existing literature were used to further understand how *lyn1* rescues the light response.

5.2.4 Association of *jmj14* function with early development

In search for a common biological role of genes whose expression is affected by the mutation, I next considered the fact that JARID-family proteins have been found to play developmental roles in animal cells (see section 3.3.3). Genes with roles early or late in development are not annotated as such, however it is possible to identify genes expressed at different stages of development making use of publicly-available data. The expression of a total of 23257 genes was quantified by Affymetrix microarray analysis to generate an *Arabidopsis* gene expression map (Schmid et al., 2005). The data obtained are referred to as the *Arabidopsis* Gene Atlas. The overlap between those Affymetrix-represented genes with the total expressed genes (21612) of the RNA-Seq processed result was sought. 18644 genes overlapped, so about 85% genes of RNA-Seq data are covered in the study mentioned (as well as other analyses which are based on Affymetrix microarray). The *Arabidopsis* Gene Atlas includes samples representing the youngest shoot apices, containing the apical meristem, young leaf primordia, multiple stages of leaf development, all the way to senescing leaves. Multiple lists of genes whose expression changed during organ development could be generated, by comparing sets of data from within the Gene Atlas (see methods, section 2.10.3). For example, the expression values of all

genes in sample 6 (early seedling, meristem-containing shoot apex, see Table 5.1) were divided by the expression values of the same genes in sample 5 (young leaf primordia) or 7 (whole green shoots of seedlings) to obtain two lists of genes whose expression changed between early and later development in 7-day old seedlings. The ratios 6/5, 6/7 and 17/13 represent Early or Ongoing vs. Late development (of shoot meristem-derived organs or leaves). The ratio 6/1 and 17/12 represents Early or Ongoing vs. Completed Development. The ratio 7/3 represents Photosynthetic vs. Non-Photosynthetic organs. The lists of differentially expressed genes resulting from each ratio were selected by setting up a threshold of 2 fold change. For example genes whose expression ratio (6/1, 6/5, 6/7) was greater than two were considered to show elevated expression early in development. The genes resulting from this selection first underwent the functional gene classification analysis. All 3 ratios (6/5, 6/7 and 6/1) of 7-day seedlings represent very similar biological processes. They are overrepresented in ribosome/translation, transcription, cell cycle and development functional genes (Figure 5.9). Notably, neither a single gene from the functional gene categories of chloroplast developmental (global) genes nor chloroplast translation-related genes is present in 6/5 and 6/7. This suggests that genes involved in chloroplast development are not more highly expressed in the shoot apex or meristem before leaf initiation than they are later in very young leaves or in green shoots.

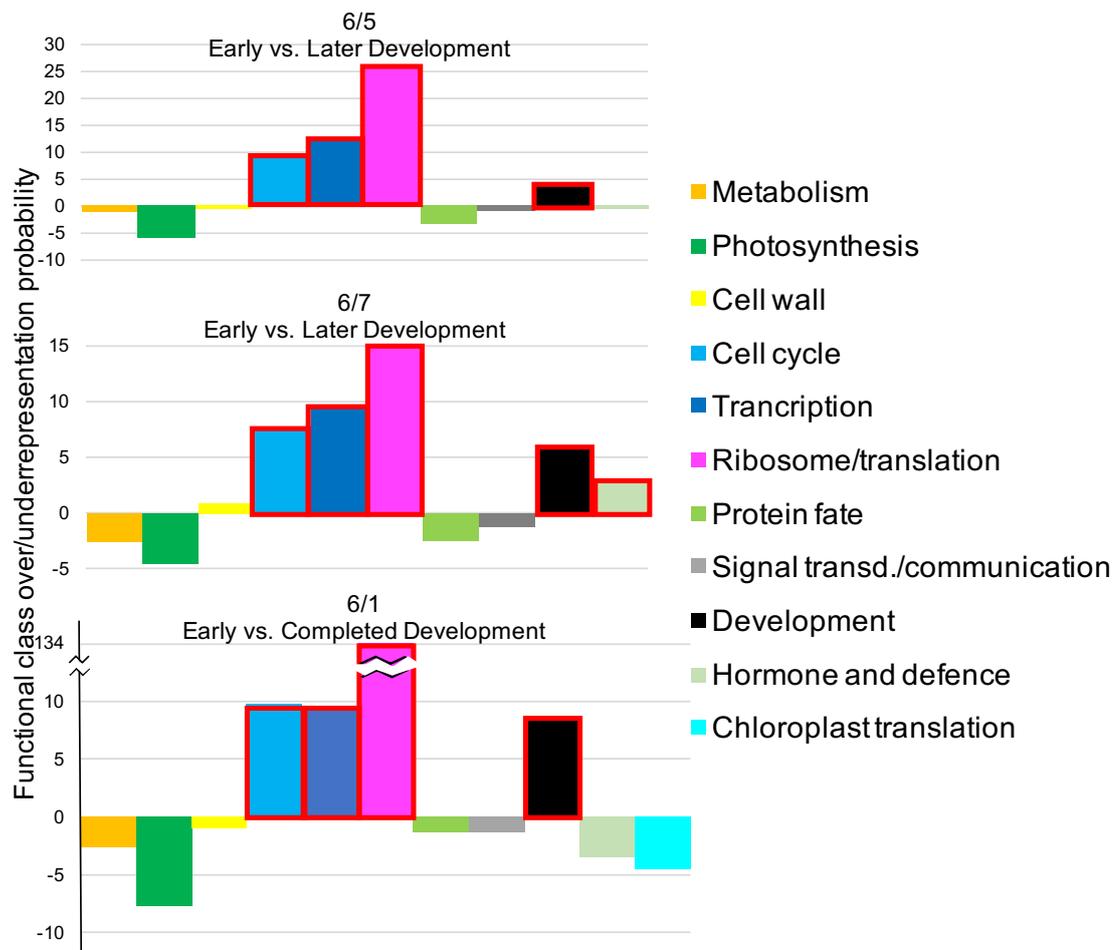


Figure 5. 9 Functional gene classification of the differentially expressed genes during the early development of 7-day old seedlings

Probability of over/underrepresentation of selected functional categories for the genes whose expression differed between stages 6/5, 6/7 and 6/1. 6/5 and 6/7 represent early vs. later development (or entry into differentiation) and 6/1 represents early vs completed development of shoot tissues of 7-day old seedlings. The overrepresented functional categories are highlighted by red outlines. Function categories of cell cycle, cell wall, ribosome/translation and development are overrepresented in common.

17/13 and 17/12 of the 17-day old plant samples also show very similar overrepresentation pattern of biological functions, different to the one from seedling samples. They both exhibit overrepresentation of ribosome/translation, cell wall, cell cycle and chloroplast translation-related genes (Figure 5.10). Only

the photosynthesis function is slightly overrepresented in 17/12, but not in 17/13. The overrepresented function of chloroplast translation suggests that chloroplasts have started to develop after leaf initiation.

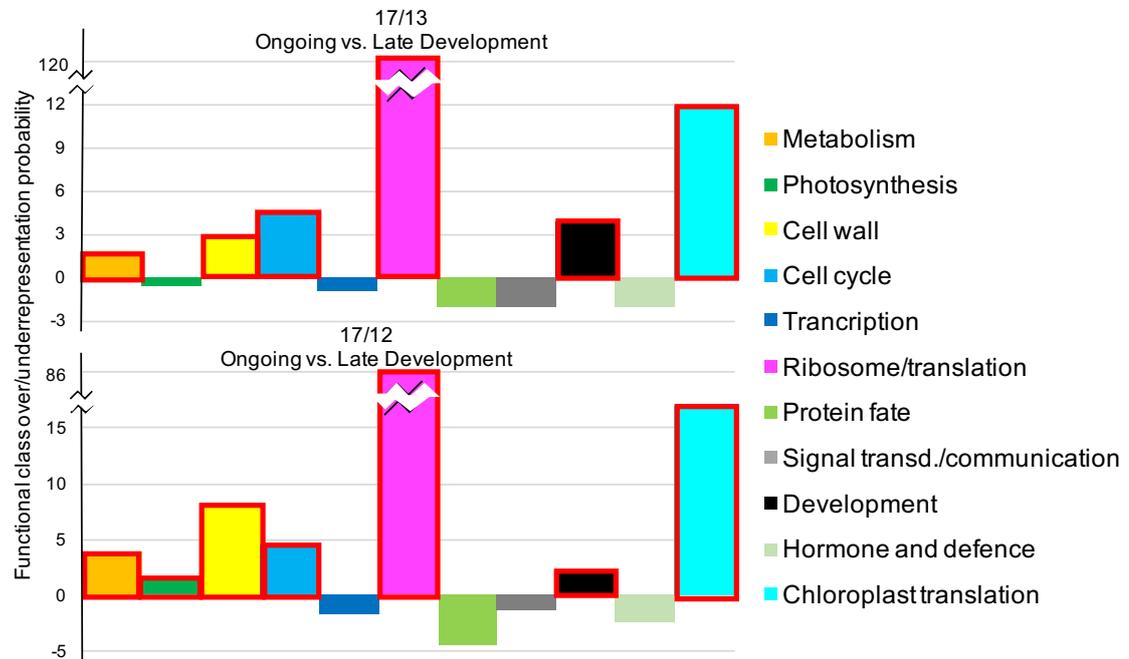


Figure 5.10 Functional gene classification of differentially expressed genes during early development of 17-day young plants

Probability of over/underrepresentation of selected functional categories for the genes whose expression differed between stages 17/13 and 17/12. 17/13 represent early vs. late development and 17/12 represents early vs completed development of 17-day old young plants. The overrepresented functional categories are highlighted by red outlines. Function categories of metabolism, cell wall, cell cycle, ribosome/translation, development and chloroplast translation are overrepresented in common.

The 7/3 ratio selection shows a completely different function which contains an overrepresented number of photosynthetic genes, ribosome/translation and chloroplast translation-related genes (Figure 5.11). Development of chloroplasts is, as expected, more active in photosynthetic organs as the function of chloroplast translation genes is highly overrepresented in this comparison.

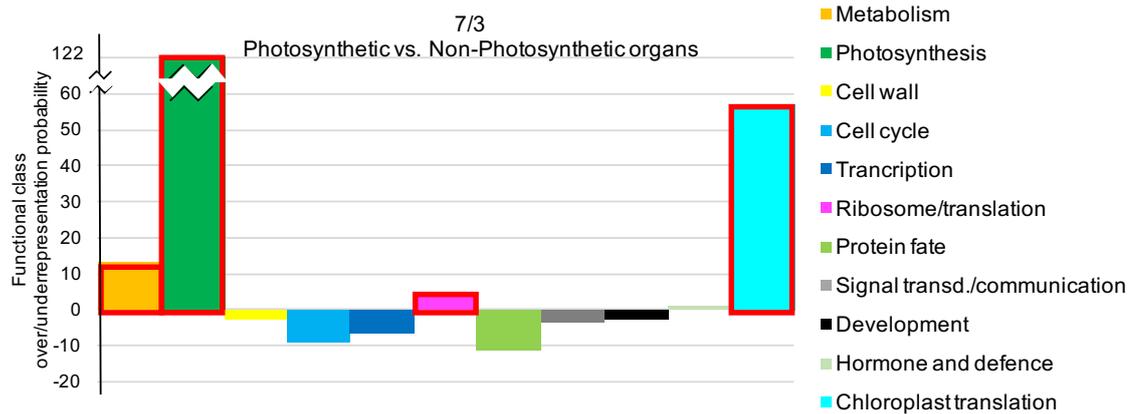


Figure 5. 11 Functional gene classification of the differentially expressed genes between photosynthetic vs. non-photosynthetic organs (7/3) of 7-day old seedlings

Probability of over/underrepresentation of selected functional categories for the genes whose expression differed between organs present in samples 7/3. Functional categories of metabolism, photosynthesis, ribosome/ translation, and chloroplast translation are overrepresented and highlighted by red outlines.

The differentially expressed genes contained in these six ratios should include the regulators of cell entry into differentiation at early developmental stages, as long as such regulators themselves change in expression, and should contain their targets, the agents of development. The overrepresentation of genes in these lists was examined among the *lyn1* up-regulated, differentially expressed genes clusters from processed RNA-Seq data. The result shows that all the ratios are underrepresented in cluster 5 (Figure 5.12 A). Subcluster 1.1 shows overrepresentation of 17/13, 6/7 and 7/3-selected genes (Figure 5.12 B). Subcluster 1.10 shows overrepresentation of 6/5 and 7/3-selected genes (Figure 5.12 C). Because the genes of cluster 1 are light-dependent genes, the ratio 7/3 which represents photosynthetic vs. non-photosynthetic organs is always significantly overrepresented.

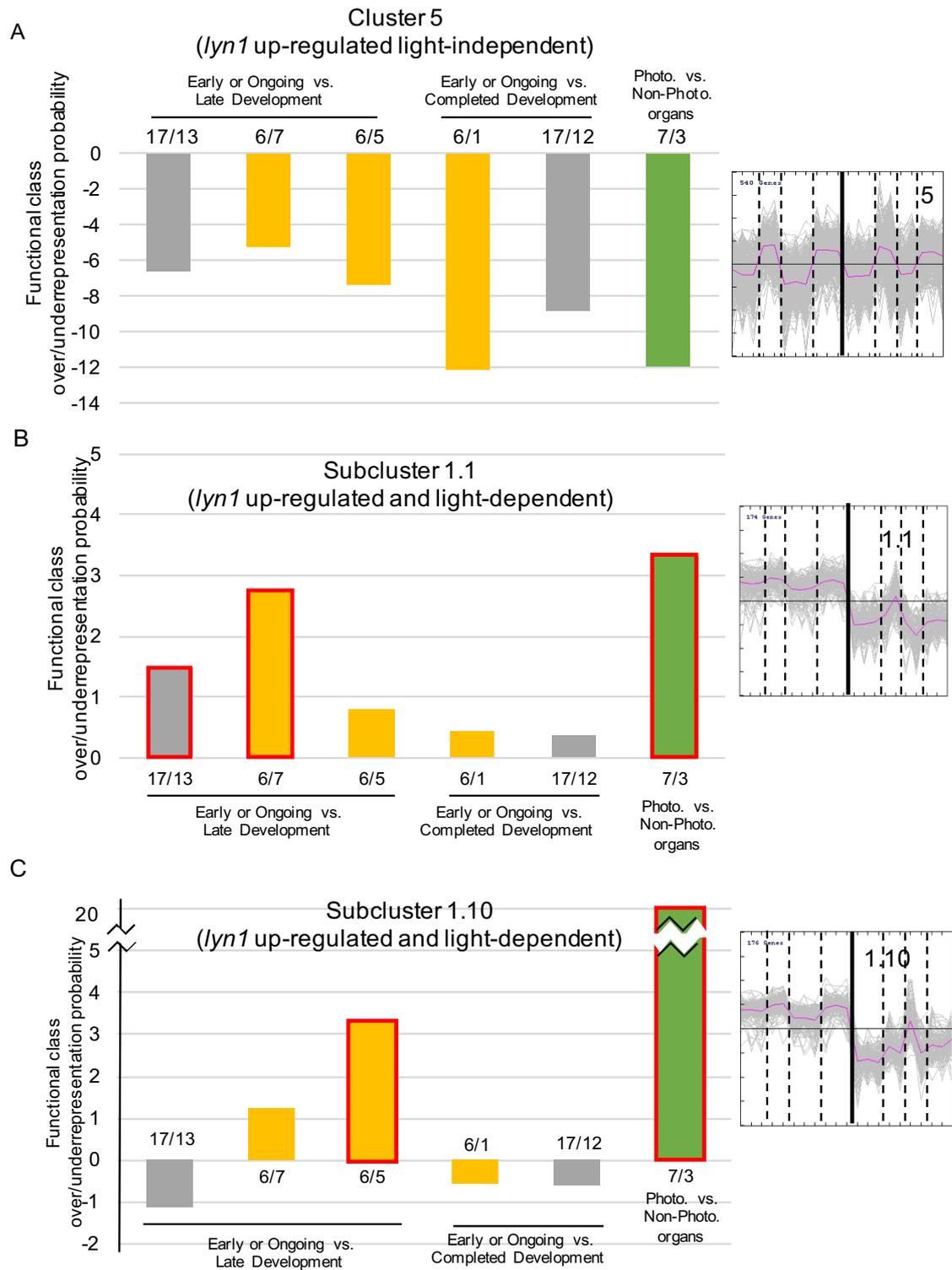


Figure 5. 12 Overrepresentation of genes which are active at a particular developmental stage among *lyn1*-dependent gene clusters
6/7 6/5 and 6/1 (yellow) are genes whose expression differed between stages of early development of 7-day old seedlings. 17/13 and 17/12 (grey) are genes

whose expression differed between stages of development of 17-day old young plants. 7/3 are genes whose expression differed between photosynthetic and non-photosynthetic organs. Probability of over/underrepresentation of genes representing each of those developmental comparisons, among genes in *lyn1* up-regulated, light-independent gene cluster 5 (A) or *lyn1* up-regulated, light-dependent subclusters 1.1 (B) and 1.10 (C). The overrepresented developmental comparisons are highlighted by red outlines. Early vs. later development (6/5 and 6/7) in 7-day old seedlings is overrepresented. The genes which are highly expressed in photosynthetic relative to non-photosynthetic organ are overrepresented in both light-dependent clusters (subclusters 1.1 and 1.10).

To find out whether the overrepresentation of 17/13, 6/7 and 6/5-selected genes in subclusters 1.1 and 1.10 were a consequence of light induction, the overrepresentation of these ratios was also analysed in the rest of subclusters 2~9. If the overrepresentation of 17/13, 6/7 and 6/5 in subclusters 1.1 and 1.10 is a simple consequence of light induction, then these three ratios should also be overrepresented in the other clusters. The result shows that 17/13, 17/12 and 7/3-selected genes (those which select for ongoing development or photosynthetic organs) are overrepresented in the rest of subclusters, but 6/7, 6/5 and 6/1-selected genes (those which select for expression in young cells, entering differentiation, at the shoot apex), are not. (Figure 5.13).

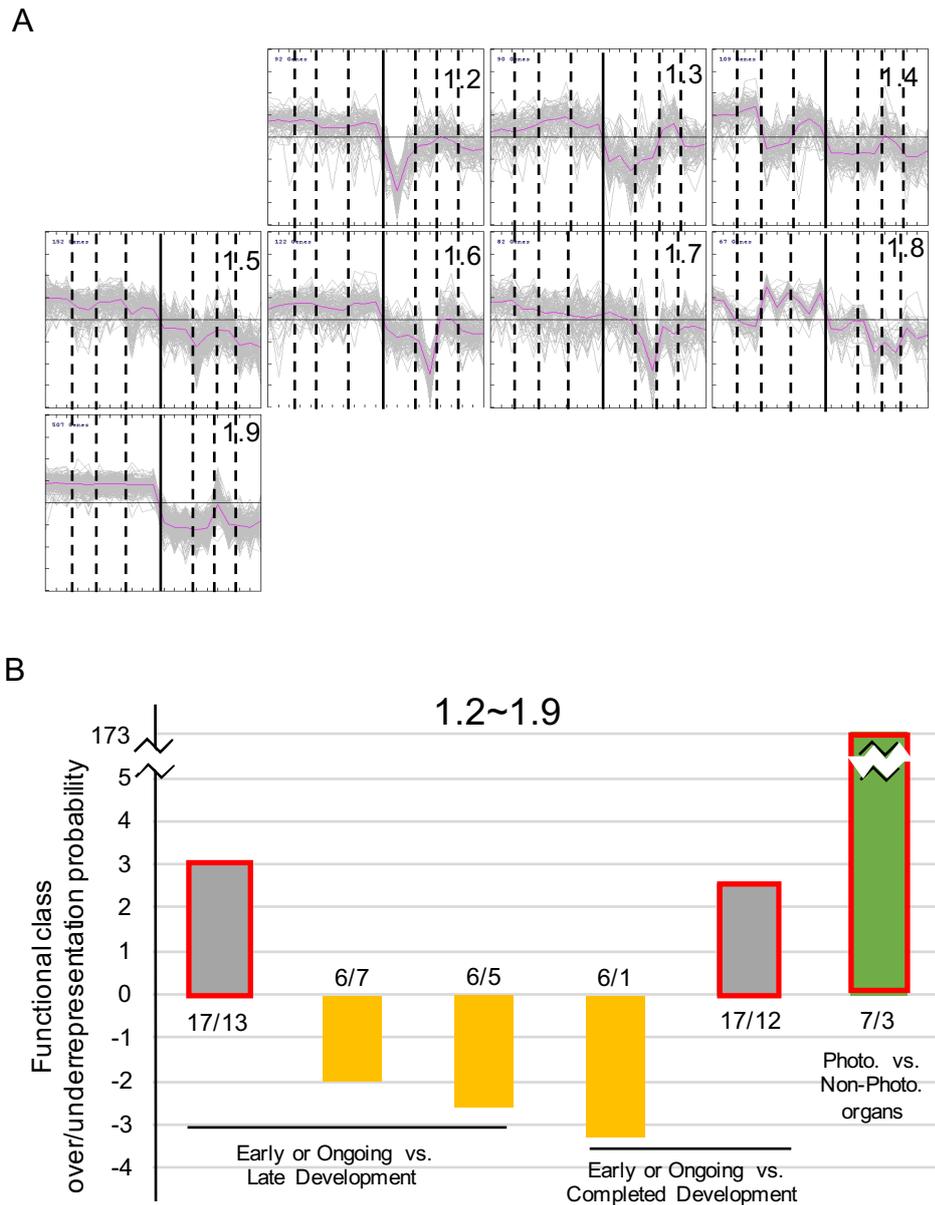


Figure 5. 13 Overrepresentation of genes which are active at a particular developmental stage among light-dependent but *lyn1*-independent gene clusters. A. Subcluster 1.2~ 1.9, which are *lyn1*-independent and light-dependent, were obtained by K-means clustering of cluster 1 using MeV. x and y axes are labelled as in Figure 5.5.

B. Probability of over/underrepresentation of genes which accompany cell entry into differentiation or represent ongoing development or photosynthetic activity, among *lyn1*-independent, light-dependent gene subclusters 1.2~1.9 of cluster 1. The overrepresented developmental comparisons are highlighted by red outlines.

The genes overrepresented are only those whose expression differed between stages of development in 17-day old young plants. In addition, genes which are highly expressed in photosynthetic relative to non-photosynthetic organs are always overrepresented among light-dependent clusters.

5.2.5 Association of *jmj14* function with H3K4me3-marked global genes

Roudier and collaborators (2011) detected the presence of H3K4me3 mark along genes in the DNA of whole seedlings or roots using custom-designed whole genome tiling arrays. 17861 H3K4me3 genes were found to be marked in whole seedlings' DNA and 13117 H3K4me3 genes were found to be marked in the roots' DNA (Roudier et al., 2011). As an initial guide, the representation of the H3K4me3 chromatin marked state of a small number of selected genes was observed online using the *Arabidopsis* epigenome (Epigara) viewer, and compared between 7-day old whole seedlings and roots (Figure 5.14). Clearly, the photosynthetic genes (*LHCB1.2*, *LHCB1.3*, *APRR2* and *GLK1*) have significant overrepresentation of the H3K4me3 mark (the brown bars representing mark above background) in the DNA from whole seedlings compared to that from roots. The constitutive genes (*ACT2* and *UBQ10*) present significant H3K4me3 mark in both whole seedlings and roots. *CHS* is a key enzyme involved in the flavonoid biosynthesis process (Dao et al., 2011). It exhibits significant H3K4me3 mark in the roots, but barely so in the whole seedlings. In summary, the H3K4me3 chromatin marking on my small selection of chloroplast-associated genes appears more frequently in green tissue, where photosynthesis occurs, than in roots. Following these initial observations, a more in depth examination appeared justified. After genes marked exclusively in green tissue were identified, a functional classification analysis of those genes was carried out. The occurrence of those genes among the gene clusters showing *lyn1*-dependence was

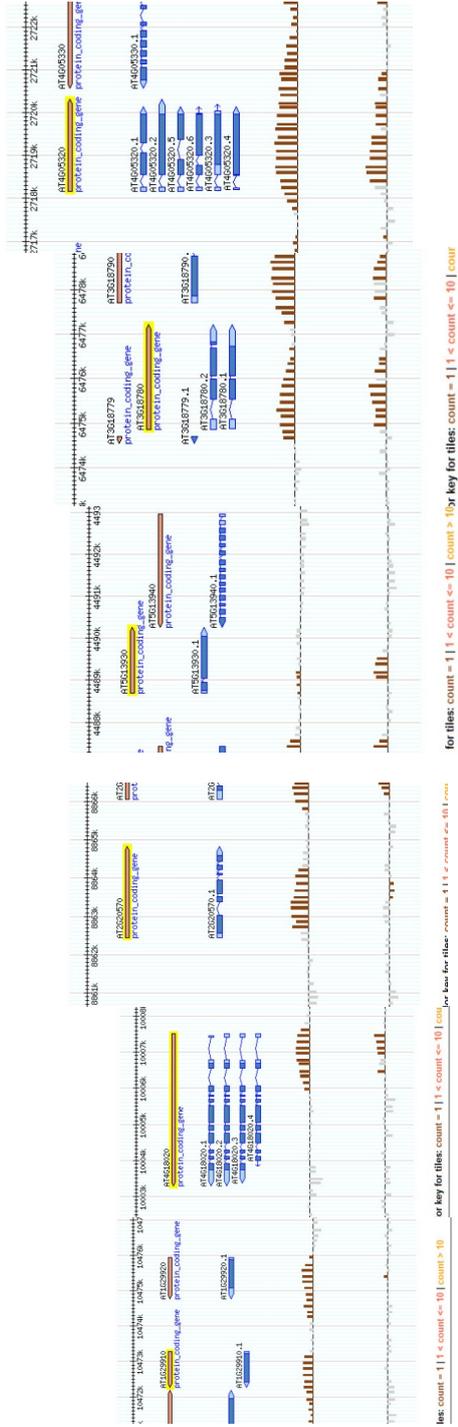
quantified. Finally, the frequency of presence of the mark among genes within photosynthesis or chloroplast development functional classes was assessed.

Figure 5. 14 (in the following page) Apparent differential extent of H3K4me3 chromatin marking associated with photosynthesis related genes in the DNA of whole seedlings and in the DNA of roots

Roudier and collaborators (2011) extracted DNA samples from 7-day whole seedlings and separately from roots. The ChIP-Chip technique was used to quantify the amount of H3K4me3-marked DNA copies among the total genomic DNA (i.e. relative to total, input DNA) by the intensity of hybridisation. Each bar represents the frequency of H3K4me3 chromatin modification and the position of each bar corresponds to the histone marked location on the gene (the genomic position of the oligonucleotide to which hybridisation occurs). The height of each bar represents the intensity of hybridisation after normalisation, background subtraction, etc. Brown indicates that the frequency is statistically significantly above the background (control DNA hybridisation). Grey indicates that the frequency is not statistically different from background. Genes of interest were selected for visualisation, and the images displaying their map of marks downloaded from the Generic Genome Browser at Epigara (<http://epigara.biologie.ens.fr/cgi-bin/gbrowse/a2e/>).

7 day old Whole seedlings ↗

7 day old roots ↗



LHCB1.2 LHCB1.3 APRR2

GLK1

CHS

ACT2

UBQ10

Photosynthetic

Other / constitutive

According to the analysis of association of *jmj14* function with early development, *jmj14*-regulated genes show an overrepresentation of genes expressed during early development in green tissue. Therefore, it is interesting to know whether *jmj14* regulates the H3K4me3-marked genes in shoots, typically in the green organs of seedlings. The overlap between the list of H3K4me3-marked whole seedling genes with root genes was sought to obtain the H3K4me3-marked genes in green tissue. The non-overlapping whole seedling genes represent the H3K4me3-marked genes in shoots (or “green tissue”) only. 5291 of the H3K4me3-marked genes are found to present the mark in green tissue only (Figure 5.15).

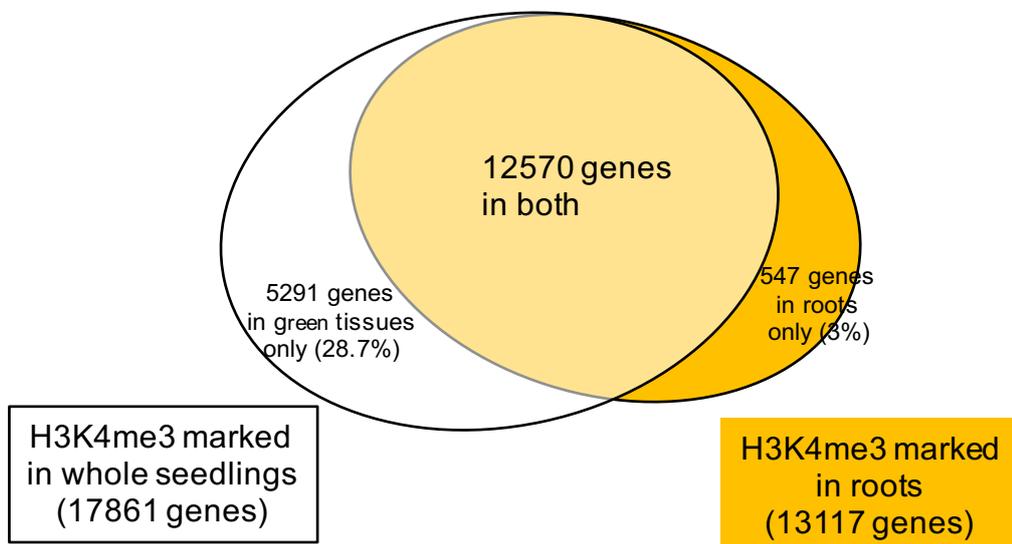


Figure 5. 15 Venn diagram showing the number of genes which are H3K4me3-marked in common or uniquely in genomic DNA from whole seedlings or roots. The group of genes marked in whole seedlings after subtracting those marked in both whole seedlings and roots represents genes marked in “green tissues” only.

The H3K4me3-marked genes in the whole seedlings and, separately, in the “green tissues” underwent functional gene classification to reveal the biological process they represented. Using a Venn diagram, a total of 17016 H3K4me3 genes marked in the DNA of whole seedlings and 4668 H3K4me3 genes marked in the DNA of “green tissue” are found among the expressed genes from the processed RNA-Seq data, respectively. These were then treated as a group of samples to undergo the functional gene classification analysis. The H3K4me3-marked genes in the whole seedlings exhibit overrepresentation of ribosome/translation, cell cycle, protein fate and signal transduction/communication-related genes (Figure 5.16 A). H3K4me3-marked genes in the “green tissues” exhibit overrepresentation of photosynthetic genes and cell cycle-related genes. (Figure 5.16 B).

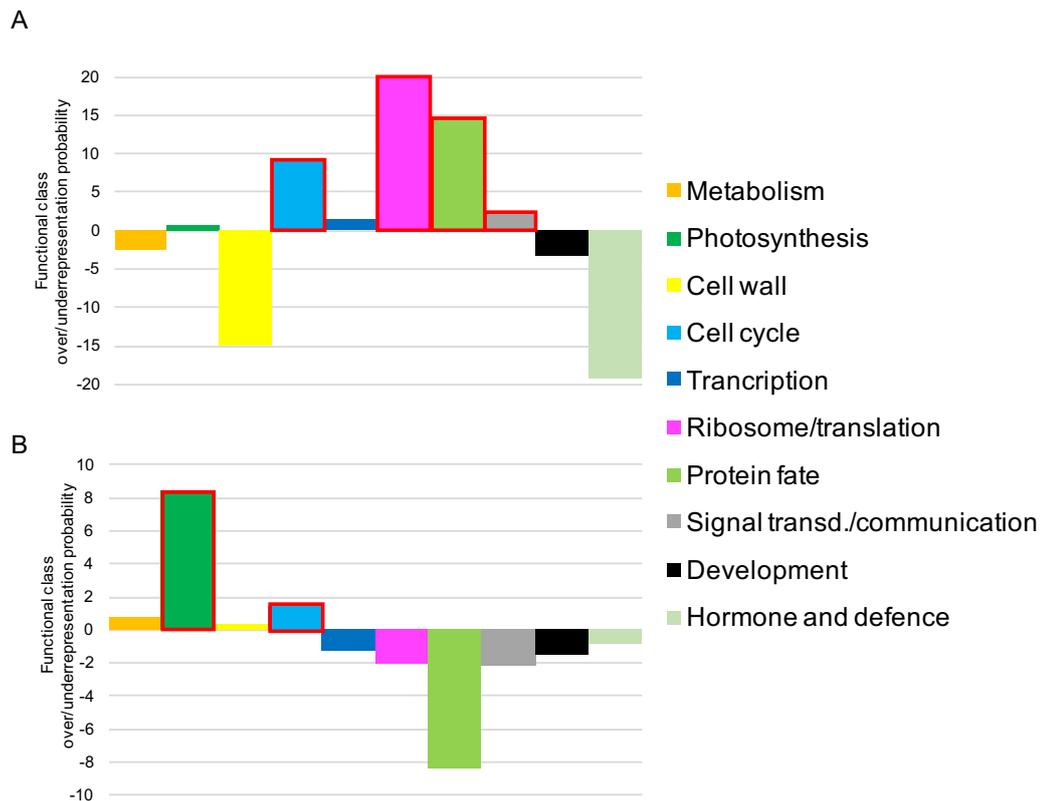


Figure 5.16 Functional classification analysis of H3K4me3-marked genes

y-axis represents the probability of over/underrepresentation of selected functional categories for the H3K4me3-marked genes in the DNA of whole seedlings (A) and in the DNA of green tissues only (B). The overrepresented classes are highlighted by red outlines. Functional categories of cell cycle, ribosome/translation, protein fate and signal transd./communication are overrepresented in the DNA of whole seedlings. Functional categories of photosynthesis and cell cycle are overrepresented in the DNA of green tissues only.

Because *JMJ14* encodes a H3K4 trimethyl demethylase, the clusters which are defined as *lyn1* up-regulated genes are probably the H3K4me3-marked if they are *JMJ14* primary affected genes. The overrepresentation of total H3K4me3-marked gene among all the gene clusters of interest was analysed to confirm if they were H3K4me3-marked genes. The result shows that all the gene clusters

of interest have significant overrepresentation on H3K4me3-marked genes in DNA of green tissues, with the exception of cluster 5 (*lyn1*-dependent light-independent genes) (Figure 5.17 B). However, none of these clusters show overrepresentation on H3K4me3-marked genes in DNA of whole seedlings (Figure 5.17 A).

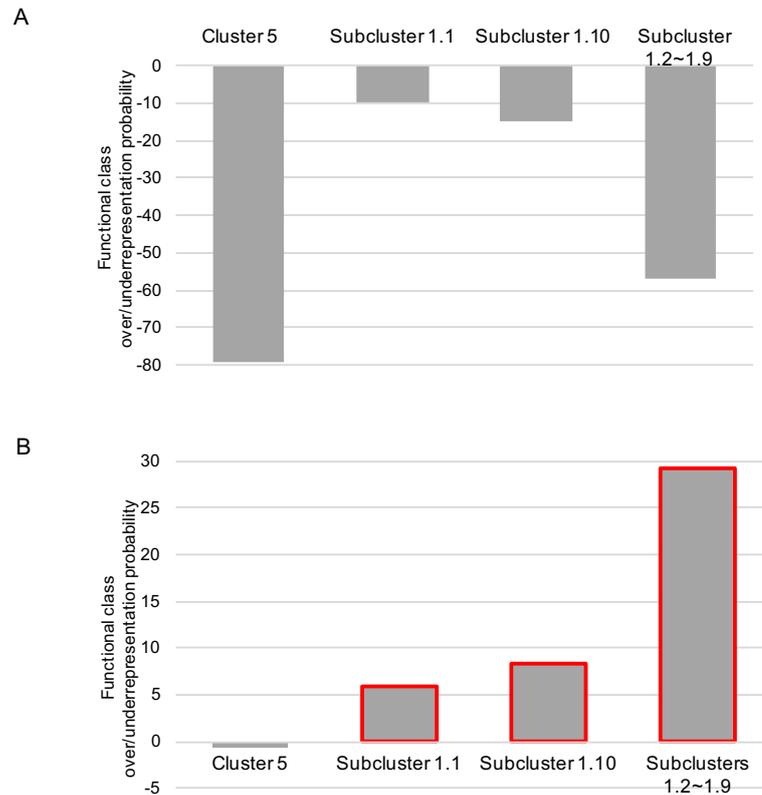


Figure 5. 17 Over/underrepresentation of H3K4me3-marked genes among all the differentially expressed gene clusters of interest

The gene clusters of interest include cluster 5 (*lyn1* up-regulated and light-independent), subclusters 1.1 and 1.10 (*lyn1* up-regulated and light-dependent) and subclusters 1.2~1.9 (light-dependent only). y-axis represents the probability of over/underrepresentation of genes part of the gene clusters of interest among the H3K4me3-marked genes in the DNA of whole seedlings (A) and in the DNA of green tissues only (B). The overrepresented clusters are highlighted by red outlines. All subclusters of cluster 1 (all light-dependent genes) are overrepresented among the H3K4me3-marked genes in the DNA of green tissues only.

5.2.6 Association of H3K4me3-marked global genes with early development

So far, I associated, through overrepresentation, the *lyn1*-regulated genes with genes active during cell entry into differentiation at an early developmental stage and with H3K4me3-marked chromatin. Are the genes active during cell entry into differentiation at an early developmental stage frequently (disproportionately) H3K4me3-marked genes? This was answered by analysing the overrepresentation of H3K4me3-marked genes in the DNA of whole seedlings and in that of green tissues among genes selected as the expression is increased at particular developmental stages, to see if the H3K4me3 mark is particularly present in genes expressed at any particular developmental stages. The results show that H3K4me3-marked genes in the DNA of whole seedlings and in that of roots are greatly overrepresented among the genes upregulated in all early developmental stages (Figure 5.18).

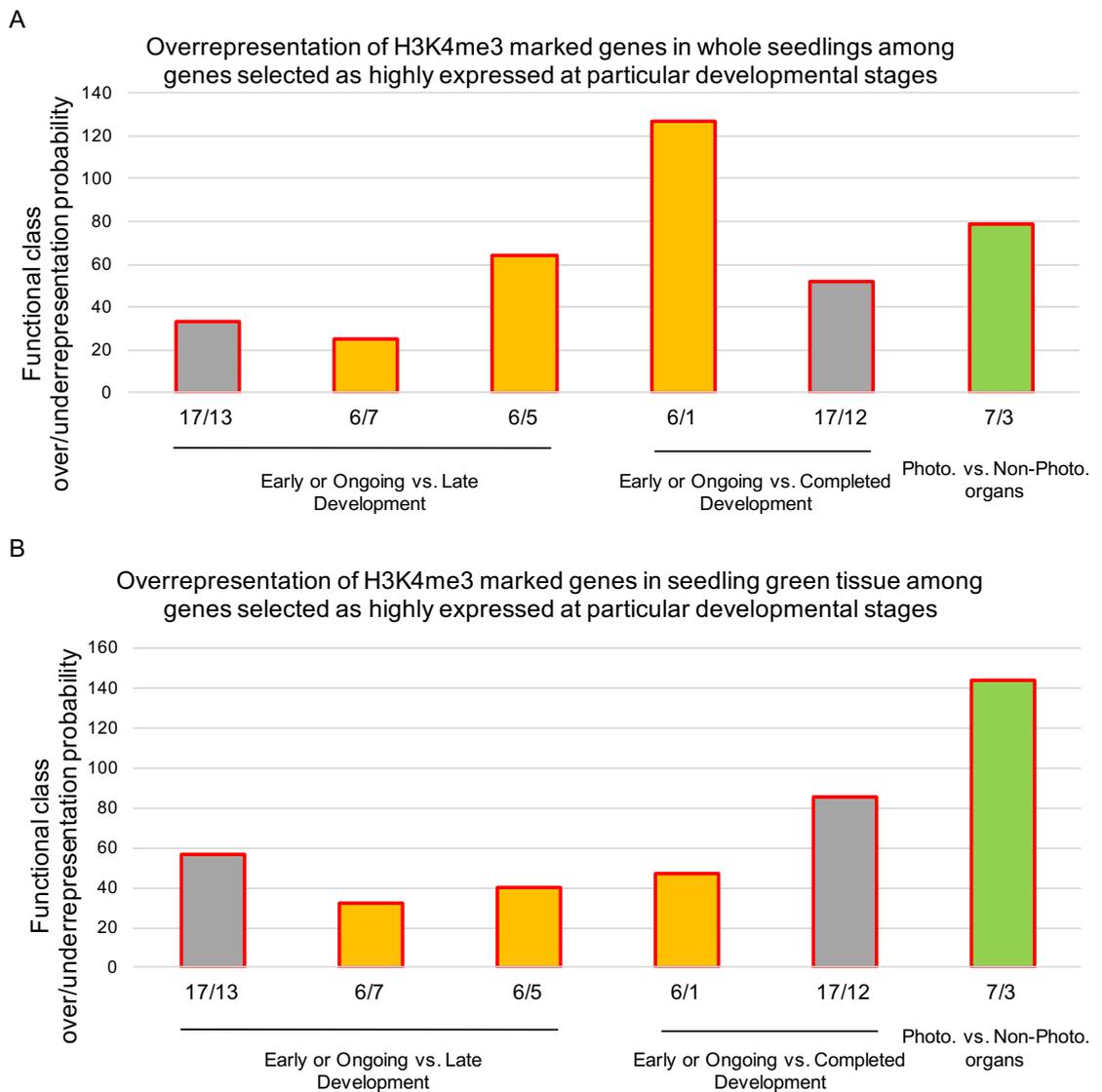


Figure 5. 18 Overrepresentation of H3K4me3-marked genes among genes selected for their increased expression at particular developmental stages y-axis represents the probability of over/underrepresentation of genes which accompany cell entry into differentiation or represent ongoing development or photosynthetic activity among the H3K4me3-marked genes in the DNA of whole seedlings (A) and in the DNA of green tissues only (B). The overrepresented developmental comparisons are highlighted by red outlines. All the groups of genes are very strongly overrepresented among H3K4me3-marked genes in the both DNA of whole seedlings and green tissues only.

Additionally, the overrepresentation of H3K4me3-marked genes in the DNA of both whole seedlings and green tissues in the photosynthesis-related genes, global chloroplast developmental genes (a group of genes largely involved in chloroplast gene transcription, protein import and division) and chloroplast translation-related genes were analysed to reveal if it is possible that *jmj14* promotes early development by particularly regulating chloroplast-related genes through H3K4me3 histone modification. The results show that the chloroplast translation genes are always underrepresented among the H3K4me3-marked genes in the DNA of whole seedlings and in that of green tissues (Figure 5.19). The global chloroplast development-associated genes are overrepresented among H3K4me3-marked genes in the DNA of whole seedlings but not in that of green tissue only. The photosynthesis-associated genes are overrepresented among H3K4me3-marked genes in DNA of green tissues but not of whole seedlings.

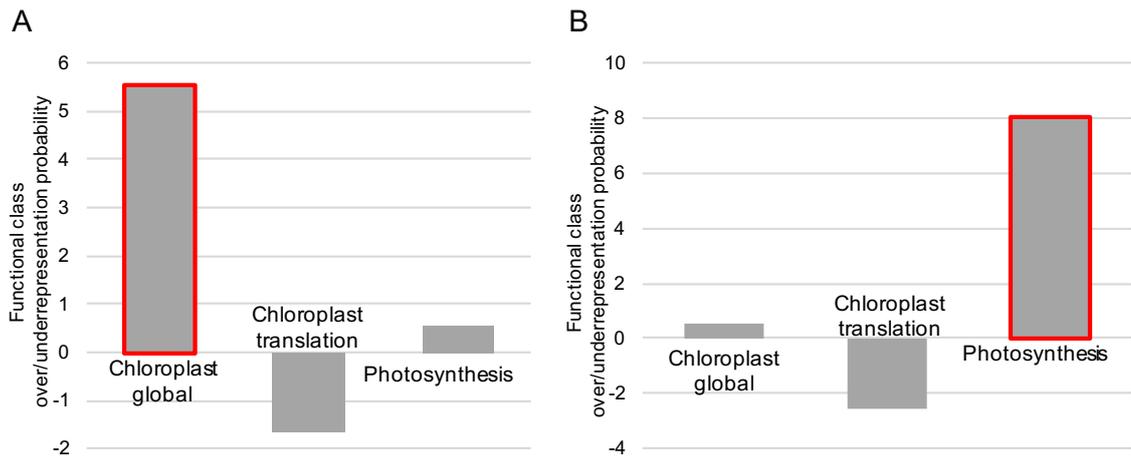


Figure 5. 19 Overrepresentation of H3K4me3-marked genes among genes classified as chloroplast developmental (global) genes, chloroplast translation-related genes and photosynthesis-related genes

y-axis represents the probability of over/underrepresentation of the H3K4me3-marked genes in the DNA of whole seedlings (A) and in the DNA of green tissues only (B) among genes within the following functional classes: chloroplast developmental function (global) genes, chloroplast translation-related genes and photosynthesis-related genes. The overrepresented functional categories are highlighted by red outlines. Chloroplast developmental function (global) genes are overrepresented in the DNA of whole seedlings and photosynthesis-related genes are overrepresented in the DNA of green tissues only.

5.3 Discussion

5.3.1 Association of *jmj14* regulation with H3K4me3 demethylation

Gene expression is closely related to chromatin state and it can be regulated by chromatin modification (Sequeira-Mendes et al., 2014). H3K4 methylation, which is histone post-translational modification, has been found to couple with gene expression (He et al., 2011; Black et al., 2013). Some previous studies have

found that the up-regulation of *FT* gene expression in *jmj14* mutants is coupled with the elevated H3K4 methylation (Jiang et al., 2008; Yang et al., 2010). The association between expression level of *FT* gene and the enrichment of H3K4 methylation was assessed *in vivo*. The qPCR result confirms that *FT* gene expression is significantly increased in *jmj14* compared to WT. To confirm whether the increased *FT* gene expression level accompanies an increased level of H3K4me1/2/3 in *FT* chromatin, the enrichment of H3K4me1/2/3 on *FT* chromatin at the transcriptional start site and middle region of the gene was determined by ChIP-qPCR. It found that H3K4me3 and H3K4me2 at the 5' end of the transcribed region of *FT* chromatin were increased in *jmj14-1* compared to WT. H3K4me1 was not increased at this same region. In the central region of *FT* chromatin, only the H3K4me2, in contrast to H3K4me3 and H3K4me1, was enriched in *jmj14-1* (Yang et al., 2010). Finally, whether JMJ14 directly bind to *FT* gene and repress *FT* gene expression was confirmed using transgenic plants expressing the Jmj4::FLAG fusion protein. Jmj4::FLAG protein was detected by ChIP assay using FLAG-specific antibody. It found that JMJ14 binding is significantly enriched at the 5' end of the transcribed region compared to other regions. Therefore, the JMJ14 is determined to directly bind and repress *FT* expression by H3K4me2 and H3K4me3 demethylation. A similar assay was also performed on JMJ18 and confirmed that JMJ18 repress *FLC* gene expression by directly binding *FLC* chromatin and demethylating H3K4me2/3 (Yang et al., 2010). The H3K4 methylation activity was also assessed *in vitro*. There is substantial evidence to show that JMJ14 is much more efficient at H3K4me3 demethylation and less efficient at H3K4me2 and H3K4me1 demethylation *in vitro* (Lu et al., 2010; Yang et al., 2009; Jeong et al., 2009). Probably, JMJ14 can also readily demethylate H3K4me2 *in vivo*. Although the majority of genes are marked by H3K4me2, H3K4me3 or both. H3K4me3 methylation occurred at the promoter region of genes that leads to a positive correlation with transcript abundance and this does not show in H3K4me2 (Zhang et al., 2009; Roudier et al., 2011).

H3K4me2 is highly enriched when genes were intermediately expressed, whereas it is less enriched when genes were either less or highly expressed. Therefore, it is very unlikely that *jmj14*-regulated genes for rescuing light response would be H3K4me2-marked, although this cannot be completely ruled out. Analysis of the correlation between *jmj14*-regulated genes and H3K4me2-marked genes will be more complicated, because when changing the marking state, gene expression can be either up- or down-regulated and the expected gene expression pattern will become unpredictable. If H3K4me2 is also potentially regulated by *jmj14*, then it will be necessary to study the H3K4 marking state on *lyn1* directly by ChIP-Seq. According to all the studies of the association of H3K4 methylation with gene expression, I predicted that the *jmj14* mutant causes a defect of H3K4me3 demethylation that increases the expression level of JMJ14-regulated genes in *jmj14* compared to WT.

According to the results shown in the last chapter, I realized that *jmj14* enhances plastid development, even in the dark. A hypothesis immediately followed. *jmj14* may rescue light responses, defective in *hy1*, by primarily regulating plastid genes, in turn increasing chloroplast development later in the light. Based on this hypothesis, the study was initially focused on finding the *lyn1*-enhanced expression of chloroplast biogenesis regulatory genes. The lists of H3K4me3-marked genes in the WT DNA from whole seedlings and roots were obtained from the existing literature (Roudier et al., 2011). Several representative genes were selected to study their H3K4me3 chromatin marking states in the DNA of whole seedlings or of roots. As a very preliminary assessment, the hypothesis was supported by comparing the H3K4me3 marking state between different categories of gene. The photosynthesis-related genes are all significantly marked by H3K4me3 compared to other types of genes. As mentioned in chapter 1, the level of H3K4me3 peaks at the 5'-ends of the coding region of actively transcribed genes and gradually decreases towards the 3'-ends of the genes (Roudier et al.,

2011; Dong and Weng, 2013). This characteristic is clearly observed in Figure 5.14. *LHCB1.2*, *LHCB1.3*, *APRR2* and *GLK1* are typical light-induced and nuclear-encoded chloroplast marker protein or biogenesis regulators. *CHS* catalyses the initial step of flavonoid biosynthesis. Flavonoids are important plant secondary metabolites that provide floral pigments. *CHS* gene expression is influenced by UV light (Dao et al., 2011). Unlike these photosynthesis-related genes, the light-induced *CHS* gene is not involved in chloroplast development. Additionally, the comparisons of H3K4me3 chromatin marking state between whole seedlings and roots among different categories of genes also support the hypothesis. Because all four photosynthesis-related genes are significantly marked in DNA of the whole seedlings and are not in the roots, photosynthesis-related genes may be regulated by H3K4me3 demethylation in green tissues where chloroplasts are present. In contrast, *CHS* and constitutive genes do not have a significantly higher frequency of H3K4me3 chromatin marking in the whole seedlings. This observation suggests it is worth exploring the possibility that *jmj14* rescues the light response by directly regulating chloroplast biogenesis regulators. Accordingly, studying the H3K4me3-marked gene in the green tissue may be critical.

Unfortunately, the study of Roudier and collaborators only examined whole seedlings and roots and not just the green part of seedlings. The H3K4me3-marked genes in the green organs were manually processed by searching for the overlap between these two H3K4me3-marked gene lists of whole seedlings and roots. The list of H3K4me3-marked genes in the green region of seedlings was expected to show very similar biological functions to the list of whole seedlings, because the majority of 7-day whole seedlings sample materials are shoots. Although the shoots do not just contain the green organs but also the hypocotyl, the majority of tissue of 7-day light-grown seedling's samples is from the green

tissues as the hypocotyls are short. The shoots' data were assumed as green tissue data because the small amount of mixture of hypocotyl tissue is negligible.

The analysis returned to the RNA-Seq data. Once the H3K4me3-marked chromatin genes in the whole seedlings and green tissues were obtained, all the clusters of interest (*lyn1* up-regulated cluster 5 and subclusters 1.1 and 1.10) were analysed as to whether they showed overrepresentation of H3K4me3-marked genes. If these clusters show significant overrepresentation on H3K4me3-marked genes, then they would be more likely to be *jmj14* up-regulated the genes. Unfortunately, the H3K4me3-marked genes are always underrepresented in cluster 5 in both the DNA of whole seedlings and green tissues (Figure 5.13). Since cluster 5 actively excluded any light regulated genes, this may have selected against the mark. Probably, the genes of cluster 5 are not JMJ14 directly targeted. They are downstream of JMJ14-regulated pathway and have no effect of rescuing light response. The gene expression patterns of subcluster 1.1 and 1.10 exactly support this idea. Apart from the light influence, the genes of these two subclusters are up-regulated in *lyn1*. In the light, the expression of these genes is strongly increased in all genotypes. However, the increased expression in *lyn1* is still retained in the light. The overrepresentation of H3K4me3 green tissue marked genes among the genes of subcluster 1.1 and 1.10 suggests that the genes of these two subclusters are more likely to be *jmj14*-dependent (Figure 5.17 B). The overrepresentation of H3K4me3-marked genes also on other gene subclusters (1.2~1.9) suggests that it is possible that the impact of the mark on expression is an important factor in the light regulation of gene expression, because H3K4me3 has been found as an important component for representative light-regulated gene transcription during *Arabidopsis* seedling photomorphogenesis (Guo et al., 2008). In addition, there is not even a single chloroplast development-associated gene and chloroplast translation gene is

present in the *jmj14* up-regulated clusters. Therefore, it is very unlikely that *jmj14* rescues light response by directly regulating chloroplast-associated genes.

However, H3K4me3-marked genes do not show a similar overrepresentation in the whole seedlings. The possible reason is that not all of the genes in the H3K4me3-marked gene list of the Roudier's data are JMJ14-regulated. In *Arabidopsis*, JMJ15 and JMJ18 also display H3K4me3 demethylase activities (Liu et al., 2010; Yang et al., 2012a,b). Using the global H3K4me3-marked genes as a biological function class to analyse the overrepresentation of *JMJ14*-regulated H3K4me3-marked genes in the hypergeometric test cause absence of significant overrepresentation. Unfortunately, the overrepresentation of H3K4me3-marked genes in green tissues cannot fully indicate whether such genes are significantly overrepresented in the gene clusters of interest. This is because the non-overlapping genes in the whole seedlings list are not the total H3K4me3-marked genes in the green tissues. Some genes are H3K4me3-marked in both roots and shoots. Removing all the overlapping H3K4me3-marked genes from the list of whole seedlings also removes many genes which are H3K4me3-marked in both of roots and shoots. The remaining whole seedling genes are the H3K4me3-marked genes in green tissues only. Using the list of H3K4me3-marked genes in green tissue only as a biological function class reduces the size of the class substantially, resulting in a reduced expected ratio of H3K4m3- marked genes to the population. The hypergeometric test may now show a significant overrepresentation of H3K4me3-marked genes. It is more convincing if the H3K4me3-marked genes are significantly overrepresented in both whole seedlings and green tissues. Another disadvantage of using the H3K4me3-marked gene list from the Roudier's data is that it is not certain that all the *jmj14* regulated H3K4me3-marked genes are present on the lists. Because JMJ14 is predicted to remove the H3K4me3 marks, WT shows the H3K4me3 chromatin marking state when the markers have already been removed. *jmj14*

mutants would fail to remove the marks, so the H3K4me3 chromatin marking state of *jmj14* may contain additional genes. Even the genes of cluster 5 show no overrepresentation of H3K4me3-marked genes, but they still have the possibility to be regulated by *jmj14*. In summary, the association between H3K4me3-marked genes and *jmj14*-regulated gene clusters 1.1 and 1.10 indicates to some extent that *jmj14*-dependent genes can be modified by H3K4me3 demethylation.

Surprisingly, the presence of *HO3* and *HEMA3* in cluster 5 was observed when *jmj14*-dependent genes of each cluster of interest were ranked by fold change and considered for possible ChIP assay. The elevated expression of *HO3* in particular in *jmj14* mutants is very relevant to the initial hypothesis by which *jmj14* may rescue light responses by directly regulating chloroplast development, as *HO3* has a redundant function with *HO1* (Davis, 2001; Emborg et al., 2006). This clearly requires further study.

One further issue deserves consideration. The “hormone and defence” functional category is overrepresented in both *lyn1*-dependent clusters 1.1 and 1.10. Recent studies reveal that chromatin modification plays a role in plant defence responses. Plants have a potential ability to resist biotic and abiotic stress by activating defence genes in a stereotypic defence response (Boller and Felix, 2009). It has been observed that defence genes become primed after plants undergo stress, in order to respond to a subsequent stress, even a low level one, with a more rapid and robust activation of defence response or stress tolerance. Histone modification has been found present after stress, and has been predicted to act as a memory for defence response in *Arabidopsis* (van den Burg and Takken, 2009). The mark poises the expression of defence gene from enhanced activation by gene priming. In particular H3K4 methylation has been found to strongly correlate with the occurrence of priming (Conrath, 2011; Jaskiewicz et al., 2011). Stress induces a memory of the primary infection on defence genes by H3K4 methylation, and as a result the expression of these genes can be

amplified during a second stress stimulus. Therefore, the overrepresentation of genes which represent “hormone and defence” function in *lyn1* up-regulated clusters 1.1 and 1.10 means they are probably involved in defence priming by H3K4 trimethylation. Given that mutants that are attenuated in pathogen defence are often compromised in gene priming (Kohler et al., 2002; Durrant and Dong, 2004), *jmj14* is very likely to play a role in defence gene priming in *Arabidopsis*.

5.3.2 Association of *jmj14* regulation with early development

The big question of this project remains which biological process *jmj14* regulates to rescue the light response. As mentioned above, *jmj14* was suspected to enhance light response by increasing chloroplast development. Thus, the study was forced on finding the *lyn1*-increased chloroplast genes expression at the beginning. However, all the *lyn1* up-regulated gene clusters (5, 1.1 and 1.10) do not show a significant overrepresentation on any biological function related to chloroplast biogenesis except the overrepresentation of metabolism in subcluster 1.10 (Figure 5.7). According to the functional gene classification results and the information contained in the existing literature, another hypothesis emerged: *jmj14* may regulate genes involved in (or at least active during) cell entry into differentiation at an early developmental stage; this would lead to an increased chloroplast development, resulting in an enhanced light response.

Fortunately, Schmid and collaborators studied how transcriptional programmes control *Arabidopsis* development by analysing global gene expression during development at multiple stages (Schmid et al., 2005). These data were manually processed to obtain several lists of genes which were actively expressed during cell entry into differentiation. The lists selected for genes which are active at the shoot apical region or in early leaf developmental stages. These lists of genes are defined as the representative genes of different stages of development. By analysing the overrepresentation of these lists of genes on subclusters of cluster

1, the overrepresentations of genes particularly active at the shoot apex (6/7 and 6/5) appear on the *lyn1*-dependent subclusters 1.1 and 1.10 (Figure 5.9). Analysing the rest of subclusters (1.2~1.9) suggests that those overrepresentations are not just a general consequence of the response to light, rather the rest of subclusters include a disproportionate number of genes selected as expressed specifically in young leaves (17/13 and 17/12). Sample 12, 13 and 17 are collected from leaves of 17-day old plants. Although 17/13 and 17/12 represent different development stages (17/13 represents the ongoing development vs late development and 17/12 represents the ongoing development vs completed development), they probably have already passed the point of observation during which the influence of *lyn1* regulating gene expression changes is apparent. Additionally, the expression of chloroplast-associated genes is not significantly elevated during cell entry into differentiation in 7-day old seedlings and increased later in 17-day old plants. This also suggests that *jmj14* does not rescue light response by regulating chloroplast development directly. In summary, the association of the *jmj14*-regulated genes with the early development (activity in the tissues in the close vicinity of the shoot meristem) suggests that light-dependent *jmj14*-dependent genes are specifically enriched in meristem or primordia-active genes and may therefore be involved in cell entry into differentiation at an early developmental stage.

My analysis suggests that it may not be easy to identify individual genes making a major contribution to explaining the *lyn1* phenotype. However some potentially significant genes can be observed. "Cell cycle" was an overrepresented function in cluster 5. By reviewing the cell cycle genes present in cluster 5 (genes dependent only on *lyn1*) it is difficult to associate their functions with the impact of *jmj14*, because only 11 cell cycle genes are present in this cluster. One among them which may be of particular interest is *CDC48B*, the Arabidopsis homologue of *CDC48*. *Cell division cycle 48 (CDC48)* was first identified in yeast (Moir et al.,

1982). *cdc48* mutation was found to cause an apparent block of nuclear division in yeast (Moir et al., 1982). CDC48 was also found to be involved in cell division and growth during cell proliferation processes particularly at low temperature in zebrafish embryonic cells (Imamura et al., 2012). Rapid cell division is often the first sign of exit from the stem cell state, and conversion into a progenitor cell. In Arabidopsis, cell division rate is limited by CLAVATA3 (CLV3) in the central zone of the meristem (Clark et al., 1995). When cells exit stemcell-ness and enter a differentiation stage, cells move to primordia. Cells in primordia can be argued to be progenitor cells. The cell division rate is no longer repressed by CLV3, so cells rapidly divide and expand. Evidence showed that *clv3* mutant has enlarged meristems (Grandjean et al., 2004). Therefore, the expression of genes which regulate cell division rate is very likely to be increased as cells initiate a differentiation stage. Because *jmj14* is predicted to be involved in cell entry into differentiation, the simultaneously increased expression of genes which associate with cell proliferation would be explained.

5.3.3 Association of H3K4me3 demethylation with early development

Through the study of *jmj14*, the H3K4me3 demethylation and early development are indirectly connected, assuming that the function of JMJ14 is indeed that in its genomic annotation. Whether H3K4me3 demethylation and early development can be directly connected, a prediction of such a function of JMJ14, was also analyzed. The H3K4me3-marked genes in whole seedlings and green tissues are both significantly overrepresented among all the lists of genes which active during cell entry into differentiation in different early developmental stages (Figure 5.18). This strongly suggests that the genes which are active during cell entry into early differentiation at early development is very likely to be regulated by H3K4me3 demethylation.

Because there is not even a single chloroplast associated gene is present in neither *jmj14*-dependent cluster of interest (5, 1.1 and 1.10) nor the genes which are involved in cell entry into differentiation at the early developmental stage of 6/7 and 6/5, the initial hypothesis which JMJ14 directly regulates chloroplast biogenesis can be rule out. Before this hypothesis was completely rejected, the overrepresentations of H3K4me3-marked genes among global chloroplast development-associated genes and chloroplast translation genes were analysed (Figure 5.14). The overrepresentation among H3K4me3-marked genes of the chloroplast development (global) functional gene class suggests that such genes are disproportionately marked, and therefore may be regulated by H3K4me3 demethylation. The fact that H3K4me3-marked genes in green tissue only show no overrepresentation of either global chloroplast development-associated genes or chloroplast translation genes is probably a result of the fact that active plastid development is as important during root as it is during shoot development.

5.3.4 Future work

Before the RNA-Seq experiment was carried out, expression of several chloroplast-associated nuclear genes was quantified by qPCR for finding out if any chloroplast development TFs are *lyn1*-regulated. After the global gene expression analysis and association with published literature, I realized that the study should particularly focus on the H3K4me3 modified genes which are also involved in cell entry into differentiation at the early developmental stages from cluster 1.1, 1.10 and 5. The ideal follow-up experiment, instead of examining expression of individual genes by qPCR, should be the ChIP followed by qPCR. ChIP antibodies are available which will bind and help select modified histones carrying the mark of interest, together (if crosslinking has been carried out) with the genomic region they are associated with. Antibodies selecting for H3K4me3 are expected to precipitate a greater number of copies of the target genes from *jmj14* than from WT chromatin. However, JMJ14 is histone demethylase, the

histone methylation state associated with the expression of genes reported by Schmid et al. (2005) corresponds to the WT, in which methyl groups may have already been removed from the histone. This means some genes whose expression can be modified by H3K4me3 are missed. If the histone methylation state is tested in *jmj14*, more genes should appear within the list. Therefore, one potential further experimental approach would be to examine the H3K4me3-marked genes in samples dissected from shoot apices, young leaf primordia and cotyledons as carried out by Schmid et al. (2005), but this time comparing mutant and WT.

5.4 Conclusion

Through studies on two existing publications, the mechanism through which *jmj14* rescues the light response defect has been, to some extent, revealed. In summary, *jmj14* causes a defective H3K4me3 demethylation of target nuclear-encoded genes which in turn keeps these genes active beyond their level when JMJ14 is functional. These *jmj14* activated genes are fundamental genes involved in cell entry into differentiation at the early developmental stages, among which are those involved in ribosome/translation, transcription, cell cycle and metabolism, but not directly involved in chloroplast development. The overall increased development also encompasses the development of plastids, not only in the light but also in the dark. Although these genes are also light-induced genes, the increased expression of these genes is not due to the light signal directly. The light signal is the secondary factor that enlarges the initial light response compensation from the enhanced chloroplast development in the differentiation stages, resulting in a major rescue of the light response in the *hy1* mutant.

Chapter 6 General Discussion

6.1 How does *lyn1* rescue the light response defect?

6.1.1 Is enhanced chloroplast development in *lyn1* a primary or secondary effect of the mutation?

Since *lyn1* was isolated, I have been assessing the fundamental impact of *jmj14* in rescuing the light response deficiency of *hy1* through diverse assays. The initial question was whether *lyn1* rescues light responses by (a) direct amplification of the weak light response in *hy1* or (b) due to *lyn1* primarily causing enhanced PΦB level or etioplast development, and subsequently enhancing chloroplast biogenesis, resulting in a rescued light response. Ultimately, this is to question whether the primary impact of *lyn1* is or is not based on an alteration of light signalling. Although *hy1* cannot produce a sufficient level of phys, a very small amount of phys is still synthesised. This has been supported by the observation of *hy1* has an intermediate phenotype between that of dark- and light-grown seedlings. This implies a leaky synthesis of haem in *hy1*, due to the fact that HO1 is not the only enzyme to catalyse the conversion of haem to biliverdin IX α . In *Arabidopsis*, the *HO* gene family contains 4 members, *HO1* to *HO4* (Davis, 2001; Emborg et al., 2006). It can also be categorized into 2 subfamilies, one including *HO1*-like genes and the other constituted by the *HO2* gene alone, according to their cellular functions and phylogenetic relationships (Davis, 2001; Emborg et al., 2006). Although the *HO2* gene product has been found to also affect biliverdin IX α synthesis and has a significant effect of photomorphogenesis (Emborg et al., 2006), *HO2* does not play the same role as *HO1* (Davis, 2001). The *HO2* cellular function is currently unknown. Its unique function, which is different to that of other HOs, could be determined by its poor solubility (Emborg et al., 2006). *HO1* has been found to be responsible for the majority of biliverdin IX α synthesis because

6.1.2 How does *lyn1* increase the extent of plastid development regardless of light?

Once the primary effect of *lyn1* had been identified, a second question immediately arises: how does *lyn1* enhance plastid development? According to the analysis of plastid to nuclear genome copy number ratio, the plastid compartment was increased by *lyn1* not only in the light but also in the dark. Furthermore, chloroplast compartment analysis suggests that *lyn1* increases the chloroplast compartment by increasing chloroplast size. I proposed that the *lyn1* mutation increases plastid development in the dark by enlarging plastid size, which ultimately requires enhanced or prolonged growth. The enhanced plastid development may, as a consequence, simultaneously enhance tetrapyrrole synthesis. Indeed, the assays quantifying tetrapyrrole synthesis support this hypothesis. However, I do not know whether *lyn1* primarily increases plastid development in general, which leads to an increased tetrapyrrole synthesis, or whether *lyn1* primarily regulates the expression of genes of the tetrapyrrole synthetic pathway-related enzymes, which in turn leads to an increased plastid development via retrograde signalling. In the second scenario, some enzymes, such as HO family enzymes, could possibly be affected by *jmj14*. I tried to coordinate all the evidence to explain the function of *lyn1*, but some contradictions remain between these conclusions and some current models of processes involved in chloroplast development, as will be described in the following section.

6.1.2.1 *lyn1* boosts general plastid biogenesis during cell entry into differentiation very early in development

Plant development, or the development of a plant organ such a leaf, is initiated by originally-identical cells acquiring individual identities, and them doing so while growing and at the right locations and times. Leaf development can be analysed

through this "cellular perspective" (Kalve et al., 2014). According to such a perspective, one can consider a cell as going through a series of consecutive stages. Initially, fully undifferentiated cells in the central zone of the shoot apical meristem undergo slow division. During this stage, expressed genes would primarily be considered housekeeping genes. In a second stage, cells exit stemcell-ness and enter a differentiation route, which initially involves rapid proliferation, to generate the number of future mesophyll cells (or those of other identities). Meanwhile, together with organelle growth, plastids rapidly proliferate to increase in number. Cells would, at this stage, be at the flanks of the shoot meristem and have entered leaf primordia. Although the leaf primordia can still be considered part of the meristem, cells are no longer considered to be stem cells. At this time, cells in the subepidermal layer of the *Arabidopsis* meristem contain abundant plastids but very little sign of photosynthetic chloroplast development (Charuvi et al., 2012). In a third stage, cells in those very early primordia initiate full differentiation, including greening and full chloroplast development. A green organ forms. It is at this stage that one would expect the most abundant expression of photosynthesis-associated genes. Finally, I could consider a fourth stage in which cells no longer build new organelles, and eventually enter senescence in order to recycle cellular materials. Such a sequence is approximately the one seen in a differentiating maize leaf. Although the first, "stem cell" stage is almost by definition not present (since a leaf is being analysed), stages two, three and four can be observed, involving the expression of plastid housekeeping genes, then that of photosynthesis-associated genes, then that of senescence-associated genes (Chotewutmontri and Barkan, 2016).

The analysis of the transcript profile changes in the *lyn1* mutant, together with that of presence of the associated histone modification in genes expressed at different stages of development, allow an overall hypothesis in which the mutation enhances the expression of genes active in the transition between stages 1 and

2. Such genes are active at the meristem but are associated with exit from stemcell-ness, and their action accompanies early plastid development. The most promising explanation of how *jmj14* rescues the light response defect is that *jmj14* increases the overall development and simultaneously encourages plastid development in these early developmental stages, which subsequently leads to a somewhat enhanced building of chloroplasts. This slightly enhanced chloroplast development can promote, via photoreceptor rescue, a slightly enhanced light response, which in turn elevates chloroplast biogenesis in a virtuous cycle, resulting in a substantially increased light response. This scenario is supported by the majority of assays, especially the global gene expression analysis. Notably, the increased tetrapyrrole synthesis (Pchl_{ide} and PΦB) can be explained as the unavoidable result of the generally boosted plastid biogenesis. In other words, *lyn1* enhances plastid biogenesis and simultaneously increases the capacity of the tetrapyrrole synthesis pathway. However here a difficulty arises. As *hy1* has a defect in the oxidative cleavage of haem, increasing the capacity of the tetrapyrrole synthesis pathway can lead to a greater accumulation of haem. Haem acts as a positive regulator of chloroplast development through its role in retrograde signalling (Woodson et al., 2011). Elevated haem would be the easiest explanation for the *gun* phenotype of the *hy1* mutant. Stronger retrograde signalling should lead to greater *PhANGs* gene expression even in *lyn1 hy1* under treatment with NF or other plastid-toxic drugs, or at least the same level of gene expression as that in *hy1*. However, the *gun* phenotype assays which were performed by the previous researchers contradict this prediction. *PhANGs* gene expression is elevated under NF in *hy1* relative to the wild type. In other words, *hy1* is a *gun* mutant. One previous study (as mentioned in chapter 1.4.4) showed *PhANGs* expression decreased in *lyn1 hy1* compared to *hy1*. This phenomenon was also supported by microscopy observations. The GFP reporter which is driven by *LHCB* promoter was present in the seedlings among which the *lyn1* mutant was originally identified, since it formed the basis for the genetic

screen. In the presence of NF, the fluorescence of the reporter was elevated in *hy1*, and reduced in WT and *lyn1 hy1* (López-Juez, personal communication). Because of the observation that *lyn1* eliminates the *gun* phenotype of *hy1*, according to these two assays, my initial hypothesis was that *lyn1* impacts plastid development directly and somehow increases the PΦB level.

6.1.2.2 *lyn1* eliminates accumulation of haem

Another scenario is my initial hypothesis that *lyn1* directly regulates nuclear-encoded chloroplast proteins, especially the chloroplast-associated TFs and the enzymes which mediate the tetrapyrrole synthesis pathway, and in turn increases plastid development. Although qPCR and RNA-Seq analyses do not show any enhanced chloroplast TFs, remarkably *HO3* is found among the genes of cluster 5 which are light-independent and *lyn1*-dependent, from the analysis of RNA-Seq data. *HO3* gene expression is significantly increased by the *lyn1* mutation compared to the control. This immediately raises the possibility that *lyn1* may increase HO3 production, which in turn unblocks the PΦB synthesis pathway. In this way, more PΦB is synthesised and less haem is accumulated, resulting in an enhanced phytochrome production and reduced retrograde signalling. The reduced accumulation of haem would reduce the expression of *PhANGs*. This is in accordance with the observation that *lyn1* eliminates the *gun* phenotype of *hy1* in the presence of NF. However, this observation has to be understood together with another one resulting from previous research (described in section 1.4.5). Previous researchers in my laboratory determined that PhyA apoprotein remained stable in the light, undegraded in the *lyn1 hy1* mutant. PΦB synthesised by HO3 and HO4 can bind multiple phytochrome apoproteins (Shekhawat and Verma, 2010). The continued presence of PhyA apoprotein in the cytoplasm in the light suggests that while some extra HO activity is probably present in *lyn1 hy1*, that amount is small. In addition, the *HO3* gene carries only H3K4me2 and not H3K4me3 modifications (according to data from Roudier et al., 2011), making

it less likely to be a direct target of JMJ14 action. Although *jmj14* is more efficient at H3K4me3 demethylation, it is still possible that *jmj14* regulates *HO3* by H3K4me2 demethylation. Alternatively, *lyn1* might carry at *HO3* a mark which is absent in the WT. Finally, *lyn1* might be a higher hierarchical regulator that indirectly regulates *HO3* expression. To further confirm whether *lyn1* increases PΦB level in *lyn1* mutants by increasing the *HO3* gene expression, my laboratory is currently generating a *lyn1 hy1 ho3* triple mutant for further research. Any extent of rescue of *hy1* by *lyn1* due to elevated *HO3*, should disappear in the triple mutant (i.e. *ho3* would reverse the suppression of *hy1* by *lyn1*).

Another attractive gene of interest whose expression is slightly elevated by the *lyn1* mutation is *HEMA3*. *HEMA3* is also found in the above-mentioned cluster 5. *HEMA1* and *HEMA2* are the two most common glutamyl-tRNA reductases that have been widely studied. As mentioned before, glutamyl-tRNA reductases are the rate-limiting enzymes in the initial step of the tetrapyrrole synthesis pathway. Increasing *HEMAs* can enhance tetrapyrrole synthesis by boosting upstream intermediates. *HEMA1* is expressed in all regions of the plant, but *HEMA2* is only expressed at low levels in roots and flowers (Kumar et al., 1996). Under normal conditions *HEMA3* expression is very low (Matsumoto et al., 2004), so *HEMA3* probably acts like *HO3* and *HO4*, to complement glutamyl-tRNA reductases' activity if the dominant *HEMA1* and *HEMA2* enzymes are not present. Interestingly, *HEMA3* is an H3K4me3-marked gene, hence potentially directly regulated by JMJ14.

In all truth, I cannot at present conclude whether or not the enhanced chloroplast development in *lyn1* is related to a reduced tendency to accumulate haem, because I do not understand whether, or how, haem accumulation impacts chloroplast development in *hy1*. On the one hand *hy1* exhibits reduced plastid development, even in the dark. On the other, in the presence of NF, it exhibits enhanced *PhANG* expression, beyond that observed in the wild type.

Understanding whether tetrapyrrole retrograde signals explain the phenotype of *lyn1* needs a better understanding of the role of tetrapyrrole retrograde signals in chloroplast development.

An investigation from my laboratory has shown that *lyn1* suppresses *cue8*, which is a chloroplast protein import defective mutant (Loudya and López-Juez, personal communication). It was found that *lyn1 cue8* has a significantly higher chlorophyll content than *cue8*. *CUE8* encodes TIC100, which is a subunit of a Translocon of the Inner membrane of Chloroplasts (TIC) complex. Because this complex has channel activity when it is reconstituted *in vitro*, it is proposed to be a general TIC translocon (Jarvis and López-Juez, 2013). Chloroplast protein import requires TIC and Translocon of the Outer membrane of Chloroplasts (TOC) components on the inner and outer membrane, respectively. Currently, some evidence indicates that photosynthetic proteins and non-photosynthetic (housekeeping) proteins are imported separately through different TOC complexes on the outer membrane. It is not clear whether these different types of proteins are imported into the stroma by sharing the same TIC complex channel or through a different channel on the inner membrane. If the imports of two types of proteins share the same TIC complex channel, then *lyn1* cannot simply rescue the chloroplast defect of *cue8* by importing more non-photosynthetic proteins. Therefore, HO3 and HEMA3 cannot be import into plastid and subsequently convert haem to Biliverdin IX α even the production of HO3 and HEMA3 is increased. If non-photosynthetic proteins are imported through a different TIC complexes channel which is established without the CUE8 subunit on the inner membrane, then *lyn1* mediating tetrapyrrole synthesis, by regulating *HO3* and *HEMA3* gene expression, becomes possible. The possibility of this modal gradually reduced during the progression of the research but is still under the consideration. The fact that *lyn1* is able to rescue a general chloroplast

defect in *cue8* becomes one of the stronger indications that the primary effect of the *lyn1* mutation is a primary boost of plastid biogenesis.

6.2 Importance of the analysis of epigenetic mechanisms beyond this study

Epigenetic mechanisms have been proposed to play crucial roles in a variety of organisms' development. Typically, they are important in embryonic stem cell differentiation and self-renewal and somatic cell reprogramming in mammals. The earliest development stage of multicellular organisms is a zygote, which is a totipotent stem cell containing a single genome and can give rise to hundreds of cell types (Mitalipov and Wolf, 2009). The inner cell mass of the blastocyst, which develops from the zygote, is made of embryonic stem cells (Evans and Kaufman, 1981; Thomson et al., 1998). Embryonic stem cells are pluripotent, having the capacity to generate all cell lineages of the developing and adult organism by differentiation (Gaspar-Maia et al., 2011). They are also self-renewing cells that have the ability to proliferate in the same state. A gene expression programme of embryonic stem cells responds to developmental cues to make the decision of remaining at a self-renew stage or entering into a differentiation stage (Young, 2011). Compared to multipotent somatic cells which usually generate cells of the tissue in which they are, embryonic cells have a more unrestricted differentiation capacity (Dzierzak, 2005; Kriegstein and Alvarez-Buylla, 2009). Although the differentiation capacity of the somatic cell is restricted, pluripotent stem cells can be technically induced from differentiated somatic cells to regain the pluripotency through somatic cell reprogramming (Jaenisch and Young, 2008). Studying of stem cell pluripotency is important for regenerative medicine, tissue engineering and disease modelling in the clinic.

Determination of cellular identity is dependent on the lineage-specific regulatory gene expression programme which is closely controlled by epigenetic marks. These epigenetic marks participate in the transcriptional activation, establishing a poised state for activation in response to development cues or keeping a silent state of the lineage-specific regulatory gene expression by chromatin modifications (Young, 2011). Chromatin exists in either heterochromatin or euchromatin form according to their differential compaction. Heterochromatin is intensely stained and more compact, a “closed” chromatin form, compared to euchromatin, an “open” form (Heitz, 1928). The ratio between euchromatin and heterochromatin being much higher in pluripotent stem cells than in differentiated cells reveals that chromatin is naturally “opened” in pluripotent stem cells (Meshorer and Misteli, 2006). In other words, chromatin is “closed” and more heterochromatin appears in differentiated cells. Epigenetic marks such as DNA methylation and histone modification are important factors to maintain chromatin state. DNA methylation is more stable and inheritable (Gaspar-Maia et al., 2011). DNA methylation at promoters has been shown to silence the expression of “lineage-specific regulatory genes” during cell differentiation (Bird, 2002). However, DNA methylation is probably dispensable for initial lineage specification in early embryos. Lacking all three DNA methyltransferases, mouse embryonic stem cells still can self-renew and can even begin to differentiate (Jackson et al., 2004; Tsumura et al., 2006). Like DNA methylation, histone modification is also important in regulating the expression of lineage-specific regulatory genes. The active histone marks, H3K9ac, H3K27ac, H4K5ac, H3K4me3 and H3K36me3, are highly correlated with pluripotency of undifferentiated cells (Cota et al., 2013; Li et al., 2015). In addition to the H3K4me3 histone mark which has been well described in this thesis, H3K36me3 is another interesting active histone methylation mark which is sharply enriched after promoter and consistently present along the transcribed region in mammals (Bannister et al., 2005; Barski et al., 2007). This indicates that, unlike H3K4me3 promoting the transcriptional

initiation, H3K36me3 is associated with transcriptional elongation. H3K9me3, H3K27me3 and H2AK119u (H2A-K119 monoubiquitination) are the most common of the well known repressive marks (Cota et al., 2013). H3K9me3 is located in the promoter region to repress the self-renewal and pluripotency of lineage-specific regulatory genes by silencing their expression during cell differentiation. This has been further supported by western blot and immunofluorescence analyses showing that H3K9me3 are enriched in heterochromatin (Meshorer et al., 2006). Although H3K27me3 and DNA methylation are both involved in the establishment and maintenance of gene silencing in the promoter region, H3K27me3 is more often recruited onto CG-rich promoters and DNA methylation preferentially occurs in the promoters with low CG density (Hagarman et al., 2013; Xie et al., 2013; Rose and Klose, 2014). Interestingly, functionally opposite histone marks can synchronously exist at the same promoter regions. The activating mark (H3K4me3) and repressive mark (H3K27me3) are both found on the bivalent domains to poise expression of lineage-specific regulatory genes. Bivalent histone modifications prevent premature differentiation through H3K27me3 repression and simultaneously keep genes poised for rapid expression upon differentiation by H3K4me3 activation (Bernstein et al., 2006; Pan et al., 2007).

6.3 Applications of this work

6.3.1 Silencing of *JMJ14* homologues in rice bundle sheath cells.

The study of *JMJ14* function has revealed that rather than acting as the chloroplast developmental-associated TFs (such as *HY5*) or their negative regulators (as *COP1*), which directly regulate chloroplast biogenesis, *JMJ14* can be determined as a distinctive chloroplast biogenesis regulator which acts through an alternative route to boost chloroplast development. As mentioned in chapter 1.3.1, *HY5* has been identified as a chloroplast biogenesis positive

regulator (Jiao et al., 2007; Li et al., 2011; Arsovski et al., 2012), whereas COP1, DET1 and DDB1 are identified as chloroplast biogenesis negative regulators (Jiao et al., 2007; Li et al., 2011). The genes encoding their homologues, *LeHY5*, *LeCOP1LIKE*, *HP2* (the homologue of *DET1*) and *HP1* (the homologue of *DDB1*) have been studied in tomato (Liu et al., 2004). All of them showed the same function, regulating light responses, as their homologues in *Arabidopsis*, respectively. Because HY5 is required for *LHCB* transcriptional expression in *Arabidopsis* (Lee et al., 2007; Waters et al., 2008), repressing *LeHY5* leads to normal plastid numbers but clear deficiencies in both abundance and organization of thylakoids in tomato. *hp1* mutation induces an increased plastid number and a typical thylakoid structure and grana stacking. Meanwhile, the study of Liu and collaborators (2004) found that repressing *LeHY5* reduces chlorophyll accumulation and repressing *LeCOP1LIKE* enhances chlorophyll accumulation in tomato leaves and fruits (Liu et al., 2004). This reveals that manipulation of light signal transduction can modulate activity of the chloroplast biogenesis pathway that may ultimately improve the efficiency of crop yields. However, HY5, COP1, DET1 and DDB1 all function in a same biological pathway. Overexpressing or repressing these chloroplast biogenesis regulators does not enhance chloroplast development to a dramatic extent, with the exception of *DET1*, but silencing of *DET1* has multiple other effects. In contrast, JMJ14 enhances chloroplast development through a completely different pathway. It would be interesting to investigate whether repressing JMJ14, while simultaneously enhancing *HY5* or repressing *COP1*, *DET1* and *DDB1*, can generate a dramatically increased chloroplast development mutant which may be worthy of study in other species for commercial utilization.

JMJ704 is the closest homologue of *JMJ14* in rice. There have been very few studies of *JMJ704* function. The most recent study found that *JMJ704* is a positive regulator of rice bacterial blight (BB) resistance (Hou et al., 2015). *Xanthomonas*

oryzae pv. oryzae (*Xoo*) is a BB pathogen which can cause serious disease in rice. *JMJ704* gene expression level is up-regulated after *Xoo* infection and reaches the maximum level after 24 hours. The *jmj704* knock-down mutant is more susceptible to *Xoo* compared to WT. The preliminary result from Hou et al., (2015) showed that global methylation levels of H3K4me2/3 increased in *jmj704* mutants. This, as expected, suggests that *JMJ704* is a potential H3K4me2/3 demethylase. Moreover, it was found that the expression of several negative master regulators of BB resistance were up-regulated in *jmj704* whereas H3K4me2/3 marks were enriched in those genes. Therefore, *JMJ704* is proposed to repress BB resistance negative regulators via H3K4me2/3 demethylation. If I want to enhance chloroplast development in rice by silencing *JMJ704*, an enhanced disease susceptibility may also occur.

As shown in chapter 4.1.1, chloroplast development is almost saturated in mesophyll cells, so it is difficult to further increase the chloroplast development in those cells. In contrast, bundle sheath cells have less developed chloroplast compared to mesophyll cells, so there is a greater possibility to increase the chloroplast development in bundle sheath cells. My lab collaborators who are working on the C4 rice project are interested in using this gene in their study. The majority of plants, including *Arabidopsis* and crop plants such as rice, are C3 plants. Compared to C3 plants, C4 plants carry out more efficient photosynthesis, specifically faster carbon fixation because they almost completely avoid photorespiration. My lab collaborators are developing a “cutting-edge” rice for the future by converting C3 rice to C4 rice and resulting in a better yielding C4 rice (Hibberd et al., 2008). Additionally, they are using different strategies to further increase the photosynthesis in C4 rice. *jmj14* has shown that it directly or indirectly up-regulates chloroplast development, which leads to an enhanced greening in *Arabidopsis*. According to the alignment of *JMJ14* homologues, *OsJMj704* has the closest phylogenetic relationship in rice with *AtJMj14*. The

Osjmj704 mutation may also enhance chloroplast development and photosynthesis in the rice bundle sheath cells. Studying *OsJMJ704* may help the C4 rice project to roll into the next phase. Clearly the development of converting C3 rice into C4 rice and enhancing photosynthesis will surely lead to considerable economic benefit to the grower.

6.3.2 Stress tolerance and *JMJ14* homologues

Since *JMJ14* shows a close phylogenetic relationship with *JMJ15* and *JMJ18*, *JMJ14* may regulate the targets of the other paralogues. There is one study which focused on the stress tolerance of *JMJ15*-overexpressing plants (Shen et al., 2014). The authors used the Affymetrix *Arabidopsis* ATH1 Genome Array to quantify the gene expression in those plants. They found that a third of the down-regulated genes are stress-responsive signalling protein genes when *JMJ15* is overexpressed. They confirmed that 85% of down-regulated genes are H3K4me2/3 marked genes. Therefore, *JMJ15* was proposed to down-regulate stress-responsive signalling protein genes by H3K4me2/3 demethylation. Additionally, the authors found that half of the *JMJ15* down-regulated and H3K4me2/3 marked genes are TFs. Thus, *JMJ15* was proposed to act as a higher hierarchical regulator. They particularly assessed the salt stress tolerance of overexpressed *JMJ15* plants and *jmj15* loss-of-function mutants by growing these two types of mutants on a $\frac{1}{2}$ MS media containing NaCl. The overexpressed *JMJ15* plants have an enhanced plant tolerance to salt stress. In contrast, the loss-of-function *jmj15* mutants have a reduced stress resistance.

In the study of rice *JMJ704* regulating BB resistance, the GO analysis of significantly expressed genes between rice *jmj704* mutant and WT also reveals that “response to stress” was significantly enriched (Shen et al., 2014). As *OsJMJ704* and *AtJMJ15* are the closest homologues of *JMJ14* in rice and *Arabidopsis*, respectively, *jmj14* may have a similar function of mediating stress

response and regulating TFs. Indeed, the GO term of total significantly expressed genes from the RNA-Seq analysis has shown that the categories of “response to stress” were significantly enriched by the *jmj14* mutation. This suggests that *jmj14* might be involved in regulating stress response. Probably, *jmj14* is also a higher hierarchical regulator that enhances light response by regulating some TFs which might not have been detected by the RNA-Seq data analysis. Although this project is not focused on the stress response of *jmj14*, this could be further researched as a separate project.

6.4 Conclusion

The work shown in this thesis is a project continuing from the work of previous researchers and my MSc study, and which focused on the *lyn1* impact on chloroplast development. Through the work in this project, the conclusions achieved, and the further studies of potential promise are summarized below:

1. The *LYN1* gene was confirmed to be *JMJ14*. *JMJ14* has been studied with regards to controlling flowering time and mediating gene silencing by other research groups, but none of those studies had noticed a function in regulating chloroplast development. Fortunately, my laboratory identified *lyn1* by screening a suppressor of a light response defective mutant.
2. More studies were carried out to assess the *lyn1* suppression on *hy1* and the function of *LYN1*. Interestingly, *lyn1* was found to boost plastid development by increasing the plastid compartment and the tetrapyrrole synthesis pathway regardless of light.
3. The *jmj14*-regulated cellular function was investigated by global gene expression analysis. The most likely scenario is that *jmj14* enhances plastid development by elevating meristem or primordia-active genes which modulate the transition towards cell differentiation in the early development stages.

4. The histone marking state of these meristem or primordia-active genes can be further studied in future by CHIP and expression of these genes can be quantified by CHIP-qPCR.
5. The study of *JMJ14* homology helped me to consider other potential functions of *JMJ14*. Among the *JMJ14* homologues, *OsJMJ704* would be the primary candidate to have conserved the same function which is regulating chloroplast development in rice. The study of *OSJMJ704* in the C4 rice project may make a great contribution for future life on planet Earth.

References

- APEL, K. (1981) The Protochlorophyllide Holochrome of Barley (*Hordeum Vulgare* L.): Phytochrome-Induced Decrease of Translatable mRNA Coding for the NADPH: Protochlorophyllide Oxidoreductase. *European Journal of Biochemistry*, **120**, 89–93.
- Arsovski, A.A., Galstyan, A., Guseman, J.M. & Nemhauser, J.L. (2012) Photomorphogenesis. *The Arabidopsis Book*, **10**, e0147.
- Balciunas, D. & Ronne, H. (2000) Evidence of domain swapping within the jumonji family of transcription factors. *Trends in Biochemical Sciences*, **25**, 274–276.
- Bannister, A.J., Schneider, R., Myers, F.A., Thorne, A.W., Crane-Robinson, C. & Kouzarides, T. (2005) Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *Journal of Biological Chemistry*, **280**, 17732–17736.
- Barnes, S., Nishizawa, N., Quaggio, R., Whitelam, G. & Chua, N. (1996) Far-Red Light Blocks Greening of Arabidopsis Seedlings via a Phytochrome A-Mediated Change in Plastid Development. *THE PLANT CELL ONLINE*, **8**, 601–615.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. & Zhao, K. (2007) High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell*, **129**, 823–837.
- Barton, K.A., Schattat, M.H., Jakob, T., Hause, G., Wilhelm, C., Mckenna, J.F., Máthé, C., Runions, J., Van Damme, D. & Mathur, J. (2016) Epidermal

Pavement Cells of Arabidopsis Have Chloroplasts. *Plant Physiology*, **171**, 723 LP-726.

Van Bel, M., Diels, T., Vancaester, E., Kreft, L., Botzki, A., Van de Peer, Y., Coppens, F. & Vandepoele, K. (2017) PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Research*.

Benevolenskaya, E. V., Murray, H.L., Branton, P., Young, R.A. & Kaelin, W.G. (2005) Binding of pRB to the PHD protein RBP2 promotes cellular differentiation. *Molecular Cell*, **18**, 623–635.

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. & Sayers, E.W. (2012) GenBank. *Nucleic Acids Research*, **41**, D36–D42.

Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. & Huala, E. (2015) The arabidopsis information resource: Making and mining the “gold standard” annotated reference plant genome. *genesis*, **53**, 474–485.

Bergé-Lefranc, J., Jay, P., Massacrier, A., Cau, P., Mattei, M., Bauer, S., Marsollier, C., Berta, P. & Fontes, M. (1996) Characterization of the human jumonji gene. *Human Molecular Genetics*, **5**, 1637–1641.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S.L. & Lander, E.S. (2006) A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell*, **125**, 315–326.

Berr, A., McCallum, E.J., Ménard, R., Meyer, D., Fuchs, J., Dong, A. & Shen, W.-H. (2010) Arabidopsis SET DOMAIN GROUP2 Is Required for H3K4

Trimethylation and Is Crucial for Both Sporophyte and Gametophyte Development. *The Plant Cell*, **22**, 3232–3248.

Biehl, A., Richly, E., Noutsos, C., Salamini, F. & Leister, D. (2005) Analysis of 101 nuclear transcriptomes reveals 23 distinct regulons and their relationship to metabolism, chromosomal gene distribution and co-ordination of nuclear and plastid gene expression. *Gene*, **344**, 33–41.

Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev*, **16**, 6–21.

Black, J., Van Rechem, C. & Whetstine, J. (2013) Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact. **48**, 1–32.

Black, J.C., Van Rechem, C. & Whetstine, J.R. (2012) Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact. *Molecular Cell*, **48**, 491–507.

Blair, L.P., Cao, J., Zou, M.R., Sayegh, J. & Yan, Q. (2011) Epigenetic regulation by lysine demethylase 5 (KDM5) enzymes in cancer. *Cancers*, **3**, 1383–1404.

Blankenship, R.E. (2001) Molecular evidence for the evolution of photosynthesis. *Trends in Plant Science*, **6**, 4–6.

Boller, T. & Felix, G. (2009) A Renaissance of Elicitors: Perception of Microbe-Associated Molecular Patterns and Danger Signals by Pattern-Recognition Receptors. *Annual Review of Plant Biology*, **60**, 379–406.

Briggs, W.R. & Christie, J.M. (2002) Phototropins 1 and 2: versatile plant blue-light receptors. *Trends in Plant Science*, **7**, 204–210.

- Bruggemann, W., Handwerger, K., Essex, C. & Storz, G. (1996) Analysis of fast neutron-generated mutants at the *Arabidopsis thaliana* HY4 locus. *Plant Journal*, **10**, 755–760.
- van den Burg, H.A. & Takken, F.L.W. (2009) Does chromatin remodeling mark systemic acquired resistance? *Trends in Plant Science*, **14**, 286–294.
- Cartagena, J.A., Matsunaga, S., Seki, M., Kurihara, D., Yokoyama, M., Shinozaki, K., Fujimoto, S., Azumi, Y., Uchiyama, S. & Fukui, K. (2008) The *Arabidopsis* SDG4 contributes to the regulation of pollen tube growth by methylation of histone H3 lysines 4 and 36 in mature pollen. *Developmental Biology*, **315**, 355–368.
- Casal, J.J. & Boccalandro, H. (1995) Co-action between phytochrome B and HY4 in *Arabidopsis thaliana*. *Planta*, **197**, 213–218.
- Casal, J.J., Candia, A.N. & Sellaro, R. (2014) Light perception and signalling by phytochrome A. *Journal of Experimental Botany*, **65**, 2835–2845.
- Cellot, S., Hope, K.J., Chagraoui, J., Sauvageau, M., Deneault, É., MacRae, T., Mayotte, N., Wilhelm, B.T., Landry, J.R., Ting, S.B., Krosli, J., Humphries, K., Thompson, A. & Sauvageau, G. (2013) RNAi screen identifies Jarid1b as a major regulator of mouse HSC activity. *Blood*, **122**, 1545–1555.
- Charuvi, D., Kiss, V., Nevo, R., Shimoni, E., Adam, Z. & Reich, Z. (2012) Gain and Loss of Photosynthetic Membranes during Plastid Differentiation in the Shoot Apex of *Arabidopsis*. *The Plant Cell*, **24**, 1143–1157.
- Cheminant, S., Wild, M., Bouvier, F., Pelletier, S., Renou, J.-P., Erhardt, M., Hayes, S., Terry, M.J., Genschik, P. & Achard, P. (2011) DELLAs Regulate Chlorophyll and Carotenoid Biosynthesis to Prevent Photooxidative

Damage during Seedling Deetiolation in *Arabidopsis*. *The Plant Cell*, **23**, 1849–1860.

Chen, H. (2006) Arabidopsis CULLIN4 Forms an E3 Ubiquitin Ligase with RBX1 and the CDD Complex in Mediating Light Control of Development. *THE PLANT CELL ONLINE*, **18**, 1991–2004.

Cheng, C.Y., Krishnakumar, V., Chan, A.P., Thibaud-Nissen, F., Schobel, S. & Town, C.D. (2017) Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *Plant Journal*, **89**, 789–804.

Chiang, Y.-H., Zubo, Y.O., Tapken, W., Kim, H.J., Lavanway, A.M., Howard, L., Pilon, M., Kieber, J.J. & Schaller, G.E. (2012) Functional Characterization of the GATA Transcription Factors GNC and CGA1 Reveals Their Key Role in Chloroplast Development, Growth, and Division in Arabidopsis. *PLANT PHYSIOLOGY*, **160**, 332–348.

Chotewutmontri, P. & Barkan, A. (2016) Dynamics of Chloroplast Translation during Chloroplast Differentiation in Maize. *PLoS Genetics*, **12**.

Clark, S., Running, M. & Meyerowitz, E. (1995) *CLAVATA3* is a specific regulator of shoot and floral meristem development affecting the same processes as *CLAVATA1*. *Development*, **121**, 2057–2067.

Colombo, M., Tadini, L., Peracchio, C., Ferrari, R. & Pesaresi, P. (2016) GUN1, a Jack-Of-All-Trades in Chloroplast Protein Homeostasis and Signaling. *Frontiers in Plant Science*, **7**.

Conrath, U. (2011) Molecular aspects of defence priming. *Trends in Plant Science*, **16**, 524–531.

- Cota, P., Shafa, M. & E., D. (2013) *Stem Cells and Epigenetic Reprogramming. Pluripotent Stem Cells* (ed. by D. Bhartiya) and N.B.T.-P.S.C. Lenka), p. Ch. 09. InTech, Rijeka.
- Dandekar, A.M., McGranahan, G.H., Leslie, C.A. & Uratsu, S.L. (1989) Agrobacterium-mediated transformation of somatic embryos as a method for the production of transgenic plants. *Journal of tissue culture methods*, **12**, 145–150.
- Dao, T.T.H., Linthorst, H.J.M. & Verpoorte, R. (2011) Chalcone synthase and its functions in plant resistance. *Phytochemistry Reviews*, **10**, 397–412.
- Davis, S.J. (2001) The Heme-Oxygenase Family Required for Phytochrome Chromophore Biosynthesis Is Necessary for Proper Photomorphogenesis in Higher Plants. *PLANT PHYSIOLOGY*, **126**, 656–669.
- Davis, S.J., Kurepa, J. & Vierstra, R.D. (1999) The Arabidopsis thaliana HY1 locus, required for phytochrome-chromophore biosynthesis, encodes a protein related to heme oxygenases. *Proceedings of the National Academy of Sciences*, **96**, 6541–6546.
- Debrieux, D., Trevisan, M. & Fankhauser, C. (2013) Conditional Involvement of CONSTITUTIVE PHOTOMORPHOGENIC1 in the Degradation of Phytochrome A. *PLANT PHYSIOLOGY*, **161**, 2136–2145.
- Devi, M.T., Rajagopalan, A. V. & Raghavendra, A.S. (1995) Predominant localization of mitochondria enriched with glycine-decarboxylating enzymes in bundle sheath cells of *Alternanthera tenella*, a C3–C4 intermediate species. *Plant, Cell & Environment*, **18**, 589–594.
- Dey, B.K., Stalker, L., Schnerch, A., Bhatia, M., Taylor-Papadimitriou, J. & Wynder, C. (2008) The Histone Demethylase KDM5b/JARID1b Plays a

- Role in Cell Fate Decisions by Blocking Terminal Differentiation. *Molecular and Cellular Biology*, **28**, 5312–5327.
- Dieterle, M., Buche, C., Schafer, E. & Kretsch, T. (2003) Characterization of a Novel Non-Constitutive Photomorphogenic cop1 Allele1. *Plant Physiology*, **133**, 1557–1564.
- Dong, X. & Weng, Z. (2013) The correlation between histone modifications and gene expression. *Epigenomics*, **5**, 113–116.
- Durrant, W.E. & Dong, X. (2004) Systemic acquired resistance. *Annual review of phytopathology*, **42**, 185–209.
- Dyall, S.D., Brown, M. & Johnson, P. (2004) Ancient Invasions: From Endosymbionts to Organelles. *Science*, **304**, 253–257.
- Dzierzak, E. (2005) The emergence of definitive hematopoietic stem cells in the mammal. *Current Opinion in Hematology*, **12**, 197–202.
- Edwards, G. & Walker, D. (1983) *C3, C4: mechanisms, and cellular and environmental regulation, of photosynthesis*, Univ of California Press.
- Edwards, K., Johnstone, C. & Thompson, C. (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*, **19**, 1349.
- Egea, I., Bian, W., Barsan, C., Jauneau, A., Pech, J.-C., Latché, A., Li, Z. & Chervin, C. (2011) Chloroplast to chromoplast transition in tomato fruit: spectral confocal microscopy analyses of carotenoids and chlorophylls in isolated plastids and time-lapse recording on intact live tissue. *Annals of Botany*, **108**, 291–297.

- Eickbush, T.H. & Moudrianakis, E.N. (1978) The histone core complex: an octamer assembled by two sets of protein-protein interactions. *Biochemistry*, **17**, 4955–4964.
- Emborg, T.J., Walker, J.M., Noh, B. & Vierstra, R.D. (2006) Multiple Heme Oxygenase Family Members Contribute to the Biosynthesis of the Phytochrome Chromophore in Arabidopsis. *PLANT PHYSIOLOGY*, **140**, 856–868.
- Enfissi, E.M.A., Barneche, F., Ahmed, I., Lichtlé, C., Gerrish, C., McQuinn, R.P., Giovannoni, J.J., Lopez-Juez, E., Bowler, C., Bramley, P.M. & Fraser, P.D. (2010) Integrative Transcript and Metabolite Analysis of Nutritionally Enhanced DE-ETIOLATED1 Downregulated Tomato Fruit. *The Plant Cell*, **22**, 1190–1215.
- Evans, M.J. & Kaufman, M.H. (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature*, **292**, 154–156.
- Fankhauser, C. & Casal, J.J. (2004) Phenotypic characterization of a photomorphogenic mutant. *The Plant Journal*, **39**, 747–760.
- Fankhauser, C. & Chen, M. (2008) Transposing phytochrome into the nucleus. *Trends in Plant Science*, **13**, 596–601.
- Feng, J. & Shen, W.-H. (2014) Dynamic regulation and function of histone monoubiquitination in plants. *Frontiers in Plant Science*, **5**.
- Fernando, V.C.D. & Schroeder, D.F. (2016) Shedding light on plant development: light signalling in the model plant *Arabidopsis thaliana*. *Ceylon Journal of Science*, **45**, 3.

- Fisher, A.J. & Franklin, K.A. (2011) Chromatin remodelling in plant light signalling. *Physiologia Plantarum*, **142**, 305–313.
- Fitter, D.W., Martin, D.J., Copley, M.J., Scotland, R.W. & Langdale, J.A. (2002) GLK gene pairs regulate chloroplast development in diverse plant species. *Plant Journal*, **31**, 713–727.
- Fornieris, F., Binda, C., Adamo, A., Battaglioli, E. & Mattevi, A. (2007) Structural Basis of LSD1-CoREST Selectivity in Histone H3 Recognition. *Journal of Biological Chemistry*, **282**, 20070–20074.
- Fujii, S., Kobayashi, K., Nagata, N., Masuda, T. & Wada, H. (2017) Monogalactosyldiacylglycerol Facilitates Synthesis of Photoactive Protochlorophyllide in Etioplasts. *Plant Physiology*, **174**, 2183–2198.
- Gaspar-Maia, A., Alajem, A., Meshorer, E. & Ramalho-Santos, M. (2011) Open chromatin in pluripotency and reprogramming. *Nature Reviews Molecular Cell Biology*, **12**, 36–47.
- Genoud, T., Schweizer, F., Tscheuschler, A., Debrieux, D., Casal, J.J., Schäfer, E., Hiltbrunner, A. & Fankhauser, C. (2008) FHY1 Mediates Nuclear Import of the Light-Activated Phytochrome A Photoreceptor. *PLoS Genetics*, **4**, e1000143.
- Gil, P., Kircher, S., Adam, E., Bury, E., Kozma-Bognar, L., Schafer, E. & Nagy, F. (2000) Photocontrol of subcellular partitioning of phytochrome-B:GFP fusion protein in tobacco seedlings. *The Plant Journal*, **22**, 135–145.
- Gollnick, K. (1968) Type II photooxygenation reactions in solution. *Adv Photochem.*

- Goodstadt, L. & Ponting, C.P. (2001) CHROMA: consensus-based colouring of multiple alignments for publication. *Bioinformatics*, **17**, 845–846.
- Gorman, a a & Rodgers, M. a (1992) Current perspectives of singlet oxygen detection in biological environments. *Journal of photochemistry and photobiology. B, Biology*, **14**, 159–176.
- Graham, B. & Karen, L. (2016) *Nutrition, Epigenetics and Health*, World Scientific.
- Grandjean, O., Vernoux, T. & Laufs, P. (2004) In vivo analysis of cell division, cell growth, and differentiation at the shoot apical meristem in Arabidopsis. *The Plant Cell ...*, **16**, 74–87.
- Gray, J.C., Sullivan, J.A., Wang, J.-H., Jerome, C.A. & MacLean, D. (2003) Coordination of plastid and nuclear gene expression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **358**, 135–145.
- Guo, L., Yu, Y., Law, J.A. & Zhang, X. (2010) SET DOMAIN GROUP2 is the major histone H3 lysine 4 trimethyltransferase in Arabidopsis. *Proceedings of the National Academy of Sciences*, **107**, 18557–18562.
- Guo, L., Zhou, J., Elling, A.A., Charron, J.-B.F. & Deng, X.W. (2008) Histone Modifications and Expression of Light-Regulated Genes in Arabidopsis Are Cooperatively Influenced by Changing Light Conditions. *PLANT PHYSIOLOGY*, **147**, 2070–2083.
- Hagarman, J.A., Motley, M.P., Kristjansdottir, K. & Soloway, P.D. (2013) Coordinate Regulation of DNA Methylation and H3K27me3 in Mouse Embryonic Stem Cells. *PLOS ONE*, **8**, e53880.

- He, G., Elling, A.A. & Deng, X.W. (2011) The Epigenome and Plant Development. *Annual Review of Plant Biology*, **62**, 411–435.
- Heitz, E. (1928) Das Heterochromatin der Moose I. *Jahrb. wise. Bot.*, 69, 761-818. *Quoted by Callan.*
- Hibberd, J.M. & Quick, W.P. (2002) Characteristics of C4 photosynthesis in stems and petioles of C3 flowering plants. *Nature*, **415**, 451–454.
- Hibberd, J.M., Sheehy, J.E. & Langdale, J.A. (2008) Using C4 photosynthesis to increase the yield of rice-rationale and feasibility. *Current Opinion in Plant Biology*, **11**, 228–231.
- Higa, T., Suetsugu, N., Kong, S.-G. & Wada, M. (2014) Actin-dependent plastid movement is required for motive force generation in directional nuclear movement in plants. *Proceedings of the National Academy of Sciences*, **111**, 4327–4331.
- Hofgen, R. & Willmitzer, L. (1988) Storage of competent cells for *Agrobacterium* transformation. *Nucleic Acids Research*, **16**, 9877–9877.
- Hou, Y., Wang, L., Wang, L., Liu, L., Li, L., Sun, L., Rao, Q., Zhang, J. & Huang, S. (2015) JMJ704 positively regulates rice defense response against *Xanthomonas oryzae* pv. *oryzae* infection via reducing H3K4me2/3 associated with negative disease resistance regulators. *BMC Plant Biology*, **15**, 286.
- Howe, E.A., Sinha, R., Schlauch, D. & Quackenbush, J. (2011) RNA-Seq analysis in MeV. *Bioinformatics*, **27**, 3209–3210.
- Huang, F., Chandrasekharan, M.B., Chen, Y.-C., Bhaskara, S., Hiebert, S.W. & Sun, Z.-W. (2010) The JmjN Domain of Jhd2 Is Important for Its Protein

Stability, and the Plant Homeodomain (PHD) Finger Mediates Its Chromatin Association Independent of H3K4 Methylation. *Journal of Biological Chemistry*, **285**, 24548–24561.

Huson, D.H. & Bryant, D. (2006) Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution*, **23**, 254–267.

Imamura, S., Yabu, T. & Yamashita, M. (2012) Protective role of cell division cycle 48 (CDC48) protein against neurodegeneration via ubiquitin-proteasome system dysfunction during zebrafish development. *Journal of Biological Chemistry*, **287**, 23047–23056.

Inaba, T. & Ito-Inaba, Y. (2010) Versatile roles of plastids in plant growth and development. *Plant and Cell Physiology*, **51**, 1847–1853.

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J. & Ramsahoye, B. (2004) Severe Global DNA Hypomethylation Blocks Differentiation and Induces Histone Hyperacetylation in Embryonic Stem Cells. *Molecular and Cellular Biology*, **24**, 8862–8871.

Jaenisch, R. & Young, R. (2008) Stem Cells, the Molecular Circuitry of Pluripotency and Nuclear Reprogramming. *Cell*, **132**, 567–582.

Jang, I., Yang, J., Seo, H. & Chua, N. (2005) HFR1 is targeted by COP1 E3 ligase for post-translational proteolysis during phytochrome A signaling. *Genes & Development*, **19**, 593–602.

Jarvis, P. & López-Juez, E. (2013) Biogenesis and homeostasis of chloroplasts and other plastids. *Nature Reviews Molecular Cell Biology*, **14**, 787–802.

- Jaskiewicz, M., Conrath, U. & Peterhänsel, C. (2011) Chromatin modification acts as a memory for systemic acquired resistance in the plant stress response. *EMBO reports*, **12**, 50–55.
- Jeong, J.H., Song, H.R., Ko, J.H., Jeong, Y.M., Kwon, Y.E., Seol, J.H., Amasino, R.M., Noh, B. & Noh, Y.S. (2009) Repression of FLOWERING LOCUS T chromatin by functionally redundant histone H3 lysine 4 demethylases in Arabidopsis. *PLoS ONE*, **4**.
- Jiang, D., Wang, Y., Wang, Y. & He, Y. (2008) Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the Arabidopsis Polycomb Repressive Complex 2 Components. *PLoS ONE*, **3**, e3404.
- Jiao, Y., Lau, O.S. & Deng, X.W. (2007) Light-regulated transcriptional networks in higher plants. *Nature Reviews Genetics*, **8**, 217–230.
- Johansson, C., Tumber, A., Che, K., Cain, P., Nowak, R., Gileadi, C. & Oppermann, U. (2014) The roles of Jumonji-type oxygenases in human disease. *Epigenomics*, **6**, 89–120.
- Kakizaki, T., Matsumura, H., Nakayama, K., Che, F., Terauchi, R. & Inaba, T. (2009) Coordination of Plastid Protein Import and Nuclear Gene Expression by Plastid-to-Nucleus Retrograde Signaling. *PLANT PHYSIOLOGY*, **151**, 1339–1353.
- Kakizaki, T., Yazu, F., Nakayama, K., Ito-Inaba, Y. & Inaba, T. (2012) Plastid signalling under multiple conditions is accompanied by a common defect in RNA editing in plastids. *Journal of Experimental Botany*, **63**, 251–260.
- Kalve, S., De Vos, D. & Beemster, G.T.S. (2014) Leaf development: a cellular perspective. *Frontiers in Plant Science*, **5**.

- Kang, X. & Ni, M. (2006) Arabidopsis SHORT HYPOCOTYL UNDER BLUE1 contains SPX and EXS domains and acts in cryptochrome signaling. *The Plant cell*, **18**, 921–934.
- Karimi, M., Inzé, D. & Depicker, A. (2002) GATEWAY™ vectors for Agrobacterium-mediated plant transformation. *Trends in Plant Science*, **7**, 193–195.
- Kessler, F. & Schnell, D. (2009) Chloroplast biogenesis: diversity and regulation of the protein import apparatus. *Current Opinion in Cell Biology*, **21**, 494–500.
- Kevei, E., Schafer, E. & Nagy, F. (2007) Light-regulated nucleo-cytoplasmic partitioning of phytochromes. *Journal of Experimental Botany*, **58**, 3113–3124.
- Kidder, B.L., Hu, G. & Zhao, K. (2014) KDM5B focuses H3K4 methylation near promoters and enhancers during embryonic stem cell self-renewal and differentiation. *Genome Biology*, **15**, R32.
- Kinsman, E.A. a & Pyke, K.A.A. (1998) Bundle sheath cells and cell-specific plastid development in Arabidopsis leaves. *Development (Cambridge, England)*, **125**, 1815–1822.
- Kircher, S., Gil, P., Kozma-Bognár, L., Fejes, E., Speth, V., Husselstein-Muller, T., Bauer, D., Ádám, É., Schäfer, E. & Nagy, F. (2002) Nucleocytoplasmic Partitioning of the Plant Photoreceptors Phytochrome A, B, C, D, and E Is Regulated Differentially by Light and Exhibits a Diurnal Rhythm. *THE PLANT CELL ONLINE*, **14**, 1541–1555.
- Klose, R.J., Kallin, E.M. & Zhang, Y. (2006) JmjC-domain-containing proteins and histone demethylation. *Nature Reviews Genetics*, **7**, 715–727.

- Knoche, K. & Kephart, D. (1999) Cloning Blunt-End Pfu DNA Polymerase-Generated PCR Fragments into pGEM-T Vector Systems. *Promega Notes*, **71**, 10.
- Kobayashi, K., Baba, S., Obayashi, T., Sato, M., Toyooka, K., Keranen, M., Aro, E.-M., Fukaki, H., Ohta, H., Sugimoto, K. & Masuda, T. (2012a) Regulation of Root Greening by Light and Auxin/Cytokinin Signaling in Arabidopsis. *The Plant Cell*, **24**, 1081–1095.
- Kobayashi, K., Obayashi, T. & Masuda, T. (2012b) Role of the G-box element in regulation of chlorophyll biosynthesis in Arabidopsis roots. *Plant Signaling & Behavior*, **7**, 922–926.
- Kohchi, T., Mukougawa, K., Frankenberg, N., Masuda, M., Yokota, A. & Lagarias, J. (2001) The Arabidopsis HY2 Gene Encodes Phytochromobilin Synthase, a Ferredoxin-Dependent Biliverdin Reductase. *THE PLANT CELL ONLINE*, **13**, 425–436.
- Kohler, A., Schwindling, S. & Conrath, U. (2002) Benzothiadiazole-Induced Priming for Potentiated Responses to Pathogen Infection, Wounding, and Infiltration of Water into Leaves Requires the NPR1/NIM1 Gene in Arabidopsis. *Plant Physiology*, **128**, 1046–1056.
- Koussevitzky, S., Nott, A., Mockler, T.C., Hong, F., Sachetto-Martins, G., Surpin, M., Lim, J., Mittler, R. & Chory, J. (2007) Signals from chloroplasts converge to regulate nuclear gene expression. *Science*, **316**, 715–719.
- Kriegstein, A.R. & Alvarez-Buylla, A. (2009) The glial nature of embryonic and adult neural stem cells. *Annual review of neuroscience*, **32**, 149–184.

- Kumar, A.M., Csankovszki, G. & Soll, D. (1996) A second and differentially expressed glutamyl-tRNA reductase gene from *Arabidopsis thaliana*. *Plant Mol Biol*, **30**, 419–26.
- Kuźbicki, L., Lange, D., Strączyńska-Niemiec, A. & Chwirot, B.W. (2013) JARID1B expression in human melanoma and benign melanocytic skin lesions. *Melanoma research*, **23**, 8–12.
- Larkin, R., Alonso, J., Ecker, J. & Chory, J. (2003) GUN4, a Regulator of Chlorophyll Synthesis and Intracellular Signaling. *Science*, **299**, 902–906.
- Lau, O.S. & Deng, X.W. (2012) The photomorphogenic repressors COP1 and DET1: 20 years later. *Trends in Plant Science*, **17**, 584–593.
- Lee, J., He, K., Stolc, V., Lee, H., Figueroa, P., Gao, Y., Tongprasit, W., Zhao, H., Lee, I. & Deng, X.W. (2007) Analysis of Transcription Factor HY5 Genomic Binding Sites Revealed Its Hierarchical Role in Light Regulation of Development. *THE PLANT CELL ONLINE*, **19**, 731–749.
- Leister, D. & Kleine, T. (2016) Definition of a core module for the nuclear retrograde response to altered organellar gene expression identifies GLK overexpressors as gun mutants. *Physiologia Plantarum*, **157**, 297–309.
- Letunic, I., Doerks, T. & Bork, P. (2015) SMART: recent updates, new developments and status in 2015. *Nucleic Acids Research*, **43**, D257–D260.
- Li, J., Li, G., Gao, S., Martinez, C., He, G., Zhou, Z., Huang, X., Lee, J.-H., Zhang, H., Shen, Y., Wang, H. & Deng, X.W. (2010) Arabidopsis Transcription Factor ELONGATED HYPOCOTYL5 Plays a Role in the Feedback Regulation of Phytochrome A Signaling. *THE PLANT CELL ONLINE*, **22**, 3634–3649.

- Li, J., Li, G., Wang, H. & Wang Deng, X. (2011) Phytochrome Signaling Mechanisms. *The Arabidopsis Book*, **9**, e0148.
- Li, J., Terzaghi, W. & Deng, X.W. (2012) Genomic basis for light control of plant development. *Protein & Cell*, **3**, 106–116.
- Li, Y., Mukherjee, I., Thum, K.E., Tanurdzic, M., Katari, M.S., Obertello, M., Edwards, M.B., McCombie, W.R., Martienssen, R.A. & Coruzzi, G.M. (2015) The histone methyltransferase SDG8 mediates the epigenetic modification of light and carbon responsive genes in plants. *Genome Biology*, **16**.
- Lin, C. & Todo, T. (2005) The cryptochromes. *Genome Biology*, **6**, 220.
- Liu, C., Lu, F., Cui, X. & Cao, X. (2010) Histone Methylation in Higher Plants. *Annual Review of Plant Biology*, **61**, 395–420.
- Liu, Y., Roof, S., Ye, Z., Barry, C., van Tuinen, A., Vrebalov, J., Bowler, C. & Giovannoni, J. (2004) Manipulation of light signal transduction as a means of modifying fruit nutritional quality in tomato. *Proceedings of the National Academy of Sciences*, **101**, 9897–9902.
- Lopez-Juez, E. (2007) Plastid biogenesis, between light and shadows. *Journal of Experimental Botany*, **58**, 11–26.
- López-Juez, E., Dillon, E., Magyar, Z., Khan, S., Hazeldine, S., de Jager, S.M., Murray, J.A.H., Beemster, G.T.S., Bogre, L. & Shanahan, H. (2008) Distinct Light-Initiated Gene Expression and Cell Cycle Programs in the Shoot Apex and Cotyledons of Arabidopsis. *the Plant Cell Online*, **20**, 947–968.

- López-Juez, E. & Hills, A. (2011) *Screening or Selection for Chloroplast Biogenesis Mutants of Arabidopsis, Following Chemical or Insertional Mutagenesis. Methods in Molecular Biology*, pp. 3–18.
- López-Juez, E. & Pyke, K.A. (2005) Plastids unleashed: Their development and their integration in plant development. *International Journal of Developmental Biology*, **49**, 557–577.
- Lu, F., Cui, X., Zhang, S., Liu, C. & Cao, X. (2010) JMJ14 is an H3K4 demethylase regulating flowering time in Arabidopsis. *Cell Research*, **20**, 387–390.
- Lu, F., Li, G., Cui, X., Liu, C., Wang, X.-J. & Cao, X. (2008) Comparative Analysis of JmjC Domain-containing Proteins Reveals the Potential Histone Demethylases in Arabidopsis and Rice. *Journal of Integrative Plant Biology*, **50**, 886–896.
- Maere, S., Heymans, K. & Kuiper, M. (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics*, **21**, 3448–3449.
- Makino, S., Kiba, T., Imamura, A., Hanaki, N., Nakamura, A., Suzuki, T., Taniguchi, M., Ueguchi, C., Sugiyama, T. & Mizuno, T. (2000) Genes Encoding Pseudo-Response Regulators: Insight into His-to-Asp Phosphorelay and Circadian Rhythm in Arabidopsis thaliana. *Plant and Cell Physiology*, **41**, 791–803.
- Mara, C.D. & Irish, V.F. (2008) Two GATA transcription factors are downstream effectors of floral homeotic gene action in Arabidopsis. *Plant physiology*, **147**, 707–18.

- Martin, C. & Zhang, Y. (2005) The diverse functions of histone lysine methylation. *Nature Reviews Molecular Cell Biology*, **6**, 838–849.
- Martin, G., Leivar, P., Ludevid, D., Tepperman, J.M., Quail, P.H. & Monte, E. (2016) Phytochrome and retrograde signalling pathways converge to antagonistically regulate a light-induced transcriptional network. *Nature Communications*, **7**.
- Le Masson, I., Jauvion, V., Bouteiller, N., Rivard, M., Elmayan, T. & Vaucheret, H. (2012) Mutations in the Arabidopsis H3K4me2/3 Demethylase JMJ14 Suppress Posttranscriptional Gene Silencing by Decreasing Transgene Transcription. *The Plant Cell*, **24**, 3603–3612.
- Masuda, T. & Fujita, Y. (2008) Regulation and evolution of chlorophyll metabolism. *Photochemical & Photobiological Sciences*, **7**, 1131.
- Matsumoto, F., Obayashi, T., Sasaki-Sekimoto, Y., Ohta, H., Takamiya, K. & Masuda, T. (2004) Gene Expression Profiling of the Tetrapyrrole Metabolic Pathway in Arabidopsis with a Mini-Array System. *PLANT PHYSIOLOGY*, **135**, 2379–2391.
- McCormac, A.C., Fischer, A., Kumar, A.M., Söll, D. & Terry, M.J. (2001) Regulation of HEMA1 expression by phytochrome and a plastid signal during de-etiolation in Arabidopsis Thaliana. *Plant Journal*, **25**, 549–561.
- McCormac, A.C. & Terry, M.J. (2002) Loss of nuclear gene expression during the phytochrome A-mediated far-red block of greening response. *Plant physiology*, **130**, 402–414.
- Meshorer, E. & Misteli, T. (2006) Chromatin in pluripotent embryonic stem cells and differentiation. *Nature Reviews Molecular Cell Biology*, **7**, 540–546.

- Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T. & Misteli, T. (2006) Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Developmental Cell*, **10**, 105–116.
- Mitalipov, S. & Wolf, D. (2009) Totipotency, pluripotency and nuclear reprogramming. *Advances in Biochemical Engineering/Biotechnology*, **114**, 185–199.
- Mochizuki, N., Brusslan, J.A., Larkin, R., Nagatani, A. & Chory, J. (2001) Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. *Proceedings of the National Academy of Sciences*, **98**, 2053–2058.
- Mochizuki, N., Tanaka, R., Grimm, B., Masuda, T., Moulin, M., Smith, A.G., Tanaka, A. & Terry, M.J. (2010) The cell biology of tetrapyrroles: a life and death struggle. *Trends in Plant Science*, **15**, 488–498.
- Mockler, T.C., Guo, H., Yang, H., Duong, H. & Lin, C. (1999) Antagonistic actions of Arabidopsis cryptochromes and phytochrome B in the regulation of floral induction. *Development (Cambridge, England)*, **126**, 2073–2082.
- Moir, D., Stewart, S.E., Osmond, B.C. & Botstein, D. (1982) Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies. *Genetics*, **100**, 547–563.
- Monte, E., Tepperman, J.M., Al-Sady, B., Kaczorowski, K.A., Alonso, J.M., Ecker, J.R., Li, X., Zhang, Y. & Quail, P.H. (2004) The phytochrome-interacting transcription factor, PIF3, acts early, selectively, and positively in light-induced chloroplast development. *Proceedings of the National Academy of Sciences*, **101**, 16091–16098.

- Moon, J., Zhu, L., Shen, H. & Huq, E. (2008) PIF1 directly and indirectly regulates chlorophyll biosynthesis to optimize the greening process in Arabidopsis. *Proceedings of the National Academy of Sciences*, **105**, 9433–9438.
- Moulin, M., McCormac, A.C., Terry, M.J. & Smith, A.G. (2008) Tetrapyrrole profiling in Arabidopsis seedlings reveals that retrograde plastid nuclear signaling is not due to Mg-protoporphyrin IX accumulation. *Proceedings of the National Academy of Sciences*, **105**, 15178–15183.
- Muramoto, T., Kohchi, T., Yokota, A., Hwang, I. & Goodman, H. (1999) The Arabidopsis Photomorphogenic Mutant hy1 Is Deficient in Phytochrome Chromophore Biosynthesis as a Result of a Mutation in a Plastid Heme Oxygenase. *THE PLANT CELL ONLINE*, **11**, 335–348.
- Murphy, J.T. & Lagarias, J.C. (1997) The phytofluors: a new class of fluorescent protein probes. *Current Biology*, **7**, 870–876.
- Nagatani, A. (2004) Light-regulated nuclear localization of phytochromes. *Current Opinion in Plant Biology*, **7**, 708–711.
- Neff, M.M. & Chory, J. (1998) Genetic Interactions between Phytochrome A, Phytochrome B, and Cryptochrome 1 during Arabidopsis Development. *Plant Physiology*, **118**, 27–35.
- Nolte, C. & Staiger, D. (2015) RNA around the clock – regulation at the RNA level in biological timing. *Frontiers in Plant Science*, **6**, 1–15.
- Nott, A., Jung, H.-S., Koussevitzky, S. & Chory, J. (2006) PLASTID-TO-NUCLEUS RETROGRADE SIGNALING. *Annual Review of Plant Biology*, **57**, 739–759.

- O'Malley, R.C. & Ecker, J.R. (2010) Linking genotype to phenotype using the Arabidopsis unimutant collection. *The Plant Journal*, **61**, 928–940.
- Okazaki, K., Kabeya, Y., Suzuki, K., Mori, T., Ichikawa, T., Matsui, M., Nakanishi, H. & Miyagishima, S. -y. (2009) The PLASTID DIVISION1 and 2 Components of the Chloroplast Division Machinery Determine the Rate of Chloroplast Division in Land Plant Cell Differentiation. *THE PLANT CELL ONLINE*, **21**, 1769–1780.
- Okkels, F.T. & Pedersen, M.G. (1988) THE TOXICITY TO PLANT TISSUE AND TO AGROBACTERIUM TUMEFACIENS OF SOME ANTIBIOTICS. *Acta Horticulturae*, 199–208.
- op den Camp, R.G.L., Przybyla, D., Ochsenbein, C., Laloi, C., Kim, C., Danon, A., Wagner, D., Hideg, E., Göbel, C., Feussner, I., Nater, M. & Apel, K. (2003) Rapid Induction of Distinct Stress Responses after the Release of Singlet Oxygen in Arabidopsis. *THE PLANT CELL ONLINE*, **15**, 2320–2332.
- Page, M.T., Kacprzak, S.M., Mochizuki, N., Okamoto, H., Smith, A.G. & Terry, M.J. (2017) Seedlings Lacking the PTM Protein Do Not Show a genomes uncoupled (gun) Mutant Phenotype. *Plant Physiology*, **174**, 21–26.
- Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G.A., Stewart, R. & Thomson, J.A. (2007) Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell*, **1**, 299–312.
- Pan, Y., Bradley, G., Pyke, K., Ball, G., Lu, C., Fray, R., Marshall, A., Jayasuta, S., Baxter, C., van Wijk, R., Boyden, L., Cade, R., Chapman, N.H., Fraser, P.D., Hodgman, C. & Seymour, G.B. (2013) Network Inference Analysis

Identifies an APRR2-Like Gene Linked to Pigment Accumulation in Tomato and Pepper Fruits. *PLANT PHYSIOLOGY*, **161**, 1476–1485.

Pandey, R., Müller, A., Napoli, C.A., Selinger, D.A., Pikaard, C.S., Richards, E.J., Bender, J., Mount, D.W. & Jorgensen, R.A. (2002) Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana* suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucleic Acids Research*, **30**, 5036–5055.

Parks, B.M. & Quail, P.H. (1991) Phytochrome-Deficient hy1 and hy2 Long Hypocotyl Mutants of *Arabidopsis* Are Defective in Phytochrome Chromophore Biosynthesis. *The Plant cell*, **3**, 1177–1186.

Paul, M. (2013) Photosynthesis. Plastid biology, energy conversion and carbon assimilation. *Annals of Botany*, **111**, ix–ix.

Pearson, W.R. (2013) An introduction to sequence similarity (“homology”) searching. *Current Protocols in Bioinformatics*.

Peter, E. & Grimm, B. (2009) GUN4 Is Required for Posttranslational Control of Plant Tetrapyrrole Biosynthesis. *Molecular Plant*, **2**, 1198–1210.

Pyke, K. (2011) Analysis of plastid number, size, and distribution in *Arabidopsis* plants by light and fluorescence microscopy. *Methods in Molecular Biology*, **774**, 19–32.

Pyke, K. (2012) *Mesophyll*. eLS, John Wiley & Sons, Ltd, Chichester, UK.

Pyke, K. (2009) *Plastid Biology*, Cambridge University Press.

R Core Team (2014) R Core Team (2014). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna*,

Austria. URL <http://www.R-project.org/>, R Foundation for Statistical Computing.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A. & Young, R.A. (2010) C-Myc regulates transcriptional pause release. *Cell*, **141**, 432–445.

Ramel, F., Birtic, S., Ginies, C., Soubigou-Taconnat, L., Triantaphylides, C. & Havaux, M. (2012) Carotenoid oxidation products are stress signals that mediate gene responses to singlet oxygen in plants. *Proceedings of the National Academy of Sciences*, **109**, 5535–5540.

Rasmussen, P.B. & Staller, P. (2014) The KDM5 family of histone demethylases as targets in oncology drug discovery. *Epigenomics*, **6**, 277–286.

Rausenberger, J., Tscheuschler, A., Nordmeier, W., Wüst, F., Timmer, J., Schäfer, E., Fleck, C. & Hiltbrunner, A. (2011) Photoconversion and nuclear trafficking cycles determine phytochrome A's response profile to far-red light. *Cell*, **146**, 813–825.

Rauwolf, U., Golczyk, H., Greiner, S. & Herrmann, R.G. (2010) Variable amounts of DNA related to the size of chloroplasts III. Biochemical determinations of DNA amounts per organelle. *Molecular Genetics and Genomics*, **283**, 35–47.

Reed, J.W., Nagatani, A., Elich, T.D., Fagan, M. & Chory, J. (1994) Phytochrome A and Phytochrome B Have Overlapping but Distinct Functions in Arabidopsis Development. *Plant physiology*, **104**, 1139–1149.

Reed, J.W., Nagpal, P., Poole, D.S., Furuya, M. & Chory, J. (1993) Mutations in the gene for the red/far-red light receptor phytochrome B alter cell

elongation and physiological responses throughout Arabidopsis development. *The Plant cell*, **5**, 147–157.

Reinbothe, S., Reinbothe, C., Lebedev, N. & Apela, K. (1996)

Protochlorophyllide-Reducing Enzymes of Angiosperm Chlorophyll Biosynthesis. **8**, 763–769.

Richly, E., Dietzmann, A., Biehl, A., Kurth, J., Laloi, C., Apel, K., Salamini, F. &

Leister, D. (2003) Covariations in the nuclear chloroplast transcriptome reveal a regulatory master-switch. *EMBO Reports*, **4**, 491–498.

Richter, R., Behringer, C., Müller, I.K. & Schwechheimer, C. (2010) The GATA-

type transcription factors GNC and GNL/CGA1 repress gibberellin signaling downstream from DELLA proteins and phytochrome-interacting factors.

Genes and Development, **24**, 2093–2104.

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., -Z., C., Jiang, Keddie, J.,

Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J.Z., Ghandehari, D., Sherman, B.K. & -L. Yu, G.

(2000) Arabidopsis Transcription Factors: Genome-Wide Comparative Analysis Among Eukaryotes. *Science*, **290**, 2105–2110.

Rizzini, L., Favory, J.J., Cloix, C., Faggionato, D., O'Hara, A., Kaiserli, E.,

Baumeister, R., Schafer, E., Nagy, F., Jenkins, G.I. & Ulm, R. (2011)

Perception of UV-B by the Arabidopsis UVR8 Protein. *Science*, **332**, 103–106.

Rockwell, N.C., Su, Y.-S. & Lagarias, J.C. (2006) PHYTOCHROME

STRUCTURE AND SIGNALING MECHANISMS. *Annual Review of Plant Biology*, **57**, 837–858.

- Rodrigues, M.A., Bianchetti, R.E. & Freschi, L. (2014) Shedding light on ethylene metabolism in higher plants. *Frontiers in Plant Science*, **5**, 665.
- Rose, N.R. & Klose, R.J. (2014) Understanding the relationship between DNA methylation and histone lysine methylation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1839**, 1362–1372.
- Rossetto, D., Avvakumov, N. & Côté, J. (2012) Histone phosphorylation: A chromatin modification involved in diverse nuclear events. *Epigenetics*, **7**, 1098–1108.
- Roudier, F., Ahmed, I., Bérard, C., Sarazin, A., Mary-Huard, T., Cortijo, S., Bouyer, D., Caillieux, E., Duvernois-Berthet, E., Al-Shikhley, L., Giraut, L., Després, B., Drevensek, S., Barneche, F., Dèrozier, S., Brunaud, V., Aubourg, S., Schnittger, A., Bowler, C., Martin-Magniette, M.-L., Robin, S., Caboche, M. & Colot, V. (2011) Integrative epigenomic mapping defines four main chromatin states in *Arabidopsis*. *The EMBO Journal*, **30**, 1928–1938.
- Rowan, B.A. & Bendich, A.J. (2011) *Isolation, Quantification, and Analysis of Chloroplast DNA. Methods in molecular biology (Clifton, N.J.)*, pp. 151–170.
- Ruijtenberg, S. & van den Heuvel, S. (2016) Coordinating cell proliferation and differentiation: Antagonism between cell cycle regulators and cell type-specific gene expression. *Cell Cycle*, **15**, 196–212.
- Saleh, A., Alvarez-Venegas, R., Yilmaz, M., Le, O., Hou, G., Sadler, M., Al-Abdallat, A., Xia, Y., Lu, G., Ladunga, I. & Avramova, Z. (2008) The Highly Similar *Arabidopsis* Homologs of *Trithorax* ATX1 and ATX2 Encode

Proteins with Divergent Biochemical Functions. *THE PLANT CELL ONLINE*, **20**, 568–579.

Santel, H. -J & Apel, K. (1981) The Protochlorophyllide Holochrome of Barley (*Hordeum vulgare* L.): The Effect of Light on the NADPH: Protochlorophyllide Oxidoreductase. *European Journal of Biochemistry*, **120**, 95–103.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E. & Tabata, S. (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA research: an international journal for rapid publication of reports on genes and genomes*, **6**, 283–90.

Schlicke, H., Richter, A., Rothbart, M., Brzezowski, P., Hedtke, B. & Grimm, B. (2015) Function of Tetrapyrroles, Regulation of Tetrapyrrole Metabolism and Methods for Analyses of Tetrapyrroles. *Procedia Chemistry*, **14**, 171–175.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D. & Lohmann, J.U. (2005) A gene expression map of *Arabidopsis thaliana* development. *Nature Genetics*, **37**, 501–506.

Schmitz, S.U., Albert, M., Malatesta, M., Morey, L., Johansen, J. V., Bak, M., Tommerup, N., Abarategui, I. & Helin, K. (2011) *Jarid1b* targets genes regulating development and is involved in neural differentiation. *EMBO Journal*, **30**, 4586–4600.

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, **9**, 671–675.

Searle, I.R., Pontes, O., Melnyk, C.W., Smith, L.M. & Baulcombe, D.C. (2010) JM14, a JmjC domain protein, is required for RNA silencing and cell-to-

cell movement of an RNA silencing signal in Arabidopsis. *Genes & Development*, **24**, 986–991.

Sequeira-Mendes, J., Araguez, I., Peiro, R., Mendez-Giraldez, R., Zhang, X., Jacobsen, S.E., Bastolla, U. & Gutierrez, C. (2014) The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. *The Plant Cell*, **26**, 2351–2366.

Shannon, P. (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, **13**, 2498–2504.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. & Ideker, T. (2003) Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Research*, **13**, 2498–2504.

Sharrock, R.A. & Quail, P.H. (1989) Novel phytochrome sequences in Arabidopsis thaliana: structure, evolution, and differential expression of a plant regulatory photoreceptor family. *Genes & Development*, **3**, 1745–1757.

Shaver, J.M., Oldenburg, D.J. & Bendich, A.J. (2006) Changes in chloroplast DNA during development in tobacco, Medicago truncatula, pea, and maize. *Planta*, **224**, 72–82.

Shekhawat, G.S. & Verma, K. (2010) Haem oxygenase (HO): an overlooked enzyme of plant metabolism and defence. *Journal of Experimental Botany*, **61**, 2255–2270.

- Shen, Y., Conde e Silva, N., Audonnet, L., Servet, C., Wei, W. & Zhou, D.-X. (2014) Over-expression of histone H3K4 demethylase gene JMJ15 enhances salt tolerance in Arabidopsis. *Frontiers in Plant Science*, **5**.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A. & Shi, Y. (2004) Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. *Cell*, **119**, 941–953.
- Shilatifard, A. (2008) Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Current Opinion in Cell Biology*, **20**, 341–348.
- Shin, J., Kim, K., Kang, H., Zulfugarov, I.S., Bae, G., Lee, C.-H., Lee, D. & Choi, G. (2009) Phytochromes promote seedling light responses by inhibiting four negatively-acting phytochrome-interacting factors. *Proceedings of the National Academy of Sciences*, **106**, 7660–7665.
- Stephenson, P.G., Fankhauser, C. & Terry, M.J. (2009) PIF3 is a repressor of chloroplast development. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 7654–9.
- Strand, Å., Asami, T., Alonso, J., Ecker, J.R. & Chory, J. (2003) Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. *Nature*, **421**, 79–83.
- Summons, R.E., Jahnke, L.L., Hope, J.M. & Logan, G.A. (1999) 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis. *Nature*, **400**, 554–557.
- Sun, X., Feng, P., Xu, X., Guo, H., Ma, J., Chi, W., Lin, R., Lu, C. & Zhang, L. (2011) A chloroplast envelope-bound PHD transcription factor mediates chloroplast signals to the nucleus. *Nature Communications*, **2**, 477.

- Susek, R.E., Ausubel, F.M. & Chory, J. (1993) Signal transduction mutants of arabidopsis uncouple nuclear CAB and RBCS gene expression from chloroplast development. *Cell*, **74**, 787–799.
- Tadini, L., Pesaresi, P., Kleine, T., Rossi, F., Guljamow, A., Sommer, F., Mühlhaus, T., Schroda, M., Masiero, S., Pribil, M., Rothbart, M., Hedtke, B., Grimm, B. & Leister, D. (2016) GUN1 Controls Accumulation of the Plastid Ribosomal Protein S1 at the Protein Level and Interacts with Proteins Involved in Plastid Protein Homeostasis. *Plant physiology*, **170**, 1817–30.
- Takeuchi, T., Watanabe, Y., Takano-Shimizu, T. & Kondo, S. (2006) Roles of jumonji and jumonji family genes in chromatin regulation and development. *Developmental Dynamics*, **235**, 2449–2459.
- Takeuchi, T., Yamazaki, Y., Katoh-Fukui, Y., Tsuchiya, R., Kondo, S., Motoyama, J. & Higashinakagawa, T. (1995) Gene trap capture of a novel mouse gene, jumonji, required for neural tube formation. *Genes and Development*, **9**, 1211–1222.
- Tepperman, J.M., Hudson, M.E., Khanna, R., Zhu, T., Chang, S.H., Wang, X. & Quail, P.H. (2004) Expression profiling of phyB mutant demonstrates substantial contribution of other phytochromes to red-light-regulated gene expression during seedling de-etiolation. *Plant Journal*, **38**, 725–739.
- Tepperman, J.M., Hwang, Y.-S. & Quail, P.H. (2006) phyA dominates in transduction of red-light signals to rapidly responding genes at the initiation of Arabidopsis seedling de-etiolation. *The Plant Journal*, **48**, 728–742.
- Terry, M.J. (1997) Phytochrome chromophore-deficient mutants. *Plant Cell And Environment*, **20**, 740–745.

- Terry, M.J. & Kendrick, R.E. (1999) Feedback Inhibition of Chlorophyll Synthesis in the Phytochrome Chromophore-Deficient aurea and yellow-green-2 Mutants of Tomato. *Plant Physiology*, **119**, 143–152.
- Terry, M.J. & Smith, A.G. (2013) A model for tetrapyrrole synthesis as the primary mechanism for plastid-to-nucleus signaling during chloroplast biogenesis. *Frontiers in Plant Science*, **4**, 14.
- Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S. & Jones, J.M. (1998) Embryonic stem cell lines derived from human blastocysts. *Science (New York, N.Y.)*, **282**, 1145–7.
- To, J.P.C., Deruere, J., Maxwell, B.B., Morris, V.F., Hutchison, C.E., Ferreira, F.J., Schaller, G.E. & Kieber, J.J. (2007) Cytokinin Regulates Type-A Arabidopsis Response Regulator Activity and Protein Stability via Two-Component Phosphorelay. *THE PLANT CELL ONLINE*, **19**, 3901–3914.
- Toufighi, K., Brady, S.M., Austin, R., Ly, E. & Provart, N.J. (2005) The Botany Array Resource: e-Northern, Expression Angling, and promoter analyses. *The Plant Journal*, **43**, 153–163.
- Trerotola, M., Relli, V., Simeone, P. & Alberti, S. (2015) Epigenetic inheritance and the missing heritability. *Human genomics*, **9**, 17.
- Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S.I., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R., Nakayama, J.I. & Okano, M. (2006) Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes to Cells*, **11**, 805–814.
- vanTuinen, A., Hanhart, C.J., Kerckhoffs, L.H.J., Nagatani, A., Boylan, M.T., Quail, P.H., Kendrick, R.E. & Koornneef, M. (1996) Analysis of

- phytochrome-deficient yellow-green-2 and aurea mutants of tomato. *Plant Journal*, **9**, 173–182.
- Verdone, L., Caserta, M. & Di Mauro, E. (2005) Role of histone acetylation in the control of gene expression. *Biochem Cell Biol*, **83**, 344–353.
- Wagner, D., Przybyla, D., Op den Camp, R., Kim, C., Landgraf, F., Lee, K.P., Würsch, M., Laloi, C., Nater, M., Hideg, E. & Apel, K. (2004) The genetic basis of singlet oxygen-induced stress responses of *Arabidopsis thaliana*. *Science (New York, N.Y.)*, **306**, 1183–1185.
- Waters, M.T. & Langdale, J.A. (2009) The making of a chloroplast. *EMBO Journal*, **28**, 2861–2873.
- Waters, M.T., Moylan, E.C. & Langdale, J.A. (2008) GLK transcription factors regulate chloroplast development in a cell-autonomous manner. *Plant Journal*, **56**, 432–444.
- Waters, M.T., Wang, P., Korkaric, M., Capper, R.G., Saunders, N.J. & Langdale, J.A. (2009) GLK Transcription Factors Coordinate Expression of the Photosynthetic Apparatus in *Arabidopsis*. *THE PLANT CELL ONLINE*, **21**, 1109–1128.
- Weiler, J.L., Terry, M.J., Reid, J.B. & Kendrick, R.E. (1997) The phytochrome-deficient *pcd2* mutant of pea is unable to convert biliverdin IX α to 3(Z)-phytochromobilin. *Plant Journal*, **11**, 1177–1186.
- Whitelam, G., Johnson, E., Peng, J., Carol, P., Anderson, M., Cowl, J. & Harberd, N. (1993) Phytochrome A Null Mutants of *Arabidopsis* Display a Wild-Type Phenotype in White Light. *the Plant Cell Online*, **5**, 757–768.

- Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V. & Provart, N.J. (2007) An “Electronic Fluorescent Pictograph” Browser for Exploring and Analyzing Large-Scale Biological Data Sets. *PLoS ONE*, **2**, e718.
- Wise, R.R. & Hooper, J.K. (2006) *The structure and function of plastids*,.
- Woodson, J.D., Perez-Ruiz, J.M. & Chory, J. (2011) Heme Synthesis by Plastid Ferrochelatase I Regulates Nuclear Gene Expression in Plants. *Current Biology*, **21**, 897–903.
- Xie, W., Schultz, M.D., Lister, R., Hou, Z., Rajagopal, N., Ray, P., Whitaker, J.W., Tian, S., Hawkins, R.D. & Leung, D. (2013) Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell*, **153**, 1134–1148.
- Xu, M., Soloveychik, M., Ranger, M., Schertzberg, M., Shah, Z., Raisner, R., Venkatasubrahmanyam, S., Tsui, K., Gebbia, M., Hughes, T., van Bakel, H., Nislow, C., Madhani, H.D. & Meneghini, M.D. (2012) Timing of Transcriptional Quiescence during Gametogenesis Is Controlled by Global Histone H3K4 Demethylation. *Developmental Cell*, **23**, 1059–1071.
- Xu, X., Paik, I., Zhu, L., Bu, Q., Huang, X., Deng, X.W. & Huq, E. (2014) PHYTOCHROME INTERACTING FACTOR1 Enhances the E3 Ligase Activity of CONSTITUTIVE PHOTOMORPHOGENIC1 to Synergistically Repress Photomorphogenesis in Arabidopsis. *The Plant Cell*, **26**, 1992–2006.
- Yamane, K., Tateishi, K., Klose, R.J., Fang, J., Fabrizio, L.A., Erdjument-Bromage, H., Taylor-Papadimitriou, J., Tempst, P. & Zhang, Y. (2007) PLU-1 Is an H3K4 Demethylase Involved in Transcriptional Repression and Breast Cancer Cell Proliferation. *Molecular Cell*, **25**, 801–812.

- Yang, H., Han, Z., Cao, Y., Fan, D., Li, H., Mo, H., Feng, Y., Liu, L., Wang, Z., Yue, Y., Cui, S., Chen, S., Chai, J. & Ma, L. (2012a) A companion cell-dominant and developmentally regulated H3K4 demethylase controls flowering time in Arabidopsis via the repression of FLC expression. *PLoS Genetics*, **8**.
- Yang, H., Mo, H., Fan, D., Cao, Y., Cui, S. & Ma, L. (2012b) Overexpression of a histone H3K4 demethylase, JMJ15, accelerates flowering time in Arabidopsis. *Plant Cell Reports*, **31**, 1297–1308.
- Yang, J., Lin, R., Sullivan, J., Hoecker, U., Liu, B., Xu, L., Deng, X.W. & Wang, H. (2005) Light Regulates COP1-Mediated Degradation of HFR1, a Transcription Factor Essential for Light Signaling in Arabidopsis. *THE PLANT CELL ONLINE*, **17**, 804–821.
- Yang, W., Jiang, D., Jiang, J. & He, Y. (2010) A plant-specific histone H3 lysine 4 demethylase represses the floral transition in Arabidopsis. *The Plant Journal*, **62**, 663–673.
- Young, R.A. (2011) Control of the embryonic stem cell state. *Cell*, **144**, 940–954.
- Zhang, B., Holmlund, M., Lorrain, S., Norberg, M., Bakó, L., Fankhauser, C. & Nilsson, O. (2017) BLADE-ON-PETIOLE proteins act in an E3 ubiquitin ligase complex to regulate PHYTOCHROME INTERACTING FACTOR 4 abundance. *eLife*, **6**, 1–19.
- Zhang, S., Zhou, B., Kang, Y., Cui, X., Liu, A., Deleris, A., Greenberg, M.V., Cui, X., Qiu, Q., Lu, F., Wohlschlegel, J.A., Jacobsen, S.E. & Cao, X. (2015) C-terminal domains of histone demethylase JMJ14 interact with a

pair of NAC transcription factors to mediate specific chromatin association.
Cell Discovery, **1**, 15003.

Zhang, X., Bernatavichute, Y. V, Cokus, S., Pellegrini, M. & Jacobsen, S.E.
(2009) Genome-wide analysis of mono-, di- and trimethylation of histone
H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biology*, **10**, R62.

Zhang, Y. (2003) Transcriptional regulation by histone ubiquitination and
deubiquitination. *Genes and Development*, **17**, 2733–2740.

Zheng, Y., Jiao, C., Sun, H., Rosli, H.G., Pombo, M.A., Zhang, P., Banf, M., Dai,
X., Martin, G.B., Giovannoni, J.J., Zhao, P.X., Rhee, S.Y. & Fei, Z. (2016)
iTAK: A Program for Genome-wide Prediction and Classification of Plant
Transcription Factors, Transcriptional Regulators, and Protein Kinases.
Molecular Plant, **9**, 1667–1670.

Appendix

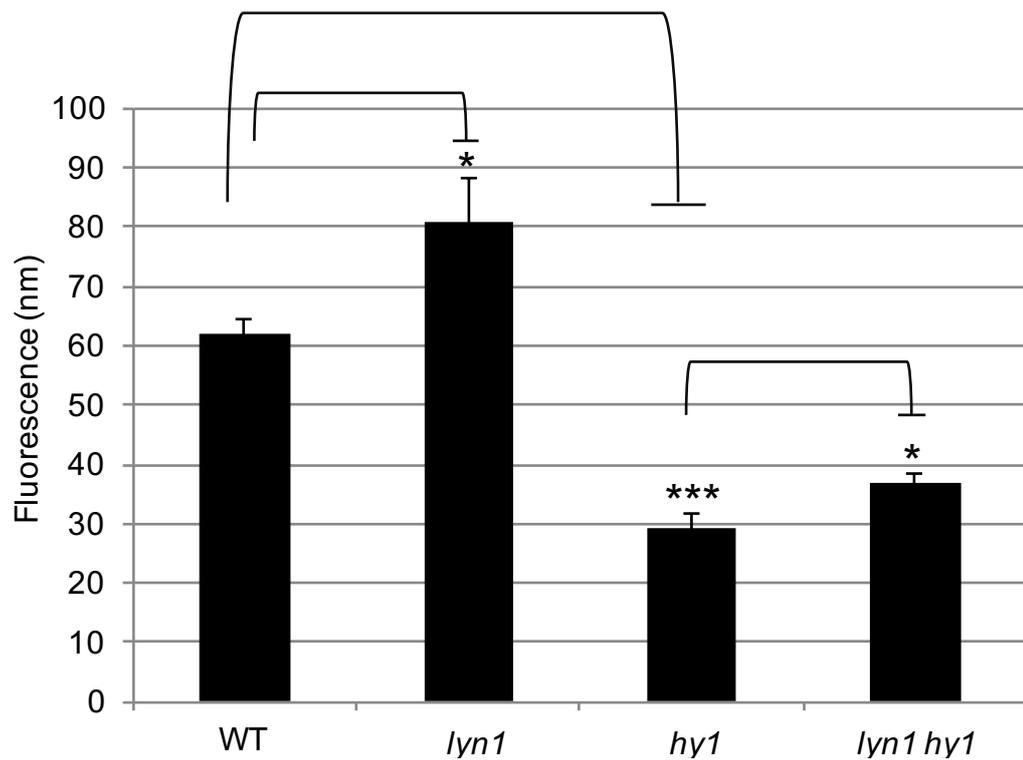


Figure 7. 1 *lyn1* effect of protochlorophyllide level in the dark.

Protochlorophyllide was extracted from 0.025g of 5-day-old dark grown seedlings and incubated in dimethyl formamide (DMF) in the dark. The fluorescence emission spectra recorded using 433nm excitation. The peak of emission at approximately 631nm corresponds to protochlorophyllide. n=6.

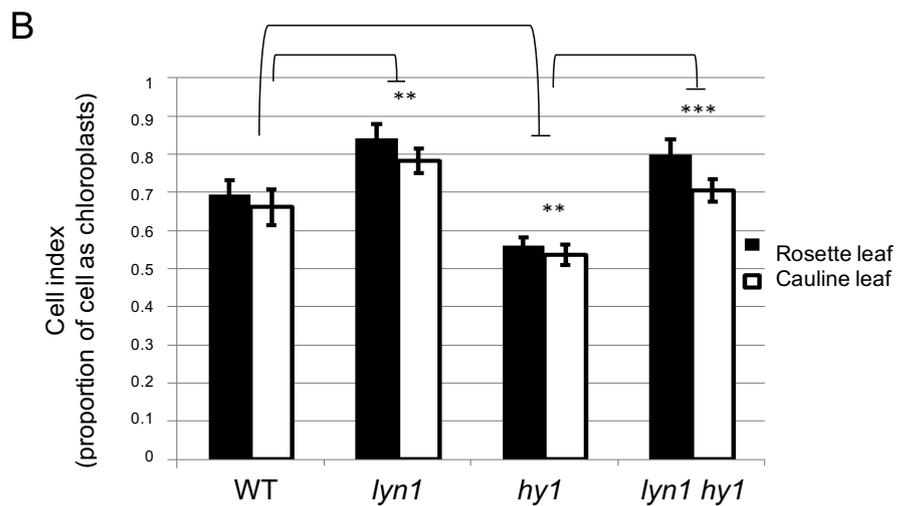
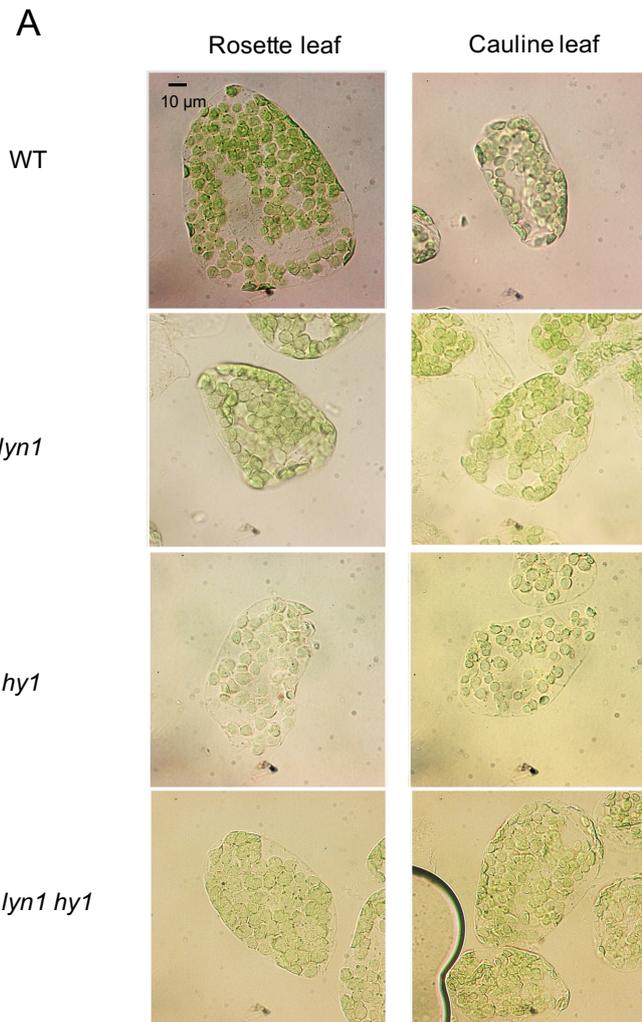
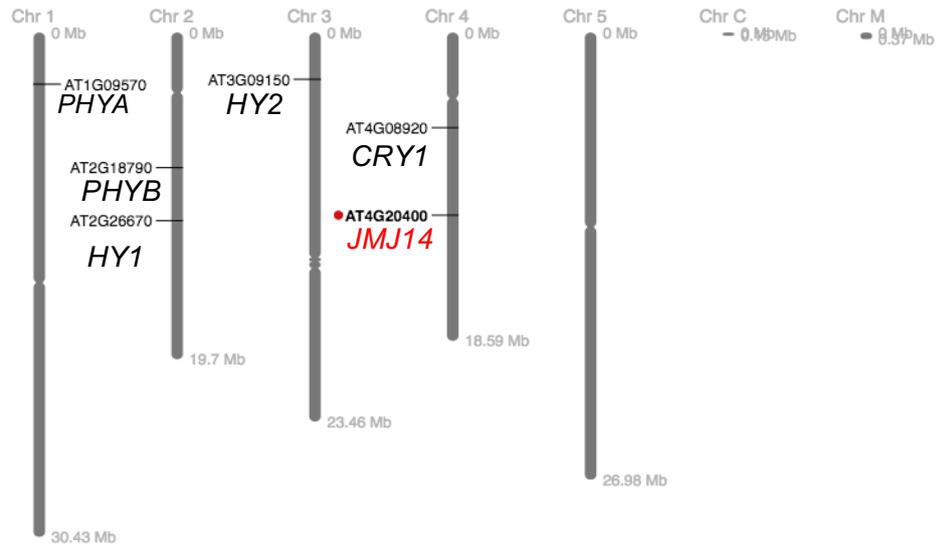


Figure 7. 2 Chloroplast compartment quantification in mesophyll cells of rosette and cauline leaves

A. Observation of chloroplast development under DIC microscopy

B. Quantifying the cell index of WT, *lyn1*, *hy1* and *lyn1 hy1*.



This image was generated with the Chromosome Viewer at bar.utoronto.ca/eplant by Waese et al. 2017

Figure 7. 3 *JMJ14* location on *Arabidopsis* genetic map

This figure shows the location of *JMJ14* and the genetic linkage of *JMJ14* with other genes which are involved in this project. The doubles mutants can be generated because the mutations are not linked except for *cry1*. The Figure was generated from chromosome viewer at The Bio-Analytic Resource for Plant Biology.

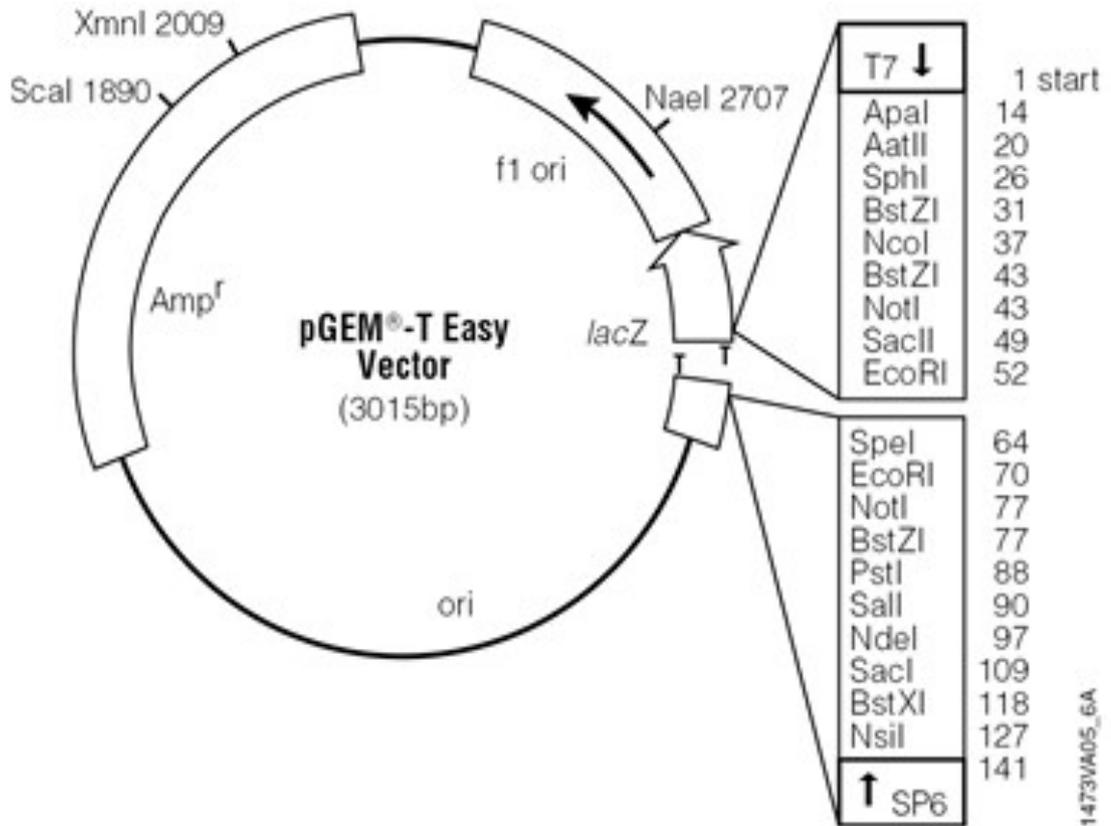


Figure 7. 4 pGEM®-T Easy Vector map

Figure showing pGEM®-T Easy Vector Systems was reproduced from Promega manual script.

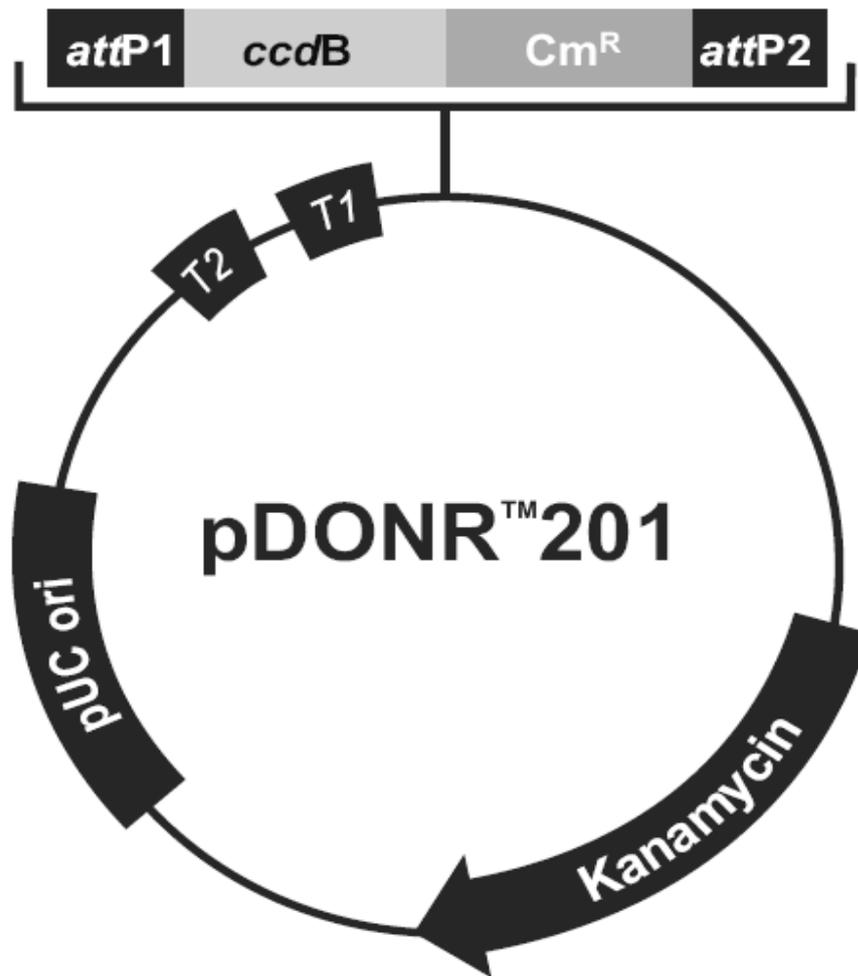


Figure 7. 5 pDONR™201 Vector map

Figure reproduced from Invitrogen Manual "Gateway pDONR Vectors" vers. E,
27 June 2007, 25-0531

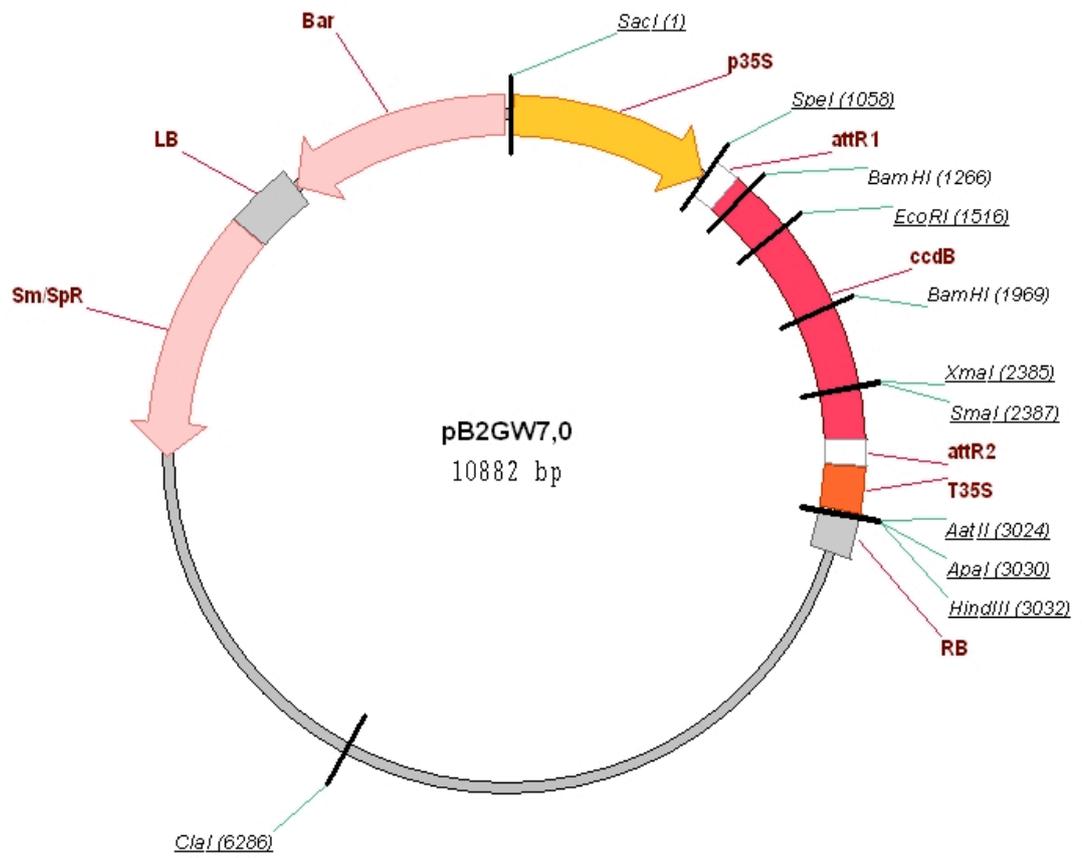


Figure 7. 6 Gateway™ Destination Vector pB2GW7 vector circle map
 Figure from (Karimi et al., 2002).

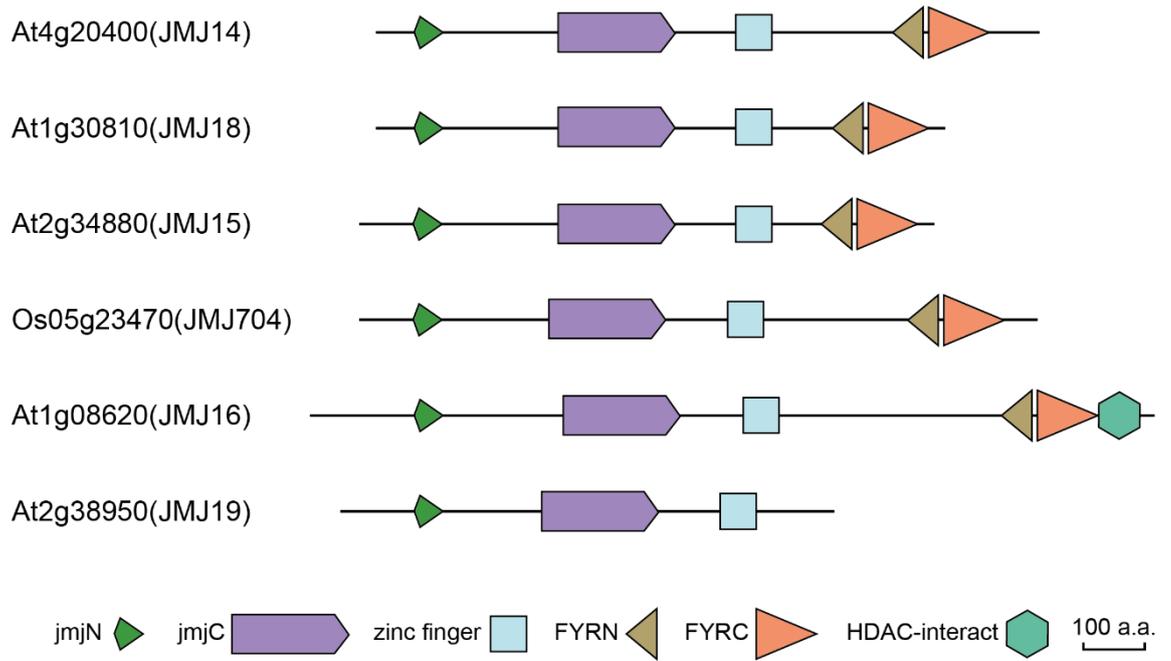


Figure 7. 7 Schematic domain structure of *JMJ14* homologues

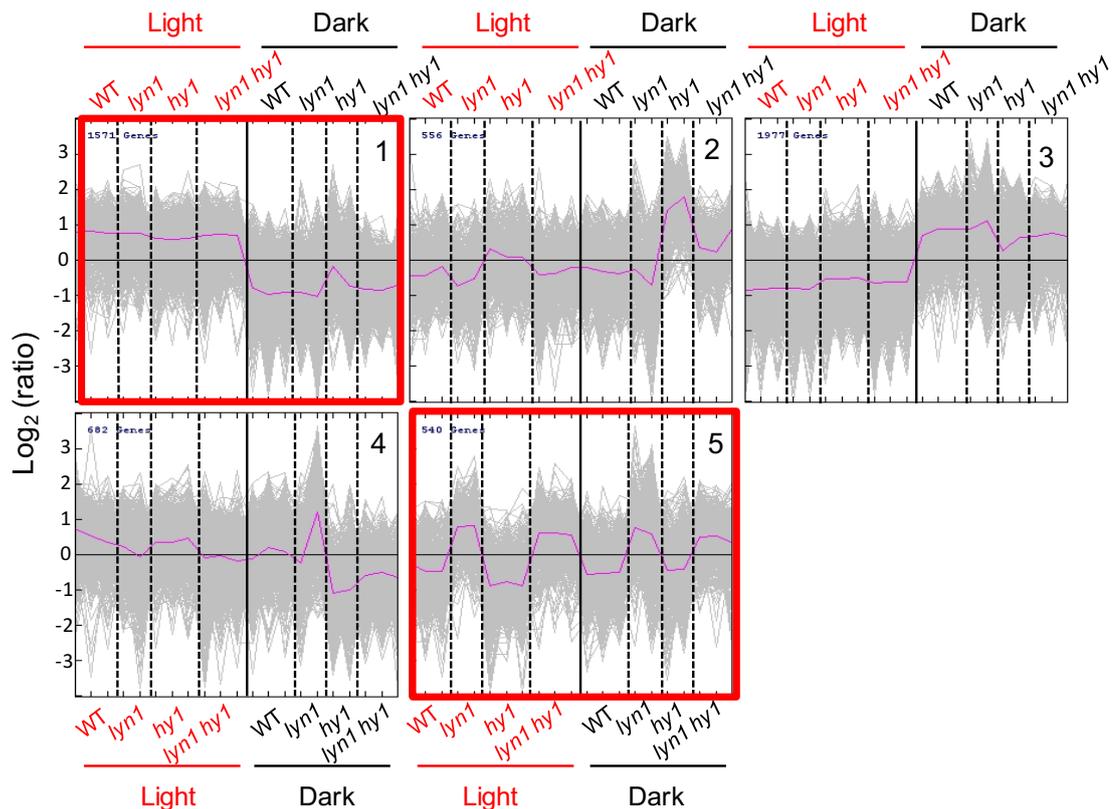


Figure 7. 8 All clusters resulting from the k-means clustering of all differentially expressed genes

Clustering of significantly, differentially expressed genes according to the gene expression pattern of 4 genotypes (WT, *lyn1*, *hy1* and *lyn1 hy1*) and both light conditions (light and dark). Clustering used k-means algorithm, Euclidean Distance as a similarity metric, and a limit of 5 clusters for clustering total differentially expressed genes. x-axis shows genotypes and light condition. Light and Dark samples are coloured in red and black respectively and separated by solid lines. Four genotypes are separated by dotted lines. Except *lyn1*_light, *lyn1*_dark and *hy1*_dark which have 2 biological replicates, the rest of samples all include 3 biological replicates. The y-axis shows expression values (\log_2 scale) centred around the median for each gene. The pink lines represent the average of gene expression values in \log_2 scale after normalization. Cluster 1 is light up-regulated. Cluster 5 is light independent and *lyn1* up-regulated. Cluster 1 and 5, highlighted by red outlines, are the gene clusters of interest.

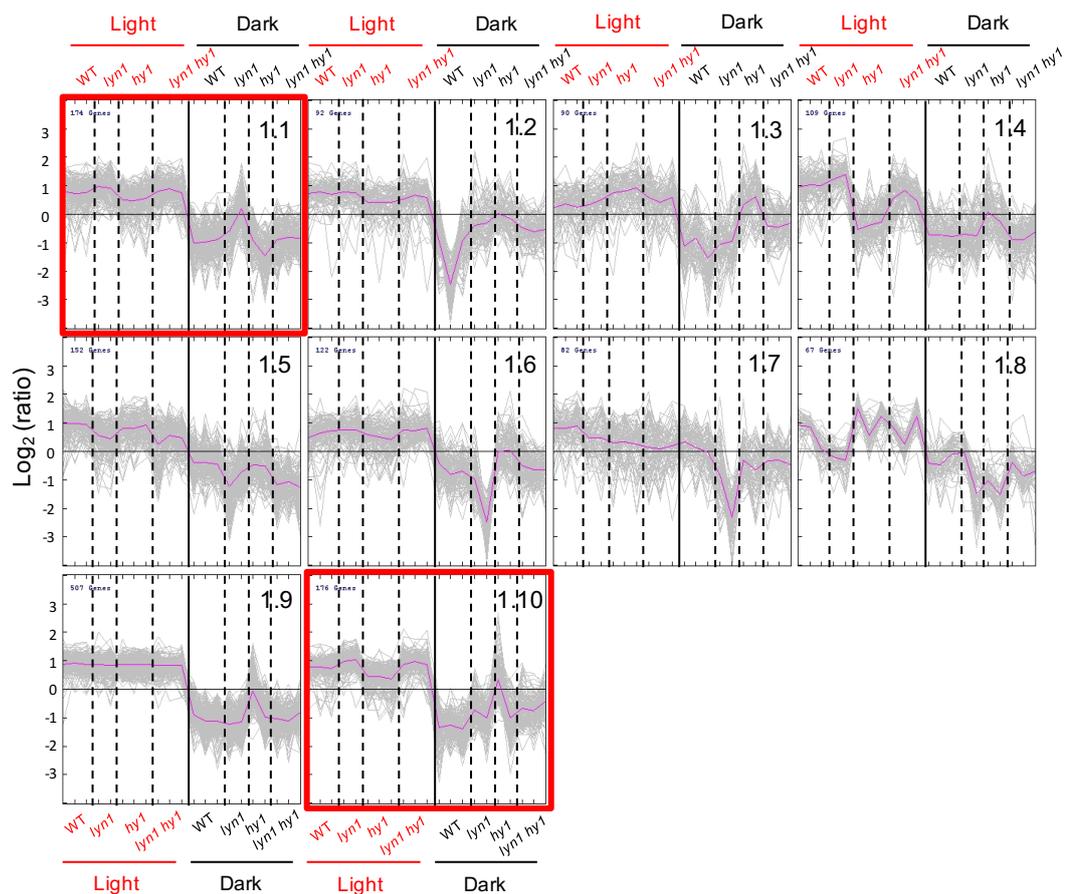


Figure 7. 9 Sub-clusters resulting from the k-means clustering of differentially expressed genes present in cluster 1

Re-clustering of cluster 1 of significantly, differentially expressed genes according to the gene expression pattern of 4 genotypes (WT, *lyn1*, *hy1* and *lyn1 hy1*) and both light conditions (light and dark). Clustering used k-means algorithm, Euclidean Distance, and a limit of 10 clusters for re-clustering genes included in cluster 1 of differentially expressed genes. Axes and legend as for Figure 7.8. Subclusters 1.1 and 1.10, highlighted by red outlines, are light up-regulated and *lyn1* up-regulated, and are the gene clusters of interest.

Table 7.1 Genotyping and sequencing primer details

Project	Primer Pair	Forward Primer (5'-3')	Reverse Primer (3'-5')	Product size (bp)	Annealing Temp (°C)
<i>jmj14-1</i> genotyping	<i>jmj14_1</i>	GAATTCATAATTTCCCACCCG	TGAAATGAAAAACGAATTGGG	1128	55
	LBb1.3 (border)	ATTTGCCGATTTCGGAAC	TGAAATGAAAAACGAATTGGG	574-874	55
Presentation of <i>JMJ14</i>	<i>JMJ14</i> full length cDNA	ATGGATCAGCTTGCATCTCTAGC	TTAAGGACTTATCTCCATCTTATCAACC	2865	55
	S6RP	TCTCGACCTCTCTTTGCC	TCAACGTTGCCAATCCAAC	707	54
<i>hy1-100</i> genotyping	<i>hy1-100</i> dCAP	TCTTGGATTGAGTTGGTTGGTTGGTGT	TTTCCAGCCCCGTTCTTGAACCTCGGAAT	289	63
<i>hy2-gun3</i> genotyping	<i>hy2-gun3</i>	CTTCAAGGCACCAAAACCCAC	GTCATCTCAA GCCCATGCCTG	1147	55
Presentation of <i>HY2</i>	<i>HY2</i> full length cDNA	ATGGCTTATCAATGGAGTTTGG	TTAGCCGATAAATTTGTCCTGTTAAAT	990	55
<i>phyA-211</i> genotyping	<i>phyA-211</i>	GCTGTGCCGGTGCTCTAAGG	GACACGATGATTCCCTGCATCTGT	1237	60
Cloning	<i>JMJ14</i>	ATGGATCAGCTTGCATCTCTAGC	TTAAGGACTTATCTCCATCTTATCAACC	2865	54
	<i>JMJ14+attB</i>	GGGACAAGTTGTACAAAAAGCAGGCT ATGGATCAGCTTGCATCTCTAGC	GGGACCACTTTGTACAAGAAAAGCTGGGT TTAAGGACTTATCTCCATCTTATCAACC	2923	65
Cloning sequencing	pDONR 201 sequencing primers	TCGCGTTAACCGTAGCATGGATCTC	GTAACATCAGAGATTTTGAAGACAC	n/a	n/a
	pB2GW7 sequencing primers	ACGCACAATCCCACATATCCT	CAACACATGAGCGGAAACCCCT	3427	54
	<i>JMJ14</i> sequencing primer2	CATCCTGGTTCCAAAGCGTC		n/a	54
	<i>JMJ14</i> sequencing primer3	TGCGAAAGATGGAGGGTGAT		n/a	54

Table 7.2 Restriction enzyme details

Project	Enzyme name	Restriction site	Enzyme incubation Temp. (°C)	Cutting sample	Digested products (bp)
<i>hy1</i> -100 dCAPs genotyping assay	ApoI	R/AATTY YTTAA/R	50	Mutant	258 and 31
<i>lyn1</i> dCAPs genotyping assay	HaeIII	GG/CC CC/GG	37	WT	256 and 29
<i>hy2</i> -105 CAPs genotyping assay	MfeI-HF	C/AATTG GTTAA/C	37	WT	533, 324 and 290
				Mutant	832 and 290

Table 7. 3 Multiple sequence alignment of *JMJ14* homologues

Sc <i>JHD2</i>	-----
Dm <i>lid</i>	MSAKTEADNTTAANS GGGGV SGTSSGGGASANGTATPARRLRTRNSTGNGTNSGSESVK
Hs <i>JARID1B</i>	-----
At <i>JMJ17</i>	-----
Oi <i>JMJ3520</i>	-----
At <i>JMJ19</i>	-----MGIEGV
At <i>JMJ14</i>	-----
Gm <i>JMJ14-like</i>	-----
At <i>JMJ15/MEE27</i>	-----
At <i>JMJ18</i>	-----
Sm <i>SELMODRAFT_170257</i>	-----
Os <i>JMJ704</i>	-----
Pp <i>PHYPADRAFT_136290</i>	-----
At <i>JMJ16/PKDM7D</i>	-----MGTELMR
Sc <i>JHD2</i>	-----
Dm <i>lid</i>	KSNANDEPSTPVT PAGATG SHTHAPPGISP-----AVM-----
Hs <i>JARID1B</i>	-----
At <i>JMJ17</i>	-----
Oi <i>JMJ3520</i>	-----
At <i>JMJ19</i>	TYLKS GNMDTIS APP GFVSQTS FVLRNVPRDKESP-----RSVSRQEQTTGFGTDDKD
At <i>JMJ14</i>	-----MDQLASLAESVAM
Gm <i>JMJ14-like</i>	-----MEQLKLAADSEAK
At <i>JMJ15/MEE27</i>	-----M--EPFSAAQNK
At <i>JMJ18</i>	-----ME--NPPLSEIK
Sm <i>SELMODRAFT_170257</i>	-----
Os <i>JMJ704</i>	-----MVSSRDPG
Pp <i>PHYPADRAFT_136290</i>	-----
At <i>JMJ16/PKDM7D</i>	ICVKEDSDDLPSVPPGFESYATFTLKRVPATTSDKAKTPAIESVSATEQAKMEVESDEA
Sc <i>JHD2</i>	-----
Dm <i>lid</i>	-----ERPMPSPVMNHASSV-SASKKYHN SCPHPT PTPAPT-----GHKKS VHTQ PHSS
Hs <i>JARID1B</i>	-----
At <i>JMJ17</i>	-----
Oi <i>JMJ3520</i>	-----MFGCEKCGFEP---DGCEACLGGP
At <i>JMJ19</i>	SCNMFLKSRPWI VHGTIP SSEALR-----PKKT
At <i>JMJ14</i>	EED-SEK-----Q-----SIK ESS -----LEPDSTPSSP
Gm <i>JMJ14-like</i>	EDK-SLGHKP-----KNNNA-----LESSDSL RNK
At <i>JMJ15/MEE27</i>	EDK-DTSVEP--PRRRCH-----RKNKGTN-----VEPPSSPYHP
At <i>JMJ18</i>	EDM-SLKNHP--PD--K-----DKDKDTI-----MEQPSSPRHR
Sm <i>SELMODRAFT_170257</i>	-----
Os <i>JMJ704</i>	EEASAPPPPP--PRR-----GEKRRMRG--RTPSP EPASAPQ DLCP S GAC--GDNVAGAT
Pp <i>PHYPADRAFT_136290</i>	-----
At <i>JMJ16/PKDM7D</i>	KAARALRRRPWINHSGCDDGDCAANN DN AASQNPQNC VDK PALPKGVV RG CEECKDCQ
Sc <i>JHD2</i>	-----MEEIPALYPT EQEF -KNPIDYLSNPHIKRLG VRYGM VKVVPPNGF
Dm <i>lid</i>	NKFDQ GKNEEF HFDTPPECPVFRPTTEEF-KNPLAYISK--IR SIAEK CGI AKIL PPATW
Hs <i>JARID1B</i>	-----
At <i>JMJ17</i>	-----MLLCD-----
Oi <i>JMJ3520</i>	PIRSACAWDEARARDVPPVKTYRPT EQEWAGD PLEYINS--IRPEAEKYGV CNI IP PA SW
At <i>JMJ19</i>	E VRRRR PLK VSETK VLEEAPVFNPT EEEF -RDTLSYISS--LR DRAE PYGIC CV PPPSW
At <i>JMJ14</i>	KITARWNPSEACRPLVDDAPIFYP TNE DF-DDPLGYIEK--LR S KAESY GIC RI V PPVAV
Gm <i>JMJ14-like</i>	KISARWDPVEASRP IEE APVFYPT IEEF -DDT LSY IAK--IR PLA EPHG IC RI V PPACW
At <i>JMJ15/MEE27</i>	KVLARWDPANEKRPDI GEAPV FHP TSEEF -EDTLAYIEK--IR PLA ESY GIC RI V PPSNW
At <i>JMJ18</i>	KVVARWLPDEAQRPIINDAPVF TPS LEEF-VDPLAYIEK--IR PLA EPY GIC RI IP PS TW
Sm <i>SELMODRAFT_170257</i>	--MTKYRGEAGR APNLD PAPVFFPT EEEF -QDTLKYIEK--IR PL VEPY GIC RI V PPKSW
Os <i>JMJ704</i>	TTNGK WHP HE SYR PEIDDAPVF TP TEEF-KDPIRYITS--IR QA EKY GIC RI V PPSSW
Pp <i>PHYPADRAFT_136290</i>	-VLATWRPDAGRRPCIDEAPVFYPT EEEF -KDPLRYIAS--IR ARA EPY G VC R VPPQLW
At <i>JMJ16/PKDM7D</i>	KVTARWHPDEARRPDLEADAPVFYPS EEEF -EDTLNYIAK--IR PE A E KY G IC RI VPPPSW

Sc JHD2 CPPLSIDME----NFTFQPRIQNLENLDDLKNCRLFFMKQLNNFKRSVKDPSKLILREP
Dm lid SPPFAVDVD----KLRFPVRVQRLNELEAKTRVKLNFLDQIAKFWELQGSSLKIPMVER
Hs JARID1B -----
At JMJ17 -----SCNKGW-----HIYCLSPPLKHI-----
Ol JMJ3520 QPEFCLPGK---EKLRFRTRIQALNELQNRPAQPSA-----
At JMJ19 KPPCLLKEKQIW-EASTFFPQVQLFGIQTENRKIKKE-----
At JMJ14 RPPCPLKEKKIW-ENSKFPTRIQFIDLLQNRPIKKS-----
Gm JMJ14-like APPCPLKEKDLW-ENTFFPTRIQIIDLLQNRPFMRKK-----
At JMJ15/MEE27 SPPCRLKGDSIW-KNKNFPTRVQFVDLLQNRGPVKKK-----
At JMJ18 KPPCRLKEKSIW-EQTKFPTRIQTVDLLQNRPEMKKK-----
Sm SELMODRAFT_170257 RPPCSLKDAG--ETVRFSTRVQKIHKLQVREPTTSS-----
Os JMJ704 RPPCSLKEKNFW-ECTEFNTRVQVQDKLQNRPEPTKKK-----
Pp PHYPADRAFT_136290 RPPCPLRGDSVEAQNMEFPTRVQVHKLQIRQPTTKV-----
At JMJ16/PKDM7D KPPCPLKEKQVW-EGSKFTTRVQRVQDKLQNRSSMKKI-----

Sc JHD2 YTIVEYSDSTHASEILKKKVYFYDVFSELIKDNRTLTD QSFRRKLFKFRDISQLRGDIS
Dm lid KALDLY-----TL-HRIVQEEGMEQTTKD-----R
Hs JARID1B -----
At JMJ17 -----
Ol JMJ3520 -----
At JMJ19 -----
At JMJ14 -----
Gm JMJ14-like -----
At JMJ15/MEE27 -----
At JMJ18 -----
Sm SELMODRAFT_170257 -----
Os JMJ704 -----
Pp PHYPADRAFT_136290 -----
At JMJ16/PKDM7D -----

Sc JHD2 LWRTISKKFNVPI----GLLKEIFEKYIASYYIFLHSLNE-----
Dm lid KWAKVANRMQYPPSSKSVGATLKAHYERILHPFEVYTSKVLGPTPTSSGSGSTPVKLEDG
Hs JARID1B -----MGFAPGKAVGSHIRGHYERILNPNLFLSGDSLRLCQ-----KPNLTDT
At JMJ17 -----
Ol JMJ3520 -RA-----
At JMJ19 -VD-----
At JMJ14 -T-----
Gm JMJ14-like -S-----
At JMJ15/MEE27 -TP-----
At JMJ18 -P-----
Sm SELMODRAFT_170257 -HG-----
Os JMJ704 -SQ-----
Pp PHYPADRAFT_136290 -WS-----
At JMJ16/PKDM7D -SK-----

Sc JHD2 -----
Dm lid GGTDYKAHEIPTRQOIAPPNETNTR-RSKRFGNSNASCGLSGVTPTTK-----PSAGVF
Hs JARID1B KDKEYKPHDIPQRQSVQPSSETCPPARRAKRMRAEAMNIKIEPEETTEARTHNLRRRMGCP
At JMJ17 -----
Ol JMJ3520 -----RA---KMLEEEK-----NGVK
At JMJ19 -----A-----
At JMJ14 -----K-----TKRKR
Gm JMJ14-like -----S-----GRKRKR
At JMJ15/MEE27 -----K-----GRKRKR
At JMJ18 -----K-----SRKRKR
Sm SELMODRAFT_170257 -----KK---S-----
Os JMJ704 -----PR---VQKRKR-----
Pp PHYPADRAFT_136290 -----PT---KLASKRR-----
At JMJ16/PKDM7D -----LP---NQMRKKK-----

Sc JHD2 -----NVHTALHADQYPKSLSDDEDDFDL-----GPDNSGSDFEEDDDDAC
Dm lid -VKTETKEEFKRDLLSSFNNAVNSGGSPLATGTTANTRGASQKKGGEPALIVDPLMKYIC
Hs JARID1B TPKCENEKEM---KSSIK---QEPIERKDYI--VENEKEPKRSKKNATNAVDLYVC
At JMJ17 -----PLGN-----
Ol JMJ3520 STKNQ-----GVASGGRMSGGR-----
At JMJ19 -----
At JMJ14 -RRIS-----KIGYTRRKRDSG-----
Gm JMJ14-like -RKHS-----KTGTCRRKPANA-----
At JMJ15/MEE27 -GKYS-----RTVAPKKRNGSV-----
At JMJ18 -RRNS-----RMGSSKRRSGSS-----
Sm SELMODRAFT_170257 -----RPKVS-----
Os JMJ704 -RKRL-----RFGMTHRRP-----
Pp PHYPADRAFT_136290 -RGRA-----TIGR---MGGL-----
At JMJ16/PKDM7D -RKCM-----KMGMSV TNGMG-----

Sc JHD2
Dm lid
Hs JARID1B
At JMJ17
Ol JMJ3520
At JMJ19
At JMJ14
Gm JMJ14-like
At JMJ15/MEE27
At JMJ18
Sm SELMODRAFT_170257
Os JMJ704
Pp PHYPADRAFT_136290
At JMJ16/PKDM7D

IVCRKTNDPKRTILCDSCDKPFHIYCLSPPLERVPSGDWICNTCIVG-----NGYYGFTQ
HICNRGDVEESMLLCDGCDSDSYHTFCLLPPLTSIPKGEWLCPRCVVEVSKPQEAFFGFEQ
LLCGSGNDEDRLLLCDGCDSDSYHTFCLIPPLHDVPGKDWRCPKLAQECQKQEAFFGFEQ
-----WYCLECLNTEDETFGFVP
-----MGASAQADADAVAKEYGFQQ
-----DSNDAASEGVLQCRVER
-----CDTASSGSSDSEKGFQFQT
-----A-SEAKNASESEKFGFQS
-----SKSVSTPKATEEENFGFES
-----P-AESTSSPEAEKFGFNS
-----KILTFTPQAAQQQEFFGFEP
-----SANTSEDCADAEKFGFQS
-----AACTTSPPIINDQPEYFGFWP
-----DPCSASTG-MNELETFGFEP

Sc JHD2
Dm lid
Hs JARID1B
At JMJ17
Ol JMJ3520
At JMJ19
At JMJ14
Gm JMJ14-like
At JMJ15/MEE27
At JMJ18
Sm SELMODRAFT_170257
Os JMJ704
Pp PHYPADRAFT_136290
At JMJ16/PKDM7D

DTHDYSLPFEQYCKRQNSRLLPAR-----KLSIDLEEMFWSLVTKN
AEREYTLQQFGQMADQFKQYFRKPVHLV-----PTMVEREFWRIVSSI
AARDYTLRTFGEMADAFKSDYFNMPVHMV-----PTELVEKEFWRLVSTI
G-KCLLLEDFKRIADRAKRWFGSGT-----VSRTQIEKKFWEIVIEGS
G-QRHNLATLERYSKYFKRKYFSKNGKPVV-----NVTVKMEGEFWRLIEDN
G-PGYTLKSFKNFADTYKKSHPGKDEVLGSENS-SPSLKPNELIVADIEKEYRQIVESP
G-PDFTLEEFQKYDEYFKECYFQSEDHPGSKA---SENKKFKPKVKDLEGEYWRIVEQA
G-SDFTLKDFQLYADDFKCYFGLRDTNGDRIVSDNNHQKIWEPSSEEEIEGEYWRIVIEQP
G-PEFTLEKFEKYAQDFKDSYFERKDNVGD-----PSVEEIEGEYWRIVIEKE
G-SDFTLDEFEKYALHFKDSYFEKKDSSGGDI-----VKWTPSVDDIEGEYWRIVEQP
G-PSFTIKFEFEAYADELKEKYFQAGEEGD-----TSRLDPSVEQIEREFWRIVERP
G-SDFTLDEFQKYADEFKQYFQIGKSDIPLSEIKKKKNWQPSVDEIEGEYWRIVVCP
G-DPFPLRAFENYANDFKSQYFRIPERQS-----SEPDEPTVNMIEGEYWRIVEQA
G-PGFTLKDFQKYADEFKAQYFKKSETSTDDKCKVDNSIDCWEPALDEVEGEYWRIVDKA
: : : : : * : : :

Sc JHD2
Dm lid
Hs JARID1B
At JMJ17
Ol JMJ3520
At JMJ19
At JMJ14
Gm JMJ14-like
At JMJ15/MEE27
At JMJ18
Sm SELMODRAFT_170257
Os JMJ704
Pp PHYPADRAFT_136290
At JMJ16/PKDM7D

RRSS-----LTTVKYGADIHNELPQGITGFPTRFIPKNINGDELIDYLK
DEDV-----TVEYGADLHTMDHG--SGFPFKSS-LYL-----LPGDQE
EEDV-----TVEYGADIASKEFG--SGFPVRDGKIKL-----SPEEEE
GGEV-----EVMYGNLDLTSVYG--SGFPRIQDQPE--SVEADIWDE
KGRS-----VEVIYGAD IATMDVG--SGFAKKGASACP-----PGQER
LIEI-----GVLYGNLDLTATFG--SGFPPLSAPSE-----S-S
TDEV-----EVYYGADLETKKFG--SGFPKYKPG-YP-----ISEADQ
TDEV-----EVYYGADLETGALG--SGFPKASS--LT-----KSESDQ
TNEV-----KVLVYGTDLLENPILG--SGFSKGVKIPTR-----RNDMDK
TDEV-----EVYYGADLENGVLG--SGFYKRAEK-FT-----GSDMEQ
SEQIEARLLRLCYHLCLTPDFQVLYGAD IETNVFK--SGFPKLATVA-----NKQATP
TDEV-----EVDYGADLDTSMFS--SGFSKLSSDS-----NRRDP
TEQI-----EVLYGADVETGKFG--SGFPKAPLGS-----EAAATH
TEEI-----EVLYGADLETGVFG--SGFPKISSHNA-----SSSEDK
* * * * : . : **

Sc JHD2
Dm lid
Hs JARID1B
At JMJ17
Ol JMJ3520
At JMJ19
At JMJ14
Gm JMJ14-like
At JMJ15/MEE27
At JMJ18
Sm SELMODRAFT_170257
Os JMJ704
Pp PHYPADRAFT_136290
At JMJ16/PKDM7D

```

-----
ELLEFVQEIIDNLCCTIIDEGASVRELLVLGKQFVERSESQQLSLESLESELETLINEGS
ELRQFVTQLYALPCVLSQTPLLKDLLNRVEDFQQHSQK--LLSEETPSAAEQDLLDVSF
--ELLR--VNPVPCFNNSGYLKLKDYAEEARKLSEKIDSALS---SSPTITQLELLHSEVS
-----
--QKQ-----
--ERLQ-----SDLT--SNK--EVQLKQDG--
--EIVQFESHPTCAAYDSIDSRHD-----NNNDKNLI--TDS--KDKVDQAG--
-----SGII--V--KEKQVQEE--
-----SSVI--S--EEKKLEK--
-----GLSE-----SRSW-----
-----YQQEKGIQI-----ASE-----KH-----
-----QPSEGGIICMT--AVKSASG-----KKNSQSLP--N-----DVIILLSD--
-----

```

Sc JHD2
Dm lid
Hs JARID1B
At JMJ17
Ol JMJ3520
At JMJ19
At JMJ14
Gm JMJ14-like
At JMJ15/MEE27
At JMJ18
Sm SELMODRAFT_170257
Os JMJ704
Pp PHYPADRAFT_136290
At JMJ16/PKDM7D

```

-----
SLRIELQQLDLLQKRLKQCKWYKRSQGLRETSSKLTQDVKNLLHIAAADLDPDTPYVVK
EFDVELPQLAEMRIRLEQARWLEEVQQACLDPSSITLDDMRRLIDLGVG-LAP-YSAVEK
RSPISLKKHEILSKKISSAKMLAKRAKRYLTDAPPGIEMDALFKLNSEMLELHVQLPE-
-----
DSDVNR-----HGHE-----SER-----NHVHGITDKSAVTDVVKL--
SLDLNL-----DVIS-----GENE-----NHLHLIADNHNKGVSVVEE-
CFDLNG-----ECNK-----SS-----
SFDLNI-----DLEM-----DY-----QE-
-----DKN-----KMOVNLESPAT-----ASN-
-----
EYDIPRKRGSVRRDAISSGKKLEIRERP-----THVLALEASAK-----IAA-
-----

```

Sc JHD2
Dm lid
Hs JARID1B
At JMJ17
Ol JMJ3520
At JMJ19
At JMJ14
Gm JMJ14-like
At JMJ15/MEE27
At JMJ18
Sm SELMODRAFT_170257
Os JMJ704
Pp PHYPADRAFT_136290
At JMJ16/PKDM7D

```

-----
EMRKLQQIGADIEAWESQAAKYFRRLTQQHELGEIEQFLKSASDINGQVPSHGLLKD---
AMARLQELLTVSEHWDDKAKSLK-ARPRHSLNSLATAVKEIEEIPAYLPNGAALKD---
TEG-ILDLVKKS---ESARDKSNKVLGTGSLLENVEELLHEFDSSFSINVELNLIRQYH-
-----
--G-----
KVC-CSEAKKEE---DIMELCGEGNLSNLFVSLKT-----DFSSCSRG--
EIC-----
DVK-----EAA-----
-----FAACG-----
PSR-S-----KSDCS-----GSLSLNHS-----SELPSSRIQTGNST
PIC-Q---REG---DSLDRTR---NTISLPTN---DQKTMRRDVPSSSTSHA---
-----

```

Sc JHD2
Dm lid
Hs JARID1B
At JMJ17
Ol JMJ3520
At JMJ19
At JMJ14
Gm JMJ14-like
At JMJ15/MEE27
At JMJ18
Sm SELMODRAFT_170257
Os JMJ704
Pp PHYPADRAFT_136290
At JMJ16/PKDM7D

```

-----
--ALRKAREWLRAVE-----QLQQNNHVITYCH-----
--SVQRARDWLQDVE-----GLQAGGRVPVLD-----
--V--DTLSWISR--FNDVMVDVREGKDQ-----
-----
-----VGKGFDEKKISVESQNPHSV-----SDVGCSELAKKVDGCL-----
--VRNYCTFDGSKI-EKDLQVDSDSGKQH-----SNLFEREVIVTTHTST-----
-----S-----
--EVLNPAPGKGL-----MVVEPRS-----
LASITTE-KLFGVDIK-SNLAQ-----SSDGQVSQLAKPSSS
--EVNAEATGLTQDIC-NRMATNSHGGKPTSCSKNSGGLAIVDQVVDGTRSSSGTPSCS
-----

```

Sc JHD2
Dm lid
Hs JARID1B
At JM17
Ol JMJ3520
At JM19
At JM14
Gm JMJ14-like
At JM15/MEE27
At JM18
Sm SELMODRAFT_170257
Os JMJ704
Pp PHYPADRAFT_136290
At JM16/PKDM7D

-----TLEAMIERGLNIPIQLEELSRMQGHLNSAHQWKDN---TACAFLLK
-----TLIELVTRGRSIPVHLNSLPRLETVAEVQAWKEC---AVNTFLTE
-----RKLIS---DLSLLRDGASLGIQVEGLPLV---EVELKKASCREKARTVYTARKS

-----GGKDQNAATNRLSLSVELLSSG---SLVVKKLWCSKQAIYPKGFKSR
-----SLMDESCLVQMFGTSVKLVSLG---SVVYKGLWCSKHTLYPKGFKTQ
-----EDASIMDLAAYHVEPINLG---FLVVGKLVKCNKHAIFPKGFKSR
-----TSGGELTASENLGVSEVINLG---FLIFGKLVKCNKYAIFPKGFRRS
-----GGGPRVARVIRPSKAFSPNIELVRTG---RLVLKPGWHTKHAHPAGFRTR
QTDE-----VSKPAI-----AKYTVELLDSE---TMMIGKKWCNQAIIFPKGFKSR
QNNSPDRFIRQKGPRIAKVV---RRINCNVEPLSYG---CVLSGKSWCSRRRAIFPKGFRRS

Sc JHD2
Dm lid
Hs JARID1B
At JM17
Ol JMJ3520
At JM19
At JM14
Gm JMJ14-like
At JM15/MEE27
At JM18
Sm SELMODRAFT_170257
Os JMJ704
Pp PHYPADRAFT_136290
At JM16/PKDM7D

GTFYTLLEVLMPRSDAINIDSDLKPRFQDDFL-----KEKNPAE
NSPYSLEVLCPKRCDIGLLGLKQKRLKEPLPNGKKKSTKLESLSDLERALTESKETAS
LDF-----IEQLLSEAV---I-----LHIEEEEIFVE

VKFSLVLDPTNLTN-----YISEVLDAGLLGPL-----FRVSVEDYPTF
VNFFSIVDPKRICS-----YISEVIDAGFLGPL-----FKVTMEECNSE
VKFYNVQDPMRISY-----YVSEIVDAGLLGPL-----FKVTLEESQDE
VKFYNVLDPTRMSN-----YISEVLDAGLMGPL-----FRVTLEESDPE
VQFYDYLDLPQACY-----YMSEILDCADGKPL-----FKVSMEGRPHE
VTFHSLDPTTRTCC-----YISEVLDAGLLGPL-----FRVTVEGLPEV

VKYINILDPTNMCF-----YISEILDAGRNSPL-----FMVYLESNPSE

Sc JHD2
Dm lid
Hs JARID1B
At JM17
Ol JMJ3520
At JM19
At JM14
Gm JMJ14-like
At JM15/MEE27
At JM18
Sm SELMODRAFT_170257
Os JMJ704
Pp PHYPADRAFT_136290
At JM16/PKDM7D

IVASFKAHEEQELLDNR---ELRRQN---M-NK---NPMRDMFCLCKSEFRNL-----
AMATLGEARLREMEALQ---SLRLANEGK-LLSPLQVDIKICLCQKAPAAP-----
ISGILSTARCWEERASTILENETQMYELKD---LV-----RMSVNIDAV-----

NFSNVSAEKCWQMVTRQLKLEIIKKCDQPVSSLTSLQPLESINGLEMFGLSPHV-----
AFTDTSADNCWESVLKRLHHEIMRQKSLGELELPPFELLKLSINGHRMFGFKLPSI-----
SFSYASPKQCWEMVLLRVKKEIMRRSN---QKQDVHMLESI DGLKMFGRSPFI-----
SFFNVSAQQCWEMVRRVKDTSTSL---GLPILPQFESINGLQMFGLSPSI-----
KIVCSSIDFCWQSVQEKVNSRIKQLRESGKANLPLRPPESLKGLEMFGLFTVPSVVKVRS
SFTHTSPMQCWDVSRDRVNEEIAKQISFGKSGLPDFLSCNSLNGLEMFGLSSPI-----

VFVHMSPTRCWEMVRERVNQEI TKQH KAGKSDLPLLPSPGSPDGFEMFYSSPAI-----

Sc JHD2
Dm lid
Hs JARID1B
At JM17
Ol JMJ3520
At JM19
At JM14
Gm JMJ14-like
At JM15/MEE27
At JM18
Sm SELMODRAFT_170257
Os JMJ704
Pp PHYPADRAFT_136290
At JM16/PKDM7D

-----MFNCQLCRDWFHEDCVPPSATNQNGIVNGGSGPGTNRP
-----MIQCELCRDAFHTSCVAVPSISQ-----GLR
-----LPTLQGIENTISSAETWLKSEPFLSATS-----

-----IKVVEALDPKHOLEEYWNQKAVKLFGAEP-----
-----IQAIEAQDPShLCVEYWNHKVAPSGSV-----
-----VQATEALDPNHGQVEYWNHKNEKDSLE-----
-----VQAIEALDPNHLVEYWNHKNTSSDS-----
LALLSFSLTWRGNSLQAVETLDRHSCVEYWMERKTNDLWS-----
-----IKEIEALDPCHQCLDYWLSRVSSVGT-EL-----

-----VQAIEALDVNRVCTDYWDSRPYSRPQVQF-----

Sc JHD2 -----
Dm lid KWLCPSCVRSKRPRLEFILPLLVLQQLPIRLPEDEALRCLAERAMNWQDRARKALSSPD
Hs JARID1B IWLCPHCRRSEKPPLEKILPLLASLQIRVRLPEGDALRYMIERTVNWQHRAQQLSSGN
At JMJ17 SMASSPCSMLELPVL--KDLVTQAKLLNVQLQEPRILETLLNCERWQCDNHQLQETE
Ol JMJ3520 -----
At JMJ19 -----
At JMJ14 -----IKE
Gm JMJ14-like -----
At JMJ15/MEE27 -----
At JMJ18 -----
Sm SELMODRAFT_170257 -----
Os JMJ704 -----PSE
Pp PHYPADRAFT_136290 -----
At JMJ16/PKDM7D -----PAN

Sc JHD2 -----
Dm lid VSAAQEAIMAQQQKRRSEGGAGVGNISSPRKPRRRGSLTKEASGSTESDADDDDEDEEC
Hs JARID1B LKFVQDRVGSGLLYSRWQA--SAGQVSDTNKV-----SQPPGTTSFSLPDDWD--N
At JMJ17 DLLDNAKID-----DGTHS-----NILPKIM-----
Ol JMJ3520 -----
At JMJ19 -----
At JMJ14 GEKDDTEKG-----GASDPS-----
Gm JMJ14-like --VDNFPFG-----SSSLG-----N-----INTKIFGI-----
At JMJ15/MEE27 --MKDCFM-----SNSQS-----L-----SKARLFGV-----
At JMJ18 --KDFHIS-----SNCSAS-----L-----TKGKLFV-----
Sm SELMODRAFT_170257 -----
Os JMJ704 SVMAMVND-----ST-----NPP-----IK--LLGI-----
Pp PHYPADRAFT_136290 -----
At JMJ16/PKDM7D PLLREA-NT-----SGRSNVGNLQLNPG-----HHISPTGI-----

Sc JHD2 -----
Dm lid RLRIVEDNFSNDEDEPRTAPATSTVNSDLLKLLSDSEIENLLDLMMEGDLLEVSLDETQE
Hs JARID1B RTSYLHSPFSTGR-----SCIPLHGVSPVEVNELLMEAQLLQVSLPEIQE
At JMJ17 -----DLITRVDSA-----RRSGLALGLNFDLQPK
Ol JMJ3520 -----
At JMJ19 -----
At JMJ14 -----LDR-----DTRLLRGLLKKATPEELVM
Gm JMJ14-like -----DLIKQEKDN-----IILEEMKSIQGASPELRT
At JMJ15/MEE27 -----DLN-----
At JMJ18 -----DLM-----
Sm SELMODRAFT_170257 -----
Os JMJ704 -----E-----INRRESEQSSS
Pp PHYPADRAFT_136290 -----
At JMJ16/PKDM7D -----NSILKVLFFKKASMEELSS

Sc JHD2 -----
Dm lid LWRILETMPPTLLQAEAMERVVQYMQRRQOHTNPLPTSGAEDSNDLSMVQNSPNSNSNS
Hs JARID1B LYQTLAKPSPAQQTDRS-----SPVRP-----SSEKNDCC
At JMJ17 LRTASL-----KLGWCC
Ol JMJ3520 -----
At JMJ19 -----
At JMJ14 MHGLLC-----GETRNT
Gm JMJ14-like MHKLII-----SDAQCC
At JMJ15/MEE27 -----
At JMJ18 -----
Sm SELMODRAFT_170257 -----
Os JMJ704 FNNSCV-----R-----
Pp PHYPADRAFT_136290 -----
At JMJ16/PKDM7D LQEVLS-----ET-----

Sc <i>JHD2</i>	-----
Dm <i>lid</i>	TTTPGKQR-----AV-----QSAR-----
Hs <i>JARID1B</i>	-----
At <i>JMJ17</i>	ATEPERPSLNQRRTRMVATDAAVNDLKWKTRKH IKRTTKRSPQVHILPWFFT
Oi <i>JMJ3520</i>	-----
At <i>JMJ19</i>	-----
At <i>JMJ14</i>	-----
Gm <i>JMJ14-like</i>	-----
At <i>JMJ15/MEE27</i>	-----
At <i>JMJ18</i>	-----
Sm <i>SELMODRAFT_170257</i>	-----
Os <i>JMJ704</i>	-----
Pp <i>PHYPADRAFT_136290</i>	-----
At <i>JMJ16/PKDM7D</i>	-----