Pachycladon species evolved traits to adapt to New Zealand habitats

A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Plant Biology at Massey University, Manawatū, New Zealand

Yanni Dong

2021

Abstract

P. cheesemanii is a close relative of *A. thaliana* and is an allotetraploid perennial herb that is widespread in the South Island of New Zealand. It grows at altitudes of up to 1,000 m where it is subject to relatively high levels of UV-B radiation. However, to date the origin of this species and the mechanisms underlying its tolerance to its harsh living environmental conditions such as moderate–high UV-B radiation, cold and drought is unclear.

To gain the first insights into how *Pachycladon* copes with UV-B stress, I sequenced the *P. cheesemanii* genome and compared the UV-B tolerance of plants from Wye Creek (~300-m altitude) and Kingston (~500-m altitude) with that of *A. thaliana* from Col-0 (~100-m altitude) and Kondara (1,000–1,100-m altitude). A high-quality draft genome of *P. cheesemanii* was assembled with a high percentage of conserved single-copy plant orthologues. A synteny analysis involving genomes from other species of the Brassicaceae family suggested that the two subgenomes of *P. cheesemanii* may have the same origin as species from Brassicaceae Lineage I and EII. While UV-B radiation caused greater growth reduction in *A. thaliana* Col-0 and Kondara than in *P. cheesemanii* Wye Creek, growth was not reduced in *P. cheesemanii* Kingston. Homologues of the *A. thaliana* UV-B radiation response genes have multiple copies in *P. cheesemanii*, and an expression analysis of those genes indicated that the tolerance mechanism in *P. cheesemanii* Wye Creek and Kingston may differ from that in *A. thaliana*. Although the *P. cheesemanii* genome shows close similarity with that of *A. thaliana*, the uniqueness of the strongly UV-B-induced UVR8-independent pathway in *P. cheesemanii* may help this species to tolerate relatively high UV-B radiation.

Next, to understand the different stress responses of *A. thaliana* and *P. cheesemanii*, I designed a project to build multiple-stress transcriptomes for *A. thaliana* and *P. cheesemanii*. Since plant responses to salt and drought are related and have overlapping mechanisms, and salt stress can easily be applied in the laboratory, high salinity rather than drought stress was used to stress *A. thaliana* and *P. cheesemanii* plants in this study. Transcriptomes of *A. thaliana* and *P. cheesemanii* plants in response to cold, salt and UV-B radiation stresses were created. A high-quality *de novo* transcriptome assembly of allopolyploid *P. cheesemanii* was obtained by using multiple assemblers with further downstream processing. Differential expression analysis revealed a strong bias, in terms of the number of DEGs, towards upregulation in both *A. thaliana* and *P. cheesemanii* in responding to salt stress, as well as in *P. cheesemanii*'s cold and UV-B treatment responses. Meanwhile, in each species, a number of DEGs was shared between stresses, although the majority were unique in responding to each stress in upregulation and downregulation, respectively.

Further, GO enrichment analysis revealed that these responsive genes were involved in some biological processes shared by *A. thaliana* and *P. cheesemanii*. Immune system processes, response to stimuli, signalling, developmental processes, growth, negative regulation of biological processes, multi-organism processes, biological regulation, secondary metabolic processes, cell communication, and cellular aromatic compound metabolic processes were common in the responses of both *A. thaliana* and *P. cheesemanii* to all three stresses. In both *A. thaliana* and *P. cheesemanii*, a number of these biological processes were also stress specific. First of all, in *A. thaliana*, cold stress may easily affect photomorphogenesis in cold responses, while the majority of the *P. cheesemanii* unique cold responses occurred in root differentiation, floral whorl

development and regulation of programmed cell death. Second, *A. thaliana* responses to salt stress affected starch metabolism and lipid modification, whereas disaccharide and polysaccharide metabolism, as well as microtubule structure, were affected by salt stress uniquely in *P. cheesemanii*. Finally, *A. thaliana* responses to UV-B radiation involved a combination of physical and biological defences, including cell wall modification defence, stomatal movement, vitamin B6 metabolic processes and oxygen metabolic processes. In contrast, seed germination biological regulation was affected in *P. cheesemanii* under UV-B radiation stress. Further, *P. cheesemanii* had a larger number of unique GO enrichments in cold responses than did *A. thaliana*.

There was a wide range of crosstalk among the biological processes in responding to the three stresses in *A. thaliana*, while only one main cluster was identified in crosstalk for the three stress responses in *P. cheesemanii*. In this main cluster, the biosynthetic process for anthocyanins was in the centre position, and it was found that multiple stress-responsive biological processes probably involved anthocyanins in *P. cheesemanii*. Thus, although the *P. cheesemanii* genome shows close similarity with that of *A. thaliana*, it appears to have evolved novel strategies such as a highly UV-B-activated UVR8-independent pathway, allowing the plant to tolerate relatively high UV-B radiation. The stress process is highly conserved in plant species under various stresses, but species also develop a few unique characteristics that may help them adapt to their own ecological niche and survive particular environmental stresses.

Acknowledgments

Here, I would like to give special acknowledgements to lovely people who kindly provided me with their help and support. Without them, this project would never be completed.

First of all, I'd like to thank my main supervisor, Dr Paul Dijkwel. I cannot thank you enough for helping my project design, inspiring me, supporting my experiments and patiently helping to perfect my thesis. During this journey, I have encountered lots of personal difficulties, and without his support, I might have given up. From him, I learnt not only how to be a successful PhD but also what a scientist should be. He is my supervisor and life teacher. Meanwhile, special thanks to Erlinde Dijkwel. She kindly helped with my thesis editing process and made me feel the warmth which you get from your family.

Thank you so much my co-supervisors -- Dr Dave Wheeler, Dr Richard Macknight, and Professor Jo Putterill. They gave me a lot of precious suggestions, advice and constructive feedback for my project, experiments and thesis writing.

Many thanks to Dr Tsanko Gechev, Dr Jason J. Wargent and Professor Bernd Mueller-Roeber. They are great collaborators and provided me with lots of help to work on this project and our resulting publications. Meanwhile, I want to give special thanks to Dr Saurabh Gupta from Dr Tsanko Gechev and Professor Bernd Mueller-Roeber's laboratory, who supported me a lot while I was learning bioinformatics. He also helped a

lot with the computational part of this project. I think that just a “thank you” would never be enough to express all my gratitude.

To my lovely friends, Nikolai Kondratev, Ramadoss Dhanushkodi, and Xi Xu who gave me support, help, and companionship in the past four years. We experienced all the hardships, enjoyed the parties, and shared happiness and life difficulties. I know that I can always rely on them.

Finally, I want to thank my lab manager, Tina Sehrish, lab members, Rixta Sievers, Asmat Karim and Aakansha Kanojia. They are lovely people, and always helped me with a warm heart. I was blessed to have them around me in the past four years.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

abscisic acid	ABA
Arabidopsis Biological Resource Center	ABRC
ancestral crucifer karyotype	ACK
Amplified fragment length polymorphism	AFLP
Apetala 2	AP2
blue copper binding protein	BCB
basic helix-loop-helix	bHLH
basic region leucine zipper	bZIP
degrees Celsius	°C
C-repeat/DRE-Binding Factors	CBFs
calcium-dependent protein kinase	CDPK
CHALCONE SYNTHASE	CHS
constitutively photomorphogenic 1	COP1
Compound spring embedder	CoSE
cyclobutane pyrimidine dimers	CPD
counts per million	cpm
calcium-responsive protein kinase 1	CRLK1
constitutive triple response 1	CTR1
Cysteines	Cys
direct repeat	DR
Dithiothreitol	DTT
Ein-3-like protein 1	EIL1
ethylene insensitive protein	EIN
EARLY LIGHT-INDUCIBLE PROTEIN1	ELIP1
Endoplasmic reticulum	ER
ethylene response element binding protein	EREBP
EPITHIOSPECIFIER MODIFIER	ESM
EPITHIOSPECIFIER PROTEIN	ESP
expressed sequence tag	EST
Ethylene	ET
Endoplasmic reticulum-associated receptors	ETR
flavanone 3-hydroxylase	F3H
fold change	FC
false discovery rate	FDR
flavonol synthase 1	FLS1
fad-linked oxidoreductase	FOX1
gibberellic acid	GA
genome block	GB
Gene Ontology	GO
galactinol synthase	GolS

Geranyl pyrophosphate synthase	GPPS
G	gram
growth-regulating factor	GRF
gene-set enrichment analysis	GSEA
hookless 1	HLS1
high expression of osmotically responsive genes1	HOS1
high-performance liquid chromatography	HPLC
heat shock transcription factor	HSF
elongated hypocotyl 5	HY5
HY5-HOMOLOG	HYH
inducer of CBF expression 1	ICE1
jasmonic acid	JA
amino acid conjugate +-7-iso-jasmonoyl-l-isoleucine	JA-Ile
jasmonate ZIM domain	JAZ
low expression of osmotically responsive genes 2	LOS2
lipid transfer protein	LTP
long terminal repeat	LTR
mitogen-activated protein kinase	MAPK
Mitogen-activated protein kinase kinase kinase 1	MEKK1
methyl salicylate	MeSA
Mitogen-activated protein kinase kinase 2	MKK2
mate-pair	MP
Mitogen-activated protein kinase 4	MPK 4
Mitogen-activated protein kinase 6	MPK6
Myeloblastosis	MYB
MYB transcription factor recognition sequence	MYBRS
reduced nicotinamide adenine dinucleotide phosphate	NADPH
Next-Generation Sequencing	NGS
nonexpressor of pathogenesis-related genes 1	NPR1
New Zealand Genomics Limited	NZGL
Reduced hyperosmolality-induced $[Ca^{2+}]_i$ increase 1	OSCA1
open stomata 1	OST1
primer-binding site	PBS
plant defensin 1.2	PDF1.2
paired-end	PE
primary ethylene response element	PERE
PHOTOLYASE 1	PHR1
pollen mother cell	PMC
type 2C protein phosphatase	PP2C
polypurine tract	PPT
PATHOGENESIS-RELATED	PR
phytoene synthase	PSY
related to AP2 4a	RAP2.4a
raffinose family oligosaccharide	RFO

Reduced height 3	Rht3
receptor-like protein kinase	RLK
reactive oxygen species	ROS
salicylic acid	SA
SA-binding protein 2	SABP2
SA methyl transferase 1	SAMT1
Skp1/Cullin/F-box	SCF^{COI1}
SNF1 sucrose nonfermenting 1-related protein kinase 2	SnRK2
salt overly sensitive gene 3	SOS3
small ubiquitin-related modifier	SUMO
transcription factor	TF
TGACG-Binding	TGA
URIDINE DIPHOSPHATE GLYCOSYLTRANSFERASE 74E2	UGT74E2
UV REPAIR DEFECTIVE 3	UVR3
UVB-resistance 8	UVR8
WRKY DNA-binding protein 30	WRKY30
Weighted correlation network analysis	WGCNA
Xeroderma Pigmentosum Complementation Group B1	XPB1
XERODERMA PIGMENTOSUM GROUP D	XPD

Chapter 1 General introduction

1.1 Pachycladon

Pachycladon is a polyploid genus of Brassicaceae, deriving from the last 1-3.5 million years in New Zealand. The genus includes nine perennial species restricted to the South Island of New Zealand as well as a single species endemic to Tasmania, Australia (Heenan, 2002).

1.1.1 Distribution of Pachycladon species across New Zealand

Pachycladon species are primarily distributed across the alpine regions of the South Island, with three dominant populations evident based on morphological traits, molecular phylogenetic trees, and the Amplified fragment length polymorphism (AFLP) (Heenan & Mitchell, 2003; Yogeeswaran, Voelckel, Joly, & Heenan, 2011) (**Figure 1-1**). The first group comprises *Pachycladon cheesemanii* and *Pachycladon exile* which differ dramatically in their distribution patterns. *P. cheesemanii* has a wide latitudinal and altitudinal range (from near sea level 10 m to 1600 m) and grows on diverse rock substrate (greywacke, haast schist, and basaltic and andesitic volcanic rocks). In contrast, *P. exile* is restricted to northern Otago at altitudes of less than 500 m on calcareous substrates such as limestone and volcanic rock. The second group is composed of three species restricted to the southern end of the South Island. *Pachycladon novae-zelandiae* usually grows on high altitude rocky outcrops between 1080 and 2031 m. *Pachycladon crenatus* grows on gneiss and *Pachycladon wallii* grows on low altitude rocky outcrops and bluffs. The third group is made up of species that favor shaded rocky bluffs and cliffs composed of granite, sandstone, marble, limestone, greywacke, and volcanic rock. *Pachycladon latisiliquum* is distributed in a narrow latitudinal range at 1036-1768 m altitude. Both *Pachycladon enysii*

and *Pachycladon fastigiatum* are located in the Southern Alps at an altitude between 975 to 2492 m, with this latter species restricted to the northern and southern boundaries of this region. *Pachycladon stellatum* is restricted to the southern Marlborough at 900-1371 m altitude, where it grows on greywacke. The last species in this group, *Pachycladon fasciarium* is found in eastern Marlborough and prefers limestone-based soils.

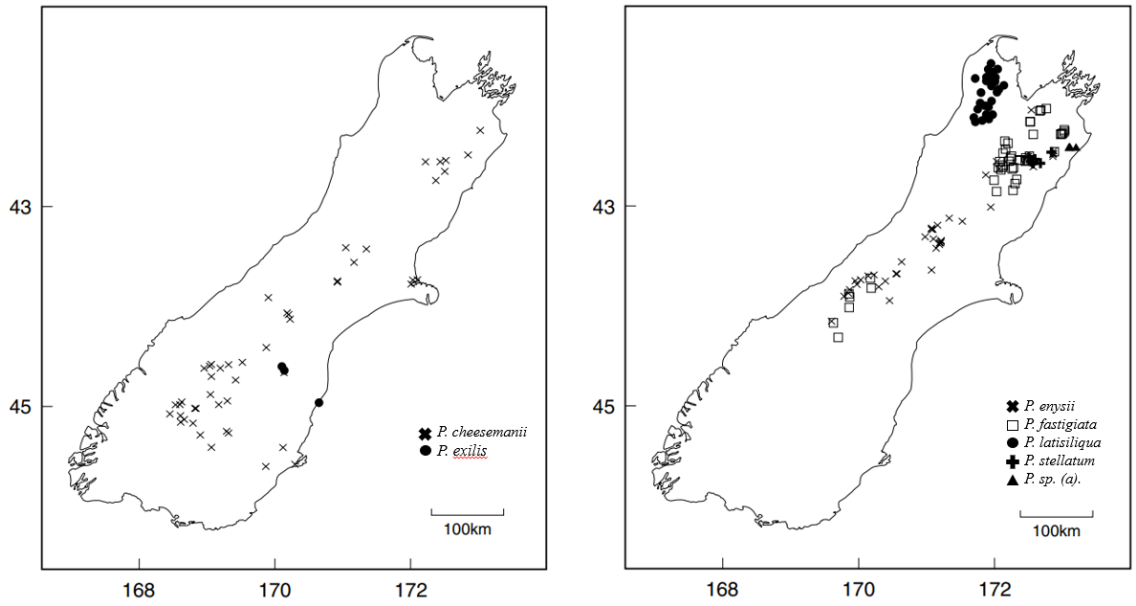

Figure 1-1. Distributions of Pachycladon species in the South Island of New Zealand.

Altitudinal range: *P. cheesemanii*, from near sea level 10 m to 1600 m; *P. exilis*, less than 500 m; *P. latisiliquum*, 1036-1768 m; *P. enysii* and *P. fastigiatum*, between 975 to 2492 m; *P. stellatum*, 900-1371 m. *P. sp. (a)* is an unnamed taxon. Modified from (Heenan & Mitchell, 2003).

1.1.2 Morphological classification of Pachycladon species

Several morphological characters have been used to define the three Pachycladon groupings in New Zealand (Yogeeswaran, Voelckel, Joly, & Heenan, 2011). The first group includes *P. exile* and *P. cheesemanii*, which have polycarpic, woody caudices, short aerial branches, heterophylly, slender and terminal inflorescences, terete siliques with uniseriate seeds that lack wings. The second group includes *P. novae-zelandiae*, *P. wallii* showing lobed leaves, lateral inflorescences, short laterally compressed siliques with biseriate seeds, and seeds with or without wings. The third group: *P. enysii*, *P. fastigiatum*,

and *P. latisiliquum*, and *P. stellatum* are monocarpic and have a stout and soft caudex with terminal inflorescences, serrate leaves, long and laterally compressed siliques with biseriate seeds, and seeds with wings. Although still part of the third group *P. fasciarium* is somewhat morphologically distinct from other members in that it has shorter and narrower leaves that are more linear and glabrous with fewer smaller teeth on the margin (**Figure 1-2**). The Tasmanian species *P. radicatum* is polycarpic, semi-woody caudex with decumbent branches, ovate leaves, lateral inflorescences, laterally compressed siliques, and winged seeds.

1.1.3 Various reproductive strategy of Pachycladon species

Almost all Pachycladon species flower in spring and mid-summer after experiencing cold weather in winter. Both polycarpic and monocarpic Pachycladon plants have six stamens and one pistil with varied flower structures, which influence the specific breeding systems of each species. For instance, the anthers of *P. cheesemanni* and *P. exilis* are short and grow away from the stigma, reducing the possibility of selfing. In contrast, the anthers of *P. novae-zalandiae* grow toward the stigma, increasing the probability of selfing. Another conservative indicator of breeding system is pollen-to-ovule ratio (the ratio of pollen grains to ovules per flower). The pollen-to-ovule ratio in typical outcrossing taxa is >3500:1, whilst this ratio is rarely greater than 1000:1 in selfing species. Based on the investigation of pollen-to-ovule ratio in six Pachycladon species, it has found that the pollen-to-ovule ratios of *P. fastigiata*, *P. stellate*, *P. latisiliqua*, *P. novae-zalandiae*, *P. exilis*, and *P. cheesemanii* are < 1000:1, suggesting that selfing is predominant reproductive system in these six Pachycladon species. Among them, *P. fastigiata* (339±156:1), *P. stellate* (525±97:1), *P. latisiliqua* (786±185:1) (±stdev) have higher

pollen-to-ovule ratios compared to other investigated selfing species (*P. novae-zalandiae* 181±79:1, *P. exilis* 74±23:1, and *P. cheesemanii* 76±15:1) coupled with their outcrossing flower structure, so they were considered having a mixed mating system (selfing and outcrossing) (P. Heenan & Mitchell, 2003). Also, the other three Pachycladon species, *P. wallii*, *P. stellatum*, and *P. fastigiatum*, have been reported exhibiting sexual dimorphism (Garnock-Jones, 1991; P. Heenan & Garnock-Jones, 1999; P. B. Heenan, Goeke, Houliston, & Lysak, 2012).

1.1.4 Metabolism divergence of Pachycladon species

Metabolomic and transcriptomic studies have also identified differences in metabolic pathways in the Pachycladon species. For example, several glucosinolate metabolism genes are differentially expressed in the in *P. fastigiata* and *P. enysii*. Two key genes that were identified in this study were *EPITHIOSPECIFIER PROTEIN GENE* (*ESP*), a marker for nitrile production, and *EPITHIOSPECIFIER MODIFIER GENE* (*ESM*), which is responsible for isothiocyanate production. In *P. enysii*, *ESP* is upregulated, whist *ESM* is upregulated in *P. fastigiata,* indicating that *P. enysii* would be more likely to produce nitriles, whist *P. fastigiata* would favor isothocyanates production (Voelckel et al., 2008). Differential expression of genes involved in glucosinolate hydrolysis have also been observed in the transcript profiling of *P. exile*, *P. novae-zelandiae* and *P. fastigiatum*, and *ESP* and related genes involved in nitrile- and epithionitirle-specifying activation were identified (Voelckel et al., 2010). In *P. novae-zelandiae*, major glucosinolates were converted into nitriles and epithionitriles rather than the isothiocyanates that are predominantly produced in *P. cheesemanii*. Through tag profiling, homologs to the genes functioning in isothiocyanate production were upregulated in *P. fastigiatum*, while the

increased expressions of homologous genes in nitrile production pathway were shown in *P. enysii* (Voelckel, Gruenheit, Biggs, Deusch, & Lockhart, 2012). The products of glucosinolate hydrolysis, nitriles and isothiocyanates affect the toxicity and palatability of plant tissues to herbivores and pathogens. This suggests interspecific biochemical differences in glucosinolate metabolism among Pachycladon species may be resulted from environmental selection pressure like herbivory, contributing to species diversification.

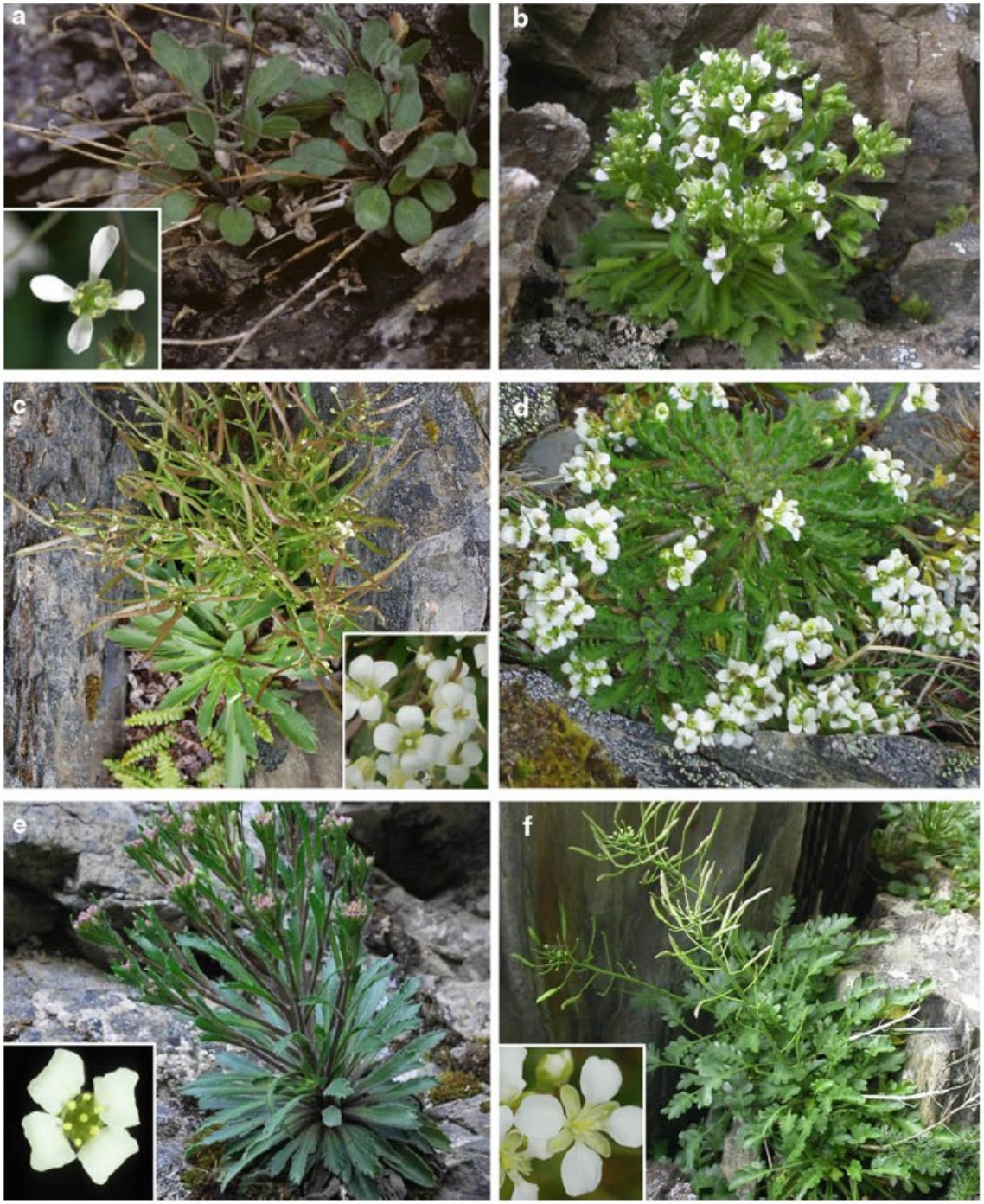

Figure 1-2. Morphological characters of Pachycladon species.

a. *P. cheesemanii*, b, *P. enysii*, c. *P. fastigiatum*, d. *P. novae-zelandiae*, e. *P. stellatum*, f. *P. wallii*. Reprinted by permission from Springer Nature: Springer, Wild crop relatives: genomic and breeding resources (pp. 227-249), Pachycladon. Yogeeswaran, K., Voelckel, C., Joly, S., & Heenan, P. B. Copyright 2011.

1.1.5 Evolution of Pachycladon species

From comparison of Pachycladon species and the derived ancestral crucifer karyotypes, it is apparent that the chromosome structure has experienced complicated rearrangement over time before allopolyploidy. FISH studies showed that karyotypes of four Pachycladon species (*P. enysii*, *P. novae-zelandiae*, *P. cheesemanii*, *P. exile*) showed high similarities to the ancestral crucifer karyotype (ACK) (n=8) (Mandakova, Heenan, & Lysak, 2010). However, dramatic chromosomal changes including inversions, translocations and centromere inactivation/losses were observed among Pachycladon species. For instance, Pachycladon chromosome 3 was generated in the way of an insertional chromosome fusion of two entire ACK chromosomes (AK2 and AK3). AK3 broke up at centromere to allow the insertion of entire AK2, which has been proposed to be a mechanism of descending dysploidy (Luo et al., 2009). Such chromosome reshuffling caused that paralogous genomic blocks did not lie on the same chromosome in Pchycladon species (**Figure 1-3**), which probably occurred even earlier than homeologous recombination of two ACK-like ancestor genomes.

The chromosome rearrangements before allopolyploidy mentioned above and the subsequent hybridization events lead to complicated attempts to define the origination of the Pachycladon karyotype. Phylogenetic analysis of Pachycladon species based on five single-copy nuclear genes indicated one of the genome copies was derived from an Arabidopsis lineage, whilst another is similar

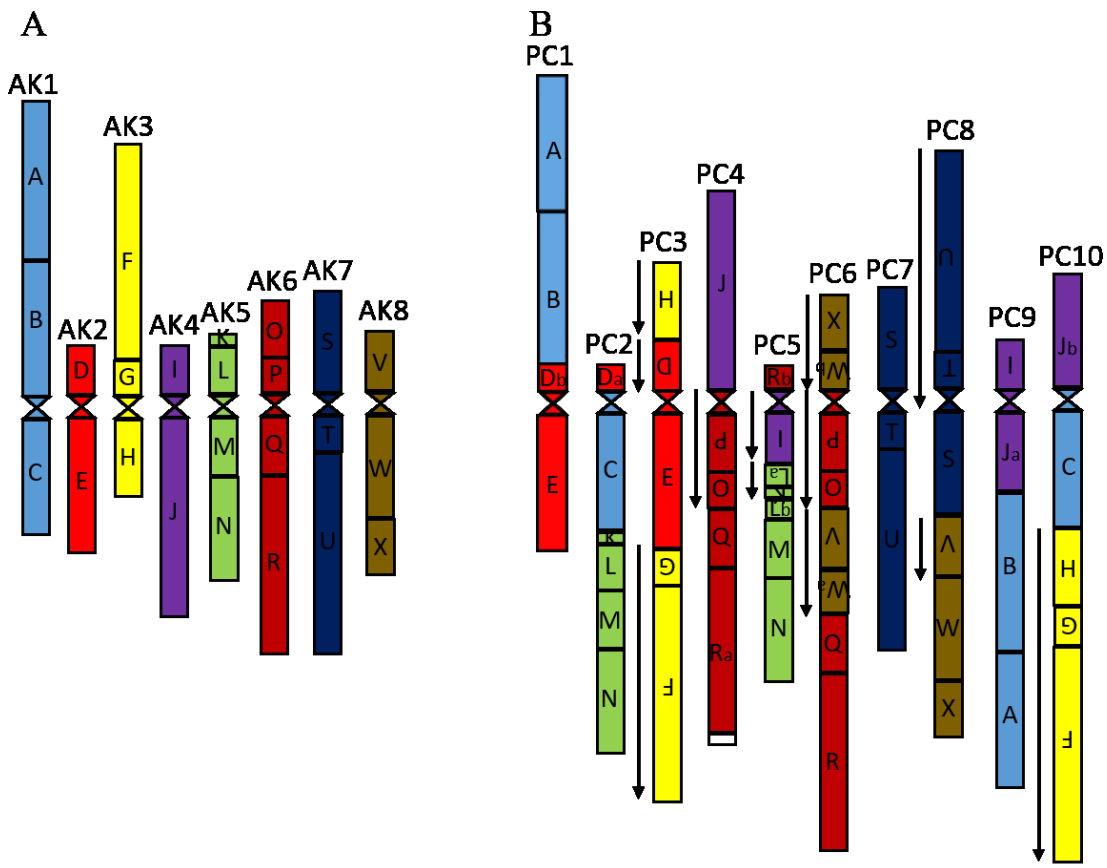

Figure 1-3. Ancestral crucifer and Pachycladon karyotypes.

A) The establishment of ancestral crucifer karyotype (ACK) (n=8) (AK1-8) is based on cytology and genetic maps of *Arabidopsis lyrata* and *Capsella rubella*, and colored genome blocks (GBs) A to X represent their positions on eight ancestral chromosomes. Modified from (Schranz et al., 2006); B) Pachycladon karyotype (PC1-10) was reconstructed by using chromosome painting on four Pachycladon species, *P. cheesemanii*, *P. enysii*, *P. exile*, *P. novae-zelandiae*. Downstream-pointing arrows show the inverse orientation of GBs relative to their positions on ACK chromosomes. Double triangles in the middle of chromosomes stand for active centromeres with different colors showing their positions on ACK chromosomes. Modified from (Mandakova et al., 2010).

to both Arabidopsis and Brassica lineages (Joly, Heenan, & Lockhart, 2009). However, after comparing 547 homeologous gene pairs of *P. cheesemanii* and *P. fastigiatum* with homologous of *Arabidopsis lyrata* and *Arabidopsis thaliana*, no set of genes showed significant higher identity to the reference genes, suggesting both copies may derive from

the same parental lineage (Gruenheit et al., 2012). Also, data from the nuclear marker *CHALCONE SYNTHASE GENE* (*CHS*) support both Pachycladon genome copies belong to Arabidopsis lineage, one of which stems from Pachycladon clade and another is from Smelowsikeae (Zhao, Liu, Tan, & Wang, 2010). Thus, if both two genome copies are from the same lineage, it is difficult to identify Pachycladon species' maternal and paternal ancestral lineages due to the high similarity of genome copies (**Figure 1-4**).

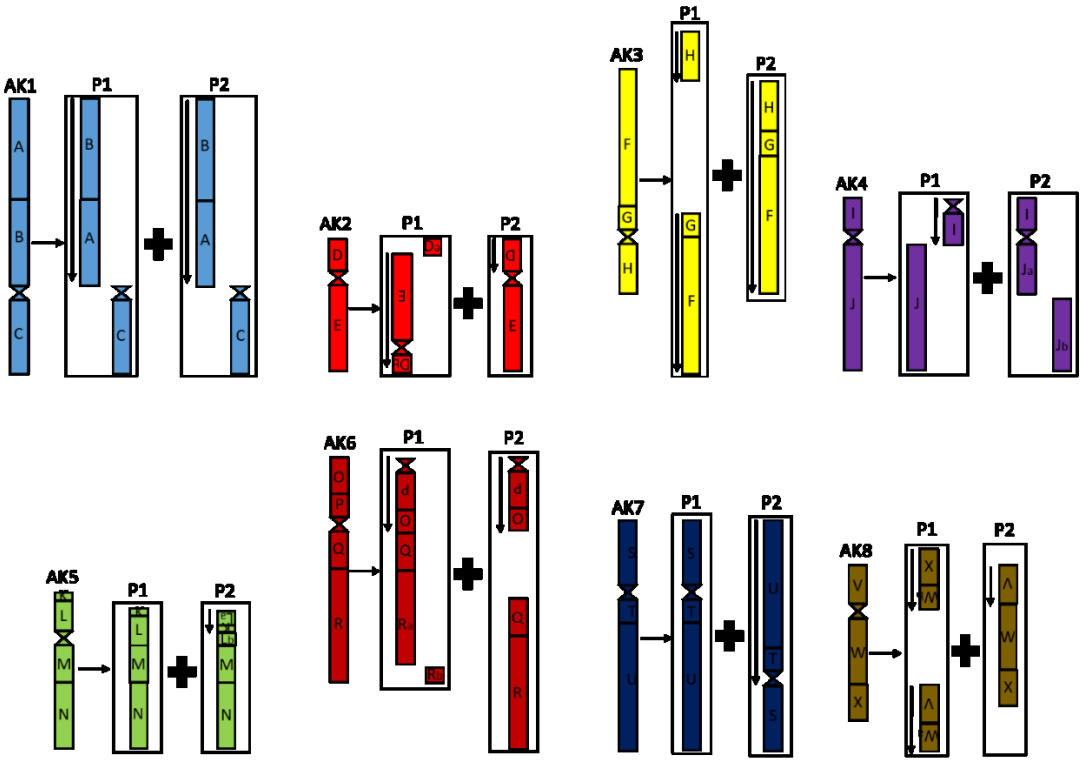

Figure 1-4. Rearrangement of ancestral crucifer karyotype in Pachycladon.

Distinct and complicated chromosomal rearrangement occurred in both two parental genomes (P1 and P2) in Pachycladon which originate from the same ancestral crucifer karyotype (ACK). Pachycladon karyotype includes exact two copies of each ancestral chromosome (AK1-8) with/without the same arrangement of GBs, and some of these GBs have been cut off and scattered on the same or different Pachycladon chromosomes (segments of one parental chromosome which were placed in two column represents that they relocate on two Pachycladon chromosomes; segments of one parental chromosome which were placed in one column stand for their colocalization on the same Pachycladon chromosome). For instance, AK6 evolved into two rearranged chromosomes in Pachycladon's parental genomes P1 and P2, respectively. However, R block in P1 broke down into Ra and Rb that were localized on different chromosomes of Pachycladon genome. In contrast, R block in P2 broke down at the bottom of centromere, and the resulting two segments were on the same Pachycladon chromosome.

1.2 Polyploidy

As mentioned before, Pachycladon is an allopolyploid. It has been suggested 47% to 100% of flowering plant species can be traced to at least an ancient polyploidy event, most of

which have evolved in locations with high environmental selection pressure (Brochmann et al., 2004; Wood et al., 2009). Also, genomics and phylogenetics have provided convincing evidence to support that polyploidization repeated throughout the evolution of flowering plants (Bowers, Chapman, Rong, & Paterson, 2003; Masterson, 1994; Vision, Brown, & Tanksley, 2000). Therefore, comprehensively understanding the mechanism and consequence of polyploidization would be vital to get insight into the power driving Pachycladon evolution and speciation.

1.2.1 Polyploidization mechanisms in plants

Polyploidy can either be in the form of autopolyploidy or allopolyploidy. Autopolyploidy is caused by duplication of a single genome, while allopolyploids are generated by intra- or interspecific crossing which results in the combination of two distinct genomes (Song & Chen, 2015) (**Figure 1-5**). Theoretically, repeated polyploidy events throughout plant evolution would tend to create giant genomes. However, large-scale chromosomal rearrangements and deletions after polyploidy reduces a genome size to a diploid-like state, which has been found in several plant species like *Oryza sativa* (Wang, Shi, Hao, Ge, & Luo, 2005), *Aegilops* and *Triticum* group (Ozkan & Feldman, 2009), and *Brassica rapa* (Yang et al., 2006). A dominant and highly efficient strategy to reduce genome size is to utilize large number of repeat elements across genomes. For instance, irregular homologous recombination between two long terminal repeats (LTR) which are at each end of one element can result in formation of a single LTR. It causes the loss of one of LTRs and the sequence between them (**Figure 1-6B**). As shown in Figure 4C, D and E, homologous recombination of two LTRs from different elements that belong to the same family could cause complicated DNA loss (Devos, Brown, & Bennetzen, 2002).

Furthermore, after genome duplication, chromosomal instability occurs prevalently in plant species, which is likely to induce abnormal chromosomal changes such as aneuploidy and in turn reduces genome size. After these polyploidy events, parental cell fusion results in extra centrosomes with multipolar spindles, and the multipolar ploidy reduction can lead to chromosome mis segregation in the daughter cells and, hence, result in aneuploidy and genomic instability. For instance, the hybridization of a doubled haploid *B. oleracea* line and a doubled haploid *B. rapa* line generated 50 allopolyploid lines in which chromosome instability was detected, as aneuploidy increased from 24.1% in the first generation to 71.4% in the fifth generation (Xiong, Gaeta, & Pires, 2011). *Aegilops tauschii* Coss. ($2n = 2x = 14$, DD) accessions were hybridized to *Triticum turgidum* L. ssp. *durum* cv Joyau ($2n = 4x = 28$, AABB) and 'Tetra Courtot' ($2n = 4x = 28$, AABB), and ten F2 seedlings were selected as synthetic wheat allohexaploids for meiotic behavior and aneuploidy analysis. It was found a significant positive correlation between the average number of univalent per pollen mother cell (PMC) in F2 plants and the frequency of aneuploids in subsequent generations, indicating meiotic irregularities across early generations leads to chromosomal aneuploidy (Mestiri et al., 2010).

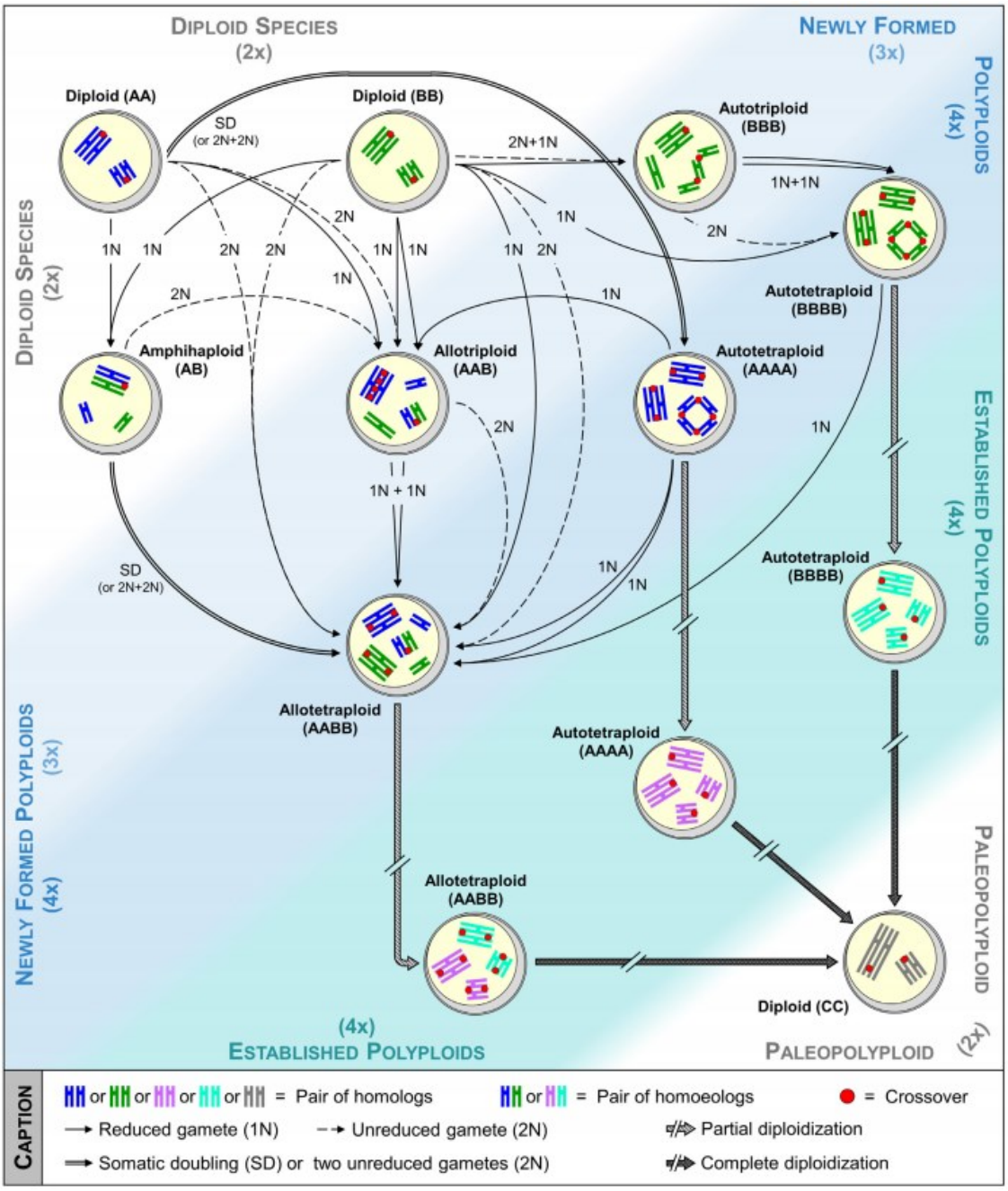

Figure 1-5. Plant polyploidization and re-diploidization events

Auto- and allotetraploids can be formed via a process referred to as partial diploidization which is characterized by the perturbed pairing and recombination as well as a diploid-like meiosis. Following millions of years, the diploidization becomes complete, and the formed polyploids return at a diploid stage which result in a chromosome number reduction. Copied from "Speciation success of polyploid plants closely relates to the regulation of meiotic recombination." by Pelé, A., Rousseau-Gueutin, M., & Chèvre, A. M. *Front Plant Sci,* 9, 907. Copyright 2018 by Pelé, Rousseau-Gueutin and Chèvre.

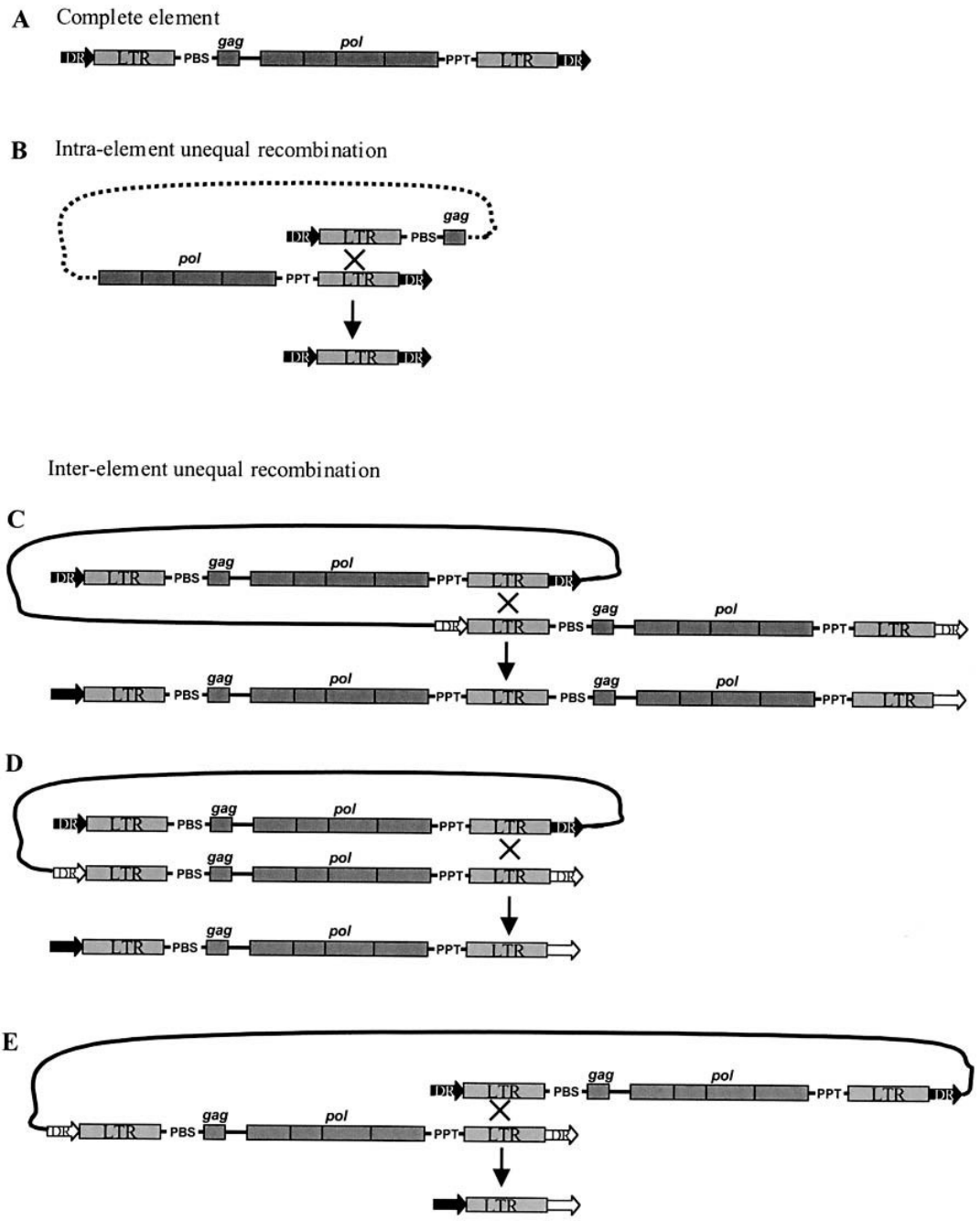

Figure 1-6. Unequal intrastrand recombination between LTR retrotransposons.

A) A complete element comprises a direct repeat (DR), two long terminal repeats (LTRs), a primer-binding site (PBS), polypurine tract (PPT), and genes (*gag* and *pol*); B) Intra-element recombination of two LTRs produces solo LTR; C) Recombination of two LTRs at 5' and 3' end of adjacent elements cause loss of one LTR; D) Recombination of two 5' LTRs of two elements result in loss of one element; E) Recombination of one 3' LTR and one 5' LTR of two elements lead to only one LTR and two DRs from two elements left. Copied from "Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis." by Devos, K.

M., Brown, J. K., & Bennetzen, J. L., 2002, *Genome Res, 12*(7), 1075-1079. Copyright 2002 by Cold Spring Harbor Laboratory Press.

1.2.2 Advantages and disadvantages of polyploidization in plant genomics

The genetic and epigenetic consequences of polyploidization drive organisms' evolution forward unpredictably. On one hand, polyploids benefit from multiple sets of chromosomes to adapt to challenging environment. In polyploids, some important genes have redundant copies due to multiple sets of genome copies. It allows the alteration of these redundant gene copies to diversify gene function, resulting in novel phenotypes against challenging environment, while the other gene copies function as usual, guaranteeing plant's regular growth and development (Comai, 2005). Redundant genes, also, are considered as the back-up plans. Once dysfunction occurs in one gene copy, another redundant gene copy could substitute to keep the regular function of plant's metabolic pathways (Huminiecki & Conant, 2012). On the other hand, polyploidy also cause difficulties in chromosome pairing during mitosis and meiosis, epigenetic instability, gene expression change (Comai, 2005). Chromosome pairing difficulties caused by polyploidy during mitosis and meiosis lead to chromosome aneuploidy, causing ploidy-conditional lethality of progeny or infertile offspring. In allopolyploids, mitosis and meiosis are even much more complicated because of inter-genomic recombination of two parental chromosomal complements (Comai, 2005). Theoretically, epigenetic changes could be positive. For instance, it has been found the adjacent genes of *TCP3* were silenced by hypermethylation and reactivated by blocking DNA methylation in an allotetraploid of *A. thaliana* and *Cardaminopsis arenosa*, *Arabidopsis suecica*, suggesting involvement of epigenetic modification in dosage effect of orthologous genes (Lee & Chen, 2001). However, after evolving under nature selection over decades, plant species have developed an optimized epigenetic modification adapting to surroundings, and new

epigenetic modification caused by epigenetic instability in allopolyploids or autopolyploids tend to alter the adaptation that has evolved in their parents. Therefore, the deleterious effects of modifying genome epigenetically frequently overwhelm its advantages in polyploidy plants (Comai, 2005).

1.2.3 Influences of polyploidization in the interactions of polyploid plant with other organisms

Genetic and epigenetic changes in polyploids can also be a driving force of speciation through the generation of cryptic reproductive barriers. This may occur due to changes in pollinator-attracting traits like flower shape, size, color, and flowering time which have been observed in polyploids (Gross & Schiestl, 2015; McCarthy et al., 2015; Schranz & Osborn, 2000; Tate et al., 2009). For example, distinct pollinators of polyploids and diploids have been reported. *Andrena carlini* was the most frequent visitor of tetraploid *Erythronium albidum*, and *Andrena erythronii* preferred diploid *Erythronium mesochoreum*, suggesting pollinator-mediated isolation plays an important role in forming reproductive isolation (Roccaforte, Russo, & Pilson, 2015). Furthermore, herbivory as a strong selective pressure functions in plant speciation and radiation. Thus, the interaction between polyploids and herbivores to some extent determine the survival and prosperity of polyploidy population. In some cases, duplicated genome copies did enhance polyploids' defense to herbivory. For example, in the research of Edger et al., whole genome duplication in Brassicales resulted in the production of novel glucosinolates that reduced ability of butterflies to feed on the plants (Edger et al., 2015). However, polyploidization does not always provide increased protection against herbivory. For example, modified secondary metabolism of *Centaurea phrygia* tetraploids produced lower concentrations of gallic acid and less categories of flavonoids compared to diploids,

which could also be a reason for higher rates of herbivore damage in some tetraploids (Münzbergová, Skuhrovec, & Maršík, 2015). However, cell polyploidization can help plants recovery from herbivory. For instance, natural herbivory was simulated in two Arabidopsis ecotypes, Columbia and Landsberg *erecta*, by clipping aboveground parts of plants. Compared to unclipped controls, clipped Columbia plants accelerated regrowth and produced significant higher dry biomass, which was associated tightly with accelerated endoreduplication, suggesting transient cell polyploidization facilitates plant's recovery from herbivory damage (Scholes & Paige, 2011).

1.3 Plant stress responses

The neo- or sub-functionalisation of the duplicated homologous genes in polyploids allows these homologues to be used differentially in various responses to stress, which increases the stress response options of polyploids (Bharadwaj, 2015). Therefore, polyploidisation could be an inner driving force for the evolution of plant adaptation to various living environments, which work as selection pressures to drive plant evolution (Jump, Marchant, & Peñuelas, 2009).

1.3.1 General process of plant stress signalling

In plants, environmental stress signals are received by membrane receptors, ion channels and kinases, activating downstream signalling cascades intracellularly. Secondary signal molecules such as Ca^{2+}, reactive oxygen species (ROS) and abscisic acid (ABA) have been shown to be generated during intercellular signalling (Zhu, 2016). These secondary signals mediate the expressions of stress-responsive genes, in turn influencing biological processes in which these genes are involved. As a consequence, plant morphology, metabolism and physiology will shift to adapt to stress tolerance directly or indirectly.

1.3.1.1 Stress signal perception

The first step in the plant stress response is the perception of an environmental signal at the cell level. However, because of functional redundancy in the genes coding sensors, mutagenesis studies on individual genes may not identify significant phenotypes in the stress response (Zhu, 2016). Reduced hyperosmolality-induced $[Ca^{2+}]_i$ increase 1 (OSCA1) is a putative sensor of hyperosmotic stress in *Arabidopsis*, and its loss-of-function mutants displayed reduced calcium increase under osmotic stress without any other stress phenotype (Yuan et al., 2014). Another hot spot that has been studied is the sensor for temperature. *COLD1* encodes a regulator that interacts with a G-protein subunit, which activates Ca^{2+} channels on the plasma membrane and endoplasmic reticulum. By mediating Ca^{2+} channels, the overexpression of *COLD1* has been shown to strengthen chilling tolerance in rice (*Oryza rufipogon*) plants (Ma et al., 2015). Histone H2A.Z is a nucleosome temperature sensor recently identified in *Arabidopsis* (Kumar & Wigge, 2010). H2A.Z occupancy at the nucleosome was found to be temperature dependent: high at 17°C but significantly lower at 27°C. H2A.Z-containing nucleosomes wrap DNA more tightly, which reduces the local unwrapping of DNA necessary for RNA Pol II to transcribe; in turn transcription is adjusted (Kumar & Wigge, 2010). Studying stress sensors is challenging because of the difficulty in demonstrating a protein as a physical signal. Therefore, most stress sensors remain unknown.

1.3.1.2 Stress signalling

There are a number of stress signals, and each take on a different role in single or multiple stress responses. In salt stress, Na^+ toxicity is one of main problems, besides hyperosmotic and oxidative stresses, and although the sensor for Na^+ has not been identified, the relevant signalling has been well studied. In plants, high Na^+ triggers the binding of cytosolic calcium and EF-hand calcium-binding protein, salt overly sensitive gene 3

(SOS3); in turn SOS2 is activated by Ca^{2+}–SOS3. Then, SOS2 phosphorylates a Na^+/H^+ antiporter on the plasma membrane, SOS1, which is able to pump out extra Na^+ and restore cellular Na^+ homeostasis (Zhu, 2016). As for Na^+ stress signalling, there is a lack of information on the sensors involved in osmotic stress signalling, but it has been reported that protein kinases from the SNF1 (sucrose nonfermenting 1)-related protein kinase 2 (SnRK2) family may be activated by osmotic stress via an unknown mechanism (Zhu, 2016). Further, nine out of 10 SnRK2s were activated by osmotic stress (Boudsocq, Barbier-Brygoo, & Laurière, 2004). In addition to SnRK2s, secondary messengers like ABA and lipid signals were involved in conducting downstream signalling, which activated osmotic-responsive TFs (Deak & Malamy, 2005).

1.3.1.3 Stress-related hormone signalling

Another important stress signalling is ABA signalling, which is at the centre of multiple stress responses (Tuteja, 2007). Because of the importance of the ABA signalling pathway in multiple stress responses, it has been widely studied, resulting in the identification of receptors (PYR/PYL/RCAR and START domain proteins) and pathways (Klingler, Batelli, & Zhu, 2010; Santiago et al., 2012; Tuteja, 2007). PYLs bind ABA in the presence of the type 2C protein phosphatases (PP2Cs) as co-receptors, which block the catalytic cleft of SnRK2 kinases in the absence of ABA (Zhang et al., 2015). In the presence of ABA, SnRK2s are released and phosphorylated and themselves then activate downstream signalling such as the mitogen-activated protein kinase (MAPK) cascade (de Zelicourt, Colcombet, & Hirt, 2016). It has been reported that, in addition to ABA, the levels of other plant hormones including salicylic acid (SA), jasmonic acid (JA) and ethylene (ET) are elevated under multiple stress conditions, and they could serve as stress signals (Devoto & Turner, 2005; Khan, Fatma, Per, Anjum, & Khan, 2015; Morgan & Drew, 1997). SA was induced by various stresses and received by SA methyl transferase

1 (SAMT1) and SA-binding protein 2 (SABP2) (Jayakannan, Bose, Babourina, Rengel, & Shabala, 2015). SAMT1 generated methyl salicylate (MeSA) from SA, while SABP2 converted MeSA to SA via its SA-inhibiting MeSA esterase activity, and SABP2-mediated SA signalling was involved in tomato salt stress, with an unknown mechanism (Jayakannan et al., 2015). Another SA receptor is nonexpressor of pathogenesis-related genes 1 (NPR1), which was reduced from an inactive oligomer to an active monomer by a SA-induced redox status shift (Fobert & Després, 2005). Then, monomeric NPR1 was transferred into the nucleus, along with TGACG-Binding (TGA) transcription factors (TFs), which determined the transcription of SA-responsive *PATHOGENESIS-RELATED* (*PR*) genes (Johnson, Boden, & Arias, 2003). In the JA signalling pathway, JA and its most bioactive compound, the amino acid conjugate (+)-7-iso-jasmonoyl-l-isoleucine (JA-Ile), are perceived by the Skp1/Cullin/F-box (SCF^{COI1})–jasmonate ZIM domain (JAZ) co-receptor complex, where JAZ proteins are repressors of TFs such as MYC2 in JA signalling (Chico, Chini, Fonseca, & Solano, 2008). The binding of JA/JA-Ile to the SCF^{COI1}–JAZ complex induced ubiquitinoylation and proteasomal degradation of JAZ proteins, which activates MYC2 and switches on transcription of JA/JA-Ile-responsive genes including MYC2 and JAZ (Chico et al., 2008). An ethylene signal was received by Endoplasmic reticulum (ER)-associated receptors (ETR) and then constitutive triple response 1 (CTR1) was released to interact with the histidine kinase domain of the ETRs (Guo & Ecker, 2004). The release of CTR1 has been proposed to activate the downstream MAPK cascade in *Arabidopsis*, *Medicago* and *Nicotiana tabacum*, among others, but the role of MAPKs in the ethylene signalling pathway is not fully understood (Guo & Ecker, 2004; Hahn & Harter, 2009; Ouaked, Rozhon, Lecourieux, & Hirt, 2003). However, downstream components of the MAPK pathway have been confirmed via transgenic studies of ethylene insensitive proteins (EINs) and

Ein-3-like protein 1 (EIL1), showing that EIN3/EIL1 binds to a primary ethylene response element (PERE) in the promoters of ethylene response element binding protein (EREBP) genes (Solano, Stepanova, Chao, & Ecker, 1998). Then, EREBPs bind to the GCC box of ethylene-responsive genes such as plant defensin 1.2 (PDF1.2) and hookless 1 (HLS1), which induces a variety of further ethylene responses under biotic and abiotic stresses (Lee et al., 2005; Lehman, Black, & Ecker, 1996).

1.3.1.4 ROS signalling

ROS signalling is another widely studied pathway in the plant stress response. Because of the multiple forms of ROS (e.g. superoxide, hydrogen peroxide and singlet oxygen), a number of mechanisms have been proposed for ROS perception (Mittler et al., 2011). These include oxidative modification of protein at cysteine, selenocysteine and methionine residues; lipid oxidation; and oxidation modification of amino acids such as proline (Wrzaczek, Brosché, & Kangasjärvi, 2013). Among these mechanisms, the oxidative modification of proteins at cysteine residues is common (Chang, Huang, Lin, Chang, & Wu, 2010; Jeong et al., 2011; Wang, Yang, & Yi, 2012). Extracellularly, H_2O_2 and superoxide interact with cell wall components such as peptides and the extracellular domains of receptors; the oxidised peptides or receptors may then function in signalling. For instance, receptor-like protein kinase (RLK) is oxidised by H_2O_2 at cysteines, leading to autophosphorylation of the kinase domain. The autophosphorylated RLK then interacts with Ca^{2+} channels to open them (Hua et al., 2012). Inside the cell, a variety of ROS signalling pathways mediate the activation of TFs and the transcription of ROS-responsive genes. NPR1 is one of the ROS-related signalling proteins in plants, and is present in the cytoplasm in an oxidised form under normal condition (Seo, Wi, & Park, 2020). Under ROS stress, NPR1 is S-nitrosylated at cysteine residues, and S-nitrosylated NPR1 has improved nucleus import and interaction with TGA TFs (Wrzaczek et al.,

2013). Also, ROS signals could be passed down to TFs via their cellular redox status. TFs such as the membrane-tethered TF ANAC089, related to AP2 4a (RAP2.4a), basic region leucine zipper (bZIP); and TF R2R3 are directly regulated by cellular redox, which affects their DNA binding activity (Heine, Hernandez, & Grotewold, 2004; Klein, Seidel, Stöcker, & Dietz, 2012; Shaikhali et al., 2008; Shaikhali et al., 2012). Under oxidising conditions, the RAP2.4a protein shifts to a dimeric state, which activates transcription of downstream genes (Shaikhali et al., 2008). Dithiothreitol (DTT) treatment could improve the DNA binding activity of AtbZIP16 by maintaining cysteines (Cys) in the reduced state (Shaikhali et al., 2012). The reduction of the single Cys residue of the plant-specific R2R3 MYB domain is crucial for DNA binding and is regulated by redox (Heine et al., 2004). Consequently, the improved DNA binding activity of these TFs increases the transcriptional efficiency of ROS-responsive genes. Also, ROS signalling is integrated with many other signalling pathways. For instance, ROS mediated ABA-induced stomatal closure via the MAPK cascade (Jammes et al., 2009). NPR1 is also a SA receptor, and, after being translocated into the nucleus, NPR1 and TGA1 regulate the expression of SA-induced genes (Wrzaczek et al., 2013). Therefore, ROS signalling is very complex, and includes various forms of ROS and multiple signalling pathways.

1.3.2 General plant stress response

Plants develop various strategies for responding to different stresses, and these responses adjust plant biochemistry, physiology and architecture to all the plant to adapt to or tolerate stress. Water deficits affect most processes throughout growth and development of plants; thus, drought stress involves highly complex networks in which ABA signalling plays a crucial role (Zhang, Jia, Yang, & Ismail, 2006). The important contribution of ABA signalling to drought stress responses includes mediation of stomatal closure in

conjunction with other hormones like ET, auxins, cytokinins, JA, SA and brassinosteroids (Burgess & Huang, 2016; Yuan et al., 2010). Consequently, water loss is prevented and in turn pressure on plant physiology from drought stress is relieved, which retains the normal functioning of plant metabolism and helps the plant adapt to water deficits (Zhang et al., 2006).

The major effect of cold stress is damage to plasma membranes, the severity of which is determined by the relative proportion of saturated and unsaturated fatty acids in the plasma membrane lipids. The higher the proportion of saturated fatty acids, the faster the plasma membrane solidifies at low temperature; membranes with less fluidity tend to be disrupted (Yadav, 2010). Cold-induced ice formation is another cause of plant damage. Ice is formed in the apoplastic space in which a vapour pressure gradient establishes between the apoplast and surrounding cells (Reaney & Gusta, 1999). Therefore, balancing cellular dehydration becomes crucial for plant tolerance to cold stress. In cold responses, the transcription of genes involved in the biosynthesis of solutes, antioxidant, and ABA, and the rearrangement of lipids in membranes are initiated in downstream cold stress signalling. The resulting changes in plant physiology and metabolism contribute to balancing cellular dehydration under cold stress, which improves plant cold tolerance (Yadav, 2010). An example of plant tolerance strategies to high temperature is the regulation of hormones in pollen sterility in response to elevated temperature. During heat stress, auxin biosynthesis is inactivated in the anthers, which could lead to male infertility, but fertility has been successfully restored using exogenous auxins (Sakata et al., 2010). This could reduce the consumption of photosynthates and in turn increase plant fitness during heatwaves.

Therefore, although various adaptive strategies are used in different stress responses, it seems that the regulation of hormones in the shift of photosynthesis and water balancing under stress is essential for plants' physiological and evolutionary adaptation. The evolution of genes in the hormone biosynthesis and regulation network is a crucial component of plant adaptation to the local stress environment. For instance, dwarfism is common in alpine populations and it has been suggested that variants of GA20ox1 in gibberellic acid (GA) biosynthesis was responsible to the dwarfism in alpine populations of *Arabidopsis*, and mutations in GA20ox1 caused dwarfism in the Green Revolution (the Third Agricultural Revolution, the set of research technology like the introduction of high-yielding varieties increased agricultural production between 1950 and the late 1960s) involving rice and barley (Barboza et al., 2013; Luo et al., 2015). The involvement of other hormones such as JA and SA in various stress responses suggests that they most likely participate in plant adaptation to stresses as well.

1.3.3 Cold stress

Recently, comprehensive transcriptomic and proteomic analyses using next-generation sequencing technologies, two-dimensional protein gel analyses, as well as shotgun proteomics approaches have provided more information about the mechanism of plant cold responses (Cui et al., 2005; Huo et al., 2016; Lee, Lee, & Kwon, 2015; Li et al., 2019). Cold stress is received by receptors in the membranes of plant cells; the signalling pathways within cells are then activated (**Figure 1-7**). There are various components involved in signalling transduction in cells, including calcium, ROS, protein kinases, protein phosphatase and lipid signalling cascades, and these signalling components activate downstream responses (Chinnusamy, Zhu, & Zhu, 2007) (**Figure 1-7**). Cold-induced elevation of cytosolic calcium level in cells is sensed by calcium-binding

proteins, which could cause protein conformation changes. These changes allow proteins to interact with other proteins to initiate a phosphorylation cascade (Guo, Liu, & Chong, 2018). Via phosphorylation, cold-related TFs could be activated, and then regulate the transcription of cold-responsive genes. One cold signalling pathway involves cold activation of SnRK2.6/ open stomata 1 (OST1), which phosphorylates the basic helix-loop-helix (bHLH) TF inducer of CBF expression 1 (ICE1) to activate the expression of Apetala 2 (AP2)-domain proteins, C-repeat/DRE-Binding Factors (CBFs), which initiates the expression of downstream cold-responsive genes (Zhu, 2016). Another pathway is via cold stress-induced calcium signalling and the MAPK cascade. The calcium-responsive protein kinase 1 (CRLK1) interacts with mitogen-activated protein kinase kinase kinase 1 (MEKK1) to cause calcium and calmodulin binding. MEKK1 may then activate downstream mitogen-activated protein kinase kinase 2 (MKK2), mitogen-activated protein kinase 4 (MPK4) and mitogen-activated protein kinase 6 (MPK6); in turn expression of related TFs such as WRKY could be activated (Mao et al., 2011; Zhao et al., 2017). The transcription of downstream cold-responsive genes is then initiated and these cold-responsive genes are involved in scavenging of cold-induced ROS, repair of cold-induced damage, synthesis of osmolytes to stabilise cellular structures, and osmotic adjustment, restructuring the plasma membrane (Jan, Majeed, Andrabi, & John, 2018; Yadav, 2010) (**Figure 1-7**). This also includes genes involved in the biosynthesis of plant stress-related hormones such as ABA, SA and ET (Ciardi, Deikman, & Orzolek, 1997; Shi & Yang, 2014; Szalai, Tari, Janda, Pestenacz, & Páldi, 2000). These hormones affect the architecture of some plants under cold stress. Reduced height 3 (Rht3) is involved in GA signalling. The wheat mutant of Rht3 displayed no increased length of the leaf extension zone when the temperature increased from 10 to 20ºC, while an increase occurred in wild type (Appleford & Lenton, 1991). Treating wild type plants with GA at

10ºC increased the length of the extension zone, suggesting an important role for GA in wheat architecture under cold conditions (Appleford & Lenton, 1991). Another stress-related hormone, SA, also influences plant growth under stress, and high concentrations of SA retard plant growth. SA-induced genes were upregulated in an *ice1* mutant that encodes a core regulator in cold signalling, and the mutant plants were small and cold sensitive, which suggested a relationship between cold and SA signalling (Miura & Ohta, 2010). Meanwhile, overexpressing *CBF2* resulted in delayed senescence in *Arabidopsis* plants via an unknown mechanism, which implied a close relationship between cold signalling and senescence (Sharabi-Schwager, Samach, & Porat, 2010). Collectively, these studies not only construct an overall picture of plant cold stress responses but also unveil the fine tuning of the expression of genes and proteins in the precise regulatory system of plant cold stress responses, which enable plants to quickly adapt to dynamic and variable low-temperature environments (**Figure 1-7**).

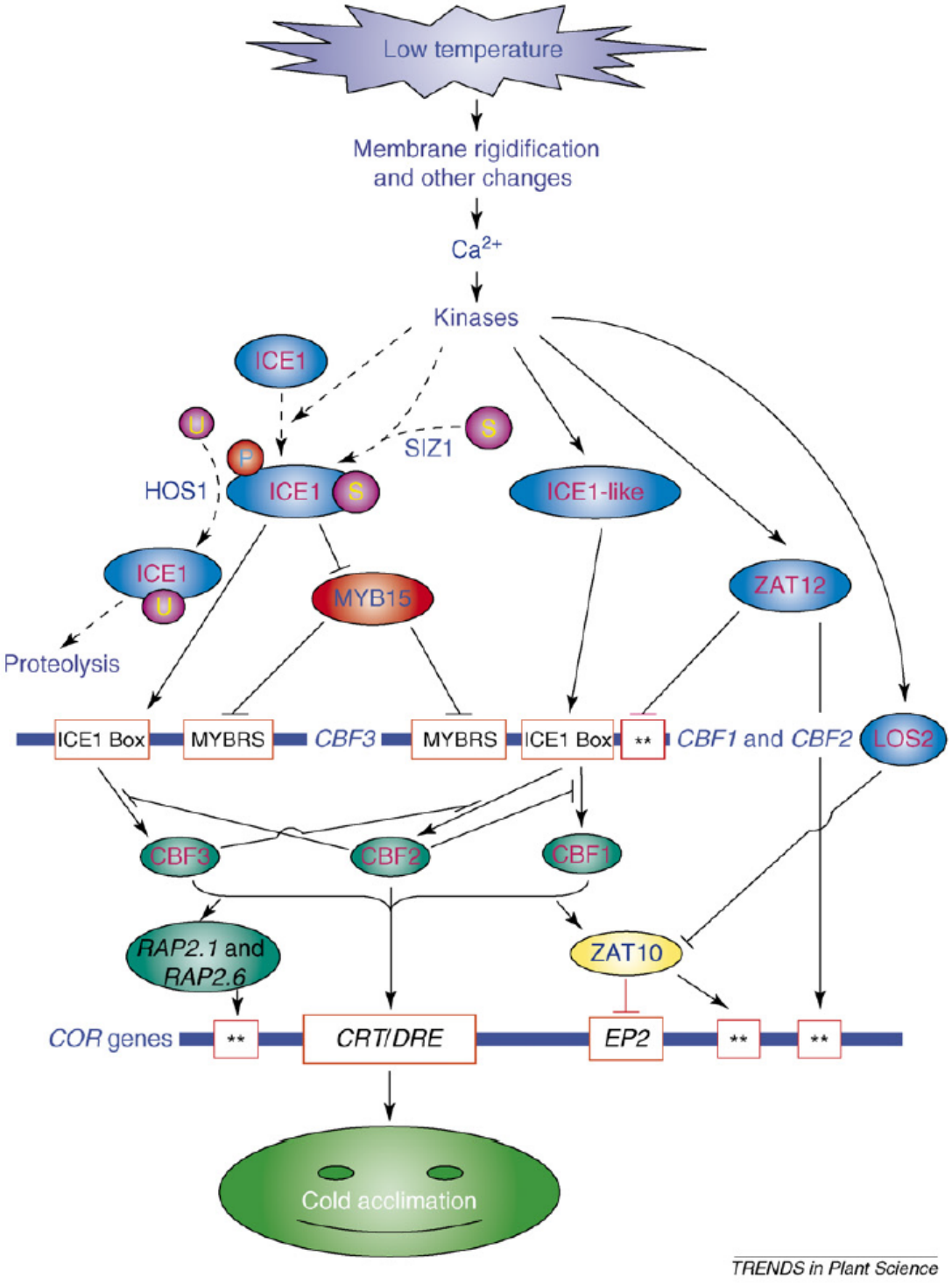

Figure 1-7. Plant cold stress signalling.

Cold signal is perceived by membrane rigidification and other cellular changes, which leads to the activation of ICE1 by phosphorylation and sumoylation. And, ubiquitination of ICE1 is mediated by HOS1. Phosphorylated and sumoylated ICE1 initiated the transcription of CBFs and supressed MYB15. CBFs regulate the expression of the freezing tolerance-related genes (COR). CBFs induced ZAT10 negatively regulate the expression of *COR* genes, while cold-induced

LOS2 represses the transcription of *ZAT10*. Broken arrows indicate post-translational regulation; solid arrows indicate activation, whereas lines ending with a bar show negative regulation; the two stars (**) indicate unknown cis-elements. Abbreviations: CBF, C-repeat binding factor (an AP2-type transcription factor); CRT, C-repeat elements; DRE, dehydration-responsive elements; HOS1, high expression of osmotically responsive genes1 (a RING finger ubiquitin E3 ligase); ICE1, inducer of CBF expression 1 (a MYC-type bHLH transcription factor); LOS2, low expression of osmotically responsive genes 2 (a bifunctional enolase with transcriptional repression activity); MYB, myeloblastosis; MYBRS, MYB transcription factor recognition sequence; SIZ1, SAP and MiZ1 (a SUMO E3 ligase); P, phosphorylation; S, SUMO (small ubiquitin-related modifier); U, ubiquitin. Copied from "Cold stress regulation of gene expression in plants." by Chinnusamy, V., Zhu, J., & Zhu, J. K., 2007, *Trends Plant Sci*, 12(10), 444-451. Copyright 2007 by Elsevier.

1.3.4 Drought and salt stresses

Plants show closely related responses to drought and salt stress, and a number of studies have suggested that the mechanisms of drought and salt responses overlap (Bartels & Sunkar, 2005; Golldack, Li, Mohan, & Probst, 2014; Zhang, Jia, Yang, & Ismail, 2006). The two stresses present two main threats to plants. One is toxicity of Na^+ and Cl^- at high concentrations. Because of the similar chemical nature of Na^+ and K^+, excess Na^+ may inhibit K^+ uptake by the root. Potassium has an important nutrient function in multiple plant biological processes and the growth of plants with a shortage of potassium was inhibited (Kant, Kafkafi, Pasricha, & Bansal, 2002; Wang, 2007). Meanwhile, sodium could inhibit the activities of many enzymes in plants as well (Zhu, 2007). Calcium is an important factor in balancing the amount of potassium and sodium inside cells. It works as a second signalling component involving pathways that regulate the expression and activity of potassium and sodium transporters (Lan et al., 2010; Luan, Lan, & Lee, 2009). Another threat is water deficit or osmotic stress caused by drought and salt stresses. These stresses induce the synthesis of ABA, which induces the closure of guard cells; photosynthesis then declines along with synthesis of osmolytes. Early in 1998, it was

shown that ABA triggered stomatal pore closure by modulating anion channel activation, and it was more recently reported that the ABA-deficient mutant *aba3-1*—in which ABA cannot be synthesised under drought—failed to close its guard cells when exposed to dry air (Bauer et al., 2013; Pei, Ghassemian, Kwak, McCourt, & Schroeder, 1998). In turn, ABA-induced stomatal closure worsened photosynthesis decline and cell wall extensibility and leaf elongation in nitrogen-stressed tomato and barley plants due to reduced carboxylation activity and electron transport. Ultimately, plant growth was limited (Chapin, Walter, & Clarkson, 1988).

Salt- and drought-responsive processes have been shown to be regulated, mediated and controlled by a large number of responsive genes, as revealed by a number of high-throughput transcriptomics studies on drought and salt stress responses in a variety of plant species (Johnson et al., 2014; Liu et al., 2016; Shen et al., 2014; Sun et al., 2010; Tang et al., 2013; Zahaf et al., 2012). Since protein levels are also regulated post-transcriptionally, salt-responsive proteomes have attracted more attention recently, and 2171 salt-responsive proteins have been identified from 34 plant species (Zhang et al., 2012). These proteome studies have shown that photosynthesis, the ROS scavenging system, ion homeostasis, osmotic homeostasis, membrane transport, signalling transduction, transcription, protein synthesis/turnover and cytoskeleton dynamics are affected and adjusted in plants to respond and adapt to drought and salinity stresses (Wani, Dutta, Neelapu, & Surekha, 2017). In drought and salt stress responses, MAPKs, phospholipases, calcium-dependent protein kinase (CDPK) and Ca^{2+} play important roles in stress signal perception and transduction, which activates downstream TFs such as heat shock transcription factors (HSFs), repeat-binding/dehydration-responsive element (CBF/DREB) and WRKY (Wani, Dutta, Neelapu, & Surekha, 2017). In turn, the

transcription of salt- and drought-responsive genes involved in supporting the integrity of cellular structures, ROS scavenging, water regulation and ion transport is initiated, which restores ion and osmotic homeostasis as well as protecting cell structure, membrane and function, resulting in plant stress tolerance (**Figure 1-8).** For instance, in responding to threats, plants synthesise organic solutes for osmotic adjustment and cell protection under salt stress, including glycerol, mannitol, glutamate and proline (Bahieldin et al., 2014; Berteli et al., 1995; Karakas, Ozias-Akins, Stushnoff, Suefferheld, & Rieger, 1997; Reddy et al., 2015). They are not harmful for plant cells even at high concentrations, so are considered compatible osmolytes, balancing the high concentrations of Na^+ and Cl^- outside the cell and inside the vacuole (Rhodes, Nadolska-Orczyk, & Rich, 2002). Coupled with other responses like compartmentalisation of toxic ions, adjustment of photosynthetic and energy metabolism, regulation of hormones and modification of cell structure, plants have evolved complex mechanisms to cope with drought and salinity stress (Wani et al., 2017).

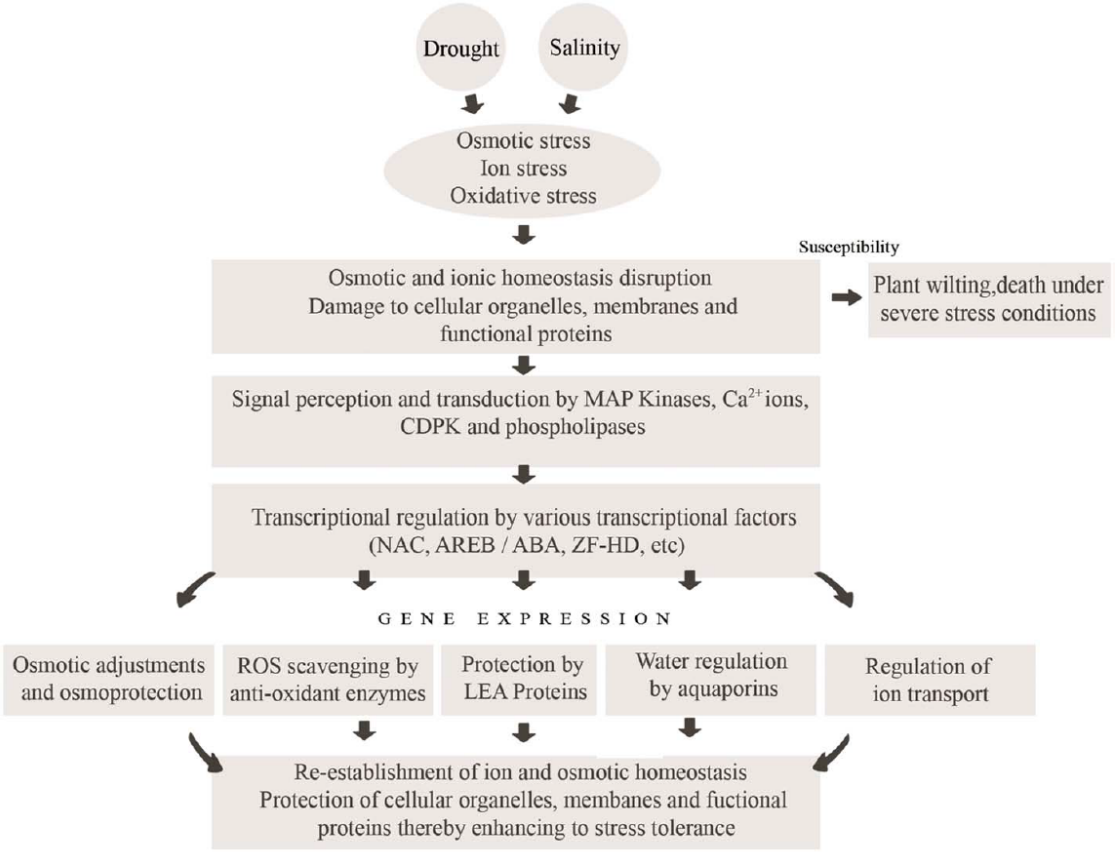

Figure 1-8. Plant salt and drought stress responses.

Copied from "Transgenic approaches to enhance salt and drought tolerance in plants." by Wani, S. H., Dutta, T., Neelapu, N. R. R., & Surekha, C., 2017, *Plant Gene, 11*, 219-231. Copyright 2017 by Elsevier.

1.3.5 UV-B radiation

UV-B radiation (280–315 nm) is a high-energy component of solar radiation. After absorption by plant, UV-B radiation may induce photoexcitation of DNA, proteins and lipids, which results in changes to a number of biological processes (Petruľová, Dučaiová, & Repčák, 2014). Some of these changes are damage, while some of them regulatory, depending on the fluence rate and duration of UV-B radiation. Low levels and short durations of UV-B radiation induce regulatory changes, while high levels and long durations trigger damage to plant biochemistry, physiology and genomes (Jenkins, 2009). Plants receive light via various photoreceptors, including phytochromes (red/far-red light

photoreceptors), cryptochromes, phototropins and Zeitlupe family members (UV-A/blue light photoreceptors). UVR8 is the only photoreceptor identified so far for UV-B radiation reception. However, UV-B radiation responses could be activated by two photomorphogenic signalling pathways, one of which is independent of, and one of which is dependent on, UVR8. UVR8-dependent signalling has been well studied and involves a number of photomorphogenesis response regulators such as constitutively photomorphogenic 1 (COP1) and elongated hypocotyl 5 (HY5) (Heijde & Ulm, 2012). UV-B radiation stimulated nuclear accumulation of UVR8 and interaction of UVR8 and COP1, which resulted in binding of UVR8 to HY5 and, in turn, transcription initiation of UV-B responsive genes in the UVR8-dependent pathway. However, the mechanism of UVR8-independent pathway remains unclear (**Figure 1-9**). A few details regarding the response to acute UV-B radiation stress have been reported in *Arabidopsis*, which is related to the activity of MAPKs (González Besteiro, Bartels, Albert, & Ulm, 2011). Further, COP1 regulates the expression of the gene encoding ARIADNE12 E3 ligase in a way that is independent of UVR8 at high fluence rates of UV-B radiation (Lang-Mladek et al., 2012). There are also nonspecific signalling pathways that are activated at low levels of UV-B radiation, and a number of UV-B-related processes regulating levels of ROS, JA, SA and C_2H_4, as well as DNA damage (**Figure 1-9**).

Both signalling pathways target downstream transcription of their own responsive genes, which results in plant photomorphogenesis and developmental reprogramming that in turn helps the plant to tolerate stress (Dotto & Casati, 2017). After UV-B exposure, ROS accumulation occurs as a result of UV-B-induced impair to photosynthetic electron transport and UV-B induced reduced nicotinamide adenine dinucleotide phosphate (NADPH) oxidase activity, which activates plant ROS scavenging enzymes like

ascorbate perosidase and superoside dismutase (Hideg, Jansen, & Strid, 2013; Rejeb et al., 2015; Wang et al., 2016). Further, UV-B radiation causes DNA damage via the formation of cyclobutane pyrimidine dimers (CPD), which may be repaired by both photoreactivation and the light-independent pathway. Other morphological and metabolic responses occur in various plant species, such as hypocotyl growth suppression in *Lepidium sativum* L., anthocyanin accumulation in sorghum (*Sorghum bicolor*) and maize (*Zea mays* L.), and cotyledon curling in *Brassica napus* (Gerhardt, Wilson, & Greenberg, 2005; Steinmetz & Wellmann, 1986; Wellmann, 1983; Yatsuhashi, Hashimoto, & Shimizu, 1982). Reduction in leaf area was shown to be a result of inhibition of cell proliferation mediated by the action of miR396 (a microRNA), which downregulates growth-regulating factor (GRF) TFs (Gómez, Falcone Ferreyra, Sheridan, & Casati, 2019), and the density of leaf trichomes was regulated by UV-B through GLABRA3 (Dotto & Casati, 2017). Leaf area reduction, downward curling, leaf number reduction, increased trichome density, reduced hypocotyl extension, reduced root growth and flowering time delay have all been found to be induced by UV-B radiation stress (Biever, Brinkman, & Gardner, 2014; Fierro et al., 2015; Hectors et al., 2010; Jiang, Wang, Bjorn, & Li, 2011; Liakoura, Stefanou, Manetas, Cholevas, & Karabourniotis, 1997; Rajendiran & Ramanujam, 2004; Wijewardana, Henry, Gao, & Reddy, 2016). Further, UV-B induced changes in hormone biosynthesis and pathways contribute to this UV-B induced plant photomorphogenesis. UV-B radiation-induced downward leaf curling required the presence of auxin, and cytokinin-induced expansion growth of cucumber cotyledons was inhibited by UV-B radiation (Ana Carolina Fierro et al., 2015; Kataria & Guruprasad, 2018). This suggests that multiple signalling pathways interact in UV-B radiation responses, most of which still remain unknown. These pathways are activated by low and/or high fluence rates and different wavelengths and duration of UV-B radiation. In

turn, the transcription of downstream responsive genes is initiated, which regulates UV protection and morphogenesis. The plant then develops adaptation and acclimation to UV-B radiation in the natural environment.

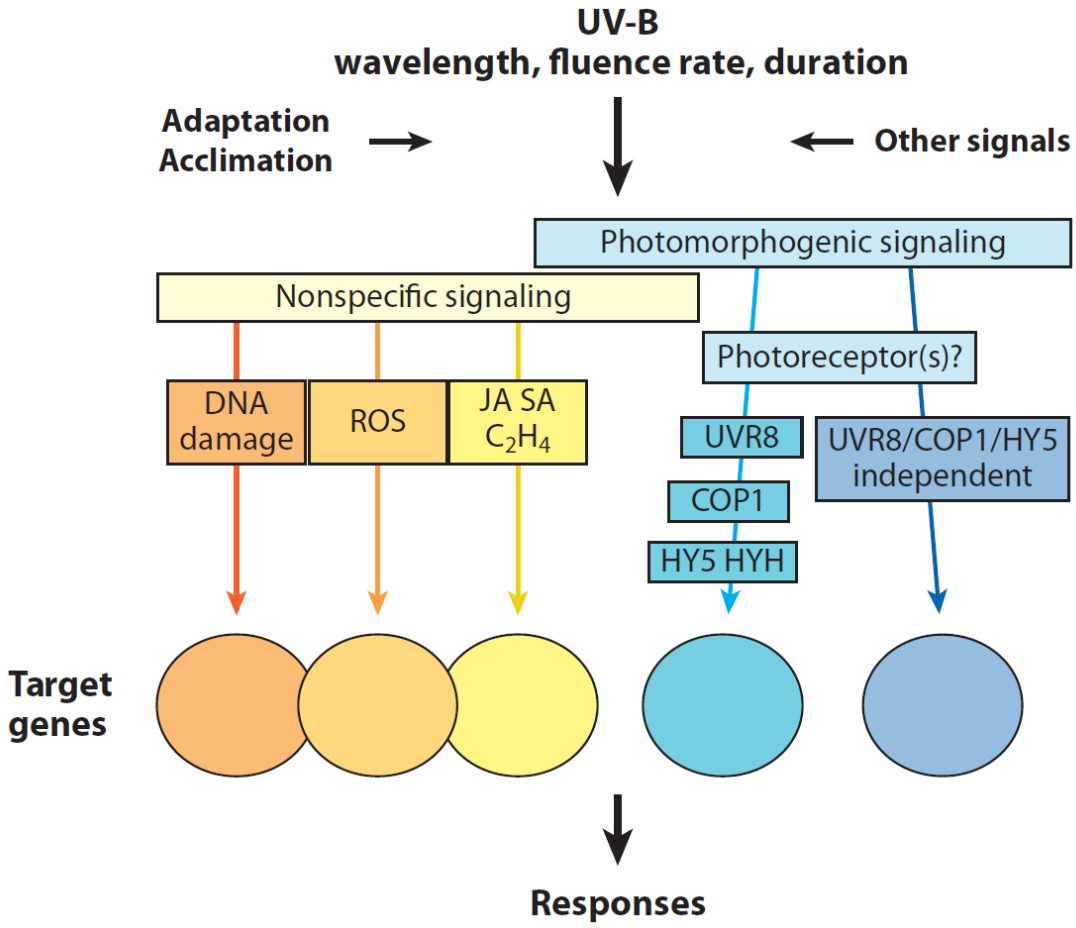

Figure 1-9. UV-B signal transduction pathways.

The wavelength, fluence rate, and duration of UV-B radiation decide whether nonspecific and/or photomorphogenic signal pathways are activated, which induce specific sets of target genes and downstream responses. ROS, reactive oxygen species; JA, jasmonic acid; SA, salicylic acid, UVR8, UV RESISTANCE LOCUS8; COP1, CONSTITUTIVELY PHOTOMORPHOGENIC1; HY5, ELONGATEDHYPOCOTYL5; HYH, HY5 HOMOLOG. Copied from "Signal transduction in responses to UV-B radiation" by Jenkins, G. I., 2009, *Annu Rev Plant Biol*, *60*, 407-431. Copyright 2009 by Annual Reviews, Inc.

1.3.6 Pachycladon and stress

Given the wide altitudinal living environment of the genus *Pachycladon* in New Zealand, a good understanding of the molecular mechanisms of plant adaptation to various environmental factors is essential to establish *Pachyladon* as a model for speciation involving different environmental niches in New Zealand. A microarray study on these

environments as abiotic driving forces in the speciation of two *Pachycladon* species, *Pachycladon enysii* and *P. fastigiate*, identified 310 and 324 upregulated genes in *P. fastigiate* and *P. enysii* respectively, that are hormone and stress responsive (Voelckel et al., 2008). In a later proteomic study on five *Pachycladon* species, it was suggested that expression differences in stress-induced genes between *P. fastigiate* and *P. enysii* resulted from differences in the environments in which they were collected (Mirzaei et al., 2011). However, it was clear that there were some stress factors in the living environments of these species, which induced their stress responses. Also, differences in glucosinolate profiles and glucosinolate hydrolysis products were identified between *P. fastigiate* and *P. enysii*, and shown to be involved in the chemical defence mechanism of Brassicaceae plants against biotic stresses (Voelckel et al., 2008). A proteomic study confirmed the relative abundance of proteins related to the myrosinase–glucosinolate system in *P. enysii* (Mirzaei et al., 2011). This system participates not only in biotic but also in abiotic stress responses, and the involvement of glucosinolates in salinity, drought, temperature, light quality and nutritional deficiencies in various plant species has been reported (Cocetta, Mishra, Raffaelli, & Ferrante, 2018; del Carmen Martínez-Ballesta, Moreno, & Carvajal, 2013; Engelen-Eigles, Holden, Cohen, & Gardner, 2006; Rybarczyk-Plonska et al., 2016; Salehin et al., 2019). Therefore, the variety of glucosinolate chemotypes in *Pachycladon* species might be related to their abiotic stress responses as well. Recently, differences in trichome morphological characteristics have been identified between *P. stellatum* and *P. enysii*, which contribute to their different leaf reflectance. It has been suggested that such variance in trichome morphology regulates *Pachycladon* leaf absorption of radiation (Mershon, Becker, & Bickford, 2015). This provides direct phenotypic evidence for how *Pachycladon* species respond to different levels of UV-B radiation in their specific living environments.

1.4 Background and aims of the project

This project aims to explore what strategy Pachycladon species have evolved to adapt to New Zealand environment and construct the crosstalk network of the responses to multiple stresses in *P. cheesemanii*.

First of all, it has been suggested that one of Pachycladon subgenomes originated from the same ancestor as the species from Arabidopsis lineage of Brassicaceae family while the origination of another subgenome was uncertain (Zhao, Liu, Tan, & Wang, 2010). Since the wide range of Pachycladon distribution in New Zealand, I hypothesized that: *P. cheesemanii* would have similar stress responses to Brassicaceae species due to the two subgenome origination, while at the same time, *P. cheesemanii* would also have evolved some unique traits which enables it to adapt to a variety of New Zealand environments like moderately high UV-B radiation. To confirm this hypothesis, I made a number of objectives:

1. To assemble high-quality *P. cheesemanii* genome draft.
2. To construct the phylogenetic relationships of *P. cheesemanii* with other Brassicaceae species based on genome data.
3. To compare the responses in *P. cheesemanii* and *A. thaliana* to moderately high UV-B radiation
4. To compare the induced expressions of the UV-B related genes between *P. cheesemanii* and *A. thaliana* in responding to moderately high UV-B radiation.

The results of genomic analysis confirmed the close relationship of *P. cheesemanii* and *A. thaliana*, while the analysis of UV-B radiation treatment displayed the different responses between *P. cheesemanii* and *A. thaliana* to moderately high UV-B radiation.

The second hypothesis was:

1. There would be the crosstalk of multiple stress responses in *P. cheesemanii* and *A. thaliana*. Differences in crosstalk of multiple stress responses would induce distinct stress responses between species.
2. *P. cheesemanii* would have evolved some unique responses to adapt to its living environment in New Zealand relative to those in *A. thaliana*.

To test this hypothesis, I set up the following objectives:

1. To generate the sequencing data of the *P. cheesemanii* and *A. thaliana* plants that are treated and nontreated with cold, salt and UV-B radiation stresses.
2. To achieve high-quality *P. cheesemanii* transcriptome and annotation.
3. To compare the differentially expressed genes and the overrepresented terms of GO biological process between stresses in *P. cheesemanii* and *A. thaliana*.
4. To construct the crosstalk network of multiple stress responses in *P. cheesemanii* and *A. thaliana*.
5. To identify the shared biological processes by *A. thaliana* and *P. cheesemanii* stress responses.
6. To identify the unique biological processes in *A. thaliana* and *P. cheesemanii* responses to three stresses.

Chapter 2 Materials and methods

2.1 Plant growth and stress treatments

Seeds of *P. cheesemanii* Kingston (geographical coordinates in decimal degrees −45.3273, 168.7078) and Wye creek (geographical coordinates in decimal degrees −45.1398, 168.7672) were provided by Dr. Claudia Voelckel (Max Planck Institute for Chemical Ecology, Jena, Germany) and Dr. Peter Heenan (Wildland Consultants Ltd., Rotorua, New Zealand). Seeds of *Arabidopsis thaliana* (L.) Heynh. accessions Col-0 and Kondara were obtained from the *Arabidopsis* Biological Resource Center (*ABRC*; https://abrc.osu.edu). Seeds of both species were sown and germinated in wet Seed Raising Mix® soil from Daltons (Matamata, New Zealand) and seedlings were grown under a 16-h light (200 µmol m^{-2} s^{-1} cool-white fluorescent tube)/8-h dark (long-day) regime at 22 °C. For UV-B treatment, the two *P. cheesemanii* accessions were sown 10 days earlier than the *A. thaliana* plants to achieve the same plant size. After 28 (*A. thaliana*) or 38 (*P. cheesemanii*) days of growth, the plants were transferred to the UV-B radiation chamber where they were subjected to normal white light (200 µmol m^{-2} s^{-1} cool-white fluorescent tube) supplemented with 5.2 µmol m^{-2} s^{-1} UV-B (290–320 nm) for UV-B treatments, while the control plants were kept under white light conditions (**Figure 2-1**). The UV-B fluorescent tubes used in the chamber were Q-Panel 313 (Q-Lab Corp, Cleveland, OH, USA), which were wrapped in 0.13-mm-thick cellulose diacetate foil (Clarifoil; Courtaulds Ltd., Derby, UK) to remove wavelengths < 290 nm. The chamber was split into a UV-B+ zone and a UV-B− zone separated by a central curtain of UV-B opaque film (Lumivar; BPI Visqueen, Ardeer, UK). For the UV-B− zone, the UV-B tubes were wrapped in the same UV-B opaque film (Wargent, Nelson, McGhie, & Barnes, 2015). UV-B treatments were quantified at plant

canopy height with an Optronics OL-756 UV-VIS Spectroradiometer (Optronic Laboratories, Gooch and Housego, FL, USA) equipped with integrating sphere. Spectroradiometric scans of the controlled environment chamber confirmed that the biologically effective UV dose was < 0.01 kJ m^{-2} d^{-1} in the UV-B− zone. UV-B-treated and nontreated plants were collected five days after the treatment. For quantification of transcript abundance and multiple stress transcriptome profiling, seedlings were grown under a 10-h light (200 µmol m^{-2} s^{-1} cool-white fluorescent lamp)/14-h dark (short-day) regime at 22 °C for six weeks (*A. thaliana*) and nine weeks (*P. cheesemanii*) to obtain a large leaf area to maximise UV-B radiation absorption. The plants were subsequently mock treated or UV-B irradiated for five hours, after which two mature leaves of each plant were collected. For cold stress, the plants were transferred to a 4°C growth chamber with otherwise the same light settings. For salt stress, the plants were treated with 250 mM NaCl solution. Treated and nontreated plants were collected five hours after the treatment (**Figure 2-2**).

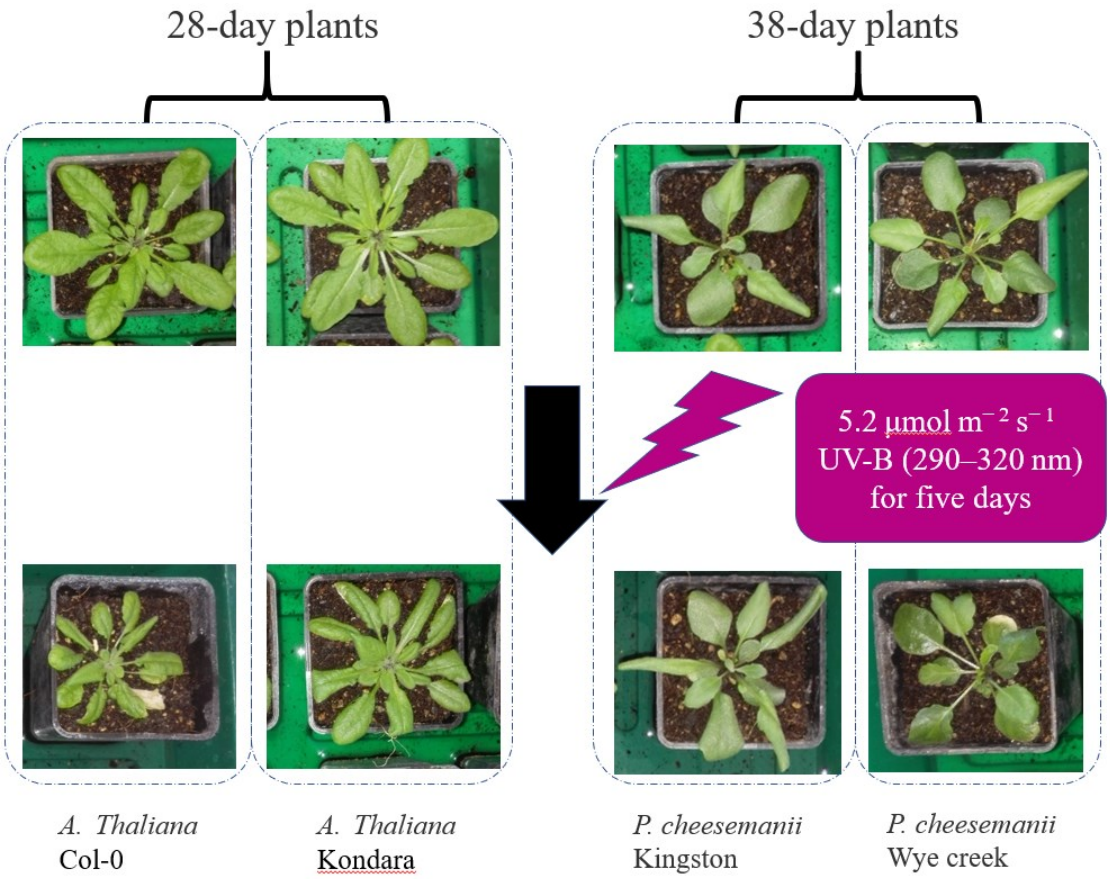

Figure 2-1 Five-day UV-B treatment of 28-day-old *A. thaliana* and 38-day-old *P. cheesemanii* plants for phenotype investgation.

28-day-old *A. thaliana* and 38-day-old *P. cheesemanii* plants after a 5-day UV-B treatment. *A. thaliana* (28 days old) and *P. cheesemanii* (38 days old) plants were grown in long day conditions and subsequently transferred to UV-B-supplemented white light for 5 days (UV-B-5-day) or to white light only (control).

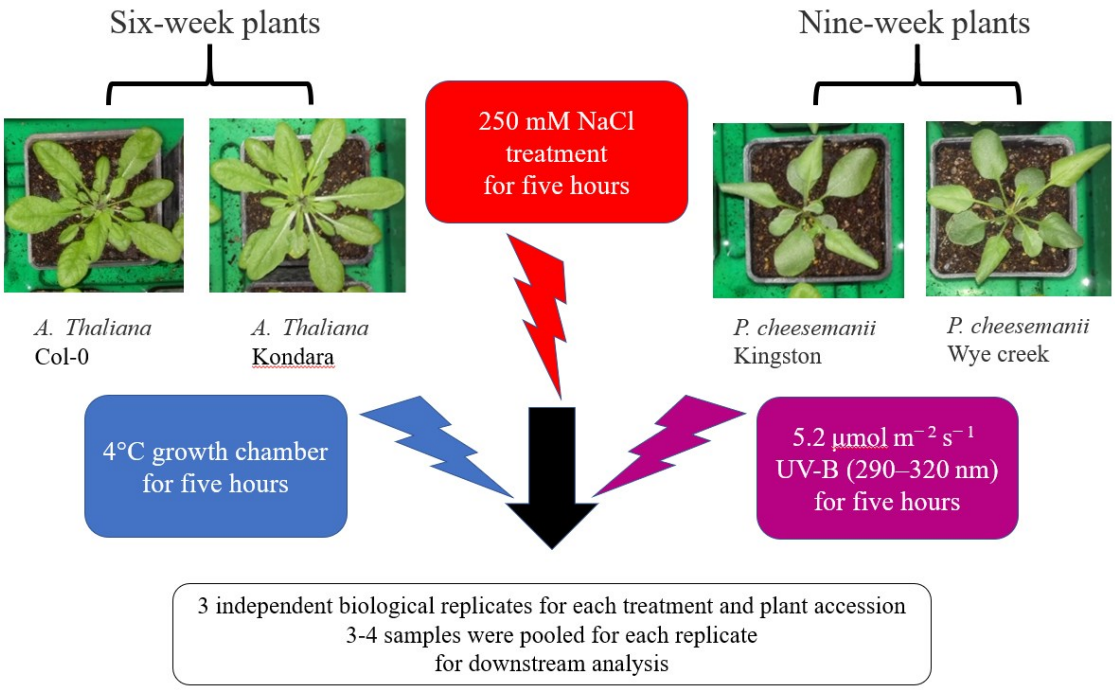

Figure 2-2 Five-hour stress treatment of six-week-old *A. thaliana* and nine-week-old *P. cheesemanii* plants for quantification of transcript abundance and multiple stress transcriptome profiling.

Six-week-old *A. thaliana* and nine-week-old *P. cheesemanii* plants after five-hour stress treatment were collected for quantification of transcript abundance and multiple stress transcriptome profiling. *A. thaliana* (six weeks old) and *P. cheesemanii* (9 weeks old) plants were grown in short-day condition and subsequently transferred to UV-B-supplemented white light, 4°C cold stress, 250 mM NaCl treatment for five hours or to white light only (control). Two mature leaves of each plant were collected after the treatments, and the leaves from three to four plants were pooled into an independent biological replicate. Three independent biological replicates for each treatment and plant accession were processed to the downstream quantification of transcript abundance and multiple stress transcriptome profiling.

2.2 Library preparation and Illumina sequencing

Genomic DNA was prepared from *P. cheesemanii* Kingston leaves. DNA was extracted using a modified CTAB DNA extraction protocol (Gawel & Jarret, 1991) with extra steps

to remove chloroplast DNA (**Supplementary File 2-1**). Purified *P. cheesemanii* genomic DNA was used to generate 550-bp fragment length paired-end (PE) and 1.5-kb to 15-kb mate-pair (MP) sequencing libraries. Following library quality control, the libraries were sequenced on two lanes of an Illumina HiSeq 2500 machine, with 2 × 125 PE and 2 × 125 MP reads generating a total of ~ 56 Gb raw sequencing data. This represents a theoretical coverage of 100-fold based on the estimated size of the *P. cheesemanii* genome (**Supplementary File 2-2**). PE and MP library construction and Illumina sequencing were performed by New Zealand Genomics Limited (NZGL, Otago, New Zealand).

2.3 Genome assembly and assessment

To assess data quality, raw reads were analyzed using SolexaQA++ v3.1.7 (Cox, Peterson, & Biggs, 2010). Adaptors were trimmed using trim_galore v0.4.1 (Krueger & Galore, 2015), and the processed reads were used for assembly. SOAPdenovo (Xie et al., 2014) and Platanus v1.2.4 (Kajitani et al., 2014) were used for assembly and different k-mer sizes were used to produce multiple assemblies for further quality assessment (**Supplementary File 2-3**). Assembly statistics were generated by QUAST v4.1 and Bowtie 2 (Gurevich, Saveliev, Vyahhi, & Tesler, 2013; Langmead & Salzberg, 2012). Leaf transcriptome data (full transcripts) of *P. cheesemanii* (Gruenheit et al., 2012) (downloaded from NCBI) were used for mapping to the genome assemblies using PASA v2.0.2 (Haas et al., 2003). BUSCO v3.0.2 (Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov, 2015) (dataset: "embryophyta_odb9" containing 1440 orthogroups, downloaded from http://busco.ezlab.org) was used to evaluate the completeness of the assemblies. The result metrics from the tools mentioned above were used to select the best assembly (**Supplementary File 2-3**). Samtools (Li et al., 2009) was used for identifying SNPs. After quality filtering, the approximate genome size was determined

using the total base number of sequencing data and sequencing depth being derived from the k-mer distribution using the formula (Zhang et al., 2012): G = (N×(L-K + 1)-B)/D; G, genome size; N, number of reads; L, length of reads; K, length of k-mer; B, low-frequency k-mers with occurrence less than four times; D, sequencing depth corresponding to selected k-mer. The heterozygosity of the genome draft was calculated by using Jellyfish v2.3.0 (Marcais & Kingsford, 2012) and GenomeScope web (http://qb.cshl.edu/genomescope/) (Vurture et al., 2017).

2.4 Identification of repeats

RepeatModeler v1.0.8 (Smit & Hubley, 2008), employing RepeatScout, Tandem Repeats Finder and RECON modules, was used to construct a *de novo* repeats library for *P. cheesemanii*. LTR_finder v1.0.5 (Xu & Wang, 2007), TransposonPSI v08222010 (Haas, 2007), and MITE-hunter v11–2011 (Han & Wessler, 2010) were used to identify Long Terminal Repeats (LTRs), retrotransposons, and Miniature Inverted repeat Transposable Elements (MITEs), respectively. Viridiplantae repeats were extracted from Repbase repeat libraries (Kapitonov & Jurka, 2008). The assembly was masked by *de novo* and Viridiplantae repeats using RepeatMasker v4.0.5 (Smit, Hubley, & Green, 2016) to estimate the non-redundant genomic repeat content. The repeats were further classified using the RepeatClassifier module of RepeatModeler.

2.5 Function annotation

A combination of *ab initio* and homology-based annotation methods was used in the Maker pipeline v2.31.8 (Campbell, Holt, Moore, & Yandell, 2014) for gene prediction. Briefly, AUGUSTUS v3.2.2 (Stanke, Diekhans, Baertsch, & Haussler, 2008) was used

with an *Arabidopsis* training set and leaf transcriptome alignments from PASA were used to improve and provide evidence for gene prediction. Gene annotation was done using BLASTP v2.6.0 (Altschul, Gish, Miller, Myers, & Lipman, 1990) and BLASTX against Uniprot (Swissprot + TrEMBL) and the NCBI nr database (only BLASTX), restricting the search to Viridiplantae using taxonomy filter, best hit, E-value cutoff of 1e-20, query coverage of ≥50% and percentage identity ≥50%. Annotations for each gene were selected as per the following order of preferences: 1) Swissprot (BLASTP), 2) Swissprot (BLASTX), 3) Uniprot (BLASTP), 4) Uniprot (BLASTX), 5) NCBI-nr (BLASTX). InterProScan v5.22–61.0 (Jones et al., 2014) was used for function annotation of the protein domains and families using all available databases in InterProScan. GO annotations were obtained from Uniprot database. Transcription factors were predicted using PlantTFDB v4.0 (Jin et al., 2016). KAAS (Moriya, Itoh, Okuda, Yoshizawa, & Kanehisa, 2007) was used to obtain KEGG pathway annotations. MISA (Thiel, Michalek, Varshney, & Graner, 2003) was used to predict SSR markers. Infernal v1.1.2 (Nawrocki & Eddy, 2013) was used to predict non-coding RNAs with an E-value cutoff of 1e-5 and removing low-scoring overlaps.

2.6 Comparative genomics

Genomes of available Brassicaceae species were downloaded from NCBI and Phytozome (Goodstein et al., 2011). Nucmer from MUMmer v3.23 (Delcher, Salzberg, & Phillippy, 2003) was used to align *P. cheesemanii* genome sequences against each Brassicaceae genome to perform synteny analysis. Delta filter was applied to select only one-to-one alignments. OrthoFinder v2.1.2 (Emms & Kelly, 2015) was used to find orthologs among the protein sequences of the selected genomes. GO annotations for *A. thaliana* were obtained from Araport11 (Cheng et al., 2017), and for *B. stricta* and *C. sativa*, the GO

annotations were obtained in the same way as for *P. cheesemanii*. Based on a novel algorithm, STAG (Species Tree inference from All Genes), OrthoFinder generated a species tree from sets of multi-copy gene trees. Then, the dendrogram was plotted using iTOL (Letunic & Bork, 2016). Overlaps between orthogroups were plotted using the ClusterVenn utility in OrthoVenn (Wang, Coleman-Derr, Chen, & Gu, 2015).

2.7 Chlorophyll content measurement

Chlorophyll content was measured by using a modified chlorophyll analysis protocol based on Ignat et al. (Ignat et al., 2013). Two fresh mature leaves were immersed in 10 ml 95% ethanol, and tubes were placed in a cool room (4 °C) for 24 h. The chlorophyll-extraction solution was then collected into new tubes, and a further 10 ml ethanol was added to the leaves for another 4 h to collect the remainder of the chlorophyll. The combined 20 ml chlorophyll-extraction solution was mixed well and 2 ml was used to quantify the chlorophyll content at 664 and 649 nm by using the following formulas (Ignat et al., 2013):

Chlorophyll a=(13.36×A664−5.19×A649)

Chlorophyll b=(27.43×A649−8.12×A664)

Total chlorophyll=(5.24×A664+22.24×A649)

Subsequently, chlorophyll contents were normalized to the corresponding total leaf areas that were used to extract chlorophyll.

2.8 RNA extraction, cDNA synthesis and RT-qPCR

Total RNA was extracted from mature leaves with a Quick-RNA MiniPrep Kit (Zymo Research, CA, USA) and treated with DNase I to remove genomic DNA contamination (**Supplementary File 2-4**). Reverse transcription was performed with an oligo (dT) primer using a Transcriptor First Strand cDNA Synthesis Kit (Roche, Basel, Switzerland). The gene sequences and primers used in RT-qPCR are listed in **Supplementary File 2-5** and **2-6**. RT-qPCR was performed using a LightCycler 480 SYBR Green I Master kit (Roche) (LightCycler 480; Roche) (**Supplementary File 2-7**). For real-time PCR quantification, *AT4G29130* was used as a reference gene in *A. thaliana*, and the homologue of *AT4G34270* was used as a reference gene in *P. cheesemanii* based on previous reports (Castelain, Le Hir, & Bellini, 2012; Castells et al., 2010). The *P. cheesemanii* homologue of *AT4G34270* was identified by applying TBLASTN.

2.9 Library preparation and Illumina transcriptome sequencing

Total RNA was extracted from mature leaves with a Quick-RNA MiniPrep Kit (Zymo Research, CA, USA) and treated with DNase I to remove genomic DNA contamination (**Supplementary File 2-4**). Purified untreated and treated plant RNA was used to generate 150-bp paired-end (PE) sequencing libraries. Following library quality control, the libraries were sequenced on Illumina NovaSeq 6000, with 2 × 150-PE reads generating a total of ~278.2 Gb raw sequencing data. PE library construction and Illumina sequencing were performed by Novogene Limited (Beijing, China).

2.10 Transcriptome assembly and gene differential expression analysis

Sequencing adaptors were removed from sequenced reads using trim_galore v0.4.1 (Krueger & Galore, 2015) and ribosomal RNA was filtered out by using SortMeRNA v2.1b (Kopylova, Noé, & Touzet, 2012). The quality filtered reads were normalised in silico by using Trinity. The normalised read sets were then independently assembled using the Trinity v2.5.1 (Grabherr et al., 2011), Velvet v1.1/Oases v0.2 (Schulz et al., 2012; Zerbino & Birney, 2008) and Trans-ABySS v2.0.1 (Robertson, et al. 2010) assemblers using a range of k-mer sizes (refer to **Table 4-2**). For each assembler, a popular range of k-mer sizes was selected (Trinity: 19-31-mer; Trans-ABySS: 51-63-mer; Velvet/Oases: 55-95-mer), and BUSCO results were used to confirm whether the best assembly was achieved in the selected range.

BUSCO (v3.0.2; dataset: 'embryophypta_odb9', containing 1,440 orthogroups, downloaded from http://busco.ezlab.org) (Simão et al., 2015) and the percentage of reads aligned using Bowtie2 were used as assembly quality metrics to select the Trans-ABySS assembly based on its superior completeness and accuracy. The final transcript assembly was clustered and further assembled using the CAP3 (Huang, & Madan, 1999) assembly program. Finally, EvidentialGene: tr2aacds (Gilbert, 2013) was applied to remove redundancies and fragments, and to identify transcript splice isoforms. To analyse differential gene expression induced by stress, *A. thaliana* and *P. cheesemanii* reads from untreated and treated samples were mapped to the *A. thaliana* transcriptome reference (GenBank CP002684–CP002688) and the *de novo* assembled *P. cheesemanii* transcriptome using kallisto v0.43.1 (Yi, Liu, Melsted, & Pachter, 2018.) Differential gene expression was then performed using edgR v3.26.1 (Robinson, McCarthy, & Smyth, 2010) with default parameters.

To generate the multiple-stress transcriptome of *A. thaliana*, six-week-old *A. thaliana* plants were treated with cold, salt and UV-B radiation stresses in the same way as was *P. cheesemanii* plants (refer to 2.1). As descried above, library construction and Illumina sequencing were then performed. The 147.1 Gb raw data with 490,400,436 raw reads from the *A. thaliana* stress response sequencing libraries was processed using trim_galore to remove adaptor contamination. SortMeRNA was then used to remove ribosomal RNA sequences from the adaptor-trimmed reads. The *A. thaliana* reference used to quantify transcripts was downloaded from www.araport.org (v10). The transcript counts and differential expression analysis were performed using kallisto/edgR with the following parameters: counts per million (cpm) >1 (removing low count), false discovery rate (FDR) <0.05 and log of fold change (FC) ≥1.

To group genes based on their expression levels I defined the following expression categories: 1, with $2 \leq |FC| < 4$, is general stress responsive; 2, with $4 \leq |FC| < 32$, is highly stress responsive; 3, with $32 \leq |FC| < 1024$, indicates an enriched stress response; and 4, with $1024 \leq |FC|$, is stress specific.

2.11 Combining weighted correlation network analysis and gene set enrichment analysis

To gain insights into stress responses in *A. thaliana*, weighted correlation network analysis (WGCNA)-based gene modules were used as gene sets for gene set enrichment analysis (GSEA). Twelve datasets from three stress responses in *A. thaliana* were utilised for WGCNA moulding. The stress-responsive genes could be moulded into 11 groups; up- and downregulated genes were identified in each. GSEA v4.0.3 (Mootha et al., 2003;

Subramanian et al., 2005) was performed on 11 WGCNA modules in each stress and the network was then generated and visualised using EnrichmentMap and the compound spring embedder (CoSE) layout in Cytoscape v3.8.0 (Shannon et al., 2003). For the interaction network analysis of stress responses in *P. cheesemanii*, GSEA was applied to the datasets. The resulting network was also generated and visualised using EnrichmentMap and the CoSE layout in Cytoscape.

All the code for bioinformatics analysis could be found in **Supplementary File 2-8**.

Chapter 3 Genome draft of the Arabidopsis relative *Pachycladon cheesemanii* reveals novel strategies to tolerate New Zealand's high UV-B radiation environment

3.1 Introduction

Pachycladon is an allopolyploid genus of the Brassicaceae family with eight perennial species endemic to the South Island of New Zealand and one species to Tasmania (Australia). These *Pachycladon* species are believed to have originated around 1–3.5 million years ago in New Zealand and are primarily distributed across the alpine regions of the South Island (Heenan & Mitchell, 2003; Yogeeswaran, Voelckel, Joly, & Heenan, 2011). *Pachycladon cheesemanii* is the most widespread of the *Pachycladon* species with a broad longitudinal distribution in New Zealand and a wide altitudinal range from 10 m to 1600 m above sea level (Heenan & Mitchell, 2003).

Pachycladon's allopolyploid genome ($2n = 20$) consists of two subgenomes which resulted from intra- or interspecific crossing (Song & Chen, 2015). Karyotype comparisons between extant *Pachycladon* species and the theoretical Ancestral Crucifer Karyotype showed that the chromosome structure had undergone multiple rearrangements prior to the allopolyploidy event taking place (Mandakova, Heenan, & Lysak, 2010), and this has hampered efforts to trace back *Pachycladon*'s progenitors. Phylogenetic analysis of *Pachycladon* species based on five single-copy nuclear genes indicated that one of the genome copies was derived from the *Arabidopsis* lineage, while another was similar to both *Arabidopsis* and *Brassica* lineages (Joly, Heenan, & Lockhart, 2009). However, a comparison of 547 homeologous gene pairs from *P. cheesemanii* and *P. fastigiatum* with the homologous genes from *Arabidopsis*

lyrata and *Arabidopsis thaliana* found that no set of genes showed significantly different identity to *A. lyrata* and *A. thaliana* homologues, suggesting the two *Pachycladon* subgenomes are derived from the same lineage (Gruenheit et al., 2012). Data from analysis of the nuclear gene *CHALCONE SYNTHASE* (*CHS*) further supported the idea that both *Pachycladon* genome copies stem from the *Arabidopsis* lineage (Zhao, Liu, Tan, & Wang, 2010).

Polyploidization has been suggested to contribute to plants' evolution and environmental adaptation under selection pressure (Bowers, Chapman, Rong, & Paterson, 2003; Brochmann et al., 2004; Wood et al., 2009). Plants with polyploid genomes can benefit from functional diversification of redundant gene copies, with one gene copy retaining the original function, guaranteeing the plant's regular growth and development, while the other can evolve to confer novel phenotypes, such as protection against challenging environmental conditions (Comai, 2005). Thus, higher levels of UV radiation in New Zealand compared with locations in the Northern Hemisphere at similar latitudes may have contributed to the evolution of the *Pachycladon* species (McKenzie, Liley, Kotkamp, Shiona, & Lopez, 2014).

UV radiation is classified into three types, UV-A, UV-B and UV-C. While UV-C does not penetrate the atmosphere, some UV-B radiation reaches Earth's surface, where it can damage important molecules like DNA. In order to acclimate to UV-B radiation, plants have developed multiple strategies, including reducing leaf area by curling of the leaves, inhibiting leaf and plant growth (Nascimento, Moreira, Leal-Costa, Costa, & Tavares, 2015; Qaderi, Yeung, & Reid, 2008) and increasing light reflection by inducing the production of a cuticular wax layer and the biosynthesis of light-absorbing secondary

metabolites (Lavola, Julkunen-Tiitto, Aphalo, Rosa, & Lehto, 2010; Steinmüller & Tevini, 1985). Nevertheless, excess UV-B radiation can cause the development of hypersensitive response-like necrotic lesions and plant death (Hectors, Prinsen, De Coen, Jansen, & Guisez, 2007; Jenkins, 2014; Kalbina & Strid, 2006).

UV-B radiation is perceived by the UVB-resistance 8 (UVR8) photoreceptor which was discovered by the UV-B hypersensitivity of the *uvr8* mutant (Kliebenstein, Lim, Landry, & Last, 2002). The crystal structure of the UVR8 protein showed that its core domain consists of a covalently bound homodimer (Wu et al., 2012). After UV-B radiation, this homodimer dissociates and monomeric UVR8 interacts with constitutive photomorphogenic 1 (COP1) and transcription factors including elongated hypocotyl 5 (HY5) and HY5-homolog (HYH) to induce the expression of UV-B-responsive genes (Heijde & Ulm, 2012). Induced genes included those that encode CHS, flavanone 3-hydroxylase (F3H) and flavonol synthase 1 (FLS1), which are core enzymes involved in the biosynthesis of flavonoids (Di Ferdinando, Brunetti, Fini, & Tattini, 2012) and are believed to function as a UV-absorbing sun screen (Day, 1993). Other induced genes include *PHOTOLYASE 1* (*PHR1*), which encodes protein phosphate starvation response 1, and *EARLY LIGHT-INDUCIBLE PROTEIN1* (*ELIP1*). ELIP1 plays a role in the interaction of UV-B-induced monomeric UVR8 with chromatin (Cloix & Jenkins, 2008). It was found that the UVR8-dependent pathway responds to a wide range of UV-B radiation (0.1–12 μmol m^{-2} s^{-1}). Another less well-understood UV-response pathway was found that functions independently of UVR8. By treating *uvr8* mutants with relatively high UV-B radiation levels (1–12 μmol m^{-2} s^{-1}), several genes induced by this pathway were identified (Brown & Jenkins, 2008).

Since *P. cheesemanii* survives in New Zealand's high UV-B radiation environments, this species may have evolved distinct UV-B-radiation response pathways. To learn how this species is able to cope in its unique environment, we first assembled a high-quality draft genome of *P. cheesemanii* and attempted to reveal the two highly similar subgenomes. The draft genome was used to identify *P. cheesemanii* candidate genes likely involved in UV-B radiation response pathways. However, interestingly, the UV-B-induced expression pattern of these genes differed from that observed in two *A. thaliana* accession with differing UV-B responses, suggesting that a distinct UV-B radiation response pathway has evolved in *P. cheesemanii* to enable adaptation to the high UV-B radiation environment in New Zealand.

3.2 Results

3.2.1 Genome assembly and assessment

We extracted *P. cheesemanii* Kingston genomic DNA for whole genome sequencing. The Illumina sequencing technology was used to obtain high coverage sequence reads to help us determine its ancestry and current gene-set. Paired-end and mate-pair libraries were sequenced and ~ 56 Gb of DNA sequence obtained. Raw reads (483,792,966 reads) were subsequently trimmed using the cutadapt algorithm that is present in the trim_galore package. Using k-mer analysis the genome size was estimated to encompass 596 Mb (**Figure S3-1**), and the genome heterozygosity between 0.063 and 0.069%. Multiple aligners (Platanus and SOAPdenovo) with different k-mer lengths were used to generate genome assemblies. Subsequently, these assemblies were further evaluated using multiple metrics, and the best one (51-k-mer assembly) was selected based on the assembly size and N50 from Platanus (P.k51) (**Supplementary File 2-3**). The assemblies using SOAP resulted in a higher scaffold size compared to Platanus, but also a much

higher number of gaps and lower percentage of complete single copy orthologues. Therefore, Platanus was used as the preferred genome assembler. The total assembly size using P.k51 was ~ 422 Mb and this represented 70.8% of the estimated genome size. The longest scaffold was 418 kb, while the number of scaffolds of length ≥ 500 and ≥ 1000 bases were 53,782 and 23,900, spanning ~ 300 Mb and ~ 280 Mb of assembly size, respectively. The N50 for the assembly (scaffolds ≥500 bp) was 24,761 bases (**Table 3-1**). This result indicated that the assembled genome draft was highly fragmented.

Table 3-1 Assembly statistics of the *P. cheesemanii* genome.

	Platanus assembly (P.k51)
Total assembly size (bp)	422,560,840
Number of scaffolds (≥500 bp)	53,782
Longest scaffold (bp)	418,003
N50 (≥500 bp)	24,761
GC (%)	36.33
Number of Ns / 100 kb (bp)	749.01
Repeats (%)	42.96
Number of variants	434,467

A high amount of repetitive DNA in the genome could be one reason for the fragmented genome assembly (Schatz, Witkowski, & McCombie, 2012). Therefore, the repeat

content in the genome draft was analyzed using different repeat identification tools, and it was estimated that ~43% of the total assembly size comprised repeat regions (**Supplementary File 3-1** and **Figure S3-2**). Among these, 15.96% were annotated as "retrotransposons", 6.84% as "DNA transposons" and 19.89% as "unclassified repeats".

BUSCO assessment revealed that 96.2% highly conserved plant orthologs were "complete", 1.5% "fragmented" and 2.3% "missing". Reads were mapped back to the assembly using Bowtie 2 to show 96.98% alignment (**Table 3-2**). The *P. cheesemanii* leaf transcriptome (Gruenheit et al., 2012) was aligned against the assembled genome using PASA, and 97.94% of transcripts could be mapped to the genome (**Table 3-2**). A total of 47,821 protein coding genes were predicted using MAKER, with an average transcript size of 1544 bp and 4.42 exons per gene. With regard to non-coding RNAs, 115 rRNA, 707 tRNA and 209 miRNA genes were predicted. In addition, in a comparison of the alleles in *P. cheesemanii,* 434,467 SNPs and 123,778 SSRs were identified, highlighting the highly polymorphic information content of its genome (**Supplementary File 3-2**). Thus, the results showed a fragmented genome draft, which may be the result of the high number of repeat elements in non-coding regions or/and having two highly similar genomes to contend with. Nevertheless, the assembly of coding regions was deemed of high quality, based on BUSCO and PASA analyses.

Table 3-2. Assessment statistics of the *P. cheesemanii* genome.

	Percentage (%)
Read alignment	96.98
Transcript alignment	97.94
BUSCO completeness	96.20

3.2.2 Genome annotation

Each of the predicted genes was functionally annotated by using BLASTX against National Center for Biotechnology Information (NCBI) non-redundant protein (Pruitt, Tatusova, & Maglott, 2006) and Uniprot databases for green plants (Viridiplantae) (**Table 3-3**). About 84% of the predicted genes had a blast hit against either NCBI nr or Uniprot databases, or against both. Among these, 63% had a hit in the manually curated Swissprot database. Based on the BLASTX result against NCBI nr, the highest number of hits was with *Camelina sativa* (24.4%), followed by *Arabidopsis lyrata* (22.7%), *Arabidopsis thaliana* (19.0%) and *Capsella rubella* (17.3%), all belonging to the Brassicaceae family (Huang et al., 2015). InterProScan identified protein signatures for 89.81% of the predicted proteins, and 2597 genes were classified as transcription factor (TF) encoding genes. Similar to *A. thaliana*, bHLH (239), MYB (212), ERF (211) and NAC (179) TFs comprised the largest TF families in *P. cheesemanii*. The predicted genes were used for classification into pathways using the KEGG database. Similar to other plant species, the terms “metabolic pathways” and “biosynthesis of secondary metabolites” were assigned to the largest numbers of the predicted genes in *P. cheesemanii* (2930 and 1594, respectively) (**Supplementary File 3-3**).

Table 3-3. Annotation statistics of the *P. cheesemanii* genome.

Number of predicted genes	47,821
Average transcript length (bp)	1,544.46
Average CDS length (bp)	941.27
Average number of exons per gene	4.42
Average exon length (bp)	212.92
Average intron length (bp)	176.32
Length of contigs (≥500 bp)	299,926,053
Length of contigs (≥1 kb)	279,782,042

3.2.3 Synteny analysis of the *P. cheesemanii* genome draft within Brassicaceae species

It has been reported that the two *Pachycladon* subgenomes originate from the hybridization of two species of the Brassicaceae family, one each from the *Arabidopsis* and *Brassica* lineages (Joly et al., 2009). Here, the *P. cheesemanii* genome was aligned against all publicly available Brassicaceae genomes using MUMmer to perform synteny analysis. Of 28 available Brassicaceae genomes, seven each were from the Brassiceae and Camelineae tribes, four from the Eutremeae tribe, three from the Arabideae tribe, two from the Cardamineae tribe, and one each from the Thlaspideae, Sisymbrieae, Euclidieae, Boechereae, and Aethionemeae tribes (the tribes of Brassicaceae Lineage I: Camelineae, Cardamineae, and Boechereae; the tribes of Lineage II: Sisymbrieae and Brassiceae (**Table 3-4**); the tribe of Lineage III: Euclidieae;

the tribes of Expanded Lineage II (EII): Thlaspideae and Eutremeae; the tribe of the basal lineage: Aethionemeae; the unassigned tribe: Arabideae) (Huang et al., 2015). *Tarenaya hassleriana* from the Cleomaceae family was selected as an outgroup (Huang et al., 2015). Species with the highest alignment percentage (Maximal Unique Matches: MUMs) against the *P. cheesemanii* genome belong to Boechereae (29%), Camelineae (~ 20%) and Eutremeae (~ 15%). All pairwise combinations of the Brassicaceae genomes were used to estimate the cumulative alignment percentage with the *P. cheesemanii* genome to determine possible ancestral genomes of *Pachycladon*. The combination of *Boechera stricta* and *Eutrema heterophyllum* had the highest cumulative alignment with *P. cheesemanii* (37.35%) at the genome level (**Figure 3-1a**).

From the species with the highest alignment percentage against the *P. cheesemanii* genome, three species from Brassicaceae Lineage I (*C. sativa*, *A. thaliana* and *B. stricta*, two from the Camelineae tribe, and one from the Boechereae tribe) and one from Lineage EII (*E. heterophyllum*, from Eutremeae tribe) (Huang et al., 2015) were selected for protein ortholog analysis. To identify orthologs, predicted proteins of all five species were blasted against each other in a pairwise manner for a total of 25 combinations. The BLAST searches were further processed using OrthoFinder to identify orthologs. A total of 182,585 genes (76%) were assigned to 20,553 orthogroups that included 14,971 orthogroups shared within the five species (**Figure 3-1b**). For *P. cheesemanii*, 66.4% of the genes (31,749) were assigned to 87% (17,881) of the total orthogroups. Among these orthogroups, 15 novel orthogroups containing 72 genes were present in *P. cheesemanii*. Based on the orthogroups, a dendrogram of the five species was constructed (**Figure 3-1c**). In accordance with the synteny analysis, *P. cheesemanii* showed the closest relationship with *B. stricta*, followed by *C. sativa* and *A.*

thaliana. Beside the orthogroups that were shared by all species, *P. cheesemanii* shared the highest number of orthogroups with *C. sativa* (2191), followed by *B. stricta* (1753), *A. thaliana* (1721) and *E. heterophyllum* (923). Thus, the data suggests that *P. cheesemanii* has a closer phylogenetic relationship with species from Lineage I of the Brassicaceae family than to those of Lineage EII.

Table 3-4. Summary of Brassiceae genomes from *Brassica* lineage.

	Assembly level	Ploidy level	n
Brassica juncea	chromosome	allotetraploid	18
Brassica napus	chromosome	allotetraploid	19
Brassica nigra	chromosome	diploid	8
Brassica oleracea	chromosome	diploid	9
Brassica rapa	chromosome	diploid	10
Raphanus sativus	scaffold	diploid	9
Raphanus raphanistrum	contig	diploid	9

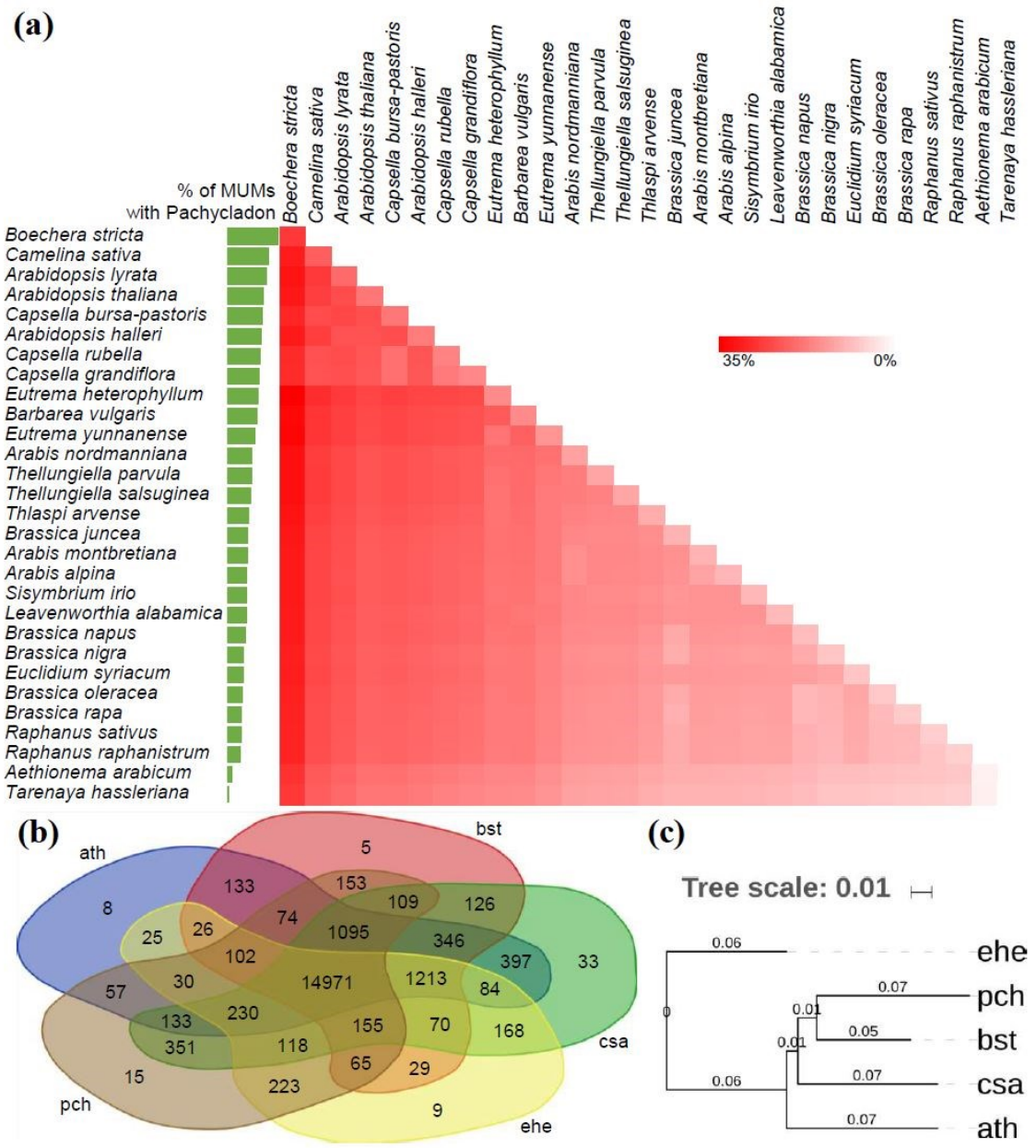

Figure 3-1 Prediction of the origin of the *P. cheesemanii* genome.

Prediction of the origin of the *P. cheesemanii* genome. **a** MUMmer alignment percentage (MUMs: Maximal Unique Matches) of *Pachycladon* against other sequenced Brassicaceae genomes. The numbers indicate cumulative percentage of MUMs for the respective pair of species against *P. cheesemanii*. **b** OrthoFinder output showing orthologous clusters between *P. cheesemanii* (pch), *A. thaliana* (ath), *B. stricta* (bst), *E. heterophyllum* (ehe) and *C. sativa* (csa). **c** Dendrogram of five species with high scores in MUMmer alignment. Numbers represent branch lengths.

Next, we used the *P. cheesemanii*, *B. stricta*, *E. heterophyllum* and *A. thaliana* genomes to analyze the GO enrichment patterns to further study the phylogenetic relationships of these species. The predicted gene annotations encompassed all major GO terms, suggesting that a core GO term set is present in the *P. cheesemanii* genome annotation (**Figure 3-2** and **Supplementary File 3-4**). A comparison with the GO enrichment distributions of *B. stricta*, *E. heterophyllum* and *A. thaliana* revealed a similar pattern across all three GO categories in *P. cheesemanii* and *B. stricta*, while the *E. heterophyllum* pattern was considerably different from the other three species of Brassicaceae Lineage I (**Figure 3-2**). Therefore, this result provides further support for the closer evolutionary grouping of *P. cheesemanii* with *B. stricta* of Brassicaceae Lineage I, than to *E. heterophyllum* of Lineage EII.

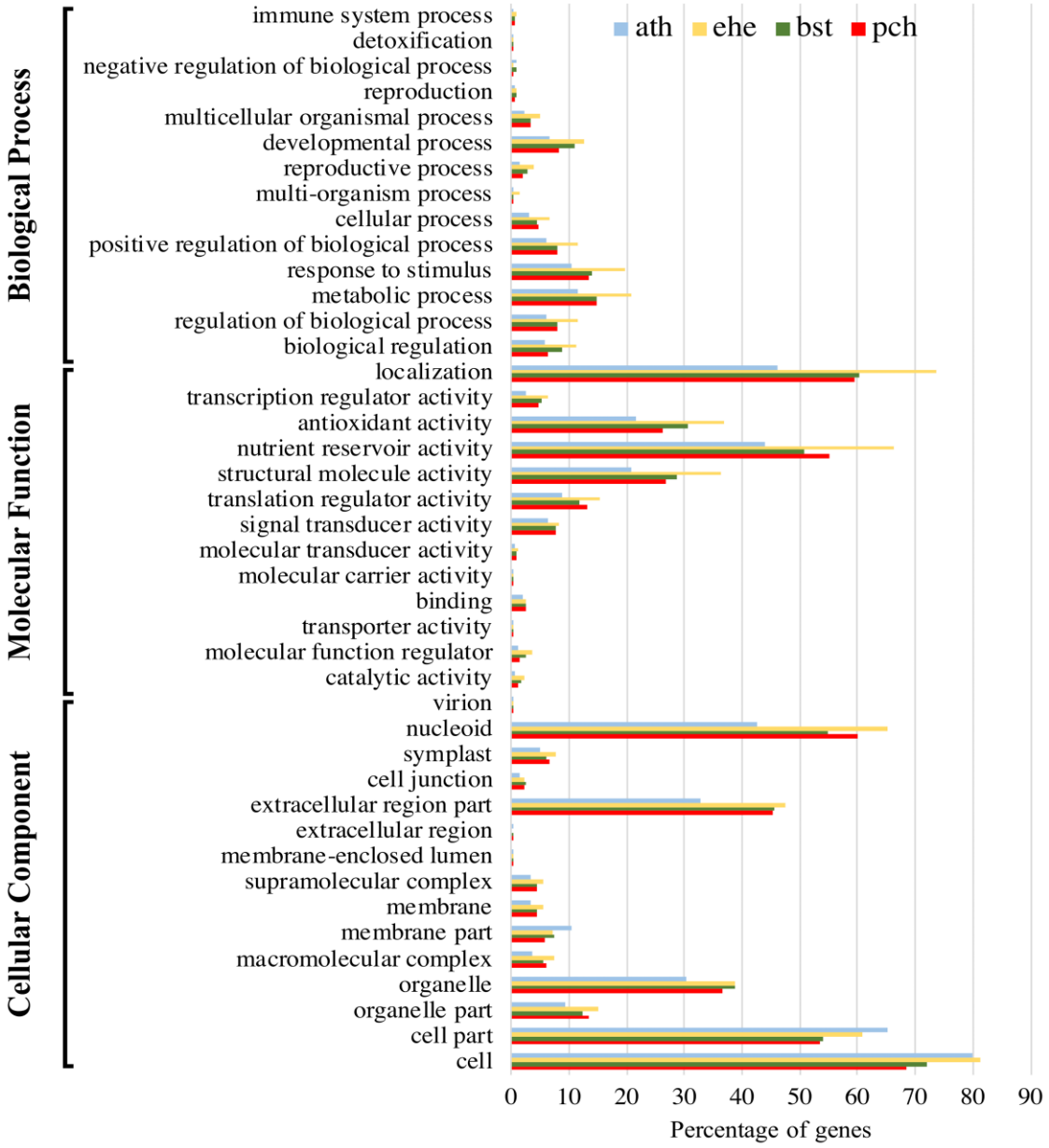

Figure 3-2 Gene Ontology (GO) annotation.

Gene Ontology (GO) annotation. Comparison of GO terms between *P. cheesemanii* (pch), *A. thaliana* (ath), *B. stricta* (bst) and *E. heterophyllum* (ehe).

3.2.4 Different UV-B responses in *P. cheesemanii* and *A. thaliana*

The New Zealand environment is prone to high UV-B radiation levels naturally (McKenzie, Connor, & Bodeker, 1999). We therefore hypothesized that *P. cheesemanii* has evolved a higher UV-B radiation tolerance than its close relative, *A.*

thaliana. Two accessions of *P. cheesemanii* were obtained from locations of relatively close proximity to each other. *P. cheesemanii* Kingston was collected just west of Kingston, New Zealand, at an altitude of ~ 500 m and *P. cheesemanii* Wye creek was collected 20 km north of Kingston at an altitude of ~ 300 m. The *P. cheesemanii* phenotypes were compared against those of the widely studied *A. thaliana* accession Col-0, which grows at an altitude of up to 100 m (www.arabidopsis.org), and the UV-B-resistant accession Kondara (distribution altitude: 1000–1100 m) (Gegas et al., 2014; Sánchez-Bermejo et al., 2014). To test for responses to UV-B radiation, 28-day-old *A. thaliana* plants and 38-day-old *P. cheesemanii* plants, of similar plant size, were treated with UV-B radiation for 5 days to allow the manifestation of typical UV-B radiation phenotypic responses. A moderately high UV-B radiation (5.2 µmol m^{-2} s^{-1}) was used to induce both UVR8-dependent and -independent responses.

Leaves of UV-B radiation-treated *A. thaliana* Col-0 and Kondara plants were significantly smaller than leaves from untreated controls, and the Col-0 accession displayed more necrotic lesions on its leaves than Kondara (**Figure 3-3a**, **b**, **e**, **f**, **i**, **j** and **Figure 3-4a**). *P. cheesemanii* Wye creek plants showed a smaller but significant decrease in leaf size upon UV-B radiation compared to untreated controls. Interestingly, the leaf size of *P. cheesemanii* Kingston was not affected by UV-B radiation (**Figure 3-4a**). All plants displayed some leaf curling and the leaves attained a glossy appearance, which was most apparent in *P. cheesemanii* Wye creek (**Figure 3-3c**, **d**, **g**, **h**, **k**, **l**).

Figure 3-3 Twenty-eight-day-old *A. thaliana* and 38-day-old *P. cheesemanii* plants after a five-day UV-B treatment.

Twenty-eight-day-old *A. thaliana* and 38-day-old *P. cheesemanii* plants after a 5-day UV-B treatment. *A. thaliana* (28 days old) and *P. cheesemanii* (38 days old) plants were grown in long day conditions and subsequently transferred to UV-B-supplemented white light for 5 days (UV-B-5-day) or to white light only (control). **a** *A. thaliana* Col-0 **b** *A. thaliana* Kondara **c** *P. cheesemanii* Kingston **d** *P. cheesemanii* Wye creek plants grown under control conditions. **e** *A. thaliana* Col-0 **f** *A. thaliana* Kondara **g** *P. cheesemanii* Kingston **h** *P. cheesemanii* Wye creek plants after UV-B treatment. **i-l** Enlarged insets are shown for UV-B-treated plants (e-h) only. Arrows indicate necrotic lesions (white), leaf curling (green) and glossy appearance (yellow), respectively. Scale bars, 3.5 cm

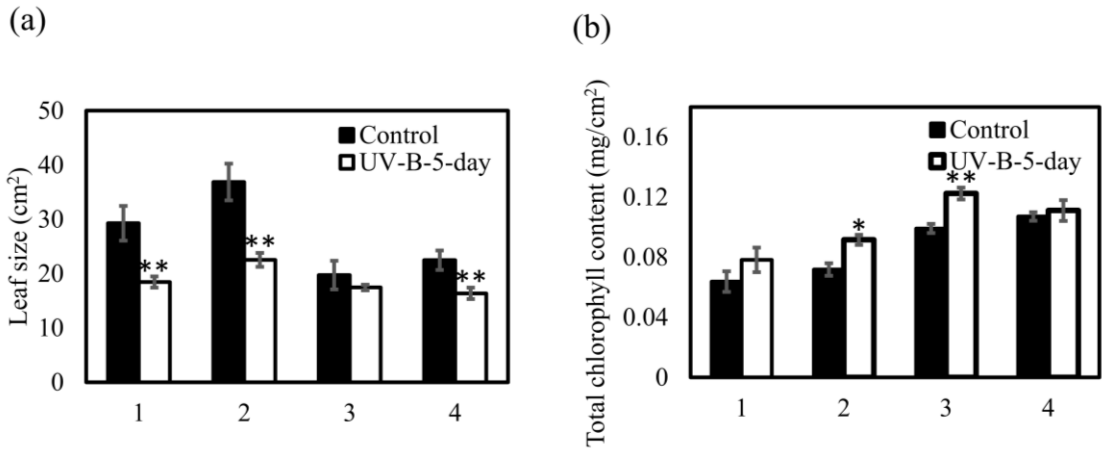

Figure 3-4 Total chlorophyll content and total leaf size of *A. thaliana* and *P. cheesemanii* plants grown with and without UV-B radiation.

Total chlorophyll content and leaf size of *A. thaliana* and *P. cheesemanii* plants grown with and without UV-B radiation. *A. thaliana* (28 days old) and *P. cheesemanii* (38 days old) plants were grown in long day conditions and subsequently transferred to UV-B-supplemented white light for 5 days (UV-B-5-day) or to white light only (control). **a** Total leaf area. **b** Total leaf chlorophyll content. 1, *A. thaliana* Col-0; 2, *A. thaliana* Kondara; 3, *P. cheesemanii* Kingston; 4, *P. cheesemanii* Wye creek. Error bars represent SEM (Student's *t*-test; *, $p < 0.05$; **, $p < 0.01$). Data were collected from 4 to 8 biological replicates

Taken together, our results support the notion that *P. cheesemanii* accessions exhibit a higher UV-B radiation tolerance than the *A. thaliana* accessions. Moreover, the two *P. cheesemanii* accessions responded to UV-B radiation in different ways.

3.2.5 Distinct expression of UV-B radiation-inducible genes in *P. cheesemanii* and *A. thaliana*

To further examine the UV-B radiation responses in *P. cheesemanii* and *A. thaliana*, we identified the *P. cheesemanii* homologues of 11 *A. thaliana* genes that function in the UVR8-dependent pathway and three homologues that play a role in the UVR8-independent pathway. The protein sequences of these genes were used to search the *P.*

cheesemanii genome draft using TBLASTN. As a result, at least two potential copies of each gene were identified (**Figure S3-3** and **Supplementary File 2-5**), consistent with the polyploid nature of the *P. cheesemanii* genome. Primers for the *P. cheesemanii* genes were designed to amplify conserved protein-coding regions, such that both copies were expected to be amplified with equal efficiency.

P. cheesemanii and *A. thaliana* plants were treated with UV-B radiation for 5 h to focus on early transcriptional effects and limit secondary responses. Gene expression of the selected genes was measured by quantitative real-time polymerase chain reaction (RT-qPCR). We initially measured 11 genes induced in *A. thaliana* by the UVR8-dependent pathway and found that eight (*HY5*, *HYH*, *CHS*, *ELIP1*, *CRYPTOCHROME 3* (*CRY3*), *GLUTATHIONE PEROXIDASE 7* (*GPX7*), *SIGMA FACTOR 5* (*SIG5*), and *WALL-ASSOCIATED RECEPTOR KINASE-LIKE 8* (*WAKL8*)) were upregulated by UV-B radiation in both *A. thaliana* accessions and three were not (*BCB*, a gene encoding a blue copper binding protein, *COP1*, and *GEM-RELATED 5* (*GER5*), which encodes a protein involved in hormone-mediated regulation of seed germination). Interestingly, while most of these genes were also upregulated in both *P. cheesemanii* accessions, the extent of upregulation was generally lower (**Figure 3-5**).

We next quantified three genes of the UVR8-independent pathway, i.e., genes encoding *Arabidopsis thaliana* WRKY DNA-binding protein 30 (*WRKY30*), uridine diphosphate glycosyltransferase 74E2 (*UGT74E2*), and FAD-linked oxidoreductase (*FOX1*), and none of those was induced significantly in the *A. thaliana* accessions by 5.2 μmol m^{-2} s^{-1} of UV. However, the *WRKY30* homologue was upregulated in both *P. cheesemanii* accessions and the transcript levels of *UGT74E2* and *FOX1* were elevated

in *P. cheesemanii* Wye creek, but not in *P. cheesemanii* Kingston. Thus, *A. thaliana* and *P. cheesemanii* accessions responded in different ways to UV-B radiation.

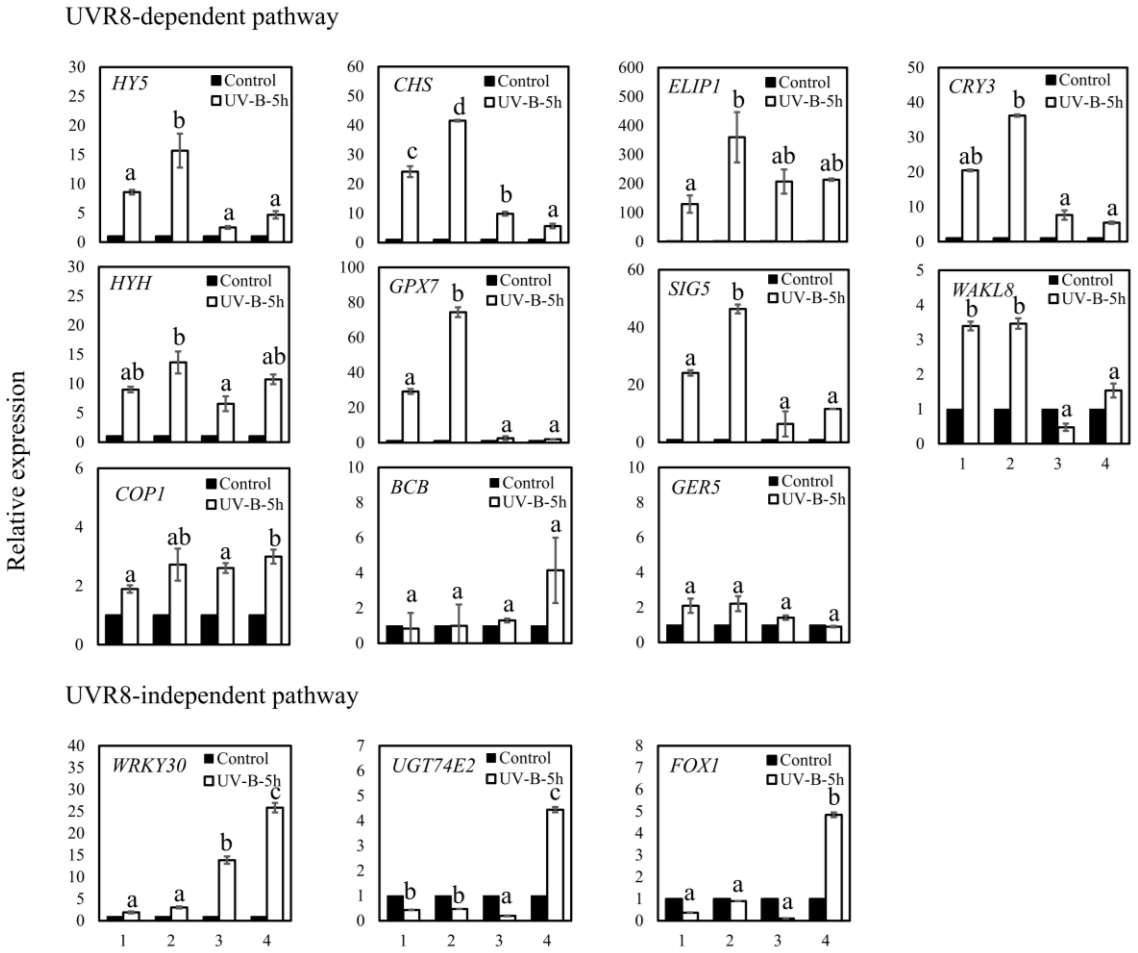

Figure 3-5 Relative expression of genes in UVR8-dependent and UVR8-independent UV-B response pathways in *A. thaliana* and *P. cheesemanii*.

Relative expression of genes in UVR8-dependent and UVR8-independent UV-B response pathways in *A. thaliana* and *P. cheesemanii*. Six-week-old *A. thaliana* and nine-week-old *P. cheesemanii* plants were treated with white light (Control) or white light supplemented with UV-B radiation for 5 h (UV-B-5 h). Mature leaves were harvested for RT-qPCR analysis. The expression of transcripts indicated in each of the graphs was normalized to the reference genes of *A. thaliana* and *P. cheesemanii*, respectively. Relative expression (fold change) was calculated by dividing expression values of UV-B-treated plants with those of control-treated plants. 1, *A. thaliana* Col-0; 2, *A. thaliana* Kondara; 3, *P. cheesemanii* Kingston; 4, *P. cheesemanii* Wye creek. Data were from three independent biological replicates. Error bars represent SEM. Means with different lowercase letters indicate values of UV-B treated samples that are significantly different between the accessions analyzed (Tukey's HSD, $p < 0.05$).

3.2.6 Similar UV-B radiation-repair systems in *P. cheesemanii* and *A. thaliana*

Plants reduce susceptibility to UV radiation-induced damage through photorepair and dark repair systems (Gill, Anjum, Gill, Jha, & Tuteja, 2015). Here, we identified *P. cheesemanii* homologues of six key genes involved in UV-B radiation-repair systems in *A. thaliana*. The UV-B radiation-induced transcript level of each gene was subsequently measured in *A. thaliana* and *P. cheesemanii* by RT-qPCR. In response to UV-B radiation, the two photorepair genes *PHOTOLYASE 1* (*PHR1*) and *UV REPAIR DEFECTIVE 3* (*UVR3*) were significantly upregulated in all four plant accessions, with *A. thaliana* Kondara showing the highest increase (**Figure 3-6**). Interestingly, four genes involved in dark repair (nucleotide excision repair) or representing the *Arabidopsis* homologue of *Xeroderma Pigmentosum Complementation Group B1* (*XPB1*) from humans, i.e., *Arabidopsis thaliana XERODERMA PIGMENTOSUM GROUP D* (*XPD*), and a homologue of the human *ERCC1* gene (*ERCC1*), did not show obvious UV-B-induced transcript level changes in *A. thaliana* Col-0 or the two *P. cheesemanii* accessions, while *XPD* and *ERCC1* were upregulated 3.5-fold in *A. thaliana* Kondara (**Figure 3-6**). Thus, the photorepair genes were upregulated in all plant accessions, while only *A. thaliana* Kondara showed an activated dark repair system.

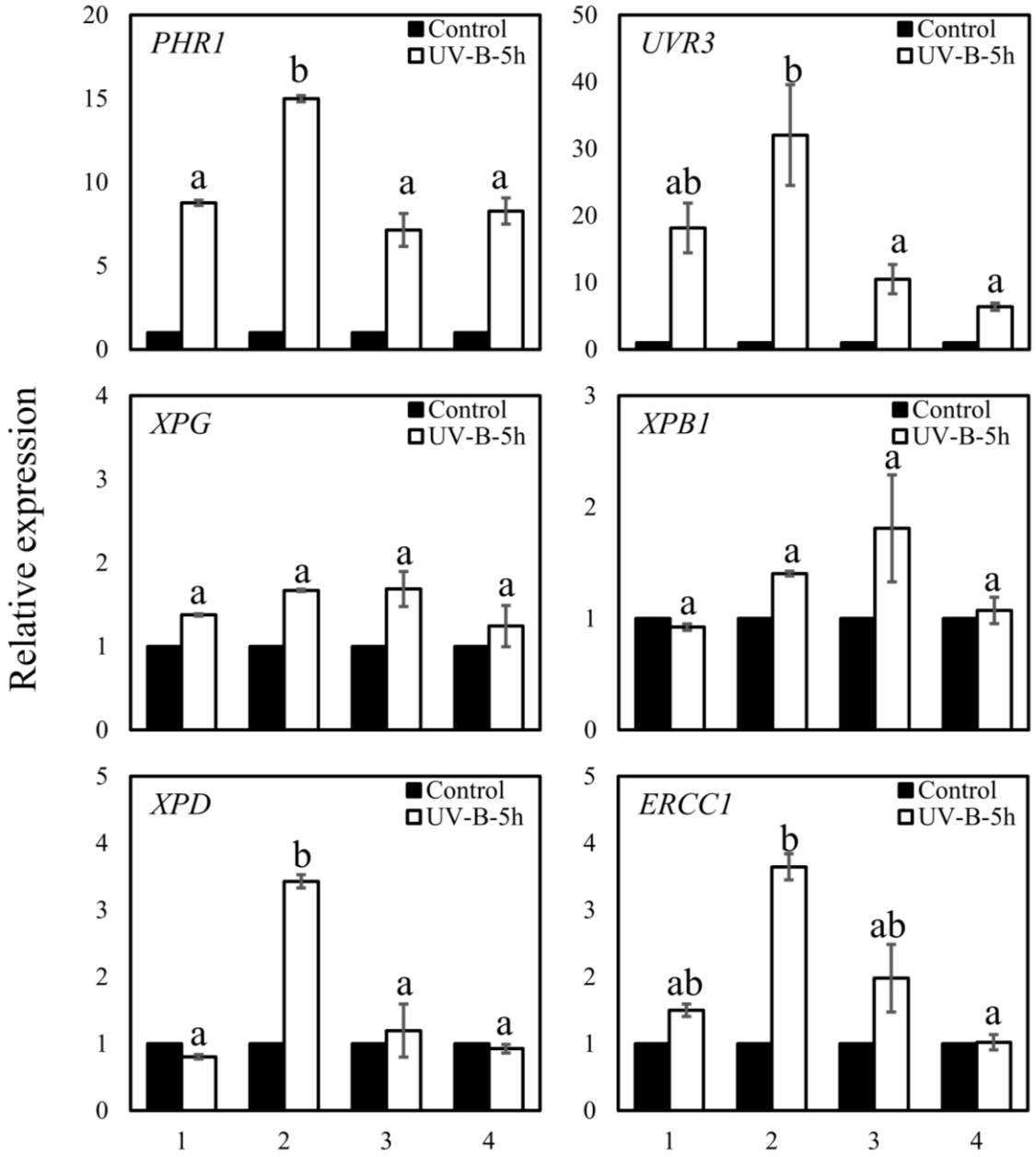

Figure 3-6 Relative expression of genes involved in DNA damage repair in *A. thaliana* and *P. cheesemanii*.

Relative expression of genes involved in DNA damage repair in *A. thaliana* and *P. cheesemanii*. Two genes involved in photorepair (*PHR1* and *UVR3*) and four genes involved in nucleotide excision repair were selected. Six-week-old *A. thaliana* and nine-week-old *P. cheesemanii* plants were treated with white light (Control) or white light supplemented with UV-B radiation for 5 h (UV-B-5 h). Mature leaves were harvested for RT-qPCR analysis. The expression of transcripts indicated in each of the graphs was normalized to the reference genes of *A. thaliana* and *P. cheesemanii*, respectively. Relative expression (fold change) was calculated by dividing expression values of UV-B-treated plants with thoes of control-treated plants. Data were from

three independent biological replicates. 1, *A. thaliana* Col-0; 2, *A. thaliana* Kondara; 3, *P. cheesemanii* Kingston; 4, *P. cheesemanii* Wye creek. Error bars represent SEM. Means with different lowercase letters indicate values of UV-B treated samples that are significantly different between the accessions analyzed (Tukey's HSD, $p < 0.05$).

3.3 Discussion

3.3.1 *P. cheesemanii* may originate from different Brassicaceae lineages

The two *Pachycladon* subgenomes have been proposed to result from a relatively young hybridization event, between unknown species, that occurred in New Zealand about 1.61 million years ago (Joly et al., 2009). Here, we sequenced the *P. cheesemanii* genome and used it to help trace back the origin of the subgenomes of *Pachycladon* and to analyze UV-B radiation responses.

In our synteny analysis of 28 available Brassicaceae genomes, all eight species from tribes Boechereae and Camelineae (genera *Arabidopsis*, *Camelina*, *Capsella* and *Boechera*) displayed consistent high cumulative genome alignment percentages, regardless as to which of the eight species was combined with any of the 27 other species. Values for Boecherea (*B. stricta*) were the highest, suggesting that at least one of the *P. cheesemanii* subgenomes originated from the *Arabidopsis* lineage (**Figure 3-1a**). Nevertheless, no cumulative values higher than 50% were found, suggesting that only one of the *Pachycladon* subgenomes has the same origin as these available Brassicaceae genomes. The findings are in general agreement with a recent Brassicaceae phylogenetic tree, established using 113 nuclear genes as markers, which proposes that *Pachycladon* is closely related to species of *Arabidopsis*, *Camelina*, *Capsella* and *Boechera* (Huang et al., 2015). Similarly, Joly et al. (Joly et al., 2009) reported that one of the predicted ancestral subgenomes of *Pachycladon* has close phylogenetic relationships

with *Boechera* and *Capsella* from the *Arabidopsis* lineage. However, the origin of the other subgenome is still unclear although phylogenetic analysis of the chloroplast *rbcL* gene in all Brassicaceae species suggested this genome to be the maternal ancestral subgenome (Joly et al., 2009). Heenan et al. (Heenan, Goeke, Houliston, & Lysak, 2012) proposed that the limited information on Brassicaceae species genomes hampered efforts to determine the origin of the second subgenome, and it may have originated from the tribe Smelowskieae or its close relatives that have subsequently gone extinct (Heenan et al., 2012). In contrast, we found that the second subgenome might have originated from Brassicaceae Lineage EII. The highest cumulative values were generated when using the combination of *B. stricta* and *E. heterophyllum*, the latter of which is from Brassicaceae Lineage EII (Huang et al., 2015). The Eutrema tribe was predicted to be phylogenetically further away from *Pachycladon* than the Boechera tribe, however, it is close to the Brassiceae tribe phylogenetically, which could explain the *Brassica* trace in the *Pachycladon* genome found previously (Huang et al., 2015; Joly et al., 2009). Unfortunately, currently little genome information is available for Brassicaceae Lineage EII species, which limits the confident identification of the second *Pachycladon* subgenome ancestor. Altogether, we concluded that one of *Pachycladon* subgenomes has an origin similar to that of *B. stricta*, and that the other possibly arose from Brassicaceae Lineage EII.

3.3.2 Distinct UV-B radiation tolerance responses in *A. thaliana* and *P. cheesemanii* species

Earlier research has reported that UV-B radiation-tolerant plants retained relatively high levels of chlorophyll after UV-B exposure (Greenberg et al., 1997). Consistent with this, a higher chlorophyll concentration was found in UV-B radiation-tolerant *A.*

thaliana Kondara as compared to the less UV-B radiation-tolerant accession *A. thaliana* Col-0 (**Figure 3-4b**) (Gegas et al., 2014). The UV-B response is regulated by at least two pathways of which one is UVR8-dependent while the other does not require UVR8 (Li et al., 2015). The UVR8-dependent pathway induces the expression of six genes (*CHS*, *ELIP1*, *CRY3*, *GPX7*, *SIG5*, *WAKL8*) which is mediated by the transcription factors HY5 and HYH (Brown & Jenkins, 2008; Jiang, Wang, Björn, & Li, 2009). Here we found that the genes encoding the transcription factors HY5 and HYH and the six above-mentioned genes were induced in response to the UV-B radiation, suggesting that the UVR8-dependent pathway was upregulated in both *A. thaliana* Kondara and Col-0 accessions. Of note, however, gene expression was generally more induced in *A. thaliana* Kondara than in Col-0 (**Figure 3-5**). In addition, accession Kondara showed a greater induction of genes involved in photorepair and nucleotide excision repair (**Figure 3-6**), which may allow it to recover from, or prevent UV-B-induced radiation stress (Muñoz et al., 2017). Thus, the induction of chlorophyll biosynthesis genes and a stronger activation of HY5/HYH-pathway genes in Kondara than in Col-0, along with an increased DNA repair capacity in Kondara could be main contributors to its UV-B radiation tolerance.

Remarkably, leaf growth of the two *P. cheesemanii* accessions studied here was less affected by UV-B radiation than that of the *A. thaliana* accessions Kondara and Col-0; *P. cheesemanii* Kingston did not show a significant reduction of leaf area upon exposure to UV-B radiation compared to plants grown at control conditions (**Figure 3-4a**). We found that the basal chlorophyll levels in the *P. cheesemanii* accessions were higher than those of the *A. thaliana* accessions. Furthermore, chlorophyll concentration in *P. cheesemanii* Kingston further elevated upon UV-B radiation and this correlated with its increased tolerance. Chlorophyll levels have been suggested to be an indicator of UV-B

radiation tolerance (Greenberg et al., 1997) and the observed correlation between UV-B radiation tolerance, chlorophyll levels (**Figure 3-4b**), and the ability to further increase chlorophyll levels upon UV-B radiation is consistent with that suggestion. The *P. cheesemanii* accessions generally induced genes of the UVR8-dependent pathway to a lower extent than the *A. thaliana* accessions, possibly as a result of a higher basal tolerance to UV-B radiation. Brown & Jenkins (Brown & Jenkins, 2008) identified three genes (*WRKY30*, *FOX1* and *UGT74E2*) in *A. thaliana* that were induced in response to 3–12 μmol m^{-2} s^{-1} UV-B radiation as part of the still poorly characterized UVR8-independent pathway. However, in contrast to the *A. thaliana* accessions, both *P. cheesemanii* accessions showed strong induction of the *WRKY30* gene at moderately high UV-B radiation levels of 5.2 μmol m^{-2} s^{-1}. Crucially, only the Wye creek accession of *P. cheesemanii* induced expression of the two other genes (*FOX1* and *UGT74E2*), suggesting that the UVR8-independent pathway may be more important in *P. cheesemanii* than in *A. thaliana*. Thus, we found a positive correlation between chlorophyll content and UV-B radiation tolerance, and while the stronger induction of the UVR8-dependent pathway in *A. thaliana* accession Kondara may be at the basis of its UV-B radiation tolerance, the UVR8-independent pathway may play a more prominent role in *P. cheesemanii* accessions. Of further interest is that the lower UV-B radiation tolerance of the Wye creek accession correlated with a much stronger upregulation of genes involved in the UVR8-independent pathway, further supporting that this pathway plays an important role in coping with UV-B radiation damage in *P. cheesemanii*. Indeed, *P. cheesemanii* accessions may be particularly well suited for further exploring the UVR8-independent UV-B response pathway helping to increase our understanding of the genomic and genetic basis of plant adaptations to environmental UV-B stress.

The polyploid nature of the *P. cheesemanii* genome may have contributed to its relatively high tolerance to UV-B radiation. We suggest that tolerance is achieved through the activation of the UVR8-independent response pathway. My results are important for understanding the allopolyploidy nature of the *P. cheesemanii* genome and may be used to improve UV-B stress tolerance in Brassicaceae crop plants.

Chapter 4 *Pachycladon cheesemanii* stress transcriptome

4.1 Introduction

Transcriptome analysis has become an essential tool to understand the function of a genome and underlying biological processes. The application of Next-Generation Sequencing (NGS) technologies in sequencing whole transcriptomes allows the achievement of base-pair-level read data. Full-length transcripts then have to be reconstructed by transcriptome assembly (Martin & Wang, 2011). Transcriptome analysis has a broader dynamic range than other high-throughput technologies such as microarrays, since it can detect low-abundance transcripts, differentiate biologically critical isoforms and identify genetic variants (One Thousand Plant Transcriptomes, 2019; S. Zhao, Fung-Leung, Bittner, Ngo, & Liu, 2014). It has been applied in various science research areas including plant stress response studies (Shulse et al., 2019; Yang et al., 2018). Accordingly, I chose to use a transcriptomic approach to study *Pachycladon* stress responses in this project.

De novo assembly of transcriptomic sequencing data of polyploid plants is complicated by the existence of homologous transcripts (Chen, Morales-Briones, Passow, & Yang, 2019). Several strategies have been applied to improve the quality of transcriptomes in a range of polyploid species. In previous studies, the transcriptomes of *P. fastigiatum* and *P. cheesemanii* were assembled separately using a combination of 19 coverage cut offs (between 2- and 20-fold) and 20 k-mer sizes (ranging from 25 to 63), resulting in 380 assemblies per species (Gruenheit et al., 2012). Based on assembly assessment, the completeness of assembled genes varied greatly with the chosen parameter values in both species (e.g. 501 genes were completely assembled using 1 as the coverage cut off and

721, using 1 as the k-mer size in *P. cheesemanii*). Thus, it was suggested that k-mer size and k-mer coverage need to be considered at the same time to assemble full-length expressed sequence tags (ESTs) and avoid chimeric assemblies of homologous and paralogous gene copies (Gruenheit et al., 2012).

In the assembly of hexaploid wheat transcriptomes, sequencing data for multiple samples collected from different plants were assembled via Trinity and Trans-AbySS with the single and multiple k-mer methods, respectively. The assembly was then optimised by merging individual sample assemblies, removing redundancy and scaffolding. As a result, the final assembly showed higher continuity and accuracy than Sanger-derived ESTs (Duan, Xia, Zhao, Jia, & Kong, 2012). Another example is the assembly of a more complete transcriptome of *Nicotiana benthamiana* (allotetraploid) by combining multiple *de novo* transcriptome assemblies with a range of k-mer sizes. Downstream optimisation steps were also taken. TGI clustering tools and the EvidentialGene: tr2aacds pipeline were used to reduce redundancy and generate a final assembly that showed maximum diversity of *de novo* assembled transcripts and optimal completeness, with low sequence redundancy (Nakasugi, Crowhurst, Bally, & Waterhouse, 2014). It was concluded that multiple samples would be an effective way to balance the spectrum of expression levels and redundancy, with this latter step being critical for generating improved *de novo* allopolyploid transcriptome assemblies.

To obtain a high-quality transcriptome for allopolyploid *P. cheesemanii* for the multiple-stress comparative transcriptomic analysis, *de novo* transcriptome assembly was performed using multiple assemblers with further downstream processing.

4.2 Results

4.2.1 Multiple stress treatments on *P. cheesemanii* plants and Illumina sequencing

To investigate changes in gene expression caused by the selected abiotic stresses, I chose to compare transcriptomes of stressed and unstressed plants using an Illumina sequencing approach. To generate a multiple-stress transcriptome, *P. cheesemanii* plants were treated with cold, salt and UV-B radiation stresses (see Chapter 2). Nine-week-old seedlings were used because they have a large leaf area that maximises UV-B radiation absorption. Treated and untreated plants for three stress treatments were collected 5 h after treatment to detect early stress responses at the transcriptional level. Total RNA was extracted from treated and untreated leaves, and a final 12 samples were collected, which included untreated, cold stress, salt stress, and UV-B radiation stress samples (each stress had three biological replicates).

Twelve high-quality Illumina libraries of 2 × 150-bp paired-end (PE) read were generated and sequenced for each control and stress response sample from *P. cheesemanii*. A total of 122.13 GB of raw data (437,155,996 raw reads) from the different stress responses of *P. cheesemanii* was trimmed using trim_galore to remove adaptor contamination and SortMeRNA to filter ribosomal RNAs. After initial filtering, the short-read datasets from the stress-treated and untreated plants were obtained for transcriptome assembly in preparation for further analysis.

4.2.2 Quality control and statistics of *P. cheesemanii* sequencing data

Since the quality of reads varies because of limitations in Illumina sequencing technology, a quality check of sequencing data and removal of low-quality reads is essential for achieving a high-quality assembly (Minoche, Dohm, & Himmelbauer, 2011).

Therefore, 437,155,996 PE raw sequence reads—each 150 bp in length—were generated, resulting in 122.13 GB data in total (**Table 4-1**). The raw read data were then processed to filter out low-quality reads with Qscore ≤5 (reads with >50% of the total bases), those containing primer/adaptor and those with N bases (the base that the sequencing platform is unable to identify) > 10%, which resulted in 122.11 GB reads with a total of 437,025,992 PE clean reads (**Table 4-1**). Finally, the high-quality read sets obtained were used in the optimisation of *de novo* assembly and analysis of the *P. cheesemanii* stress transcriptome.

Table 4-1. Summary of read data for *P. cheesemanii* stress transcriptome.

Library	Raw reads	Clean reads	Raw data (GB)	Clean data (GB)
Control	115,407,011	115,383,228	32.24	32.24
UV-B	99,224,069	99,204,290	27.72	27.72
Cold	111,242,384	111,202,228	31.08	31.07
Salt	111,282,532	111,236,246	31.09	31.08
Total	437,155,996	437,025,992	122.13	122.11

4.2.3 *P. cheesemanii* transcriptome assemblies with multiple assemblers

Once high-quality reads are obtained, full-length transcripts need to be assembled. Since a genome draft was assembled in Chapter 3, the reference genome-guided transcriptome assembly was applied. However, the results from Bowtie2 and BUSCO showed that the quality of the assembled transcriptome was not ideal. Therefore, *de novo* transcriptome

assembly was performed again using three programs with the results compared to generate a high-quality transcriptome. Trinity, Velvet/Oases and Trans-ABySS have been suggested as providing better performance for *de novo* transcriptome assembly (Zhao et al., 2011). Based on the normalised read datasets from all the libraries, I performed assembly across a range of k-mer sizes using Trinity (k-mer sizes: 19, 21, 23, 25, 27, 29, 31), Velvet/Oases (k-mer sizes: 55, 65, 75, 85, 95) and Trans-ABySS (k-mer sizes: 51, 53, 55, 57, 59, 61, 63).

The resulting 19 assemblies were evaluated using the Bowtie alignment rate and near-universal orthologue searching (**Table 4-2**). Seven assemblies produced by Trans-ABySS generated higher read alignment rates (~91.08%) than those from the other two assemblers (~87.23% from Velvet/Oases and ~88.91% from Trinity). The BUSCO results generally showed high percentages of complete BUSCO across all the assemblies (**Table 4-2**), with the exception of the 19-mer assembly from Trinity, which had higher percentages of fragmented and missing BUSCOs.

Table 4-2. Assessment of transcriptome assemblies generated by multiple assemblers.

	k-mer size	Bowtie (%)	Complete BUSCOs (%)	Fragmented BUSCOs (%)	Missing BUSCOs (%)
Velvet/Oases	k55	88.21	92.7	4.5	2.8
	k65	87.84	96.1	1.9	2
	k75	87.56	97.1	1.3	1.6
	k85	86.64	96.9	1.4	1.7
	k95	85.9	96.4	2.1	1.5
Trans-ABySS	k51	90.79	92.5	4.4	3.1
	k53	90.88	92.8	4.4	2.8
	k55	90.98	93.1	4	2.9
	k57	91.08	93.8	3.5	2.7
	k59	91.4	94.3	3.5	2.2
	k61	91.21	94	3.5	2.5
	k63	91.25	94.1	3.5	2.4
Trinity	k19	88.81	59.9	24.4	15.7
	k21	88.72	84.5	9.4	6.1
	k23	88.62	89.4	7.2	3.4
	k25	88.26	86.3	8	5.7
	k27	88.68	87	7.4	5.6
	k29	89.41	88.1	8.1	3.8
	k31	89.86	88.6	7.1	4.3

As the assemblies from the Trans-ABySS assembler outperformed those of Velvet/Oases and Trinity, based on a Bowtie evaluation, they were selected for further processing. Next, Trans-ABySS assemblies across different k-mer sizes were combined to generate the final transcript set (318,111 transcripts), which was further clustered and assembled using the CAP3 assembly program. This program removed technical redundancy with 99% overlap in percentage identity and 200 bp overlap in length, resulting in 223,341 transcripts (**Table 4-3**). Finally, EvidentialGene: tr2aacds was applied to remove redundancy and identify transcript splice isoforms, resulting in a final transcriptome of 67,905 transcripts (**Table 4-3**) (**Supplementary File 4-1**).

In the final transcriptome, the length distribution of transcripts revealed that 73.98% were between 200 and 2,000 bp, with 1% being longer than 5,000 bp (**Figure S4-1a**). The BUSCO assessment results showed that 94.7% of BUSCOs could be found in the set with a single copy (48.7%) or duplication (46.0%). Fragmented and missing BUSCOs were rare—2.7% and 2.6% (**Figure S4-1b**). These results revealed the high-quality integrity of the transcript set and showed that it could be used in downstream analysis.

Table 4-3. Summary statistics for transcriptome assembly.

Assembly	
Trans-ABySS transcripts (combined 7 assemblies)	318,111
CAP3 transcripts	223,341
EvidentialGene transcripts	67,905
EvidentialGene genes	45,911
Minimum transcript length (bp)	200
Longest transcript length (bp)	18,443
Final size of assembly (bp)	107,064,440

4.2.4 Functional Annotation

The next step of a transcriptomic analysis is homology-based functional annotation. The BLAST algorithm was used here because of its high popularity. To annotate the final transcript set, the sequences were searched against the *A. thaliana* UniProt database using BLASTX with an E-value cut off of 1e-5. Of the 45,911 genes, 39,949 had homologies in UniProt database with >50% identity. These genes could be mapped to 29,060 *Arabidopsis* proteins with 5,294 GO annotations. The overrepresented GO terms for biological processes were 'carbohydrate metabolic process' (735 members), 'cell redox homeostasis' (352 members), 'cell wall organisation' (365 members), 'defence response' (430 members), 'DNA integration' (493 members), 'intracellular protein transport' (487 members), 'protein transport' (472 members), 'regulation of transcription, DNA-templated' (1,418 members), 'signal transduction' (415 members) and 'translation" (960 members) (**Supplementary File 4-2**). In conclusion, 87% transcripts could be annotated

against the *A. thaliana* protein database, as expected, suggesting that the assembled transcriptome was suitable for downstream analysis.

4.3 Discussion

This chapter describes the assembly of the transcriptome of the allopolyploid *P. cheesemanii* using multiple assemblers and merging strategies. The final Trans-ABySS assembly using 122.11 GB of PE sequencing data was selected based on its outstanding performance in terms of completeness, continuity (**Table 4-2**) and calculated assembly statistics (**Table 4-3**) in comparison with the assemblies from the other three methods.

Multiple k-mer sizes were used in each assembler, with Trinity showing poor continuity when a short k-mer was applied: only 59.9% of complete BUSCOs were found at this k-mer size compared with ~80% with other k-mer size Trinity assemblies (**Table 4-2**). The Velvet/Oases assemblies contained more complete BUSCOs than their counterpart Trans-ABySS and Trinity assemblies (**Table 4-2**). However, Trans-ABySS assemblies had less redundancy and represented a wider range of unique transcripts (**Table 4-2 and Figure S4-1**). Comparison of the assemblies produced by each assembler revealed an apparent threshold in Trinity assembly, which decided the assembler outcome to a great extent (**Table 4-2**). Because the high read-alignment rates of highly expressed transcripts could compensate for low read-coverage when assembling with a large k-mer size, the assembly of highly expressed transcripts could be more effective with large k-mer sizes (Zhao et al., 2011). Interestingly, I observed a remarkable difference in the performance of the Trinity assembler at different k-mer sizes (**Table 4-2**), an outcome that was not observed for other assemblers. One possible explanation for this observation is that Trinity does

not include a scaffolding step, which makes it less sensitive to splice variation at low k-mer sizes (Liu, Adelman, Myles, & Zhang, 2014).

To select the best assembly, the percentage of mapped transcripts (with reference to the Viridiplantae database) using BUSCO assessment was used to rank the different assemblies. The results from BUSCO representing transcript completeness revealed that Velvet/Oases and Trans-ABySS outperformed Trinity using this criterion. When considering accuracy based on the percentage of read alignments against assemblies (Bowtie results), Trans-ABySS outperformed Velvet/Oases and Trinity (**Table 4-2**).

Plant transcriptome assemblies often combine sequencing data obtained from multiple samples (e.g. different plant tissues or the same tissues sampled under different treatment conditions). However, increasing the number of samples inflates the size of the dataset to be assembled. Thus, the first difficulty may be how to handle large read datasets. In particular, the consumption of the computing resource of some *de novo* assemblers are massive. In this study, the read sets for assembly were in silico normalised by Trinity, which would have removed a large number of the same reads among samples. All tested assemblers worked smoothly with these normalised read sets with a range of k-mer sizes (all assemblies could be done within a week). As a result, the normalised read sets also showed high alignment percentages against all assemblies, suggesting that these read sets would be suitable for assembling a relatively complete transcriptome.

It has been reported that merging multiple k-mer assemblies from Trans-ABySS can result in degradation of assembly quality, with a shorter N50 (the sum of the shortest contig lengths needed to cover 50% of the total genome), fewer number of mappable

reads, reduced proportion of bases covered from full-length cDNA and an increase in sequence redundancy (Duan et al., 2012). In this study, the normalised read sets from all samples produced by Trinity were applied to different assemblers with multiple k-mer sizes. No large difference in the number of mappable reads (Botiwe2 results) and full-length cDNA (complete BUSCOs) was observed between the assemblies from Trans-ABySS and those from the other two assemblers (**Table 4-2**). This suggested that normalising the read sets from all samples might overcome some of the natural shortcomings of Trans-ABySS.

The inherent redundancy found in transcriptomes complicates the assembly process as it is especially difficult for software to distinguish between transcript isoforms and homologous genes for polyploid plants. Thus, the correct removal of technical redundancy caused by the overlap of transcripts of multiple assemblies is crucial for minimising noise in downstream distinguishing of transcript isoforms and homologous genes in polyploidy plant transcriptome assembly (Chopra et al., 2014). After combining the assemblies from multiple k-mer sizes, the number of assembled transcripts was still relatively high (318,111), which may have resulted from overlap between assemblies (**Table 4-3**). I set an arbitrary threshold of 99% overlap percentage identity and 200 bp overlap length and highly similar transcripts considered redundant were removed correctly using CAP3, which shrunk the number of transcripts to 223,341. Subsequently, the assembly was processed using EvidentialGene: tr2aacds, which further removed fragmented transcripts and identified alternative splice transcripts based on multiple parameters. The final assembly consisted of 45,911 genes out of 67,905 transcripts, ranging from 200 to 18,443 bp. Of these identified *P. cheesemanii* transcripts, 87% could be annotated with identity > ~50% and E-value ≤ $1 \times e^{-5}$ using the *Arabidopsis* Uniprot

database, which provided evidence that a large proportion of the transcripts was unlikely to be assembly artefacts. My study as reported in this chapter provides novel *P. cheesemanii* transcriptome resources to the plant science community. Since the novel UV-B radiation response of *P. cheesemanii* was identified in Chapter 3, the multiple-stress transcriptome might help to expand understanding of the genetic basis of the stress responses in *P. cheesemanii* for adapting to the New Zealand environment.

Chapter 5 Comparative transcriptomics of multi-stress responses in *Pachycladon cheesemanii* and *Arabidopsis thaliana* to identify novel stress-responsive pathways

5.1 Introduction

The interactions between plants and their environments occur across their lifetime, and environmental influences can be seen not only in long-term plant evolution but also in short-term plant physiological responses to various biotic and abiotic stresses. Nonetheless, after millions of years of evolution, plants have developed a range of responses to environmental cues such as UV-B radiation, drought, high or low temperature and salinity (Hasanuzzaman, Nahar, & Fujita, 2013; Macková et al., 2013; Xuemin Wang, Li, Li, & Welti, 2006).

Plants survive under various environmental stresses by responding at physiological, biochemical, cellular and molecular levels (Hasanuzzaman, Nahar, Alam, Roychowdhury, & Fujita, 2013; Hasegawa, Bressan, Zhu, & Bohnert, 2000; Misra, Biswal, & Misra, 2002; Venkateswarlu & Shanker, 2011). Meanwhile, stress-responsive genes establish a complex network for each stress response, and interactions in these networks have recently been simulated in the model plant *Arabidopsis*, as well as rice. It has been found that 22 genes can be induced by drought, cold and high salinity stresses, among 53,277 transcripts and 194 genes that revealed responses to single or multiple stresses in *Arabidopsis* plants (Seki et al., 2002). A cDNA microarray analysis of drought, cold and high salinity-treated rice plants demonstrated a wide range of crosstalk involving four stress responses with 15 common stress-induced genes (Rabbani et al., 2003). In another study, it was found that cold, salt and drought responses of *Arabidopsis* plants

shared 30% of 8,100 genes induced by all three stresses in the early stress response (Kreps et al., 2002). Also, 232 clones were in common among responses to cold, heat and salt stresses in potato plants, as revealed by analysis of ~12,000-clone potato cDNA microarray (Rensink et al., 2005). These studies have revealed a number of genes involved in stress responses, some of which are stress specific and others that are shared by multiple stress responses. Therefore, high-throughput profiling could help identify comprehensive crosstalk among multiple stress responses.

In contrast, plant UV-B responses have been studied relatively independently from other stresses in *Arabidopsis* and other plant species (Dai et al., 1997; Kim, Tennessen, & Last, 1998; Raghuvanshi & Sharma, 2016; Vyšniauskienė & Rančelienė, 2014). Benefits from high-throughput biology in terms of a deeper understanding of plant UV-B response pathways have been achieved in the last couple of years. After 8 h of exposure of maize (*Zea mays*) plants to 9 kJ m^{-2} d^{-1} UV-B radiation, 285 transcripts were upregulated significantly in at least one organ and 80 were downregulated (Casati & Walbot, 2004). A microarray analysis of the gene expression responses of grape (*Vitis vinifera* L.) leaves to UV-B radiation (4.75 kJ m^{-2} d^{-1}) revealed that plant responses could be categorised as general protective responses (synthesis of UV-B absorbing compounds), antioxidant defence, pathogen defence and abiotic stress responses (Gil, Pontin, Berli, Bottini, & Piccoli, 2012). RNA sequencing analysis was performed on *Lycium ruthenicum* with and without UV-B induction, involving 1,913 upregulated and 536 downregulated genes that included antioxidant-related genes and genes involved in secondary metabolite synthesis and defence responses (Wang et al., 2018). However, there are few reports describing the interaction between UV-B and other stress responses. In the frame of the AtGenExpress project, Affymetrix ATH1 microarray analysis of multiple stress regulatory networks

consisting of heat, cold, drought, salt, high osmolarity, UV-B light and wounding treatments were performed on *Arabidopsis*. The results suggested the existence of a set of core genes in *Arabidopsis* that initially respond to multiple environmental stresses (Kilian et al., 2007). This calls for a comprehensive and overall investigation of the crosstalk network of plant UV-B responses and other stress responses.

In this chapter, I describe an analysis of the multiple-stress transcriptomes of *A. thaliana* and *P. cheesemanii*. I predict an interaction model that describes the crosstalking gene networks responding to cold, salt and UV-B radiation in both species.

5.2 Results

5.2.1 Identification of differentially expressed genes in *P. cheesemanii* and *A. thaliana* responding to multiple stresses

The activation of gene expression often implies that a particular gene contributes to a plant response. Chapter 4 describes the construction of the *P. cheesemanii* transcriptome of control and stressed plants. Here I aim to compare the *P. cheesemanii* stress transcriptome with the *A. thaliana* stress transcriptome. Therefore, *A. thaliana* plants were treated with similar stress conditions as the *P. cheesemanii* plants, namely: Six-week-old *A. thaliana* plants were treated with 5.2 µmol m^{-2} s^{-1} UV-B radiation, 4°C cold stress and 250 mM NaCl as salt stress for 5 h simultaneously. RNA was then extracted from the leaf tissue of treated and untreated *A. thaliana* plants and sequenced using the Illumina technique described in Chapter 2. Using the *A. thaliana* transcriptome (GenBank CP002684–CP002688), and using the cut-off criteria described in Chapter 2, 3,971 genes, 1,054 genes and 2,214 genes out of 27,656 *A. thaliana* genes were significantly differentially expressed (DE) in response to cold, salt and UV treatment, respectively. In

addition, differential expression analysis of *P. cheesemanii* transcriptome data in Chapter 4 identified 3,690, 806 and 1,241 DEGs in the cold, salt and UV-B radiation treatments, respectively (**Table 5-1**).

5.2.2 Differentially expressed gene number bias in the three stress responses *of A. thaliana* and *P. cheesemanii*

The upregulated loci in cold, salt and UV treatments included 2,060, 869 and 1,181 genes in *A. thaliana* and 2,509, 620 and 1005 genes in *P. cheesemanii*, respectively; while the numbers of genes downregulated in cold, salt and UV treatments were 1,911, 185, and 1,033 in *A. thaliana* and 1,181, 186 and 237 in *P. cheesemanii*, respectively (**Table 5-1**). As shown in **Figure 5-1**, consistent with a previous report (Li et al., 2012), edgeR with TMM normalization yields much lower but true FDRs in all treatments, and the response patterns of the DEGs to three treatment groups were similar for *A. thaliana* and *P. cheesemanii*. In both species, the largest number of responsive genes was observed following cold treatment, and salt treatment induced the smallest number of DEGs (3,971 and 3,690 in cold responses *v.* 1,054 and 806 genes in salt responses, and 2,214 and 1,242 genes in UV-B radiation responses) (**Figures 5-1 and Table 5-1**). To group genes based on their expression level, I defined four expression categories on the basis of FC level (Level I, $2 \leq |FC| < 4$; Level II, $4 \leq |FC| < 32$; Level III, $32 \leq |FC| < 1{,}024$; Level IV, $1{,}024 \leq |FC|$). As shown in **Table 5-1**, the DEG distribution at various levels indicated that the largest number of genes was at Level I in all three stress treatments. A small number of DEGs at Level II was found after the *P. cheesemanii* salt treatment (160), which could have contributed to the loss of the typical "V" shape in the volcano plot. The number of upregulated genes was more than that of downregulated genes in the UV-B radiation treatment, at Level III and IV in both species (**Figure 5-1**). For example, the UV-B

radiation response in *A. thaliana* at Level III included 42 upregulated and 1 downregulated gene; at Level III in *P. cheesemanii* there were only upregulated genes (**Figure 5-1**).

There was a strong bias in the direction of regulation in the response to salt, with 82.4% and 76.9% of genes being upregulated in this treatment of *A. thaliana* and *P. cheesemanii*, respectively. In contrast, aside from a similar bias in salt response, there were strong biases in cold and UV-B radiation responses in *P. cheesemanii* as well (79.6% upregulation in cold response, 69.6% upregulation in salt response and 78.9% upregulation in UV-B radiation response) (**Table 5-1**).

In conclusion, the number of DEGs in response to each stress was similar for the two species. In addition, the number of DEGs in the salt treatment displayed a strong bias in upregulation in both species; these biases in the number of *P. cheesemanii* DEGs extended to cold and UV-B treatment responses as well.

Table 5-1. Summary of differential expressed genes in multiple stresses in *A. thaliana*.

Species	Stress	Up-regulated genes	Down-regulated genes	Level I	%	Level II	%	Level III	%	Level IV	%	Total
A. thaliana	Cold	2,060	1,911	2,641	66.5	1,242	31.3	83	2.1	5	0.1	3,971
	Salt	869	185	596	56.5	423	40.1	32	3.0	3	0.3	1,054
	UV	1,181	1,033	1,763	79.6	404	18.2	43	1.9	4	0.2	2,214
P. cheesemanii	Cold	2,509	1,181	2,730	74.0	918	24.9	42	1.1	0	0	3,690
	Salt	620	186	635	78.8	160	19.9	11	1.4	0	0	806
	UV	1,005	237	937	75.5	284	22.9	20	1.6	0	0	1,242

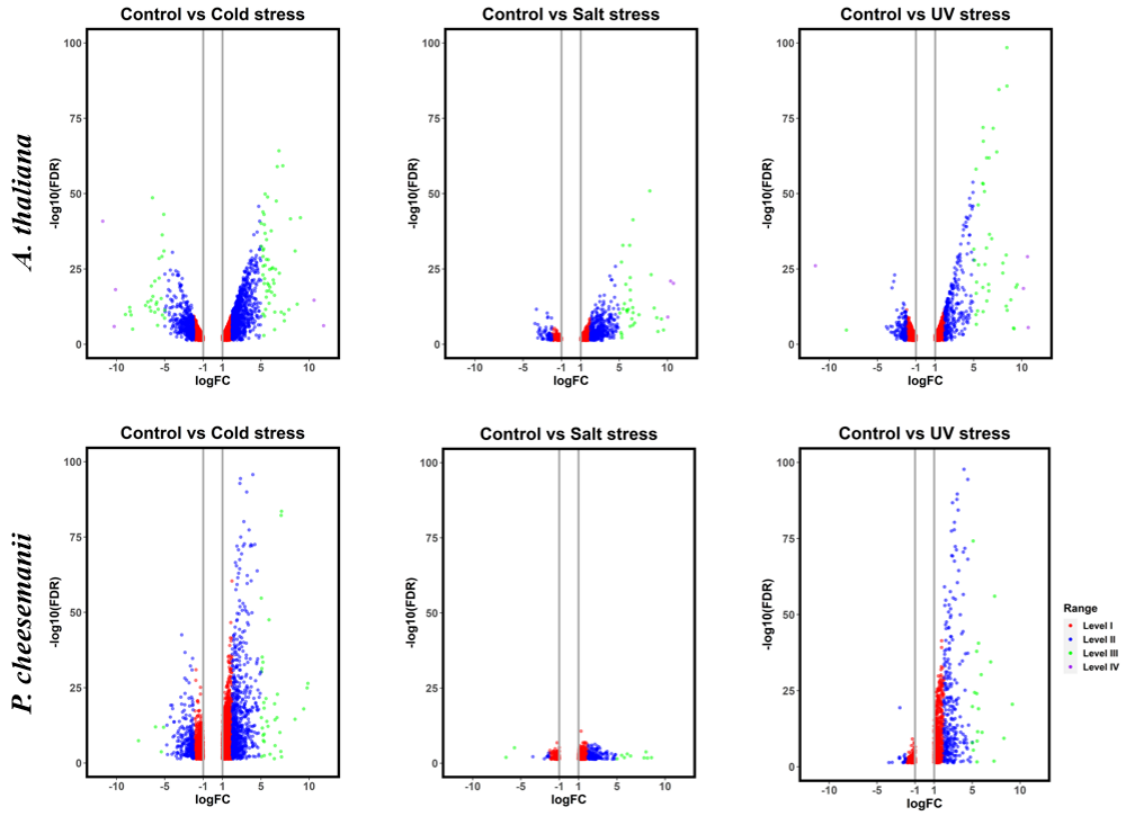

Figure 5-1 Gene differential expressions and comparisons of up-regulations and down-regulations responding to stresses in *P. cheesemanii* and *A. thaliana*.

Volcano plots of fold changes in abundance (logFC) against significance (FDR) highlight those DE genes in responding to cold, salt, and UV-B radiation. FC: fold change; FDR: false discovery ratio. Upper row: *A. thaliana*; lower row: *P. cheesemanii*. The differential expressions were grouped into 4 levels with different colours: Level I, orange; Level II, blue; Level III, green; Level IV, purple.

5.2.3 Comparative analysis of stress-responsive gene sets

The previous section provided an overall view of the multiple stress responses of two species in terms of the number of responsive genes. This section describes an examination of the relationship between stress responses by identifying responsive genes shared among stresses, as well as responsive genes unique to each stress within species. To do

so, responsive genes that were upregulated and downregulated were compared among stresses using a Venn diagram analysis.

The results indicated that some genes were shared between stress responses: 442 genes for cold and salt responses; 775 for cold and UV-B radiation responses; and 229 for salt and UV-B radiation responses in *A. thaliana*. The equivalent numbers of shared responsive genes for *P. cheesemanii* were 322, 452 and 100 (**Figure 5-2**). There were also many unique genes induced by each stress in *A. thaliana*: 3,529 for the cold response *v*. 612 for the salt response; 3,196 for cold *v*. 1,440 for UV-B radiation; and 1,986 for UV-B radiation *v*. 825 for salt. These numbers for *P. cheesemanii* were 3,368 for cold *v*. for 484 salt; 3,238 for cold *v*. 790 for UV-B radiation; and 1,142 for UV-B radiation *v*. 706 for the salt response (**Figure 5-2**).

To examine whether some of the responsive genes might potentially play a different role in different stress responses, the responsive genes that were upregulated and downregulated following the three stresses were compared, revealing a number of shared DEGs between stress treatments. For instance, 308 genes were shared between the cold_up and salt_up groups and 327 genes between the cold_down and UV_down groups in *A. thaliana*; as well as 353 genes between cold_up and UV_up and 213 between cold_up and salt_up in *P. cheesemanii* (**Figure 5-2**). However, other genes showed contrasting responses to different stresses in the two species. For instance, some upregulated genes in the cold response were downregulated in the salt response of both *A. thaliana* and *P. cheesemanii* (6 and 47 genes) (**Figure 5-2**). This suggested that some of the responsive genes played different roles in responding to different stresses although a large number of responsive genes were shared by multiple stress responses.

Interestingly, the numbers of the shared genes in both the upregulation and downregulation groups were larger than those of genes that were upregulated in response to one stress but downregulated in response to the other, in almost all pairwise comparisons. There were two exceptions, both from *P. cheesemanii*: in the comparison of salt and cold responses, the number of shared genes in the downregulation set was less (25 genes) than the numbers that were upregulated in response to one stress but downregulated for another (47 and 37 genes); in the comparison of salt and UV-B radiation responses, the number of shared genes in the downregulation set (12 genes) was also less than the numbers of genes that were upregulated in one stress but downregulated in another (20 and 13 genes).

In summary, different stress responses indeed did share a number of responsive genes, although the majority of responsive genes were unique to each stress response within species. Further, a number of responsive genes in both species responded in opposite ways to different stresses.

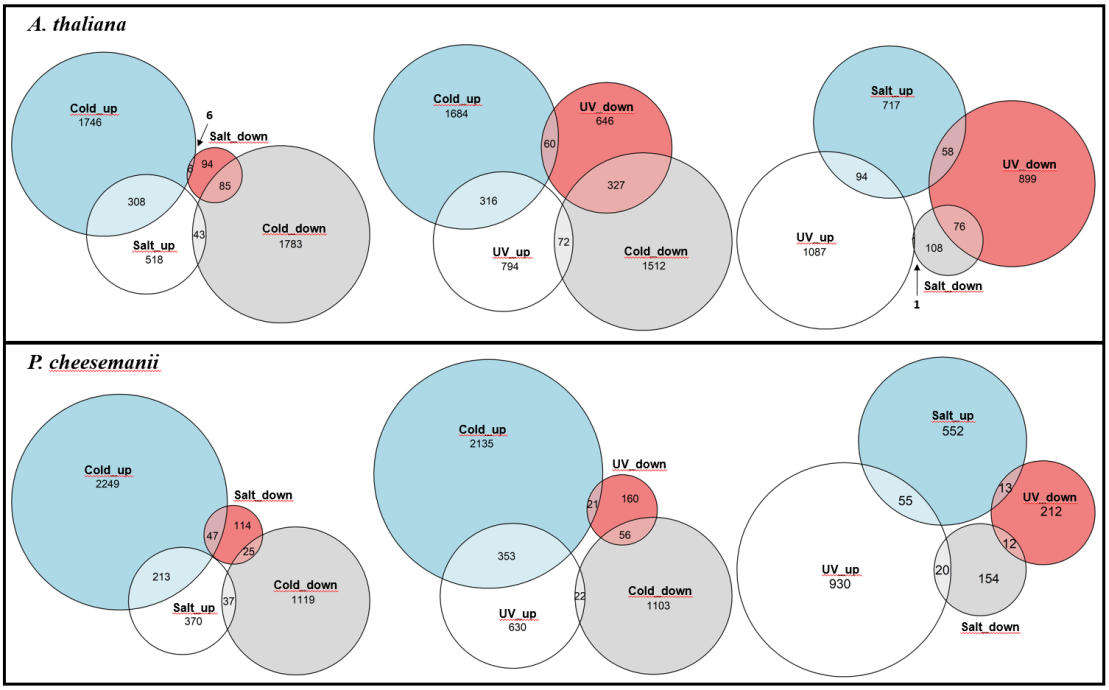

Figure 5-2 Comparisons of up-regulations and down-regulations responding to stresses in *A. thaliana* and *P. cheesemanii*.

Venn diagrams showing the up-regulations and down-regulations overlap in cold, salt, and UV-B radiation responses in both *A. thaliana* and *P. cheesemanii*. Upper row: *A. thaliana*; lower row: *P. cheesemanii*.

5.2.4 Gene ontology enrichment analysis of *A. thaliana* and *P. cheesemanii* multiple-stress transcriptomes

The last section confirmed the existence of both shared responsive genes and stress-specific responsive genes within species. The analysis described in this section aimed to discover the extent to which different stresses shared biological processes and had their own unique biological processes within species. GO includes knowledge of three aspects of biology: cellular components, biological processes and molecular functions. GO enrichment analysis aims to identify which GO terms are overrepresented (or underrepresented) using annotations for a gene set that is upregulated (or downregulated)

under certain conditions. Part of this study, GO enrichment analysis was performed (**Supplementary File 4-2**) and the overrepresented terms for GO biological processes were compared among stresses. The numbers of shared overrepresented GO terms in multiple stress responses and the unique overrepresented GO terms in each stress within species was illustrated using a Venn diagram (**Figure 5-3)**. A number of overrepresented GO terms were involved in all stress responses (82 and 80 in *A. thaliana* and *P. cheesemanii*, respectively), and a number of stress-specific responses were found as well. This result confirmed the possibility of the existence of multiple-stress crosstalk. However, the number of unique GO terms for the *P. cheesemanii* response to cold stress was four times higher than that for *A. thaliana* (152 *v.* 37). Further, the number of GO terms overrepresented in UV-B radiation and salt responses, but not the other two stress responses of *A. thaliana*, was larger than that for *P. cheesemanii* (123 *v.* 88 and 62 *v.* 23). In addition, the overlap in overrepresented GO terms between the salt and cold response, as well as that between the cold and UV-B radiation response of *P. cheesemanii*, was considerably larger than in that of *A. thaliana* (68 *v.* 21 and 53 *v.* 25). In conclusion, different stresses shared some biological processes but also had their own unique biological processes, in both *A. thaliana* and *P. cheesemanii*. *P. cheesemanii* showed more unique cold responses than did *A. thaliana*, while *A. thaliana* showed more unique UV-B radiation and salt responses than did *P. cheesemanii*.

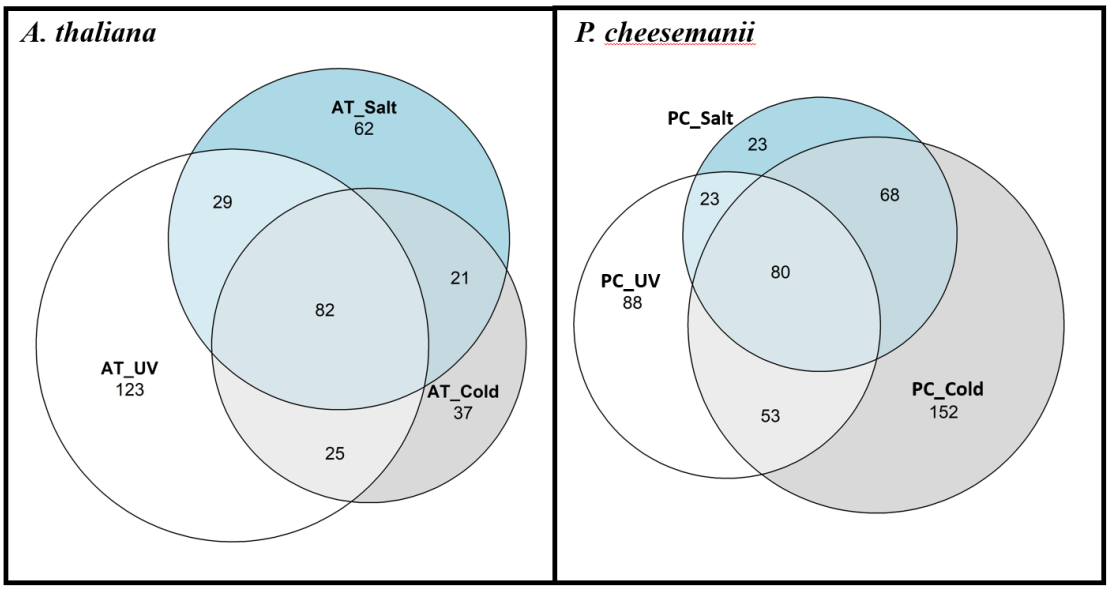

Figure 5-3 Venn diagram of comparing GO terms of biological processes between stresses in *A. thaliana* and *P. cheesemanii*, respectively.

PC: *P. cheesemanii*; AT: *A. thaliana*; UV circles: the overrepresented GO terms in UV-B radiation response; Cold circles: the overrepresented GO terms in cold response; Salt circles: the overrepresented GO terms in salt response.

5.2.5 Network analysis identifies multiple stress-responsive crosstalk in *A. thaliana* and *P. cheesemanii*

A number of the same overrepresented terms for GO biological processes between stresses within species have been confirmed. The analysis described in this section further addressed how these overrepresented terms for GO biological processes connected to form crosstalk between stress responses. To discover crosstalk involving multiple stress responses, GO enrichment networks of multiple stress responses were generated in each of *A. thaliana* and *P. cheesemanii*, as described in Chapter 2. GO terms were organised in a network where connections were created based on overlap between the gene sets. The network layout groups related GO terms into network clusters, so that the major overrepresented functional themes can be quickly identified.

In the *A. thaliana* network analysis of upregulated pathways, three stress responses illustrated apparent pathway overlap (**Figure 5-4**). Salt and cold responses shared a number of stress-responsive genes, especially those responding to hormone and ABA, with unique desiccation-responsive genes being involved in the salt response and unique auxin-responsive genes in the cold response. The biosyntheses of flavone, flavonol, carotenoid and vitamin B6 were activated in both cold and UV-B radiation responses. Another interesting overlap in the UV-B radiation and cold responses as in the GO term 'biological regulation'. These responses shared main part of regulation under UV-B radiation and cold stresses, with the exception of transcription factor import into the nucleus (unique to the UV-B radiation response) and circadian rhythm regulation (unique to the cold response). Further, UV-B radiation and salt responses overlapped in monosaccharide signalling. Conversely, each stress had its own unique response pathway. UV-B radiation had significant impacts on seed embryonic development and nucleic acid metabolic processes, which did not occur in response to the other two stresses. Wax and cuticle biosynthesis, toxin catabolic secondary, ripening and ageing were only induced by high salinity. Carbohydrate metabolism was active only in the cold response. Thus, the analysis indicates there was crosstalk between different stress responses, while each stress also had a number of unique responses.

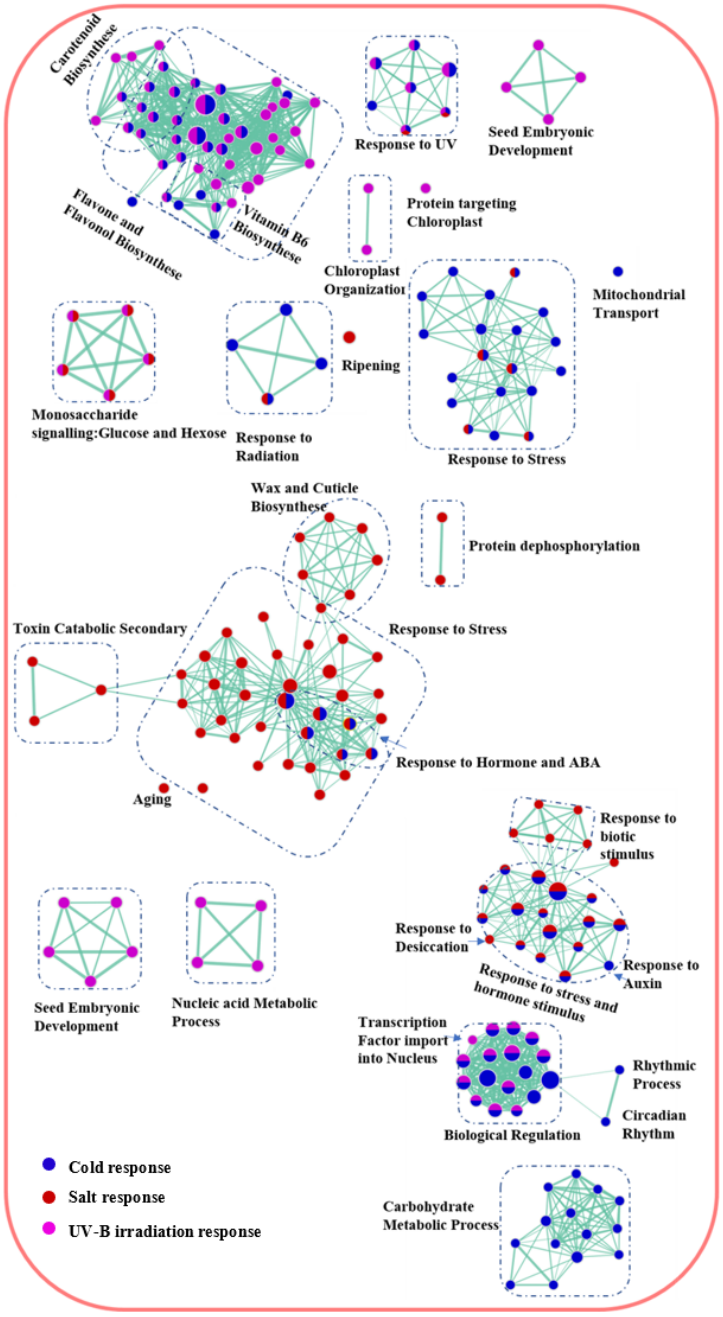

Figure 5-4 Network analysis of biological process of *A. thaliana* multiple stress-responsive transcriptomes in upregulation.

Purple circle: the overrepresented GO terms in UV-B radiation responses; red circle: the overrepresented GO terms in salt-induced responses; blue circle: the overrepresented GO terms in cold-induced responses. Box with dashed line: the cluster of the overrepresented GO terms involved in the same biological process.

In the downregulation network, both cold and salt treatments induced a response in terms of DNA geometric change (the process in which a transformation is induced in the geometry of a DNA double helix, resulting in a change in twist, writhe or both, but with no change in linking number, which is a topological invariant that requires breaking of a bond in one of the two strands of DNA to change); and UV-B radiation shared some common responses with cold treatment in terms of the defence response (**Figure 5-5**). As expected, UV-B radiation suppressed photosynthesis metabolic process and, interestingly, some salt-responsive genes were inactivated by UV-B radiation in light reaction and harvesting. Additionally, regulation of biosynthetic processes was suppressed only after salt treatment.

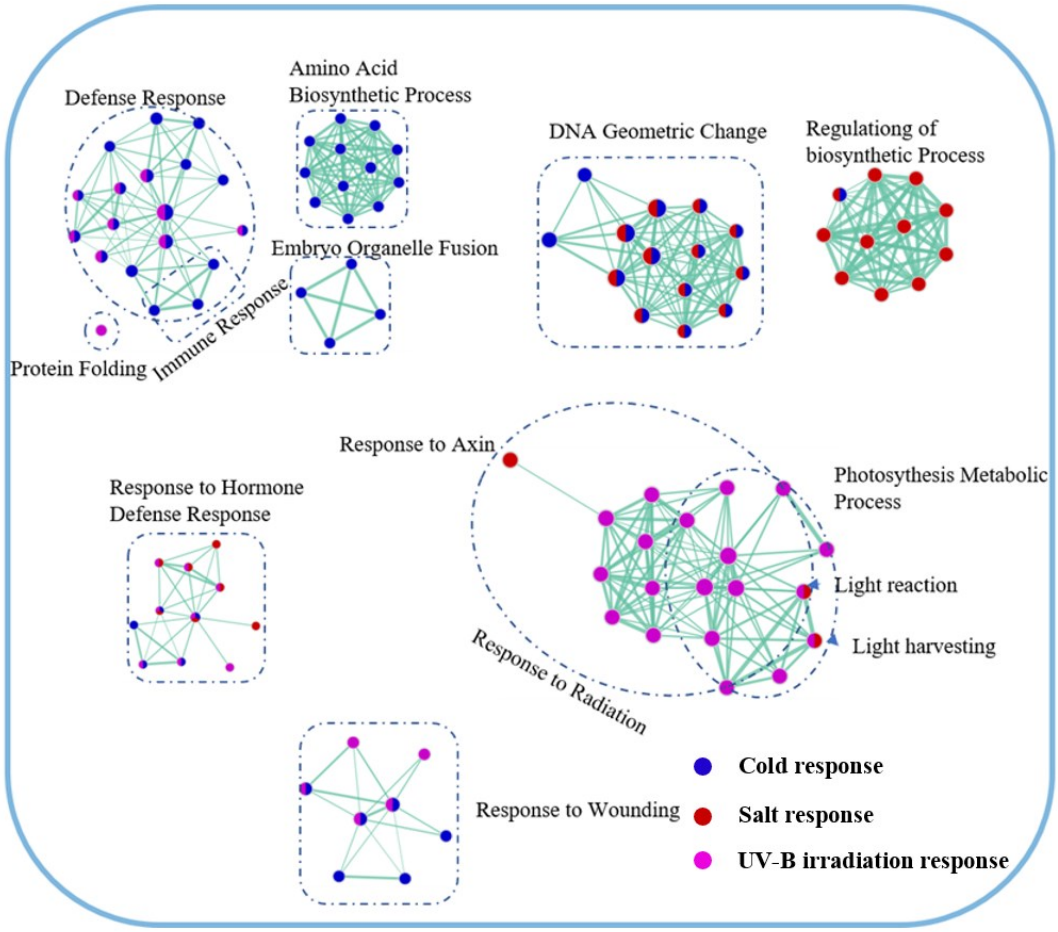

Figure 5-5 Network analysis of biological process of *A. thaliana* multiple stress-responsive transcriptomes in downregulation.

Purple circle: the overrepresented GO terms in UV-B radiation responses; red circle: the overrepresented GO terms in salt-induced responses; blue circle: the overrepresented GO terms in cold-induced responses. Box with dashed line: the cluster of the overrepresented GO terms involved in the same biological process.

Unlike the multiple-stress network in *A. thaliana*, the network in *P. cheesemanii* showed few connections (**Figure 5-6**). The main cluster of three stress responses in *P. cheesemanii* was response to stress, including some typical stress-related biological processes (response to ethylene, salicylic acid, gibberellin, ABA and so on) that were

overrepresented in the three stress responses (**Figure 5-6**). The biosynthetic process of one stress-related pigment compound, anthocyanin, occurred in this main cluster, and this process was also overrepresented in the three stress responses. Next, the GO term 'response to UV' was connected to the processes involved in mitochondrial electron transport, while the 'response to stress' cluster was connected to a small cluster of 'photosystem' in cold responses. On the other side of the main cluster, there was another small cluster relating to arogenate dehydratase activity. Additionally, there were two disconnected clusters: cell wall biogenesis in cold responses and IP3 metabolic process in UV-B radiation responses (**Figure 5-6**).

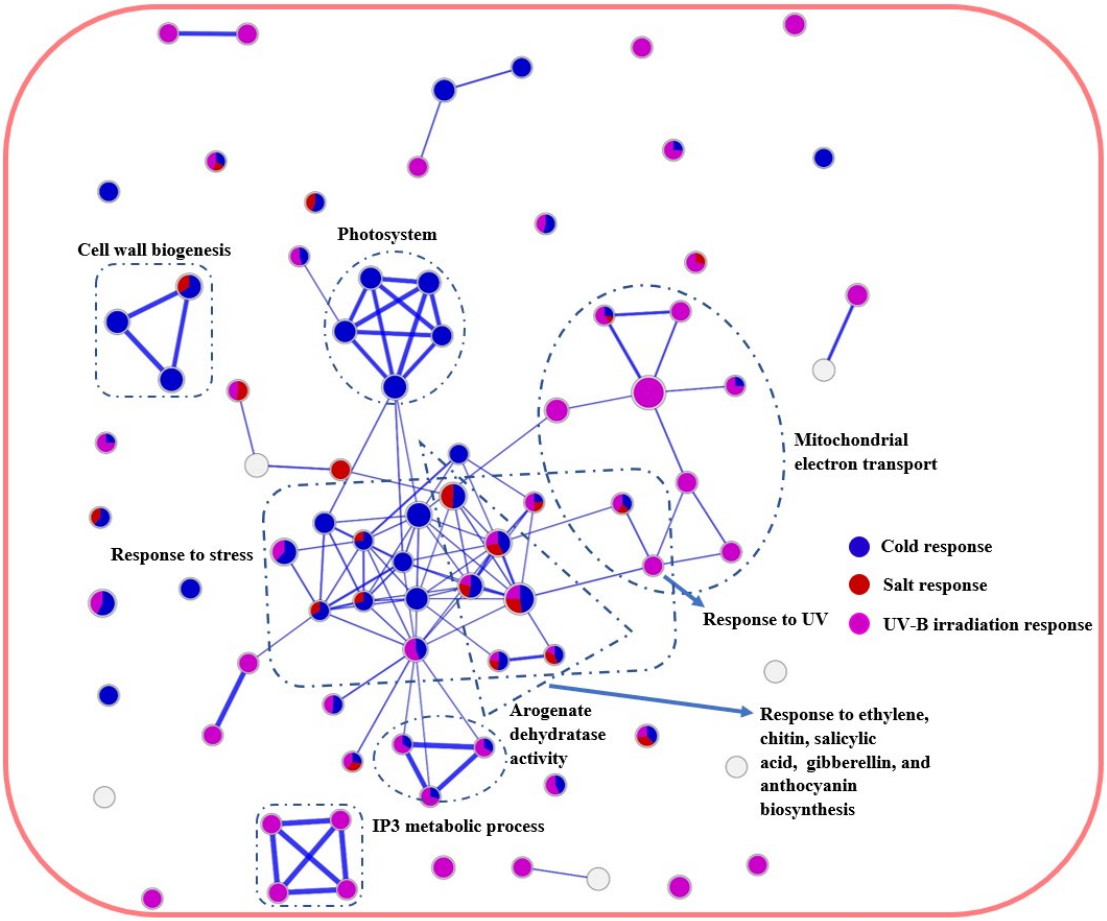

Figure 5-6 Network analysis of biological process of *P. cheesemanii* multiple stress-responsive transcriptomes.

Multiple-stress transcriptome network. Purple circle: the overrepresented GO terms in UV-B radiation responses; red circle: the overrepresented GO terms in salt-induced responses; blue circle: the overrepresented GO terms in cold-induced responses. Box with dashed line: the cluster of the overrepresented GO terms involved in the same biological process.

Overall, there was a wide range of crosstalk between stress responses in *A. thaliana* in both upregulation and downregulation, while the crosstalk network for the three stress responses in *P. cheesemanii* differed in having only one main cluster identified.

5.2.6 Comparison of overrepresented terms of gene ontology for biological process between *A. thaliana* and *P. cheesemanii* in the three stress responses

To examine whether there was a shared biological process between species in responding to each stress, as well as whether the species had some unique biological processes in responding to each stress, overrepresented terms for GO biological processes in each stress were compared between the two species. The number of overrepresented GO terms shared by the two species and that which was unique to species in each stress was then presented using a Venn diagram **(Figure 5-7)**. The results showed that a number of the GO terms were overrepresented in *P. cheesemanii* but not in *A. thaliana*: 227 in cold; 104 in salt; and 118 in UV-B radiation responses. The number of GO terms overrepresented in *A. thaliana* but not *P. cheesemanii* was 39, 104 and 133 in cold, salt and UV-B radiation responses, respectively. Interestingly, the number of novel overrepresented GO terms in responding to cold stress in *A. thaliana* (39) was much smaller than that in *P. cheesemanii* (227).

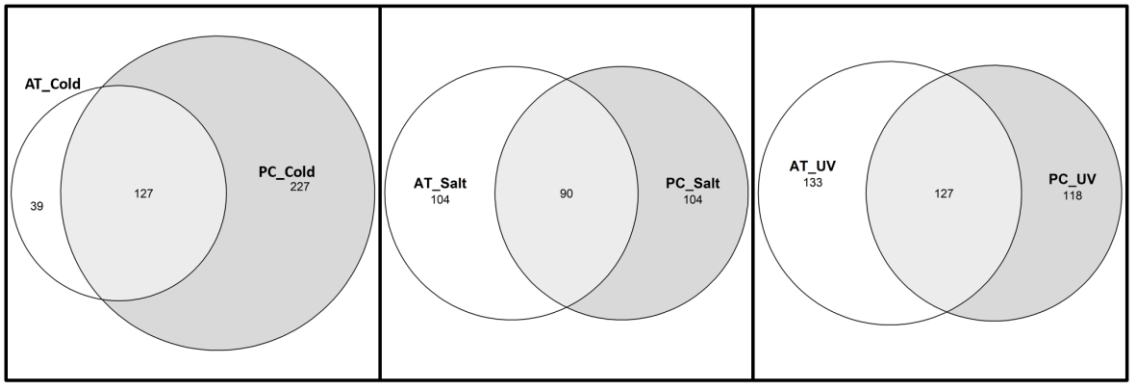

Figure 5-7 Comparisons of overrepresented GO terms of *A. thaliana* and *P. cheesemanii* in responding to three stresses.

Two species have a number of the same overrepresented GO in each stress responses, while species-specific overrepresented GO terms in responding to the stress were also present. PC: *P. cheesemanii*; AT: *A. thaliana*; UV circles: the overrepresented GO terms in UV-B radiation response; Cold circles: the overrepresented GO terms in cold response; Salt circles: the overrepresented GO terms in salt response.

In conclusion, in *A. thaliana* and *P. cheesemanii*, a number of the biological processes were shared between stress responses, while some biological processes were stress specific. Between species, a number of unique biological processes were involved in responding to each stress. Further, *P. cheesemanii* had more unique GO enrichments in its cold response than did *A. thaliana*.

5.2.7 Identification of shared biological processes between the *A. thaliana* and *P. cheesemanii* stress responses

In the previous section, a number of biological processes were reported as shared between *A. thaliana* and *P. cheesemanii* in responding to each stress. In this section, the biological processes shared by the two species are further analysed. To obtain an overall structure of responses common to the species for each stress, the overrepresented terms in common

between the *A. thaliana* and *P. cheesemanii* biological process in each stress response, with a 0.05 adjusted *p*-value (adjusted by the Benjamini and Hochberg correction), were selected as significant overrepresented terms. This resulted in three sets: a common cold response set; a common salt response set; and a common UV-B radiation response set. Because of the well-annotated *A. thaliana* transcriptome, *A. thaliana* genes from these three sets were used to generate three gene subsets for further GO ontology and clustering analysis. Each subset was scrutinised to identify clusters of overrepresented terms for GO biological processes and then the identified clusters were annotated functionally.

The genes of the common cold response set were scattered among 667 nodes (where a node is a term of GO biological process) and then these nodes were grouped into 16 clusters (**Figure S5-1**). Among these clusters, lateral primordium development (63), nucleoside metabolic process (63), regulation (68), response to stimulus (70) and terpene metabolic process (96) (**Table 5-2**) formed large clusters, suggesting that the lateral primordium development of *A. thaliana* and *P. cheesemanii* was affected by cold stress, and terpene compounds might function in both species' cold responses.

The genes of the common salt response set were scattered among 257 nodes, and these nodes were grouped into 12 clusters (**Figure S5-2**). Large clusters included organ and tissue development and developmental senescence (21), response to stimulus (30), response to hormone stimulus (31) and regulation of systemic acquired resistance (SAR) (48) (**Table 5-2**), which implied that salt stress affected tissue development and induced senescence, and led to SAR in both species.

The genes of the common UV-B radiation response set were scattered among 550 nodes, which were grouped into 17 clusters (**Figure S5-3**). Response to light (32), regulation of secondary metabolic process (40), development morphogenesis and cell fate growth (50) and vitamin B6 metabolic process (58) were four large clusters (**Table 5-2**). This showed that morphogenesis and cell growth were affected by UV-B radiation in both species, and that secondary metabolites and vitamin B6 could be candidates that protect plants from UV-B radiation damage.

Table 5-2. Number of terms of GO biological process in clusters of the gene sets of the common stress responses.

Stress	Cluster	Number of nodes
Cold responses	Plant-type cell wall organization	7
	Soluble vitamin biosynthetic process	10
	Siaminopimelate, lysine, aspartate, threonine, chorismate metabolic process	11
	Post protein ubiquitination and macromolecule modification	16
	Photoperiodism and flowering	19
	Cation volume homeostasis	24
	Protein kinase anion activity	31
	Myo-inositol phosphate rhamnose process	37
	Polysaccharide localization and organelle import callose	42
	Response to stimulus and sucrose mediated signalling	51
	Posttranscriptional regulation of secondary and DNA metabolic process	59
	Lateral primordium development	63
	Nucleoside metabolic process	63
	Regulation	68
	Response to stimulus	70
	Terpene metabolic process	96
Salt responses	Response to sucrose, chitin, disaccharide, hexose and glucose	7
	Fatty acid, wax and oxylipin biosynthetic process	9
	Auxin polar transport	11
	Signal transmission and transduction	14

	Defence bacterium and fungus	15
	Response to light	16
	Valent inorganic cation homeostasis	17
	Organ and tissue development and developmental senescence	21
	Response to stimulus	30
	Response to hormone stimulus	31
	Small molecule biosynthetic process	38
	Regulation of systemic acquired resistance	48
UV-B radiation responses	Glutamate, decarboxylation, dicarboxylic acid, glutamine, and chorismate metabolic process	10
	Quinone cofactor, oxidoreduction coenzyme, ubiquinone metabolic process	12
	Response to monosaccharide stimulus	13
	Xanthophyll tetraterpenoid, diterpenoid, carotenoid, and terpenoid metabolic process	16
	Innate immune system defence	18
	Respiratory electron transport and photosynthesis reaction	19
	Unsaturated fatty lipid and sterol metabolic process	26
	Regulation of metal ion transport	28
	Response to light	32
	Myo-inositol phosphate metabolic process	32
	Macromolecule metabolic process	33
	Protein targeting to organelle and localization	37
	Regulation of secondary metabolic process	40
	Development morphogenesis and cell fate growth	50
	Hyperosmotic salinity response	56
	Vitamin b6 metabolic process	58
	Prephenate compound metabolic process	70

Next, to identify biological processes common to the three stress responses of the two species, the overrepresented terms for GO biological processes of the two species in responding to three stresses were examined, with 125 being identified as shared. These 125 terms were summarised into 10 high-level biological processes of GO term hierarchy: immune system process, response to stimulus, signalling, developmental process, growth, negative regulation of biological process, multi-organism process, biological regulation, secondary metabolic process, cell communication and cellular aromatic compound metabolic process (**Figure 5-8**). The figure shows that all 10 biological processes were common to both *A. thaliana* and *P. cheesemanii* in responding to all three stresses.

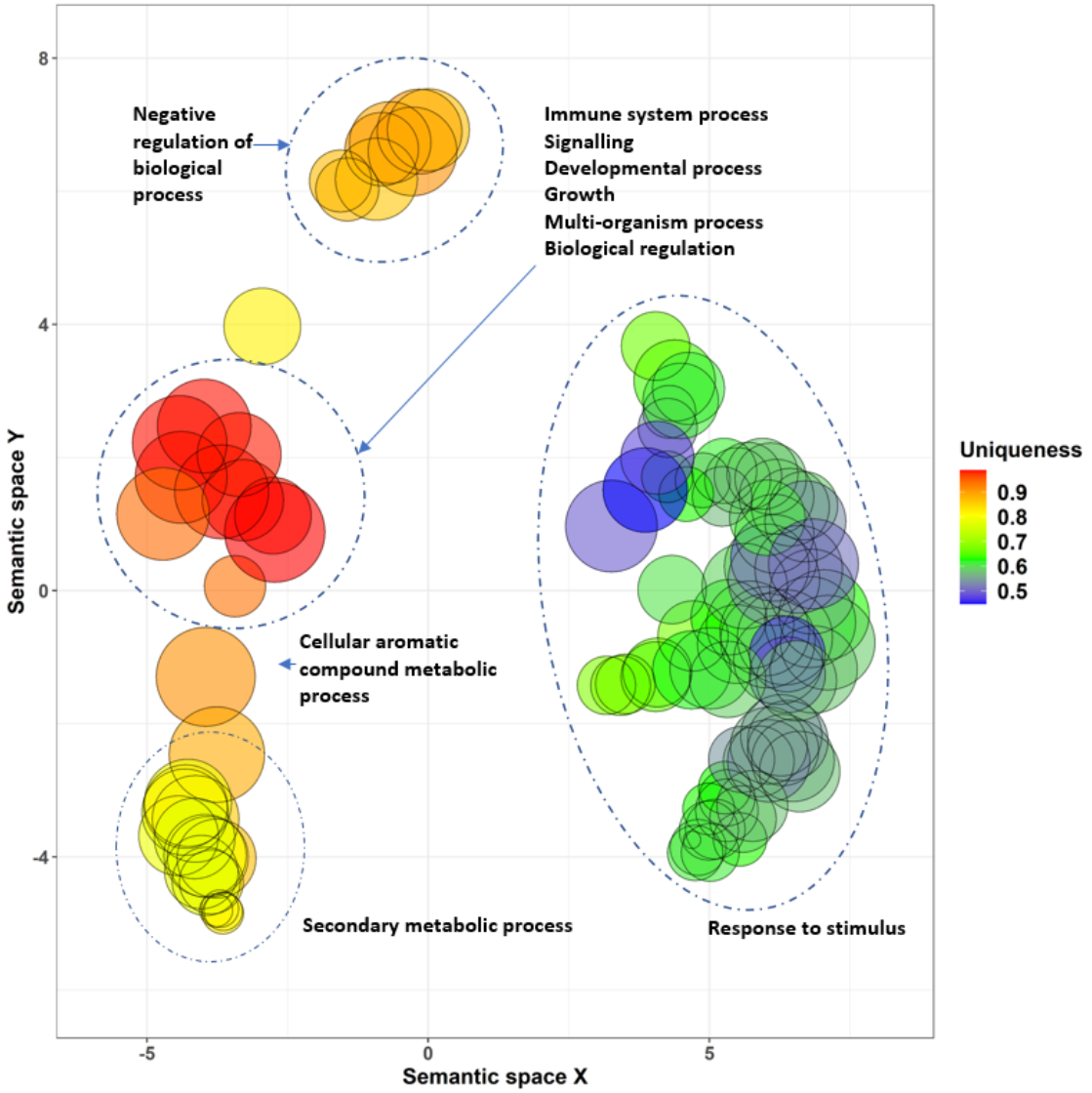

Figure 5-8 Summary of the common overrepresented terms of GO biological process of *A. thaliana* and *P. cheesemanii* in responding to all three stresses.

Two species have a number of the same overrepresented GO in all three stress responses, and they could be summarized into several high-level biological processes of GO term hierarchy. Semantic space X and Y: no intrinsic meaning; uniqueness: measure whether the term is an outlier compared to the list. Namely, the negative of average similarity of a term to all other terms. In REVIGO, multi-dimensional scaling was used to reduce the dimensionality of a matrix of the GO terms' pairwise semantic similarities. First, the terms were placed by using an eigenvalue decomposition of the terms' pairwise distance matrix. Then, a stress minimization step improved the agreement between the semantic similarities of the terms and their closeness in the two-dimensional space. Thus, the semantically similar GO terms should remain close together in the plot. Figure was generated from REVIGO web (http://revigo.irb.hr/) (Supek et al., 2011).

A number of responses were shared by *A. thaliana* and *P. cheesemanii* under each stress: among the responses common to the two species were 10 general responses of *A. thaliana* and *P. cheesemanii* that could be concluded as responding to the three stresses.

5.2.8 Identification of unique biological processes in *A. thaliana* and *P. cheesemanii* responses to three stresses

In the previous section, a number of biological processes were identified as being shared by *A. thaliana* and *P. cheesemanii* in responding to each stress. These biological processes for each species are further analysed in this section. To achieve an overall picture of each species' unique response to each stress, unique overrepresented terms for *A. thaliana* and *P. cheesemanii* biological processes in each stress response (with a 0.05 adjusted *p*-value) were selected as significantly overrepresented terms. This resulted in six sets, including a unique cold response set, salt response set and UV-B radiation response set for *A. thaliana* and a unique cold response set, salt response set and UV-B radiation response set for *P. cheesemanii*. All genes were extracted from these six sets to generate six gene subsets. Each subset was scrutinised to identify clusters of overrepresented terms for GO biological processes and then the identified clusters were annotated functionally.

The genes of the *A. thaliana* unique cold response set were scattered among 484 nodes grouped into 13 clusters (**Figure S5-4**). Among these clusters, sugar-mediated signalling (36), tissue development (40) and photomorphogenesis (55) were three large clusters (**Table 5-3**). This suggested that photomorphogenesis of *A. thaliana* plants could be affected under cold stress, and sugar-mediated signalling may play an important role in *A. thaliana* cold responses.

The genes of the *A. thaliana* unique salt response set were scattered among 358 nodes in 13 clusters (**Figure S5-5**). The large clusters included cellular lipid modification and catabolic process (33), starch biosynthesis (48) and secondary metabolic process (58) (**Table 5-3**), which implied that starch and lipid modification may have special roles in the *A. thaliana* response to salt stress.

The genes of the *A. thaliana* unique UV-B radiation response set were scattered among 729 nodes in 14 clusters (**Figure S5-6**). Vitamin metabolic process (39), cell wall modification defence (51), stomatal movement (54), regulation of starch biosynthetic process (87) and oxygen metabolic process (121) were large clusters (**Table 5-3**). This showed that both physical and biological defences were involved in *A. thaliana* responding to UV-B radiation.

Table 5-3. Number of terms of GO biological process in clusters of *A. thaliana* unique gene sets in responding to three stresses.

Stress	Cluster	Number of nodes
Cold responses	Valent inorganic cation chemical homeostasis	12
	Raffinose family of oligosaccharide metabolic process	23
	Flavone, anthocyanin, and flavonoid metabolism	24
	Nucleobase dependent DNA geometric change	25
	Callose-related cell wall defense	28
	Inositol phosphate and UDP rhamnose process	30
	Sugar mediated signaling	36
	Tissue development	40
	Secondary metabolic process	41
	Hyperosmotic salinity response	47
	Amine process	49
	Photomorphogenesis	55
	ER protein and starch catabolic process	74
Salt responses	Amino acid dephosphorylation	11
	Auxin homeostasis	14
	Cellular light signaling pathway	15
	Multicellular organism growth	19
	Sulfur compound biosynthesis	21
	Programmed cell death	22
	Terpenoid transport and arginine import ion	22
	Hormone mediated signaling pathway	24
	Positive regulation systemic acquired resistance	25
	Cellular lipid modification and catabolic process	33
	Response to salinity, wounding, and temperature	46
	Starch biosynthesis	48
	Secondary metabolic process	58
UV-B radiation responses	Cellular response red light signaling	14
	MAPK intracellular cascade phosphorylation event	14
	Myo-inositol phosphate metabolic process	15
	Anion chemical homeostasis	21
	Regulation of transmembrane anion channel activity	27
	macromolecule modification process	32
	Vitamin metabolic process	39
	Cell wall modification defense	51
	Stomatal movement	54
	Fatty acie metabolic process	59
	Seedling development	81
	Regulation of starch biosynthetic process	87
	Hyperosmotic salinity response	114
	Oxygen metabolic process	121

The genes of the *P. cheesemanii* unique cold response set were scattered among 655 nodes, which were grouped into 18 clusters (**Figure S5-7**). The large clusters, including programmed cell death (54), root differentiation and floral whorl development (63) (**Table 5-4**), suggested that cold stress uniquely affected *P. cheesemanii* in root differentiation, floral whorl development and even induced programmed cell death.

The genes of the *P. cheesemanii* unique salt response set were scattered among 391 nodes grouped into 12 clusters (**Figure S5-8**). The large clusters included glycoside and disaccharide metabolic process (33), cellular polysaccharide metabolic process (57) and regulation of microtubule polymerisation (63) (**Table 5-4**), suggesting that disaccharide and polysaccharide metabolism, as well as microtubule structure, were uniquely affected by salt stress in *P. cheesemanii*.

The genes of the *P. cheesemanii* unique UV-B radiation response set were scattered among 416 nodes in 14 clusters (**Figure S5-9**). The large clusters included seed germination biological regulation (31) and serine family (35) (**Table 5-4**). This suggested that seed germination was significantly affected in *P. cheesemanii* under UV-B radiation stress.

Table 5-4. Number of terms of GO biological process in clusters of *P. cheesemanii* unique gene sets in responding to three stresses.

Stress	Cluster	Number of nodes
Cold responses	Response to mannitol, sucrose, chitin, and disaccharide	5
	Spermidine metabolic process	5
	Photosynthesis	8
	Sterol metabolic process	8
	Glucosinolate and indole glycosinolate catabolic process	11
	Unsaturated fatty acid metabolic process	11
	systemic defense to insect, fungus, and oomycetes	19
	response to light	20
	anion channel activity	21
	anatomical structure homeostasis	26
	regulation of transcription	37
	myo-inositol catabolic process	43
	nuclear transcribed mRNA	47
	programmed cell death	54
	Root differentiation and floral whorl development	63
	Regulation	84
	Cellular response to wounding and temperature	90
	Prephenate biosynthetic pathway	103
Salt responses	Cation transport	7
	Protein glycosylation	9
	Cellular macromolecular disassembly	12
	mRNA nuclear transcribed process	15
	Response to hormone	20
	Development, reproductive, and senescence process	26
	Glycoside and disaccharide metabolic process	33
	Regulation of oxygen and reactive species metabolic process	34
	Response ion stimulus	43
	Cellular polysaccharide metabolic process	57
	Regulation of microtubule polymerization	63
	Amino acid metabolic process	72
UV-B radiation responses	Photosynthesis	8
	Cellular response stimulus to sucrose	9
	Sterol, ergosterol, and steroid metabolic process	12
	GDP-mannose, hexose, and monosaccharide metabolic process	13
	Nucleoside metabolism	24
	Organelle organization	24
	Aromatic amino acid metabolic process	25

	Regulation secondary metabolic process	29
	Seed germination biological regulation	31
	Tissue development	35
	Serine family	35
	Vitamin B6 process	50
	Protein modification process	58
	Response to zinc ion stimulus	63

Overall, in *A. thaliana*, cold stress may readily affect photomorphogenesis, while the majority of *P. cheesemanii*'s unique cold responses involved root differentiation, floral whorl development and would even induce programmed cell death. Second, *A. thaliana* responses to salt stress also affected starch metabolism and lipid modification, whereas disaccharide and polysaccharide metabolism and microtubule structure were affected by salt stress only in *P. cheesemanii*. Finally, *A. thaliana* responses to UV-B radiation were a combination of physical and biological defences, including cell wall modification defence, stomatal movement, vitamin metabolic processes and oxygen metabolic processes. In contrast, seed germination biological regulation was affected in *P. cheesemanii* under UV-B radiation stress.

5.3 Discussion

Plants must respond to a wide range of abiotic and biotic environmental stresses and different plant species have developed unique strategies for dealing with these challenges. In this study, I aimed to gain an overall view of crosstalk in multiple stress responses in *A. thaliana* and *P. cheesemanii* and identify species-specific stress-responsive processes. On the whole, the two species showed similarity in the numbers of responsive genes upregulated and downregulated under each stress, with the exception that there were many fewer downregulated genes in *P. cheesemanii* UV-B radiation responses (**Table 5-**

1 and Figure 5-1). In terms of overrepresented terms for GO biological processes, *A. thaliana* displayed more responses to UV-B radiation stress, while *P. cheesemanii* had more responses to cold stress compared with the other two stress responses (**Figure 5-3**). These findings implied that *A. thaliana* and *P. cheesemanii* would show some similar responses to each stress, and that there should also be unique stress-responsive processes in each species.

5.3.1 Stress response processes were highly conserved in both *A. thaliana* and *P. cheesemanii*

In this study, 10 identified common responsive biological processes involved in the three stresses for both *A. thaliana* and *P. cheesemanii* created an overview of the general process of responding to stress in the two plant species (**Figure 5-8**). It included the receipt of environmental stress signals and signal transduction between cells (signalling), response to stimulus, regulation of biological processes (negative regulation of biological process and biological regulation), plant responses to environmental stress signals in the production of secondary metabolites (secondary metabolic process and cellular aromatic compound metabolic process), responses by the immune system, as well as plant growth and development (multi-organism, growth and developmental processes) (**Figure 5-8**). During the general process of responding to stresses, biological regulation as a control in biological systems might be involved in various stages of stress responses (Shinozaki, Yamaguchi-Shinozaki, & Seki, 2003; Verma, Ravindran, & Kumar, 2016). In this study, negative regulation of biological processes was found to be highly unique compared to other biological processes, suggesting similarity in negative biological regulation in *A. thaliana* and *P. cheesemanii* stress responses (**Figure 5-8**). Negative regulation may act to prevent or reduce the extent of other biological processes by controlling gene

expression, protein modification or interactions between proteins or substrates (Kazan, 2006; A. Li et al., 2017; Seo et al., 2012). In *A. thaliana* and *P. cheesemanii*, the identified biological processes, which were semantically similar to the GO terms for negative regulation, included the regulation of hormone levels, plant growth and development processes, signalling processes and so on, consistent with previous reports. For example, MAPK cascades as endogenous signals negatively regulated plant defence responses, so the knockout of mpk4 in *Arabidopsis* strengthened constitutive SAR with elevated salicylate levels; CBF are a subfamily of AP2 transcription factors, and overexpression of CBF3 repressed auxin signalling while strengthening plants' freezing tolerance; mutating two homologous U-box E3 ubiquitin ligases (AtPUB18 and AtPUB19) in *Arabidopsis* affected ABA signalling positively and strengthened plants' drought tolerance (Li et al., 2017; Petersen et al., 2000; Seo et al., 2012). Therefore, negative regulation could mediate a variety of biological processes by repressing a number of intermediate steps, resulting in changes to plant physiology and morphology to adapt to environmental stresses.

Signalling includes interactions between a cell and its surroundings, such as stress signal perception and signal transduction from extracellular to intracellular spaces (Kaur & Gupta, 2005). Multiple primary sensors have been shown to be involved in initially perceiving various stress signals, while secondary signals between and within cells could be transduced by hormones and secondary messengers shared by multiple stress pathways (Smékalová, Doskočilová, Komis, & Šamaj, 2014). Further, these primary sensors and second messengers are highly conserved in plant species, which may explain the shared responses to the three stresses by *A. thaliana* and *P. cheesemanii* in terms of cell communication and signalling. To date, phospholipids have been suggested as serving as

precursors for the generation of second-messenger molecules, and ROS may be active as intermediate signals in multiple pathways such as ABA signalling (Reczek & Chandel, 2015). Consistent with this, several processes (oxidoreduction and fatty acid biosynthetic process) related to ROS production and phospholipid metabolism were observed in the stress responses of *A. thaliana* and *P. cheesemanii* (**Table 5-2**). Moreover, under various stresses, hormones such as SA, JA and ABA are believed to regulate stress signalling and plant stress tolerance (Horváth, Szalai, & Janda, 2007; Khan, Syeed, Nazar, & Anjum, 2012; Tuteja, 2007). These intercellular second messengers activate transcription factors that initiate the transcription of stress-responsive genes, which is associated with the term for the GO biological process 'response to stimulus'. This is supported by the finding in this study that a variety of responses to stress-related hormones such as ethylene, SA and auxin were activated in both species (**Figures 5-4 and 5-6**), and several hormone-related biological processes were identified as common stress responses in *A. thaliana* and *P. cheesemanii* (**Table 5-2**). This implied that hormone homeostasis was influenced under stress, and a number of the genes involved in hormone responses were activated or repressed. These induced hormone and hormone-responsive genes would have roles in biological regulation or directly mediate downstream plant growth and development processes. These findings suggested that *A. thaliana* and *P. cheesemanii* plants were highly conserved in their perception of the environmental stress signal, signal transduction between and within cells, and hormone regulation.

Plants have complex immune response pathways, which triggers rapid and long-term defence reactions that results in plant cells being insulated from environmental stress challenges (Nürnberger & Scheel, 2001). Several early cellular immune responses have been induced, such as ion flux across the membrane and the rapid production of ROS

(Nürnberger & Scheel, 2001). Further, callose deposition occurs as a long-term immune response, which isolates damaged plant tissue to create a physical barrier (Ehrhardt-Brocardo, Coelho, Souza, & Mathias, 2019). A related process, polysaccharide localisation and organelle import callose, was found to be common to the responses of the two species to cold stress in this study (**Table 5-2**). Meanwhile, as a downstream plant immune response, the metabolism of two major categories of secondary metabolites was affected by the three stresses in *A. thaliana* and *P. cheesemanii* (**Table 5-2**). Prephenate was a precursor of the biosynthesis of phenylalanine-derived flavonoids, and terpene was considered a protective secondary metabolite (Gershenzon & Dudareva, 2007; Hrazdina, 1992). Flavonoids could act as antioxidant compounds to scavenge ROS during various stresses at the cellular level, which would minimise the effects of stress on plant growth and development. For instance, enhanced photosynthesis efficiency and antioxidative capacity were observed in *A. thaliana* plants overexpressing phytoene synthase (PSY) in response to 100 mM NaCl (Han, Li, & Zhou, 2008). PSY is a rate-limiting enzyme in carotenoid biosynthesis (Ahmadabadi, 2014). Conversely, terpenes are an important group of secondary metabolites with roles in plant defence against herbivores and pathogens, and in attracting pollinators; they also serve as plant growth regulators. Geranyl pyrophosphate synthase (GPPS) is known as a key enzyme functioning in the first condensation step of terpene biosynthesis. Overexpression of the *GPPS* gene in *Camelina sativa* plants altered terpene metabolism and the transgenic plants developed larger leaves, grew more and longer internodes, and flowered earlier compared with wild-type plants (Xi, Rossi, Lin, & Xie, 2016). Further, the derivates of terpene, sterols, are components of the cell membrane (Boutte & Grebe, 2009). They interact physically with phospholipids, which are influenced under various stresses resulting in alterations to the sterol content in membranes (Rogowska & Szakiel, 2020). Under stress, free sterols of

the membrane reduced slightly in salt-tolerant dwarf-cashew (*Anacardium occidentale L.*) seedlings, while a significant increase was found in the salt-sensitive genotype, suggesting the importance of sterols for maintaining the integrity of the plasma membrane and keeping the plant healthy (Alvarez-Pizarro, Gomes-Filho, de Lacerda, Alencar, & Prisco, 2009). Overall, the plant immune responses and altered secondary metabolism observed in this study are likely to involved in scavenging ROS, maintaining membrane system stability and sealing off injured plant tissue, which would minimise the impact of stresses on plant growth and development.

Plant growth and development are subject to regulation by multiple hormones, which can result in changes to the plant morphogenesis pattern and even senescence (Goldthwaite, 1987; Holder, 1979). For instance, smaller root systems and thinner roots caused by cold stress resulted from a reduction in trafficking of the auxin efflux carrier PIN2 and inhibition of the lateral relocalisation of PIN3 in *Arabidopsis*; shoot branching was regulated by multiple factors, including scavenging of ROS, SA and JA signalling pathway; stress-induced senescence was regulated by osmotic stress signalling and multiple hormones such as strigolactones, cytokinin and auxin (Guo & Qin, 2016; Islam et al., 2013; Peleg, Reguera, Tumimbang, Walia, & Blumwald, 2011; Shibasaki, Uemura, Tsurumi, & Rahman, 2009). This may explain why a few large clusters of common responses related to plant growth and development and enrichment of multiple hormone responses were found in both *A. thaliana* and *P. cheesemanii* in this study (**Figures 5-4, 5-6 and Table 5-2**). Hormones act not only in stress signalling but also directly in plant growth and development, which tightly interconnects plant growth, development and adaptation to the environment. For example, under cold stress, limited lateral primordium development would reduce biomass, which may alleviate pressure on photosynthesis

under stress; via hormone regulation, the senescence of older leaves could be induced, which in turn may reduce the water and nutrient consumption of plants under salt stress; and UV-B-induced leaf curling may reduce leaf exposure to UV-B radiation. These morphological deformations induced by various stresses represent a balance in plant fitness and resistance that enables them to survive environmental hardship.

In conclusion, the stress process was found in this study to be highly conserved between the two plant species under various stresses; the process included stress signal perception on the cell surface, signal transduction between and within cells, generation of intracellular second messengers, activation of transcription factors, initiation of transcription of stress-responsive genes, plant immune responses with altered secondary metabolism, and reprogrammed plant growth and development. These processes are likely to minimise the impact of various stresses, which improves the survival rate of plants under a wide range of stresses.

5.3.2 *A. thaliana* unique responses present in stress signalling, immune defence, biological regulation, and plant growth and development

Starch metabolism under abiotic stresses has been investigated in a number of plant species, including *Litchi chinensis*, *Hordeum vulgare* and *A. thaliana*, some of which showed strengthened stress tolerance (Gous, Hasjim, Franckowiak, Fox, & Gilbert, 2013; Shen, Xiao, Qiu, Chen, & Chen, 2016; Yano, Nakamura, Yoneyama, & Nishida, 2005). Starch could be considered a storage compound, and its remobilisation under abiotic stress provides energy and a carbon source for seed filling (Nayyar, Bains, Kumar, & Kaur, 2005). However, since both degradation and accumulation of starch and its plasticity in different plant organs have been reported under stresses, starch should not be

considered a storage compound alone (Goyal, 2007; Nagao, Minami, Arakawa, Fujikawa, & Takezawa, 2005; Skirycz et al., 2010; Xuchu Wang et al., 2013). In this study, *A. thaliana* plants responded to cold stress in starch catabolic processes, while starch biosynthesis and regulation of the starch biosynthetic process occurred during salt and UV-B radiation responses, suggesting that there may have been different roles for starch in *A. thaliana*'s three stress responses (**Table 5-3**). This may partially explain the tight connection between starch metabolism and sugar signalling under stresses. Cuellar-Ortiz et al. (2008) found that the starch content increased in pods but decreased in leaves, coupled with an accumulation of sucrose, fructose and glucose as well as elevated sucrose synthase activity in a drought-resistant cultivar of *Phaseolus vulgaris* L. (Cuellar-Ortiz, De La Paz Arrieta-Montiel, Acosta-Gallegos, & Covarrubias, 2008). These sugars participated in stress-related sugar signalling pathways. High sugar content (sucrose, fructose and glucose) was found in jasmonate response mutants under cold stress, suggesting interactions between sucrose and jasmonate signalling in plant cold responses; global transcription profiling revealed that a large number of stress-responsive genes were induced by glucose in *Arabidopsis*; the application of exogenous sucrose improved the tolerance of *A. thaliana* seedlings to anoxia partially because the supplemented sucrose elevated the repression of auxin-responsive genes under oxygen deprivation (Loreti, Poggi, Novi, Alpi, & Perata, 2005; Price, Laxmi, Martin, & Jang, 2004; Wingler, Tijero, Müller, Yuan, & Munné-Bosch, 2020). Therefore, as found in the unique *A. thaliana* cold response, starch catabolism could increase sugar content and in turn activate sugar-mediated signalling in responding to cold stress. Starch biosynthesis could produce compounds for storage processes such as seed filling under salt and UV-B radiation stresses (**Table 5-3**).

One group of downstream products of sucrose is raffinose family oligosaccharides (RFOs), which are involved in plant stress defence mechanisms. Overexpression of the gene encoding galactinol synthase (GolS) catalysed the first step of biosynthesis of RFO, increased endogenous galactinol and raffinose, and improved drought tolerance in *A. thaliana*; overexpression of *GolS* in Arabidopsis plants also strengthened the tolerance of transgenic plants to osmotic and salinity stresses, coupled with increased galactinol and raffinose; and increased galactose and raffinose were found in a chilling-tolerant genotype of *Oryza sativa* L. under chill stress (Morsy, Jouve, Hausman, Hoffmann, & Stewart, 2007; Selvaraj et al., 2017; Sun et al., 2013). Other than sugar signalling, RFOs are likely to play a role in ROS scavenging in chloroplasts under stresses (Van den Ende & Peshev, 2013). Following RFO scavenging, the singlet oxygen produced by oxidised RFO could be scavenged by other oxidants like flavonoids (Wu & Ng, 2008). This may explain why both the RFO metabolic process and flavonoid metabolism were present simultaneously in the unique *A. thaliana* cold responses in this study (**Table 5-3**). Other ROS scavengers, including vitamin B6, seemed to be involved in UV-B radiation responses, which may have contributed to the ROS-induced stomatal movement under UV-B radiation stress (**Table 5-3**). UV-B radiation induced stomatal closure by production of NO and H_2O_2 in guard cells, which could be efficiently prevented by NO and H_2O_2 scavengers, mediated by the ethylene biosynthesis and signalling pathway (Shi et al., 2015). Further, ROS caused by UV-B radiation stress could be quenched by vitamin B6, also known as pyridoxine (Havaux et al., 2009). I propose that, in *A. thaliana*, UV-B radiation affected oxygen metabolism and the resulting production of ROS in turn induced stomatal movement. This process was regulated by ROS scavengers, including vitamins. Aside from these ROS scavengers, other protecting actions like cation and anion chemical homeostasis and callose-related cell wall defence may provide rapid and long-

term immune defence in cold and UV-B radiation responses (**Table 5-3**). This may contribute to plants' acquired resistance to biotic stresses following UV-B radiation (Kravets, Zelena, Zabara, & Blume, 2012). Further, callose-related cell wall defence may be induced by callose deposition in the cell wall which was evident in the common stress responses of *A. thaliana* and *P. cheesemanii* (**Table 5-2**). In some halophytic species, tolerance to cold, salinity and drought depends on osmotic adjustment between K^+ and Na^+ in the vacuole and compatible solutes in the cytoplasm, suggesting that cation and anion homeostasis in cells may help in stabilising membrane proteins and lipids, and in turn prevent cell disruption under cold stress (Hao, Lucero, Sanderson, Zacharias, & Holbrook, 2013; Pan et al., 2016). Although it has been suggested that the callose-related cell wall defence is correlated with pathogen tolerance in a number of plant species, callose deposition was also observed in some temperature stress cases with unknown physiological and molecular mechanisms (Stass & Horst, 2009). Thus, the mechanism of cell wall defences against cold and UV-B radiation stresses in *A. thaliana* plants may be a future research direction.

These responses, aside from the general regulation discussed in the previous section, might be regulated by stress-induced DNA geometric changes under cold stress and SAR under salt stress (**Table 5-3**). A particular degree of DNA supercoiling is essential in bacterial adaptation to environmental stress (Tse-Dinh, Haiyan, & Menzel, 1997). DNA topoisomerases could regulate DNA supercoiling at different temperatures, which may affect the interaction between transcriptional activators and RNA polymerase to promote transcription, consequently affecting the physiology of the organism; few studies in plants have been reported (Dorman & Dorman, 2016). Here, nucleobase-dependent DNA geometric change was found in the unique *A. thaliana* cold responses, which implied the

possibility of DNA topology-dependent transcription regulation functioning in *A. thaliana* cold responses. However, this needs further examination to determine if there is a similar mechanism in plants to that in bacteria. Conversely, SAR is an induced long-lasting and broad-spectrum defence that utilises SA as a signal molecule to mobilise a positive regulator to interact with a transcription factor and then initiate transcription of defence genes (Ali et al., 2018). Interestingly, recent reports demonstrated the potential role of lipid molecules in SAR signalling. The *dir1* (*DEFECTIVE IN INDUCED RESISTANCE 1*) Arabidopsis mutant, which is unable to develop SAR in systemic leaves because of a deficiency in the mobile signal demonstrates normal local resistance to pathogens (Maldonado, Doerner, Dixon, Lamb, & Cameron, 2002). The similarity between DIR1 and lipid transfer proteins (LTPs) suggested that the mutant may use lipid molecules as a mobile signal (Maldonado et al., 2002). This might be supported by the findings of the current study, in which both cellular lipid modification and catabolic processes, and positive regulation of SAR were discovered in the unique *A. thaliana* salt responses (**Table 5-3**).

Therefore, unique *A. thaliana* responses in stress signalling, immune defence and regulation resulted in plant morphology adaptation and influenced growth and development such as the affected photomorphogenesis under cold stress, multicellular organism growth under salt stress and seedling development under UV-B radiation. These changes could help plants to survive environmental stresses. For example, the UV-B radiation-induced deposition of cell wall microfibrils would alter cell shape and then change leaf morphology, and this could contribute to leaf curling and consequently reduce leaf area exposure to UV-B radiation; UV-B radiation influenced stomata to control CO_2 influx and water loss, which affected plant growth and physiology; and the limited

development of the lateral root system under cold stress could reduce the consumption of photosynthesis products (Hund, Fracheboud, Soldati, & Stamp, 2008; Staxén & Bornman, 1994; Sullivan & Teramura, 1990). In conclusions, the unique *A. thaliana* stress responses in stress signalling, immune defence and regulation, as well as plant growth and development, suggested rapid and long-term changes at different levels, which could jointly guarantee cell, organ and whole plant functioning under stress.

5.3.3 Central role of anthocyanins in *P. cheesemanii* unique stress responses

In the enrichment network analysis of *P. cheesemanii*, only one main cluster was found, which might have resulted from the small number of annotated transcripts in biological processes (**Figure S4-1**). Further, unlike in the enrichment of flavone and flavonol biosynthesis in *A. thaliana* stress responses, anthocyanins were found in the centre of this cluster, which implied its important role in *P. cheesemanii* stress responses (**Figure 5-6**). Anthocyanins are a group of water-soluble flavonoid pigments for which biosynthesis involves more enzymes and steps compared with biosynthesis of flavone and flavonol. Although three flavonoids showed similar accumulation under biotic and abiotic stresses in various plant species, they would have different biological functions. For instance, flavone and flavonol controlled plant growth and development by redirecting the polar auxin transport flow (Santelia et al., 2008). In this study, the enriched anthocyanin biosynthesis, rather than other flavonoids, in response to multiple stresses in *P. cheesemanii* also suggested that anthocyanins may play a special role in *P. cheesemanii* stress responses (**Figure 5-6**).

There were several anthocyanin-related large clusters in *P. cheesemanii*'s unique responses under the three stresses. Under cold stress, influenced floral whorl development

was a large cluster (63) in *P. cheesemanii*, and was accompanied by anthocyanin accumulation in a wide range of plant species (Martin & Gerats, 1993). Therefore, *P. cheesemanii*'s enriched anthocyanin biosynthesis could be relevant to the affected floral whorl development under cold stress. The precursor of anthocyanin biosynthesis is prephenate, for which the biosynthetic pathway was found to be the largest affected cluster in *P. cheesemanii*, which included 103 overrepresented terms for GO biological processes (**Table 5-4**). Next, increased anthocyanin content could function as antioxidant-scavenging ROS caused by cold stress, which avoided programmed cell death (another large cluster with 54 nodes in this study) (Dauphinee, Fletcher, Denbigh, Lacroix, & Gunawardena, 2017; Li et al., 2015). In contrast, under salt stress, sugar (disaccharide and polysaccharide) metabolism was affected (**Figure 5-6 and Table 5-4**). Increasing concentrations of metabolites of carbohydrate metabolism could elevate the cell osmotic potential to adjust osmotic pressure, resulting in the maintenance of water uptake (Pattanagul & Thitisaksakul, 2008). In that case, anthocyanin biosynthesis would be modulated by the changed sugar content (Solfanelli, Poggi, Loreti, Alpi, & Perata, 2006). Under UV-B radiation, there was a large cluster in regulation secondary metabolic processes (29), which included the regulation of both flavonoid biosynthesis and anthocyanin metabolism as the main overrepresented biological processes (**Table 5-4**). Meanwhile, coupled with UV-B radiation-induced sterol synthesis, anthocyanin accumulation was associated with reduced membrane damage in other plant species (Berli et al., 2010). Thus, the affected sterol metabolism in *P. cheesemanii* could stabilise the membrane system and minimise the ROS damage caused by UV-B radiation (**Table 5-4**).

The enriched anthocyanin biosynthesis under salt and UV-B radiation stresses was likely to utilise anthocyanin's antioxidative activity to scavenge ROS caused by stresses, which might differ somewhat from the role of anthocyanins in *P. cheesemanii*'s unique cold responses. Similar to salt and UV-B, under cold stress, anthocyanin biosynthesis could be induced by low temperatures to scavenge ROS, in turn avoiding the programmed cell death caused by ROS. However, the responses of *P. cheesemanii* to cold stress included many more overrepresented terms for GO biological processes relative to those in responses to salt and UV-B radiation (**Figure 5-3**), implying the complexity of *P. cheesemanii* cold responses. Increased anthocyanin levels, coupled with the large cluster of unique responses in floral whorl development, might suggest that cold stress would induce a transition from vegetative to reproductive phrase in *P. cheesemanii*. Such a transition could be an adaptation strategy for *P. cheesemanii* survival under cold stress.

In conclusion, consistencies in the stress responses of *A. thaliana* and *P. cheesemanii* reflect their close evolutionary relationship; their multiple-stress transcriptomes also displayed similarity in aspects of multiple-stress response crosstalk and common stress responses. Both species regulated their growth and development via hormone regulation under stresses. However, there were indeed unique stress responses in *A. thaliana* and *P. cheesemanii*. Photomorphogenesis in cold response, starch metabolism and lipid modification under salt stress, stomatal movement and cell wall modification defence to UV-B radiation were unique to *A. thaliana*. Anthocyanins served as a centre node to connect phenylalanine biosynthesis and programmed cell death of cold responses, disaccharide and polysaccharide metabolism of salt responses, and the regulation of secondary metabolic processes in the UV-B radiation response of *P. cheesemanii*. These similarities and uniqueness in the stress responses o *A. thaliana* and *P. cheesemanii* may

have evolved from species adaptations to their own ecological niches, which helps them survive particular environmental stresses. By analysing multiple-stress transcriptomes across species, a better understanding of the shared and unique stress-responsive pathways of *A. thaliana* and *P. cheesemanii* was achieved. Meanwhile, the cross-stress transcriptome analysis provided an insight into multiple-stress crosstalk in each species.

Chapter 6 General discussion and conclusion

P. cheesemanii is a close relative of *A. thaliana* and is an allotetraploid perennial herb widespread in the South Island of New Zealand. It grows at altitudes of up to 1,000 m where it is subject to relatively high levels of UV-B radiation. To gain the first insights into the origin of *Pachycladon* and how it copes with multiple stresses, I constructed a draft genome of *P. cheesemanii* and compared its multiple-stress transcriptome profile with that of *A. thaliana*.

A high-quality draft genome of *P. cheesemanii* was assembled with a high percentage of conserved single-copy plant orthologues. Although this draft genome was 70.8% of the estimated genome, 96.2% highly conserved single-copy plant orthologs were identified. This could be explained by the two sets of *P. cheesemanii* subgenomes, where orthologs from one genome may not have been sequenced, while they may have been successfully sequenced from the other subgenome. Based on this high-quality draft genome, a synteny analysis facilitated the establishment of a phylogenetic tree including the genomes of other species of the Brassicaceae family. This revealed a close relationship between *P. cheesemanii* and *Boechera stricta*, as well as *Arabidopsis*, in Brassicaceae Lineage I. A number of recent studies have suggested that one of the *Pachycladon* subgenomes evolved from the *Arabidopsis* lineage on the basis of several lines of evidence. However, controversy remains about the origin of another one from Pachycladon (Heenan, Mitchell, & Koch, 2002; Joly, Heenan, & Lockhart, 2009; McBreen & Heenan, 2006). Here, I have provided further evidence to support an *Arabidopsis* origin for one *Pachycladon* subgenome, by applying a comparative genome approach to analyse 28 available Brassicaceae genomes. Although I failed to identify the ancestor of another *Pachycladon*

subgenome in this study, my findings suggest that some species from Brassicaceae Lineage EII are potential candidates, based on the highest cumulative values of *B. stricta* and *E. heterophyllum* (the species from Brassicaceae Lineage EII). This result was inconsistent with speculation by Joly et al. (2009) that another *Pachycladon* subgenome may have originated from the tribe Smelowskieae or its close relatives that no longer have extant lineages, as deduced from a phylogenetic analysis based on five single-copy nuclear genes. Further, since Smelowskieae is the tribe from Brassicaceae Lineage I, this speculation cannot explain traces of the Brassica genome in the *Pachycladon* genome (Joly et al., 2009). Conversely, not only was my conclusion based on analysis of a large amount of genome data, but the identified potential candidate belongs to a tribe (Eutrema) that is a close phylogenetic relative of the Brassiceae tribe, which may explain the previous finding of traces of the Brassica genome in the *Pachycladon* genome (Joly et al., 2009). Therefore, the other *Pachycladon* subgenome would more likely have originated from the Eutrema tribe than the Smelowskieae tribe.

Supportive evidence for my conclusion would be provided by the availability of genome data from Brassicaceae Lineage EII or the Brassicaceae family. In turn, stronger clarification of the phylogenetic relationships among species in this family would help to accurately determine the origin of the two *Pachycladon* subgenomes. Previous studies focused on phylogenetic relationships between species mainly based on an orthology analysis of a single gene or several genes (Joly et al., 2009; McBreen & Heenan, 2006). In contrast, I compared the whole genomes of multiple species to determine the origin of the two *Pachycladon* subgenomes in this study. Further, analysis of a large amount of genome data enabled me to make more accurate predictions regarding the phylogenetic relationships among species compared with previous approaches based on analysis of

single or multiple genes. This approach could be widely used for phylogenetic analysis of other species for which genome data are available. Therefore, along with the increasing amount of genome data accessible for species, application of the approach presented in this study could contribute to better clarifying the evolutionary relationships among organisms.

The differences in growth reduction caused by UV-B radiation, as shown in the *A. thaliana* and the *P. cheesemanii* accessions, suggests that the two species might have different UV-B radiation responses despite their close phylogenetic relationship. Further, the genetic mechanisms underlying the different UV-B radiation response phenomena in *A. thaliana* and *P. cheesemanii* suggests that the tolerance mechanism in *P. cheesemanii* differs from that in *A. thaliana*. The UVR8-dependent pathway in the *A. thaliana* accession may be the basis of its UV-B radiation tolerance, while the UVR8-independent pathway may play a more prominent role in *P. cheesemanii* accessions (**Figure 3-5**). The UVR8-dependent pathway protects plants against harm caused by UV-B radiation by initiating a number of biological processes including morphogenic and physiological responses, and this pathway has been well characterised (Heijde & Ulm, 2012). However, it has been suggested that some UV-B responses are independent of UVR8, especially in the case of high fluence rates of UV-B radiation (Li et al., 2015). Recently, a transcriptomic analysis of the *A. thaliana* uvr8 mutant revealed that a number of genes responded to low fluence rate of UV-B independently of UVR8 as well (O'Hara et al., 2019). Interestingly, the same study mentioned that these UVR8-independent pathways induced a eustress response; that is, stress not causing permanent damage but promoting acclimation (Jansen & Urban, 2019). Such UV acclimation has been observed and widely studied in a number of higher plants including *Oenothera stricta* and timberline plants,

(Robberecht & Caldwell, 1983; Turunen & Latola, 2005). Therefore, I speculate that UV-B radiation of 2 μmol m^{-2} s^{-1} likely induces plant acclimation in *P. cheesemanii* accessions, based on its stronger UVR8-independent pathway and lack of obvious damage after UV-B radiation treatment. In contrast, in the *A. thaliana* accessions in this study, the necrotic lesions on leaves indicated that destructive damage had occurred after UV-B radiation treatment, and the UVR8-dependent pathway was activated rather than the UVR8-independent pathway as the main strategy against such destructive stress. Further, my study provides further evidence for the role of the UVR8-independent pathway in plant acclimation. If this hypothesis is correct, then I would expect that the stronger UVR8-dependent pathway in *P. cheesemanii* might be induced by UV-B radiation levels higher than 2 μmol m^{-2} s^{-1}. It could not be ruled that another UVR8-independent response would be induced by UV-B radiation of more than 2 μmol m^{-2} s^{-1} as the UVR8-independent pathway has not yet been well characterised in plant species, especially *P. cheesemanii*.

This study failed to construct a UVR8-independent pathway for the *P. cheesemanii* UV-B radiation response. Since it showed no obvious destructive damage and significantly activated its UVR8-independent pathway at moderate-to-high level of UV-B radiation, *P. cheesemanii* may serve as an excellent model to study the role of the UVR8-independent pathway in plant acclimation at high UV-B radiation. To further identify this pathway in the *P. cheesemanii* UV-B radiation response, first, classical mutagenesis may be the most efficient approach, and UVR8 mutants of *P. cheesemanii* could be generated via CRISPR technology. Second, transcriptome analysis would be helpful to identify the genes involved in the UVR8-independent pathway by using *P. cheesemanii* UVR8 mutants. These results would provide an overall view of the UVR8-independent pathway in *P.*

cheesemanii. The relationships among the genes identified from the transcriptome analysis still need to be further assessed in other experiments such as mutagenesis studies. Further, the GO enrichment and network analyses demonstrated strong similarities between the interactions in the multiple stress responses in *A. thaliana* and *P. cheesemanii* and a number of highly conserved stress response processes. Species-specific unique stress responses were also identified, which in *A. thaliana* included photomorphogenesis under cold stress; starch metabolism and lipid modification under salt stress; cell wall modification defence; stomatal movement; and the vitamin and oxygen metabolic processes under UV-B radiation. For *P. cheesemanii* these included root differentiation and floral whorl development under cold stress; disaccharide and polysaccharide metabolism under salt stress; and seed germination biological regulation under UV-B radiation. Overall, the plant species that evolved from the same ancestor were found to share most of their stress responses, while genetic diversity enabled them to develop some species-specific responses.

This study has highlighted the central role of anthocyanins in *P. cheesemanii*'s unique stress responses, implying that secondary metabolism in *P. cheesemanii* may differ from that in *A. thaliana* in responding to stresses. Secondary metabolism was considered by Deng et al., (2012) to be affected by ploidy level, and it resulted in different responses between plant species to the same environmental stress in that study (Deng et al., 2012). As the main stress-induced secondary metabolites, flavonoids have multiple roles in plant stress responses, and their composition varies among different-ploidy plants (Omezzine, Bouaziz, Daami-Remadi, Simmonds, & Haouala, 2017). Early on it was reported that 12 diploid and 8 autotetraploid plants of *Briza media* differed in which flavonoids they contained, and that there was accumulation of vitexin and isovitexin in diploids but

orientin and iso-orientin in tetraploids (Williams & Murray, 1972). Recently, it has been shown that the content of flavonoid compounds differs in diploid and mixoploid *Trigonella foenum-graecum* L. plants (Omezzine et al., 2017). In the current study, anthocyanins served as the central node in the *P. cheesemanii* response network to multiple stresses, while flavone and flavonol were the main flavonoids in *A. thaliana*'s cold and UV-B radiation responses (**Figure 5-4** and **5-6**). I hypothesised that such variation in stress-induced flavonoids between *A. thaliana* and *P. cheesemanii* would result from their different ploidy levels. In turn, its potential altered composition of flavonoids has enabled *P. cheesemanii* to survive environmental stresses in New Zealand, which may be related to its resulting altered antioxidant capacity. It has been reported that antioxidant capacity is associated with anthocyanin, polyphenol and flavonoid content in blueberries, onions and rice, among others (Li et al., 2017; Min, Gu, McClung, Bergman, & Chen, 2012; Zhang, Peng, Xu, Lü, & Wang, 2016). A plant's antioxidant capacity to a large extent determines whether it can handle the oxidative pressure caused by stress conditions (Darkó, Ambrus, Fodor, Király, & Barnabás, 2009; Gao, Feng, Bai, He, & Wang, 2019; Martins et al., 2017). One study concluded that if the oxidative pressure induced by UV-B radiation could be conducted by a plant's ROS-scavenging capacity, then eustress responses will occur; otherwise reverse distress responses will occur (Hideg, Jansen, & Strid, 2013). Coupled with a previous finding of the strengthened UV-B tolerance in *P. cheesemanii* accessions and lack of obvious destructive damage after UV-B radiation, the potential altered composition of flavonoids in *P. cheesemanii* may elevate its antioxidative capacity, helping it deal with stress-induced oxidative pressure and resulting in its eustress response. Therefore, I hypothesised that:

1. The allopolyploidy of *P. cheesemanii* changes its secondary metabolism, resulting in altered flavonoids and enhanced anthocyanin biosynthesis in responding to stresses.

2. In *P. cheesemanii*, the changed flavonoid composition under stresses improves plant oxidative capacity and in turn strengthens the scavenging of UV-B-induced ROS.

3. Consequently, eustress responses occur in *P. cheesemanii*, which enables the acclimation of *P. cheesemanii* plants to moderate–high UV-B radiation.

Therefore, as a next step, it would be interesting to study how plant ploidy influences the composition of flavonoids to examine whether the stress-induced flavonoid composition of *P. cheesemanii* differs from that of *A. thaliana* and how species-specific flavonoid composition helps plants survive stresses. First, stress-responsive phenotypes could be investigated in *A. thaliana* and *P. cheesemanii* plants with and without stress treatment. Second, a metabolomic analysis could be applied to study physiological changes underlying these phenotypes. For instance, the changed ROS homeostasis induced by stresses could be analysed by comparing antioxidant enzyme activities in treated and untreated plants; the composition of flavonoids could be profiled by using high-performance liquid chromatography (HPLC); and a comprehensive analysis of secondary metabolism could employ the isotope-labelling metabolic flux technique to compare stress-induced changes in the secondary metabolism of *A. thaliana* and *P. cheesemanii*. Third, the stress-induced expression of genes involved in the biosynthesis of flavonoids could be confirmed by applying RT-qPCR in treated and untreated plants.

The close evolutionary relationship between *A. thaliana* and *P. cheesemanii* results in their similar responses to the same environmental stresses. However, because of the unique living environment of each species, *P. cheesemanii* has evolved a number of unique environment-related characteristics at the transcriptome level that may contribute to its increased UV-B radiation tolerance. Therefore, like flavonoid composition, other

unique stress-responsive processes identified in Chapter 5 could contribute to the formation of the uniqueness of *P. cheesemanii*'s stress responses. The role of these unique processes in *P. cheesemanii* stress responses will be an interesting future research direction, which could increase understanding of how *P. cheesemanii* acclimates to its unique living environment.

References

Ahmadabadi, M. (2014). PSY is a rate-limiting enzyme in carotenoid biosynthesis pathway in maize (*Zea mays*) leaves. *New Cellular and Molecular Biotechnology Journal, 4*(13), 71-75.

Ali, A., Shah, L., Rahman, S., Riaz, M. W., Yahya, M., Xu, Y. J., . . . Cheng, B. (2018). Plant defense mechanism and current understanding of salicylic acid and NPRs in activating SAR. *Physiol Mol Plant Path, 104*, 15-22.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol, 215*(3), 403-410.

Alvarez-Pizarro, J. C., Gomes-Filho, E., de Lacerda, C. F., Alencar, N. L. M., & Prisco, J. T. (2009). Salt-induced changes on H^+-ATPase activity, sterol and phospholipid content and lipid peroxidation of root plasma membrane from dwarf-cashew (*Anacardium occidentale* L.) seedlings. *Plant Growth Regul, 59*(2), 125-135.

Appleford, N. E., & Lenton, J. R. (1991). Gibberellins and leaf expansion in near-isogenic wheat lines containing Rht1 and Rht3 dwarfing alleles. *Planta, 183*(2), 229-236.

Bahieldin, A., Sabir, J. S., Ramadan, A., Alzohairy, A. M., Younis, R. A., Shokry, A. M., . . . Al-Kordy, M. A. (2014). Control of glycerol biosynthesis under high salt stress in Arabidopsis. *Funct Plant Biol, 41*(1), 87-95.

Barboza, L., Effgen, S., Alonso-Blanco, C., Kooke, R., Keurentjes, J. J., Koornneef, M., & Alcázar, R. (2013). Arabidopsis semidwarfs evolved from independent mutations in GA20ox1, ortholog to green revolution dwarf alleles in rice and barley. *Proc Natl Acad Sci U S A, 110*(39), 15818-15823.

Bartels, D., & Sunkar, R. (2005). Drought and salt tolerance in plants. *Crit Rev Plant Sci, 24*(1), 23-58.

Bauer, H., Ache, P., Lautner, S., Fromm, J., Hartung, W., Al-Rasheid, K. A., . . . Lachmann, N. (2013). The stomatal response to reduced relative humidity requires guard cell-autonomous ABA synthesis. *Cur Biol, 23*(1), 53-57.

Berli, F. J., Moreno, D., Piccoli, P., HESPANHOL-VIANA, L., Silva, M. F., BRESSAN-SMITH, R., . . . Bottini, R. (2010). Abscisic acid is involved in the response of grape (*Vitis vinifera* L.) cv. Malbec leaf tissues to ultraviolet-B radiation by enhancing ultraviolet-absorbing compounds, antioxidant enzymes and membrane sterols. *Plant Cell Environ, 33*(1), 1-10.

Berteli, F., Corrales, E., Guerrero, C., Ariza, M. J., Pliego, F., & Valpuesta, V. (1995). Salt stress increases ferredoxin-dependent glutamate synthase activity and protein level in the leaves of tomato. *Physiol Plant, 93*(2), 259-264.

Bharadwaj, D. N. (2015). Polyploidy in crop improvement and evolution. In *Plant biology and biotechnology* (pp. 619-638): Springer.

Biever, J. J., Brinkman, D., & Gardner, G. (2014). UV-B inhibition of hypocotyl growth in etiolated *Arabidopsis thaliana* seedlings is a consequence of cell cycle arrest initiated by photodimer accumulation. *J Exp Bot, 65*(11), 2949-2961.

Boudsocq, M., Barbier-Brygoo, H., & Laurière, C. (2004). Identification of nine sucrose nonfermenting 1-related protein kinases 2 activated by hyperosmotic and saline stresses in *Arabidopsis thaliana*. *J Biol Chem, 279*(40), 41758-41766.

Boutte, Y., & Grebe, M. (2009). Cellular processes relying on sterol function in plants. *Curr Opin Plant Biol, 12*(6), 705-713. doi:10.1016/j.pbi.2009.09.013

Bowers, J. E., Chapman, B. A., Rong, J., & Paterson, A. H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature, 422*(6930), 433.

Brochmann, C., Brysting, A., Alsos, I., Borgen, L., Grundt, H., Scheen, A. C., & Elven, R. (2004). Polyploidy in arctic plants. *Biol J Linn Soc, 82*(4), 521-536.

Brown, B. A., & Jenkins, G. I. (2008). UV-B signaling pathways with different fluence-rate response profiles are distinguished in mature Arabidopsis leaf tissue by requirement for UVR8, HY5, and HYH. *Plant Physiol, 146*(2), 576-588.

Burgess, P., & Huang, B. (2016). Mechanisms of hormone regulation for drought tolerance in plants. In *Drought Stress Tolerance in Plants, Vol 1* (pp. 45-75): Springer.

Campbell, M. S., Holt, C., Moore, B., & Yandell, M. (2014). Genome annotation and curation using MAKER and MAKER-P. *Curr Protoc Bioinformatics, 48*(1), 4.11. 11-14.11. 39.

Casati, P., & Walbot, V. (2004). Rapid transcriptome responses of maize (*Zea mays*) to UV-B in irradiated and shielded tissues. *Genome Biol, 5*(3), R16. doi:10.1186/gb-2004-5-3-r16

Castelain, M., Le Hir, R., & Bellini, C. (2012). The non-DNA-binding bHLH transcription factor PRE3/bHLH135/ATBS1/TMO7 is involved in the regulation of light signaling pathway in Arabidopsis. *Physiol Plant, 145*(3), 450-460.

Castells, E., Molinier, J., Drevensek, S., Genschik, P., Barneche, F., & Bowler, C. (2010). det1-1-induced UV-C hyposensitivity through UVR3 and PHR1 photolyase gene overexpression. *Plant J, 63*(3), 392-404.

Chang, Y. C., Huang, C. N., Lin, C. H., Chang, H. C., & Wu, C. C. (2010). Mapping protein cysteine sulfonic acid modifications with specific enrichment and mass spectrometry: an integrated approach to explore the cysteine oxidation. *Proteomics, 10*(16), 2961-2971. doi:10.1002/pmic.200900850

Chapin, F. S., Walter, C. H., & Clarkson, D. T. (1988). Growth response of barley and tomato to nitrogen stress and its control by abscisic acid, water relations and photosynthesis. *Planta, 173*(3), 352-366.

Chen, L. Y., Morales-Briones, D. F., Passow, C. N., & Yang, Y. (2019). Performance of gene expression analyses using *de novo* assembled transcripts in polyploid species. *Bioinformatics, 35*(21), 4314-4320.

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 J, 89*(4), 789-804.

Chico, J. M., Chini, A., Fonseca, S., & Solano, R. (2008). JAZ repressors set the rhythm in jasmonate signaling. *Curr Opin Plant Biol, 11*(5), 486-494.

Chinnusamy, V., Zhu, J., & Zhu, J. K. (2007). Cold stress regulation of gene expression in plants. *Trends Plant Sci, 12*(10), 444-451. doi:10.1016/j.tplants.2007.07.002

Chopra, R., Burow, G., Farmer, A., Mudge, J., Simpson, C. E., & Burow, M. D. (2014). Comparisons of *de novo* transcriptome assemblers in diploid and polyploid species using peanut (Arachis spp.) RNA-Seq data. *PLoS One, 9*(12), e115055.

Ciardi, J. A., Deikman, J., & Orzolek, M. D. (1997). Increased ethylene synthesis enhances chilling tolerance in tomato. *Physiol Plant, 101*(2), 333-340.

Cloix, C., & Jenkins, G. I. (2008). Interaction of the Arabidopsis UV-B-specific signaling component UVR8 with chromatin. *Mol Plant, 1*(1), 118-128. doi:10.1093/mp/ssm012

Cocetta, G., Mishra, S., Raffaelli, A., & Ferrante, A. (2018). Effect of heat root stress and high salinity on glucosinolates metabolism in wild rocket. *J Plant Physiol, 231*, 261-270.

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nat Rev Genet, 6*(11), 836-846.

Cox, M. P., Peterson, D. A., & Biggs, P. J. (2010). SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC bioinformatics, 11*(1), 485.

Cuellar-Ortiz, S. M., De La Paz Arrieta-Montiel, M., Acosta-Gallegos, J., & Covarrubias, A. A. (2008). Relationship between carbohydrate partitioning and drought resistance in common bean. *Plant Cell Environ, 31*(10), 1399-1409.

Cui, S., Huang, F., Wang, J., Ma, X., Cheng, Y., & Liu, J. (2005). A proteomic analysis of cold stress responses in rice seedlings. *Proteomics, 5*(12), 3162-3172.

Dai, Q., Yan, B., Huang, S., Liu, X., Peng, S., Miranda, M. L. L., . . . Olszyk, D. M. (1997). Response of oxidative stress defense systems in rice (*Oryza sativa*) leaves with supplemental UV-B radiation. *Physiol Plant, 101*(2), 301-308.

Darkó, É., Ambrus, H., Fodor, J., Király, Z., & Barnabás, B. (2009). Enhanced tolerance to oxidative stress with elevated antioxidant capacity in doubled haploid maize derived from microspores exposed to paraquat. *Crop Sci, 49*(2), 628-636.

Dauphinee, A. N., Fletcher, J. I., Denbigh, G. L., Lacroix, C. R., & Gunawardena, A. H. (2017). Remodelling of lace plant leaves: antioxidants and ROS are key regulators of programmed cell death. *Planta, 246*(1), 133-147.

Day, T. A. (1993). Relating UV-B radiation screening effectiveness of foliage to absorbing-compound concentration and anatomical characteristics in a diverse group of plants. *Oecologia, 95*(4), 542-550.

Deak, K. I., & Malamy, J. (2005). Osmotic regulation of root system architecture. *The Plant J, 43*(1), 17-28.

De Zelicourt, A., Colcombet, J., & Hirt, H. (2016). The role of MAPK modules and ABA during abiotic stress signaling. *Trends Plant Sci, 21*(8), 677-685.

Del Carmen Martínez-Ballesta, M., Moreno, D. A., & Carvajal, M. (2013). The physiological importance of glucosinolates on plant response to abiotic stress in Brassica. *Int J Mol Sci, 14*(6), 11607-11625.

Delcher, A. L., Salzberg, S. L., & Phillippy, A. M. (2003). Using MUMmer to identify similar regions in large sequence sets. *Curr Protoc Bioinformatics*(1), 10.13. 11-10.13. 18.

Deng, B., Du, W., Liu, C., Sun, W., Tian, S., & Dong, H. (2012). Antioxidant response to drought, cold and nutrient stress in two ploidy levels of tobacco plants: low resource requirement confers polytolerance in polyploids? *Plant Growth Regul, 66*(1), 37-47.

Devos, K. M., Brown, J. K., & Bennetzen, J. L. (2002). Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis. *Genome Res, 12*(7), 1075-1079. *https://creativecommons.org/licenses/by-nc/4.0/*

Devoto, A., & Turner, J. G. (2005). Jasmonate-regulated Arabidopsis stress signalling network. *Physiol Plant, 123*(2), 161-172.

Di Ferdinando, M., Brunetti, C., Fini, A., & Tattini, M. (2012). Flavonoids as antioxidants in plants under abiotic stresses. In *Abiotic stress responses in plants* (pp. 159-179): Springer.

Dorman, C. J., & Dorman, M. J. (2016). DNA supercoiling is a fundamental regulatory principle in the control of bacterial gene expression. *Biophys Rev, 8*(1), 89-100.

Dotto, M., & Casati, P. (2017). Developmental reprogramming by UV-B radiation in plants. *Plant Sci, 264*, 96-101.

Duan, J., Xia, C., Zhao, G., Jia, J., & Kong, X. (2012). Optimizing *de novo* common wheat transcriptome assembly using short-read RNA-Seq data. *BMC genomics, 13*(1), 392.

Edger, P. P., Heidel-Fischer, H. M., Bekaert, M., Rota, J., Glockner, G., Platts, A. E., . . . Wheat, C. W. (2015). The butterfly plant arms-race escalated by gene and genome duplications. *Proc Natl Acad Sci U S A, 112*(27), 8362-8366. doi:10.1073/pnas.1503926112

Ehrhardt-Brocardo, N. C. M., Coelho, C. M. M., Souza, C. A., & Mathias, V. (2019). Callose accumulation in roots of soybean seedlings under water deficit. *Theor Exp Plant Physiol, 31*(4), 475-481.

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol, 16*(1), 157.

Engelen-Eigles, G., Holden, G., Cohen, J. D., & Gardner, G. (2006). The effect of temperature, photoperiod, and light quality on gluconasturtiin concentration in watercress (*Nasturtium officinale* R. Br.). *J Agric Food Chem, 54*(2), 328-334.

Fierro, A. C., Leroux, O., De Coninck, B., Cammue, B. P., Marchal, K., Prinsen, E., . . . Vandenbussche, F. (2015). Ultraviolet-B radiation stimulates downward leaf curling in *Arabidopsis thaliana*. *Plant Physiol Biochem, 93*, 9-17. doi:10.1016/j.plaphy.2014.12.012

Fobert, P. R., & Després, C. (2005). Redox control of systemic acquired resistance. *Curr Opin Plant Biol, 8*(4), 378-382.

Garnock-Jones, P. (1991). Gender dimorphism in *Cheesemania wallii* (Brassicaceae). *N Z J Bot, 29*(1), 87-90.

Gao, W., Feng, Z., Bai, Q., He, J., & Wang, Y. (2019). Melatonin-mediated regulation of growth and antioxidant capacity in salt-tolerant naked oat under salt stress. *Int J Mol Sci, 20*(5), 1176.

Gawel, N., & Jarret, R. (1991). A modified CTAB DNA extraction procedure forMusa andIpomoea. *Plant Mol Biol Rep, 9*(3), 262-266.

Gegas, V. C., Wargent, J. J., Pesquet, E., Granqvist, E., Paul, N. D., & Doonan, J. H. (2014). Endopolyploidy as a potential alternative adaptive strategy for Arabidopsis leaf size variation in response to UV-B. *J Exp Bot, 65*(10), 2757-2766.

Gerhardt, K. E., Wilson, M. I., & Greenberg, B. M. (2005). Ultraviolet wavelength dependence of photomorphological and photosynthetic responses in *Brassica napus* and *Arabidopsis thaliana*. *Photochem Photobiol, 81*(5), 1061-1068.

Gershenzon, J., & Dudareva, N. (2007). The function of terpene natural products in the natural world. *Nat Chem Biol, 3*(7), 408-414.

Gil, M., Pontin, M., Berli, F., Bottini, R., & Piccoli, P. (2012). Metabolism of terpenes in the response of grape (*Vitis vinifera* L.) leaf tissues to UV-B radiation. *Phytochemistry, 77*, 89-98. doi:10.1016/j.phytochem.2011.12.011

Gilbert, D. (2013). EvidentialGene: tr2aacds, mRNA transcript assembly software. URL *http://arthropods. eugenes. org/EvidentialGene*.

Gill, S. S., Anjum, N. A., Gill, R., Jha, M., & Tuteja, N. (2015). DNA damage and repair in plants under ultraviolet and ionizing radiations. *The Scientific World Journal, 2015*.

Goldthwaite, J. J. (1987). Hormones in plant senescence. In *Plant hormones and their role in plant growth and development* (pp. 553-573): Springer.

Golldack, D., Li, C., Mohan, H., & Probst, N. (2014). Tolerance to drought and salt stress in plants: unraveling the signaling networks. *Front Plant Sci, 5*, 151.

Gómez, M. S., Falcone Ferreyra, M. L., Sheridan, M. L., & Casati, P. (2019). Arabidopsis E2Fc is required for the DNA damage response under UV-B radiation epistatically over the micro RNA 396 and independently of E2Fe. *Plant J, 97*(4), 749-764.

González Besteiro, M. A., Bartels, S., Albert, A., & Ulm, R. (2011). Arabidopsis MAP kinase phosphatase 1 and its target MAP kinases 3 and 6 antagonistically determine UV-B stress tolerance, independent of the UVR8 photoreceptor pathway. *Plant J, 68*(4), 727-737.

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., . . . Putnam, N. (2011). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res, 40*(D1), D1178-D1186.

Gous, P. W., Hasjim, J., Franckowiak, J., Fox, G. P., & Gilbert, R. G. (2013). Barley genotype expressing "stay-green"-like characteristics maintains starch quality of the grain during water stress condition. *Journal of cereal science, 58*(3), 414-419.

Goyal, A. (2007). Osmoregulation in Dunaliella, Part II: Photosynthesis and starch contribute carbon for glycerol synthesis during a salt stress in *Dunaliella tertiolecta*. *Plant Physiol Biochem, 45*(9), 705-710. doi:10.1016/j.plaphy.2007.05.009

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., ... & Chen, Z. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol, 29*(7), 644-652.

Greenberg, B. M., Wilson, M. I., Huang, X. D., Duxbury, C. L., Gerhardt, K. E., & Gensemer, R. W. (1997). The effects of ultraviolet-B radiation on higher plants. *Plants for environmental studies*, 1-35.

Gross, K., & Schiestl, F. P. (2015). Are tetraploids more successful? Floral signals, reproductive success and floral isolation in mixed-ploidy populations of a terrestrial orchid. *Ann Bot, 115*(2), 263-273. doi:10.1093/aob/mcu244

Gruenheit, N., Deusch, O., Esser, C., Becker, M., Voelckel, C., & Lockhart, P. (2012). Cutoffs and k-mers: implications from a transcriptome study in allopolyploid plants. *BMC Genomics, 13*, 92. doi:10.1186/1471-2164-13-92

Guo, D., & Qin, G. (2016). EXB1/WRKY71 transcription factor regulates both shoot branching and responses to abiotic stresses. *Plant Signal Behav, 11*(3), e1150404.

Guo, H., & Ecker, J. R. (2004). The ethylene signaling pathway: new insights. *Curr Opin Plant Biol, 7*(1), 40-49.

Guo, X., Liu, D., & Chong, K. (2018). Cold signaling in plants: Insights into mechanisms and regulation. *J Integr Plant Biol, 60*(9), 745-756.

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics, 29*(8), 1072-1075.

Haas, B. (2007). TransposonPSI: an application of PSI-blast to mine (Retro-) transposon ORF homologies. URL *http://transposonpsi.sorceforge.net*

Haas, B. J., Delcher, A. L., Mount, S. M., Wortman, J. R., Smith Jr, R. K., Hannick, L. I., ... & Salzberg, S. L. (2003). Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res, 31*(19), 5654-5666.

Hahn, A., & Harter, K. (2009). Mitogen-activated protein kinase cascades and ethylene: signaling, biosynthesis, or both? *Plant Physiol, 149*(3), 1207-1210.

Han, H., Li, Y., & Zhou, S. (2008). Overexpression of phytoene synthase gene from *Salicornia europaea* alters response to reactive oxygen species under salt stress in transgenic Arabidopsis. *Biotechnol Lett, 30*(8), 1501-1507.

Han, Y., & Wessler, S. R. (2010). MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Res, 38*(22), e199-e199.

Hasanuzzaman, M., Nahar, K., Alam, M., Roychowdhury, R., & Fujita, M. (2013). Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. *Int J Mol Sci, 14*(5), 9643-9684.

Hasanuzzaman, M., Nahar, K., & Fujita, M. (2013). Plant response to salt stress and role of exogenous protectants to mitigate salt-induced damages. In *Ecophysiology and responses of plants under salt stress* (pp. 25-87): Springer.

Hao, G. Y., Lucero, M. E., Sanderson, S. C., Zacharias, E. H., & Holbrook, N. M. (2013). Polyploidy enhances the occupation of heterogeneous environments through hydraulic related trade-offs in *Atriplex canescens* (Chenopodiaceae). *New Phytol, 197*(3), 970-978.

Hasegawa, P. M., Bressan, R. A., Zhu, J. K., & Bohnert, H. J. (2000). Plant cellular and molecular responses to high salinity. *Annu Rev Plant Biol, 51*(1), 463-499.

Havaux, M., Ksas, B., Szewczyk, A., Rumeau, D., Franck, F., Caffarri, S., & Triantaphylidès, C. (2009). Vitamin B6 deficient plants display increased sensitivity to high light and photo-oxidative stress. *BMC Plant Biol, 9*(1), 130.

Hectors, K., Prinsen, E., De Coen, W., Jansen, M. A., & Guisez, Y. (2007). *Arabidopsis thaliana* plants acclimated to low dose rates of ultraviolet B radiation show specific changes in morphology and gene expression in the absence of stress symptoms. *New Phytol, 175*(2), 255-270.

Hectors, K., Jacques, E., Prinsen, E., Guisez, Y., Verbelen, J. P., Jansen, M. A., & Vissenberg, K. (2010). UV radiation reduces epidermal cell expansion in leaves of *Arabidopsis thaliana*. *J Exp Bot, 61*(15), 4339-4349. doi:10.1093/jxb/erq235

Heenan, P., Mitchell, A., & Koch, M. (2002). Molecular systematics of the New Zealand Pachycladon (Brassicaceae) complex: generic circumscription and relationships to Arabidopsis sens. lat. and Arabis sens. lat. *N Z J Bot, 40*(4), 543-562.

Heenan, P., & Garnock-Jones, P. (1999). A new species combination in Cheesemania (Brassicaceae) from New Zealand. *N Z J Bot, 37*(2), 235-241.

Heenan, P., & Mitchell, A. (2003). Phylogeny, biogeography and adaptive radiation of Pachycladon (Brassicaceae) in the mountains of South Island, New Zealand. *J Biogeogr, 30*(11), 1737-1749.

Heenan, P. B., Goeke, D. F., Houliston, G. J., & Lysak, M. A. (2012). Phylogenetic analyses of ITS and rbcL DNA sequences for sixteen genera of Australian and New Zealand Brassicaceae result in the expansion of the tribe Microlepidieae. *Taxon, 61*(5), 970-979.

Heenan, P. J. (2002). Melanoma: margins for error. *ANZ J Surg, 72*(4), 300-303. doi:10.1046/j.1445-2197.2002.02375.x

Heijde, M., & Ulm, R. (2012). UV-B photoreceptor-mediated signalling in plants. *Trends Plant Sci, 17*(4), 230-237.

Heine, G. F., Hernandez, J. M., & Grotewold, E. (2004). Two cysteines in plant R2R3 MYB domains participate in REDOX-dependent DNA binding. *J Biol Chem, 279*(36), 37878-37885.

Hideg, É., Jansen, M. A., & Strid, Å. (2013). UV-B exposure, ROS, and stress: inseparable companions or loosely linked associates? *Trends Plant Sci, 18*(2), 107-115.

Holder, N. (1979). Positional information and pattern formation in plant morphogenesis and a mechanism for the involvement of plant hormones. *J Theor Biol, 77*(2), 195-212.

Horváth, E., Szalai, G., & Janda, T. (2007). Induction of abiotic stress tolerance by salicylic acid signaling. *J Plant Growth Regul, 26*(3), 290-300.

Hrazdina, G. (1992). Biosynthesis of flavonoids. In *Plant polyphenols* (pp. 61-72): Springer.

Hua, D., Wang, C., He, J., Liao, H., Duan, Y., Zhu, Z., . . . Gong, Z. (2012). A plasma membrane receptor kinase, GHR1, mediates abscisic acid-and hydrogen peroxide-regulated stomatal movement in Arabidopsis. *Plant Cell, 24*(6), 2546-2561.

Huang, C.-H., Sun, R., Hu, Y., Zeng, L., Zhang, N., Cai, L., . . . Edger, P. P. (2015). Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. *Mol Biol Evol, 33*(2), 394-412.

Huang, X., & Madan, A. (1999). CAP3: A DNA sequence assembly program. *Genome Res, 9*(9), 868-877.

Huminiecki, L., & Conant, G. C. (2012). Polyploidy and the evolution of complex traits. *Int J Evol Biol, 2012*, 292068. doi:10.1155/2012/292068

Hund, A., Fracheboud, Y., Soldati, A., & Stamp, P. (2008). Cold tolerance of maize seedlings as determined by root morphology and photosynthetic traits. *Eur J Agron, 28*(3), 178-185.

Huo, C., Zhang, B., Wang, H., Wang, F., Liu, M., Gao, Y., . . . Tang, W. (2016). Comparative study of early cold-regulated proteins by two-dimensional difference gel electrophoresis reveals a key role for phospholipase Dα1 in mediating cold acclimation signaling pathway in rice. *Mol Cell Proteomics, 15*(4), 1397-1411.

Ignat, T., Schmilovitch, Z., Feföldi, J., Bernstein, N., Steiner, B., Egozi, H., & Hoffman, A. (2013). Nonlinear methods for estimation of maturity stage, total chlorophyll, and carotenoid content in intact bell peppers. *Biosyst Eng, 114*(4), 414-425.

Islam, S., Griffiths, C. A., Blomstedt, C. K., Le, T. N., Gaff, D. F., Hamill, J. D., & Neale, A. D. (2013). Increased biomass, seed yield and stress tolerance is conferred in Arabidopsis by a novel enzyme from the resurrection grass *Sporobolus stapfianus* that glycosylates the strigolactone analogue GR24. *PLoS One, 8*(11), e80035.

Jammes, F., Song, C., Shin, D., Munemasa, S., Takeda, K., Gu, D., . . . Sritubtim, S. (2009). MAP kinases MPK9 and MPK12 are preferentially expressed in guard cells and positively regulate ROS-mediated ABA signaling. *PNAS, 106*(48), 20520-20525.

Jan, N., Majeed, U., Andrabi, K. I., & John, R. (2018). Cold stress modulates osmolytes and antioxidant system in *Calendula officinalis*. *Acta Physiol Plant, 40*(4), 73.

Jansen, M. A., & Urban, O. (2019). Plant Responses to UV-B Radiation. *eLS*, 1-11.

Jayakannan, M., Bose, J., Babourina, O., Rengel, Z., & Shabala, S. (2015). Salicylic acid in plant salinity stress signalling and tolerance. *Plant Growth Regul, 76*(1), 25-40.

Jenkins, G. I. (2009). Signal transduction in responses to UV-B radiation. *Annu Rev Plant Biol, 60*, 407-431. *http://arjournals.annualreviews.org/loi/arplant.1*

Jenkins, G. I. (2014). The UV-B photoreceptor UVR8: from structure to physiology. *Plant Cell, 26*(1), 21-37.

Jeong, J., Jung, Y., Na, S., Jeong, J., Lee, E., Kim, M. S., . . . Lee, K. J. (2011). Novel oxidative modifications in redox-active cysteine residues. *Mol Cell Proteomics, 10*(3), M110 000513. doi:10.1074/mcp.M110.000513

Jiang, L., Wang, Y., Björn, L. O., & Li, S. (2009). Arabidopsis radical-induced cell death1 is involved in UV-B signaling. *Photochem Photobiol Sci, 8*(6), 838-846.

Jiang, L., Wang, Y., Bjorn, L. O., & Li, S. (2011). UV-B-induced DNA damage mediates expression changes of cell cycle regulatory genes in Arabidopsis root tips. *Planta, 233*(4), 831-841. doi:10.1007/s00425-010-1340-5

Jin, J., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J., & Gao, G. (2016). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res*, gkw982.

Johnson, C., Boden, E., & Arias, J. (2003). Salicylic acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in Arabidopsis. *Plant Cell, 15*(8), 1846-1858.

Johnson, S. M., Lim, F. L., Finkler, A., Fromm, H., Slabas, A. R., & Knight, M. R. (2014). Transcriptomic analysis of Sorghum bicolor responding to combined heat and drought stress. *BMC Genomics, 15*(1), 456.

Joly, S., Heenan, P. B., & Lockhart, P. J. (2009). A Pleistocene inter-tribal allopolyploidization event precedes the species radiation of Pachycladon (Brassicaceae) in New Zealand. *Mol Phylogenet Evol, 51*(2), 365-372.

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., . . . Nuka, G. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics, 30*(9), 1236-1240.

Jump, A. S., Marchant, R., & Peñuelas, J. (2009). Environmental change and the option value of genetic diversity. *Trends Plant Sci, 14*(1), 51-58.

Kajitani, R., Toshimoto, K., Noguchi, H., Toyoda, A., Ogura, Y., Okuno, M., . . . Maruyama, H. (2014). Efficient *de novo* assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res*, gr. 170720.170113.

Kalbina, I., & Strid, Å. (2006). Supplementary ultraviolet-B radiation reveals differences in stress responses between *Arabidopsis thaliana* ecotypes. *Plant Cell Environ, 29*(5), 754-763.

Kant, S., Kafkafi, U., Pasricha, N., & Bansal, S. (2002). Potassium and abiotic stresses in plants. *Potassium for sustainable crop production. Potash Institute of India, Gurgaon, 233*, 251.

Kapitonov, V. V., & Jurka, J. (2008). A universal classification of eukaryotic transposable elements implemented in Repbase. *Nat Rev Genet, 9*(5), 411.

Karakas, B., Ozias-Akins, P., Stushnoff, C., Suefferheld, M., & Rieger, M. (1997). Salinity and drought tolerance of mannitol-accumulating transgenic tobacco. *Plant Cell Environ, 20*(5), 609-616.

Kataria, S., & Guruprasad, K. (2018). Interaction of cytokinins with UV-B (280-315nm) on the expansion growth of cucumber cotyledons. *Horticult Int J, 2*(2), 45-53.

Kaur, N., & Gupta, A. K. (2005). Signal transduction pathways under abiotic stresses in plants. *Curr Sci*, 1771-1780.

Kazan, K. (2006). Negative regulation of defence and stress genes by EAR-motif-containing repressors. *Trends Plant Sci, 11*(3), 109-112.

Khan, M. I. R., Fatma, M., Per, T. S., Anjum, N. A., & Khan, N. A. (2015). Salicylic acid-induced abiotic stress tolerance and underlying mechanisms in plants. *Front Plant Sci, 6*, 462.

Khan, M. I. R., Syeed, S., Nazar, R., & Anjum, N. A. (2012). An insight into the role of salicylic acid and jasmonic acid in salt stress tolerance. In *Phytohormones and Abiotic Stress Tolerance in Plants* (pp. 277-300): Springer.

Kilian, J., Whitehead, D., Horak, J., Wanke, D., Weinl, S., Batistic, O., . . . Harter, K. (2007). The AtGenExpress global stress expression data set: protocols, evaluation

and model data analysis of UV-B light, drought and cold stress responses. *Plant J, 50*(2), 347-363. doi:10.1111/j.1365-313X.2007.03052.x

Kim, B. C., Tennessen, D. J., & Last, R. L. (1998). UV-B-induced photomorphogenesis in *Arabidopsis thaliana*. *Plant J, 15*(5), 667-674. doi:10.1046/j.1365-313x.1998.00246.x

Klein, P., Seidel, T., Stöcker, B., & Dietz, K.-J. (2012). The membrane-tethered transcription factor ANAC089 serves as redox-dependent suppressor of stromal ascorbate peroxidase gene expression. *Front Plant Sci, 3*, 247.

Kliebenstein, D. J., Lim, J. E., Landry, L. G., & Last, R. L. (2002). Arabidopsis UVR8 regulates ultraviolet-B signal transduction and tolerance and contains sequence similarity to human regulator of chromatin condensation 1. *Plant Physiol, 130*(1), 234-243.

Klingler, J. P., Batelli, G., & Zhu, J. K. (2010). ABA receptors: the START of a new paradigm in phytohormone signalling. *J Exp Bot, 61*(12), 3199-3210.

Kopylova, E., Noé, L., & Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics, 28*(24), 3211-3217.

Kravets, E. A., Zelena, L. B., Zabara, E. P., & Blume, Y. B. (2012). Adaptation strategy of barley plants to UV-B radiation. *Emir J Food Agri, 24*(6), 632.

Kreps, J. A., Wu, Y., Chang, H. S., Zhu, T., Wang, X., & Harper, J. F. (2002). Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. *Plant Physiol, 130*(4), 2129-2141.

Krueger, F., & Galore, T. (2015). A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. URL *https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/*

Kumar, S. V., & Wigge, P. A. (2010). H2A. Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. *Cell, 140*(1), 136-147.

Lan, W. Z., Wang, W., Wang, S. M., Li, L. G., Buchanan, B. B., Lin, H. X., . . . Luan, S. (2010). A rice high-affinity potassium transporter (HKT) conceals a calcium-permeable cation channel. *PNAS, 107*(15), 7089-7094.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods, 9*(4), 357.

Lang-Mladek, C., Xie, L., Nigam, N., Chumak, N., Binkert, M., Neubert, S., & Hauser, M. T. (2012). UV-B signaling pathways and fluence rate dependent transcriptional regulation of ARIADNE12. *Physiol Plant, 145*(4), 527-539.

Lavola, A., Julkunen-Tiitto, R., Aphalo, P., Rosa, T. D. L., & Lehto, T. (2010). The Effect of u.v.-B Radiation on u.v.-Absorbing Secondary Metabolites in Birch Seedlings Grown under Simulated Forest Soil Conditions. *New Phytol, 137*(4), 617-621.

Lee, H. S., & Chen, Z. J. (2001). Protein-coding genes are epigenetically regulated in Arabidopsis polyploids. *PNAS, 98*(12), 6753-6758.

Lee, J., Lee, W., & Kwon, S. W. (2015). A quantitative shotgun proteomics analysis of germinated rice embryos and coleoptiles under low-temperature conditions. *Proteome Sci, 13*(1), 1-11.

Lee, J. H., Kim, D. M., Lee, J. H., Kim, J., Bang, J. W., Kim, W. T., & Pai, H. S. (2005). Functional characterization of NtCEF1, an AP2/EREBP-type transcriptional activator highly expressed in tobacco callus. *Planta, 222*(2), 211-224.

Lehman, A., Black, R., & Ecker, J. R. (1996). *HOOKLESS1*, an ethylene response gene, is required for differential cell elongation in the Arabidopsis hypocotyl. *Cell, 85*(2), 183-194.

Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res, 44*(W1), W242-W245.

Li, A., Zhou, M., Wei, D., Chen, H., You, C., & Lin, J. (2017). Transcriptome Profiling Reveals the Negative Regulation of Multiple Plant Hormone Signaling Pathways Elicited by Overexpression of C-Repeat Binding Factors. *Front Plant Sci, 8*, 1647. doi:10.3389/fpls.2017.01647

Li, D., Li, B., Ma, Y., Sun, X., Lin, Y., & Meng, X. (2017). Polyphenols, anthocyanins, and flavonoids contents and the antioxidant capacity of various cultivars of highbush and half-high blueberries. *J Food Compost Anal, 62*, 84-93.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics, 25*(16), 2078-2079.

Li, J., Witten, D. M., Johnstone, I. M., & Tibshirani, R. (2012). Normalization, testing, and false discovery rate estimation for RNA-sequencing data. *Biostatistics*, 13(3), 523-538.

Li, N., Teranishi, M., Yamaguchi, H., Matsushita, T., Watahiki, M. K., Tsuge, T., . . . Hidema, J. (2015). UV-B-induced CPD photolyase gene expression is regulated by UVR8-dependent and-independent pathways in Arabidopsis. *Plant Cell Physiol, 56*(10), 2014-2023.

Li, S. J., Bai, Y. C., Li, C. L., Yao, H P., Chen, H., Zhao, H. X., & Wu, Q. (2015). Anthocyanins accumulate in tartary buckwheat (*Fagopyrum tataricum*) sprout in response to cold stress. *Acta Physiol Plant, 37*(8), 159.

Li, Y., Wang, X., Ban, Q., Zhu, X., Jiang, C., Wei, C., & Bennetzen, J. L. (2019). Comparative transcriptomic analysis reveals gene expression associated with cold adaptation in the tea plant Camellia sinensis. *BMC Genomics, 20*(1), 624.

Liakoura, V., Stefanou, M., Manetas, Y., Cholevas, C., & Karabourniotis, G. (1997). Trichome density and its UV-B protective potential are affected by shading and leaf position on the canopy. *Environ Exp Bot, 38*(3), 223-229.

Liu, M., Adelman, Z. N., Myles, K. M., & Zhang, L. (2014). A transcriptome post-scaffolding method for assembling high quality contigs. *PLoS Comput Biol, 2014*.

Liu, S. C., Jin, J. Q., Ma, J. Q., Yao, M. Z., Ma, C. L., Li, C. F., . . . Chen, L. (2016). Transcriptomic analysis of tea plant responding to drought stress and recovery. *PLoS One, 11*(1), e0147306.

Loreti, E., Poggi, A., Novi, G., Alpi, A., & Perata, P. (2005). A genome-wide analysis of the effects of sucrose on gene expression in Arabidopsis seedlings under anoxia. *Plant Physiol, 137*(3), 1130-1138.

Luan, S., Lan, W., & Lee, S. C. (2009). Potassium nutrition, sodium toxicity, and calcium signaling: connections through the CBL–CIPK network. *Curr Opin Plant Biol, 12*(3), 339-346.

Luo, M. C., Deal, K. R., Akhunov, E. D., Akhunova, A. R., Anderson, O. D., Anderson, J. A., . . . Dvorak, J. (2009). Genome comparisons reveal a dominant mechanism of chromosome number reduction in grasses and accelerated genome evolution in Triticeae. *Proc Natl Acad Sci U S A, 106*(37), 15780-15785. doi:10.1073/pnas.0908195106

Luo, Y., Dong, X., Yu, T., Shi, X., Li, Z., Yang, W., . . . Karrenberg, S. (2015). A single nucleotide deletion in gibberellin20-oxidase1 causes alpine dwarfism in Arabidopsis. *Plant Physiol, 168*(3), 930-937.

Ma, Y., Dai, X., Xu, Y., Luo, W., Zheng, X., Zeng, D., . . . Zhang, D. (2015). COLD1 confers chilling tolerance in rice. *Cell, 160*(6), 1209-1221.

Macková, J., Vašková, M., Macek, P., Hronková, M., Schreiber, L., & Šantrůček, J. (2013). Plant response to drought stress simulated by ABA application: changes in chemical composition of cuticular waxes. *Environ Exp Bot, 86*, 70-75.

Maldonado, A. M., Doerner, P., Dixon, R. A., Lamb, C. J., & Cameron, R. K. (2002). A putative lipid transfer protein involved in systemic resistance signalling in Arabidopsis. *Nature, 419*(6905), 399-403.

Mandakova, T., Heenan, P. B., & Lysak, M. A. (2010). Island species radiation and karyotypic stasis in Pachycladon allopolyploids. *BMC Evol Biol, 10*, 367. doi:10.1186/1471-2148-10-367

Mao, G., Meng, X., Liu, Y., Zheng, Z., Chen, Z., & Zhang, S. (2011). Phosphorylation of a WRKY transcription factor by two pathogen-responsive MAPKs drives phytoalexin biosynthesis in Arabidopsis. *Plant Cell, 23*(4), 1639-1653.

Marcais, G., & Kingsford, C. (2012). Jellyfish: A fast k-mer counter. *Tutorialis e Manuais*, 1, 1-8.

Martin, C., & Gerats, T. (1993). Control of pigment biosynthesis genes during petal development. *Plant Cell, 5*(10), 1253.

Martins, C. P., Neves, D. M., Cidade, L. C., Mendes, A. F., Silva, D. C., Almeida, A. A. F., . . . Costa, M. G. (2017). Expression of the citrus CsTIP2; 1 gene improves tobacco plant growth, antioxidant capacity and physiological adaptation under stress conditions. *Planta, 245*(5), 951-963.

Martin, J. A., & Wang, Z. (2011). Next-generation transcriptome assembly. *Nat Rev Genet, 12*(10), 671-682. doi:10.1038/nrg3068

Masterson, J. (1994). Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science, 264*(5157), 421-424.

Mershon, J. P., Becker, M., & Bickford, C. P. (2015). Linkage between trichome morphology and leaf optical properties in New Zealand alpine Pachycladon (Brassicaceae). *N Z J Bot, 53*(3), 175-182.

Mestiri, I., Chague, V., Tanguy, A. M., Huneau, C., Huteau, V., Belcram, H., . . . Jahier, J. (2010). Newly synthesized wheat allohexaploids display progenitor-dependent meiotic stability and aneuploidy but structural genomic additivity. *New Phytol, 186*(1), 86-101. doi:10.1111/j.1469-8137.2010.03186.x

McBreen, K., & Heenan, P. (2006). Phylogenetic relationships of Pachycladon (Brassicaceae) species based on three nuclear and two chloroplast DNA markers. *N Z J Bot, 44*(4), 377-386.

McCarthy, E. W., Arnold, S. E., Chittka, L., Le Comber, S. C., Verity, R., Dodsworth, S., . . . Leitch, A. R. (2015). The effect of polyploidy and hybridization on the evolution of floral colour in Nicotiana (Solanaceae). *Ann Bot, 115*(7), 1117-1131. doi:10.1093/aob/mcv048

McKenzie, R., Connor, B., & Bodeker, G. (1999). Increased summertime UV radiation in New Zealand in response to ozone loss. *Science, 285*(5434), 1709-1711.

McKenzie, R., Liley, B., Kotkamp, M., Shiona, H., & Lopez, L. (2014). *Long term changes in UV in New Zealand due to ozone depletion and other causes.* Paper presented at the NIWA UV Workshop, Auckland, New Zealand.

Min, B., Gu, L., McClung, A. M., Bergman, C. J., & Chen, M. H. (2012). Free and bound total phenolic concentrations, antioxidant capacities, and profiles of proanthocyanidins and anthocyanins in whole grain rice (*Oryza sativa* L.) of different bran colours. *Food Chem, 133*(3), 715-722.

Minoche, A. E., Dohm, J. C., & Himmelbauer, H. (2011). Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biol, 12*(11), R112. doi:10.1186/gb-2011-12-11-r112

Mirzaei, M., Pascovici, D., Keighley, T., George, I., Voelckel, C., Heenan, P. B., & Haynes, P. A. (2011). Shotgun proteomic profiling of five species of New Zealand Pachycladon. *Proteomics, 11*(1), 166-171.

Misra, A., Biswal, A., & Misra, M. (2002). Physiological, biochemical and molecular aspects of water stress responses in plants, and the bio-technological applications. *Proc Natl Acad Sci India Sect B Biol Sci, 72*(2), 115-134.

Mittler, R., Vanderauwera, S., Suzuki, N., Miller, G., Tognetti, V. B., Vandepoele, K., . . . Van Breusegem, F. (2011). ROS signaling: the new wave? *Trends Plant Sci, 16*(6), 300-309.

Miura, K., & Ohta, M. (2010). SIZ1, a small ubiquitin-related modifier ligase, controls cold signaling through regulation of salicylic acid accumulation. *J Plant Physiol, 167*(7), 555-560.

Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., ... & Houstis, N. (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet, 34*(3), 267-273.

Morgan, P. W., & Drew, M. C. (1997). Ethylene and plant responses to stress. *Physiol Plant, 100*(3), 620-630.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., & Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res, 35*(suppl_2), W182-W185.

Morsy, M. R., Jouve, L., Hausman, J. F., Hoffmann, L., & Stewart, J. M. (2007). Alteration of oxidative and carbohydrate metabolism under abiotic stress in two rice (*Oryza sativa* L.) genotypes contrasting in chilling tolerance. *J Plant Physiol, 164*(2), 157-167.

Muñoz, M. J., Moreno, N. N., Giono, L. E., Botto, A. E. C., Dujardin, G., Bastianello, G., . . . Blencowe, B. J. (2017). Major roles for pyrimidine dimers, nucleotide excision repair, and ATR in the alternative splicing response to UV radiation. *Cell Rep, 18*(12), 2868-2879.

Münzbergová, Z., Skuhrovec, J., & Maršík, P. (2015). Large differences in the composition of herbivore communities and seed damage in diploid and autotetraploid plant species. *Biol J Linn Soc, 115*(2), 270-287.

Nagao, M., Minami, A., Arakawa, K., Fujikawa, S., & Takezawa, D. (2005). Rapid degradation of starch in chloroplasts and concomitant accumulation of soluble sugars associated with ABA-induced freezing tolerance in the moss Physcomitrella patens. *J Plant Physiol, 162*(2), 169-180.

Nakasugi, K., Crowhurst, R., Bally, J., & Waterhouse, P. (2014). Combining transcriptome assemblies from multiple *de novo* assemblers in the allo-tetraploid plant Nicotiana benthamiana. *PLoS One, 9*(3), e91776. doi:10.1371/journal.pone.0091776

Nascimento, L. B. d. S., Moreira, N. d. S., Leal-Costa, M. V., Costa, S. S., & Tavares, E. S. (2015). Induction of wound-periderm-like tissue in *Kalanchoe pinnata* (Lam.) Pers.(Crassulaceae) leaves as a defence response to high UV-B radiation levels. *Ann Bot, 116*(5), 763-769.

Nawrocki, E. P., & Eddy, S. R. (2013). Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics, 29*(22), 2933-2935.

Nayyar, H., Bains, T. S., Kumar, S., & Kaur, G. (2005). Chilling effects during seed filling on accumulation of seed reserves and yield of chickpea. *J Sci Food Agric, 85*(11), 1925-1930.

Nürnberger, T., & Scheel, D. (2001). Signal transmission in the plant immune response. *Trends Plant Sci, 6*(8), 372-379.

O'Hara, A., Headland, L. R., Díaz-Ramos, L. A., Morales, L. O., Strid, Å., & Jenkins, G. I. (2019). Regulation of Arabidopsis gene expression by low fluence rate UV-B independently of UVR8 and stress signaling. *Photochem Photobiol Sci, 18*(7), 1675-1684.

Omezzine, F., Bouaziz, M., Daami-Remadi, M., Simmonds, M. S., & Haouala, R. (2017). Chemical composition and antifungal activity of *Trigonella foenum-graecum* L. varied with plant ploidy level and developmental stage. *Arab J Chem, 10*, S3622-S3631.

One Thousand Plant Transcriptomes, I. (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature, 574*(7780), 679-685. doi:10.1038/s41586-019-1693-2

Ouaked, F., Rozhon, W., Lecourieux, D., & Hirt, H. (2003). A MAPK pathway mediates ethylene signaling in plants. *EMBO J, 22*(6), 1282-1288.

Ozkan, H., & Feldman, M. (2009). Rapid cytological diploidization in newly formed allopolyploids of the wheat (*Aegilops Triticum*) group. *Genome, 52*(11), 926-934. doi:10.1139/g09-067

Pan, Y. Q., Guo, H., Wang, S. M., Zhao, B., Zhang, J. L., Ma, Q., . . . Bao, A. K. (2016). The photosynthesis, Na^+/K^+ homeostasis and osmotic adjustment of Atriplex canescens in response to salinity. *Front Plant Sci, 7*, 848.

Pattanagul, W., & Thitisaksakul, M. (2008). Effect of salinity stress on growth and carbohydrate metabolism in three rice (*Oryza sativa* L.) cultivars differing in salinity tolerance. URL *http://nopr.niscair.res.in/handle/123456789/4642*

Pei, Z. M., Ghassemian, M., Kwak, C. M., McCourt, P., & Schroeder, J. I. (1998). Role of farnesyltransferase in ABA regulation of guard cell anion channels and plant water loss. *Science, 282*(5387), 287-290.

Pelé, A., Rousseau-Gueutin, M., & Chèvre, A. M. (2018). Speciation success of polyploid plants closely relates to the regulation of meiotic recombination. *Front Plant Sci,* science, 9, 907.

Peleg, Z., Reguera, M., Tumimbang, E., Walia, H., & Blumwald, E. (2011). Cytokinin-mediated source/sink modifications improve drought tolerance and increase grain yield in rice under water-stress. *Plant Biotechnol J, 9*(7), 747-758. doi:10.1111/j.1467-7652.2010.00584.x

Petersen, M., Brodersen, P., Naested, H., Andreasson, E., Lindhart, U., Johansen, B., . . . Mundy, J. (2000). Arabidopsis map kinase 4 negatively regulates systemic acquired resistance. *Cell, 103*(7), 1111-1120. doi:10.1016/s0092-8674(00)00213-0

Petruľová, V., Dučaiová, Z., & Repčák, M. (2014). Short-term UV-B Dose Stimulates Production of Protective Metabolites in *Matricaria chamomilla* Leaves. *Photochem Photobiol, 90*(5), 1061-1068.

Price, J., Laxmi, A., Martin, S. K. S., & Jang, J. C. (2004). Global transcription profiling reveals multiple sugar signal transduction mechanisms in Arabidopsis. *Plant Cell, 16*(8), 2128-2150.

Pruitt, K. D., Tatusova, T., & Maglott, D. R. (2006). NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res, 35*(suppl_1), D61-D65.

Qaderi, M. M., Yeung, E. C., & Reid, D. M. (2008). Growth and physiological responses of an invasive alien species, Silene noctiflora, during two developmental stages to four levels of ultraviolet-B radiation. *Ecoscience, 15*(2), 150-159.

Rabbani, M. A., Maruyama, K., Abe, H., Khan, M. A., Katsura, K., Ito, Y., . . . Yamaguchi-Shinozaki, K. (2003). Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. *Plant Physiol, 133*(4), 1755-1767.

Raghuvanshi, R., & Sharma, R. K. (2016). Response of two cultivars of *Phaseolus vulgaris* L. (French beans) plants exposed to enhanced UV-B radiation under mountain ecosystem. *Environ Sci Pollut Res Int, 23*(1), 831-842. doi:10.1007/s11356-015-5332-7

Rajendiran, K., & Ramanujam, M. (2004). Improvement of biomass partitioning, flowering and yield by triadimefon in UV-B stressed *Vigna radiata* (L.) Wilczek. *Biol Plant, 48*(1), 145-148.

Reaney, M., & Gusta, L. (1999). Modeling sequential responses of plant cells to freezing and thawing. In *Cold-Adapted Organisms* (pp. 119-135): Springer.

Reczek, C. R., & Chandel, N. S. (2015). ROS-dependent signal transduction. *Curr Opin Cell Biol, 33*, 8-13. doi:10.1016/j.ceb.2014.09.010

Reddy, P. S., Jogeswar, G., Rasineni, G. K., Maheswari, M., Reddy, A. R., Varshney, R. K., & Kishor, P. K. (2015). Proline over-accumulation alleviates salt stress and protects photosynthetic and antioxidant enzyme activities in transgenic sorghum [*Sorghum bicolor* (L.) Moench]. *Plant Physiol Biochem, 94*, 104-113.

Rejeb, K. B., Benzarti, M., Debez, A., Bailly, C., Savouré, A., & Abdelly, C. (2015). NADPH oxidase-dependent H2O2 production is required for salt-induced antioxidant defense in *Arabidopsis thaliana*. *J Plant Physiol, 174*, 5-15.

Rensink, W. A., Iobst, S., Hart, A., Stegalkina, S., Liu, J., & Buell, C. R. (2005). Gene expression profiling of potato responses to cold, heat, and salt stress. *Funct Integr Genomics, 5*(4), 201-207. doi:10.1007/s10142-005-0141-6

Rhodes, D., Nadolska-Orczyk, A., & Rich, P. (2002). Salinity, osmolytes and compatible solutes. In *Salinity: Environment-plants-molecules* (pp. 181-204): Springer.

Robberecht, R., & Caldwell, M. M. (1983). Protective mechanisms and acclimation to solar ultraviolet-B radiation in *Oenothera stricta*. *Plant Cell Environ, 6*(6), 477-485.

Robertson, G., Schein, J., Chiu, R., Corbett, R., Field, M., Jackman, S. D., ... & Griffith, M. (2010). De novo assembly and analysis of RNA-seq data. *Nat Methods, 7*(11), 909-912.

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics, 26*(1), 139-140. doi:10.1093/bioinformatics/btp616

Roccaforte, K., Russo, S. E., & Pilson, D. (2015). Hybridization and reproductive isolation between diploid *Erythronium mesochoreum* and its tetraploid congener *E. albidum* (Liliaceae). *Evolution, 69*(6), 1375-1389. doi:10.1111/evo.12666

Rogowska, A., & Szakiel, A. (2020). The role of sterols in plant response to abiotic stress. *Phytochem Rev*, 1-14.

Rybarczyk-Plonska, A., Hagen, S. F., Borge, G. I. A., Bengtsson, G. B., Hansen, M. K., & Wold, A. B. (2016). Glucosinolates in broccoli (*Brassica oleracea* L. var.

italica) as affected by postharvest temperature and radiation treatments. *Postharvest Biol Technol, 116*, 16-25.

Sakata, T., Oshino, T., Miura, S., Tomabechi, M., Tsunaga, Y., Higashitani, N., . . . Higashitani, A. (2010). Auxins reverse plant male sterility caused by high temperatures. *PNAS, 107*(19), 8569-8574.

Salehin, M., Li, B., Tang, M., Katz, E., Song, L., Ecker, J. R., . . . Estelle, M. (2019). Auxin-sensitive Aux/IAA proteins mediate drought tolerance in Arabidopsis by regulating glucosinolate levels. *Nat Commun, 10*.

Sánchez-Bermejo, E., Castrillo, G., Del Llano, B., Navarro, C., Zarco-Fernández, S., Martinez-Herrera, D. J., . . . Paz-Ares, J. (2014). Natural variation in arsenate tolerance identifies an arsenate reductase in *Arabidopsis thaliana*. *Nat Commun, 5*, 4617.

Santelia, D., Henrichs, S., Vincenzetti, V., Sauer, M., Bigler, L., Klein, M., . . . Geisler, M. (2008). Flavonoids redirect PIN-mediated polar auxin fluxes during root gravitropic responses. *J Biol Chem, 283*(45), 31218-31226.

Santiago, J., Dupeux, F., Betz, K., Antoni, R., Gonzalez-Guzman, M., Rodriguez, L., . . . Rodriguez, P. L. (2012). Structural insights into PYR/PYL/RCAR ABA receptors and PP2Cs. *Plant Sci, 182*, 3-11.

Schatz, M. C., Witkowski, J., & McCombie, W. R. (2012). Current challenges in *de novo* plant genome sequencing and assembly. *Genome Biol, 13*(4), 243.

Scholes, D. R., & Paige, K. N. (2011). Chromosomal plasticity: mitigating the impacts of herbivory. *Ecology, 92*(8), 1691-1698. doi:10.1890/10-2269.1

Schulz, M. H., Zerbino, D. R., Vingron, M., & Birney, E. (2012). Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics, 28*(8), 1086-1092.

Schranz, M. E., & Osborn, T. C. (2000). Novel flowering time variation in the resynthesized polyploid *Brassica napus*. *J Hered, 91*(3), 242-246. doi:10.1093/jhered/91.3.242

Seki, M., Narusaka, M., Ishida, J., Nanjo, T., Fujita, M., Oono, Y., . . . Sakurai, T. (2002). Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *Plant J, 31*(3), 279-292.

Selvaraj, M. G., Ishizaki, T., Valencia, M., Ogawa, S., Dedicova, B., Ogata, T., . . . Saito, K. (2017). Overexpression of an *Arabidopsis thaliana* galactinol synthase gene improves drought tolerance in transgenic rice and increased grain yield in the field. *Plant Biotechnol J, 15*(11), 1465-1477.

Seo, D. H., Ryu, M. Y., Jammes, F., Hwang, J. H., Turek, M., Kang, B. G., . . . Kim, W. T. (2012). Roles of four Arabidopsis U-box E3 ubiquitin ligases in negative regulation of abscisic acid-mediated drought stress responses. *Plant Physiol, 160*(1), 556-568. doi:10.1104/pp.112.202143

Seo, S. Y., Wi, S. J., & Park, K. Y. (2020). Functional switching of NPR1 between chloroplast and nucleus for adaptive response to salt stress. *Sci Rep, 10*(1), 1-18.

Shaikhali, J., Heiber, I., Seidel, T., Ströher, E., Hiltscher, H., Birkmann, S., . . . Baier, M. (2008). The redox-sensitive transcription factor Rap2. 4a controls nuclear expression of 2-Cys peroxiredoxin A and other chloroplast antioxidant enzymes. *BMC Plant Biol, 8*(1), 1-14.

Shaikhali, J., Norén, L., de Dios Barajas-López, J., Srivastava, V., König, J., Sauer, U. H., . . . Strand, Å. (2012). Redox-mediated mechanisms regulate DNA binding activity of the G-group of basic region leucine zipper (bZIP) transcription factors in Arabidopsis. *J Biol Chem, 287*(33), 27510-27525.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res, 13*(11), 2498-2504.

Sharabi-Schwager, M., Samach, A., & Porat, R. (2010). Overexpression of the CBF2 transcriptional activator in Arabidopsis counteracts hormone activation of leaf senescence. *Plant Signal Behav, 5*(3), 296-299.

Shen, X., Wang, Z., Song, X., Xu, J., Jiang, C., Zhao, Y., . . . Zhang, H. (2014). Transcriptomic profiling revealed an important role of cell wall remodeling and ethylene signaling pathway during salt acclimation in Arabidopsis. *Plant Mol Biol, 86*(3), 303-317.

Shen, J., Xiao, Q., Qiu, H., Chen, C., & Chen, H. (2016). Integrative effect of drought and low temperature on litchi (*Litchi chinensis* Sonn.) floral initiation revealed by dynamic genome-wide transcriptome analysis. *Sci Rep, 6*, 32005.

Shi, C., Qi, C., Ren, H., Huang, A., Hei, S., & She, X. (2015). Ethylene mediates brassinosteroid-induced stomatal closure via Gα protein-activated hydrogen peroxide and nitric oxide production in Arabidopsis. *Plant J, 82*(2), 280-301.

Shi, Y., & Yang, S. (2014). ABA regulation of the cold stress response in plants. In *Abscisic acid: metabolism, transport and signaling* (pp. 337-363): Springer.

Shibasaki, K., Uemura, M., Tsurumi, S., & Rahman, A. (2009). Auxin response in Arabidopsis under cold stress: underlying molecular mechanisms. *Plant Cell, 21*(12), 3823-3838. doi:10.1105/tpc.109.069906

Shinozaki, K., Yamaguchi-Shinozaki, K., & Seki, M. (2003). Regulatory network of gene expression in the drought and cold stress responses. *Curr Opin Plant Biol, 6*(5), 410-417.

Shulse, C. N., Cole, B. J., Ciobanu, D., Lin, J., Yoshinaga, Y., Gouran, M., . . . Dickel, D. E. (2019). High-Throughput Single-Cell Transcriptome Profiling of Plant Cell Types. *Cell Rep, 27*(7), 2241-2247 e2244. doi:10.1016/j.celrep.2019.04.054

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics, 31*(19), 3210-3212.

Skirycz, A., De Bodt, S., Obata, T., De Clercq, I., Claeys, H., De Rycke, R., . . . Fernie, A. R. (2010). Developmental stage specificity and the role of mitochondrial metabolism in the response of Arabidopsis leaves to prolonged mild osmotic stress. *Plant Physiol, 152*(1), 226-244.

Smékalová, V., Doskočilová, A., Komis, G., & Šamaj, J. (2014). Crosstalk between secondary messengers, hormones and MAPK modules during abiotic stress signalling in plants. *Biotechnol Adv, 32*(1), 2-11.

Smit, A., & Hubley, R. (2008). RepeatModeler Open-1.0. *Available fom http://www. repeatmasker. org*.

Smit, A., Hubley, R., & Green, P. (2016). RepeatMasker Open-4.0. 2015. *Google Scholar*.

Solano, R., Stepanova, A., Chao, Q., & Ecker, J. R. (1998). Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Genes Dev, 12*(23), 3703-3714.

Solfanelli, C., Poggi, A., Loreti, E., Alpi, A., & Perata, P. (2006). Sucrose-specific induction of the anthocyanin biosynthetic pathway in Arabidopsis. *Plant Physiol, 140*(2), 637-646.

Song, Q., & Chen, Z. J. (2015). Epigenetic and developmental regulation in plant polyploids. *Curr Opin Plant Biol, 24*, 101-109.

Stanke, M., Diekhans, M., Baertsch, R., & Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. *Bioinformatics, 24*(5), 637-644.

Stass, A., & Horst, W. J. (2009). Callose in abiotic stress. In *Chemistry, biochemistry, and biology of 1-3 beta glucans and related polysaccharides* (pp. 499-524): Elsevier.

Staxén, I., & Bornman, J. F. (1994). A morphological and cytological study of Petunia hybrida exposed to UV-B radiation. *Physiol Plant, 91*(4), 735-740.

Steinmüller, D., ., & Tevini, M., . (1985). Action of ultraviolet radiation (UV-B) upon cuticular waxes in some crop plants. *Planta, 164*(4), 557-564.

Steinmetz, V., & Wellmann, E. (1986). The role of solar UV-B in growth regulation of cress (*Lepidium sativum* L.) seedlings. *Photochem Photobiol, 43*(2), 189-193.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., ... & Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *PNAS, 102*(43), 15545-15550.

Sullivan, J. H., & Teramura, A. H. (1990). Field study of the interaction between solar ultraviolet-B radiation and drought on photosynthesis and growth in soybean. *Plant Physiol, 92*(1), 141-146.

Sun, W., Xu, X., Zhu, H., Liu, A., Liu, L., Li, J., & Hua, X. (2010). Comparative transcriptomic profiling of a salt-tolerant wild tomato species and a salt-sensitive tomato cultivar. *Plant Cell Physiol, 51*(6), 997-1006.

Sun, Z., Qi, X., Wang, Z., Li, P., Wu, C., Zhang, H., & Zhao, Y. (2013). Overexpression of TsGOLS2, a galactinol synthase, in *Arabidopsis thaliana* enhances tolerance to high salinity and osmotic stresses. *Plant Physiol Biochem, 69*, 82-89.

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PloS one, 6*(7), e21800. doi:10.1371/journal.pone.0021800

Szalai, G., Tari, I., Janda, T., Pestenacz, A., & Páldi, E. (2000). Effects of cold acclimation and salicylic acid on changes in ACC and MACC contents in maize during chilling. *Biol Plant, 43*(4), 637-640.

Tang, S., Liang, H., Yan, D., Zhao, Y., Han, X., Carlson, J. E., . . . Yin, W. (2013). Populus euphratica: the transcriptomic response to drought stress. *Plant Mol Biol, 83*(6), 539-557.

Tate, J. A., Symonds, V. V., Doust, A. N., Buggs, R. J., Mavrodiev, E., Majure, L. C., . . . Soltis, D. E. (2009). Synthetic polyploids of *Tragopogon miscellus* and *T. mirus* (Asteraceae): 60 Years after Ownbey's discovery. *Am J Bot, 96*(5), 979-988. doi:10.3732/ajb.0800299

Thiel, T., Michalek, W., Varshney, R., & Graner, A. (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor Appl Genet, 106*(3), 411-422.

Tse-Dinh, Y. C., Haiyan, Q., & Menzel, R. (1997). DNA supercoiling and bacterial adaptation: thermotolerance and thermoresistance. *Trends Microbiol, 5*(8), 323-326.

Turunen, M., & Latola, K. (2005). UV-B radiation and acclimation in timberline plants. *Environ Pollut, 137*(3), 390-403.

Tuteja, N. (2007). Abscisic acid and abiotic stress signaling. *Plant Signal Behav, 2*(3), 135-138.

Van den Ende, W., & Peshev, D. (2013). Sugars as antioxidants in plants. In *Crop improvement under adverse conditions* (pp. 285-307): Springer.

Venkateswarlu, B., & Shanker, A. (2011). Abiotic stress response in plants–physiological, biochemical and genetic perspectives. In: InTech.

Verma, V., Ravindran, P., & Kumar, P. P. (2016). Plant hormone-mediated regulation of stress responses. *BMC Plant Biol, 16*(1), 86.

Vision, T. J., Brown, D. G., & Tanksley, S. D. (2000). The origins of genomic duplications in Arabidopsis. *Science, 290*(5499), 2114-2117.

Voelckel, C., Heenan, P. B., Janssen, B., Reichelt, M., Ford, K., Hofmann, R., & Lockhart, P. J. (2008). Transcriptional and biochemical signatures of divergence in natural populations of two species of New Zealand alpine Pachycladon. *Mol Ecol, 17*(21), 4740-4753. doi:10.1111/j.1365-294X.2008.03933.x

Voelckel, C., Gruenheit, N., Biggs, P., Deusch, O., & Lockhart, P. (2012). Chips and tags suggest plant-environment interactions differ for two alpine Pachycladon species. *BMC Genomics, 13*, 322. doi:10.1186/1471-2164-13-322

Voelckel, C., Mirzaei, M., Reichelt, M., Luo, Z., Pascovici, D., Heenan, P. B., . . . Lockhart, P. J. (2010). Transcript and protein profiling identify candidate gene sets of potential adaptive significance in New Zealand Pachycladon. *BMC Evol Biol, 10*, 151. doi:10.1186/1471-2148-10-151

Vurture, G. W., Sedlazeck, F. J., Nattestad, M., Underwood, C. J., Fang, H., Gurtowski, J., & Schatz, M. C. (2017). GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics*, 33(14), 2202-2204.

Vyšniauskienė, R., & Rančelienė, V. (2014). Effect of UV-B radiation on growth and antioxidative enzymes activity in Lithuanian potato (*Solanum tuberosum* L.) cultivars. *Zemdirbyste-Agric, 101*(1), 51-56.

Wang, H., Li, J., Tao, W., Zhang, X., Gao, X., Yong, J., . . . Duan, J. A. (2018). Lycium ruthenicum studies: Molecular biology, Phytochemistry and pharmacology. *Food Chem, 240*, 759-766. doi:10.1016/j.foodchem.2017.08.026

Wang, M., Zhao, X., Xiao, Z., Yin, X., Xing, T., & Xia, G. (2016). A wheat superoxide dismutase gene TaSOD2 enhances salt resistance through modulating redox homeostasis by promoting NADPH oxidase activity. *Plant Mol Biol, 91*(1-2), 115-130.

Wang, X., Chang, L., Wang, B., Wang, D., Li, P., Wang, L., . . . Guo, A. (2013). Comparative proteomics of Thellungiella halophila leaves from plants subjected to salinity reveals the importance of chloroplastic starch and soluble sugars in halophyte salt tolerance. *Mol Cell Proteomics, 12*(8), 2174-2195.

Wang, X., Li, W., Li, M., & Welti, R. (2006). Profiling lipid changes in plant response to low temperatures. *Physiol Plant, 126*(1), 90-96.

Wang, X., Shi, X., Hao, B., Ge, S., & Luo, J. (2005). Duplication and DNA segmental loss in the rice genome: implications for diploidization. *New Phytol, 165*(3), 937-946. doi:10.1111/j.1469-8137.2004.01293.x

Wang, Y., Coleman-Derr, D., Chen, G., & Gu, Y. Q. (2015). OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res, 43*(W1), W78-W84.

Wang, Y., Yang, J., & Yi, J. (2012). Redox sensing by proteins: oxidative modifications on cysteines and the consequent events. *Antioxid Redox Signal, 16*(7), 649-657. doi:10.1089/ars.2011.4313

Wang, Y. T. (2007). Potassium nutrition affects Phalaenopsis growth and flowering. *HortScience, 42*(7), 1563-1567.

Wani, S. H., Dutta, T., Neelapu, N. R. R., & Surekha, C. (2017). Transgenic approaches to enhance salt and drought tolerance in plants. *Plant Gene, 11*, 219-231. *https://doi.org/10.1016/j.plgene.2017.05.006*

Wargent, J., Nelson, B., McGhie, T., & Barnes, P. (2015). Acclimation to UV-B radiation and visible light in *Lactuca sativa* involves up-regulation of photosynthetic performance and orchestration of metabolome-wide responses. *Plant Cell Environ, 38*(5), 929-940.

Wellmann, E. (1983). UV radiation in photomorphogenesis. In *Photomorphogenesis* (pp. 745-756): Springer.

Wijewardana, C., Henry, W. B., Gao, W., & Reddy, K. R. (2016). Interactive effects on CO2, drought, and ultraviolet-B radiation on maize growth and development. *J Photochem Photobiol B, 160*, 198-209. doi:10.1016/j.jphotobiol.2016.04.004

Williams, C. A., & Murray, B. (1972). Flavonoid variation in the genus Briza. *Phytochemistry, 11*(8), 2507-2512.

Wingler, A., Tijero, V., Müller, M., Yuan, B., & Munné-Bosch, S. (2020). Interactions between sucrose and jasmonate signalling in the response to cold stress. *BMC Plant Biol, 20*, 1-13.

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., & Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. *PNAS, 106*(33), 13875-13879.

Wrzaczek, M., Brosché, M., & Kangasjärvi, J. (2013). ROS signaling loops—production, perception, regulation. *Curr Opin Plant Biol, 16*(5), 575-582.

Wu, D., Hu, Q., Yan, Z., Chen, W., Yan, C., Huang, X., . . . Wang, J. (2012). Structural basis of ultraviolet-B perception by UVR8. *Nature, 484*(7393), 214.

Wu, S. J., & Ng, L. T. (2008). Antioxidant and free radical scavenging activities of wild bitter melon (*Momordica charantia* Linn. var. abbreviata Ser.) in Taiwan. *LWT-Food Science and Technology, 41*(2), 323-330.

Xi, J., Rossi, L., Lin, X., & Xie, D. Y. (2016). Overexpression of a synthetic insect–plant geranyl pyrophosphate synthase gene in *Camelina sativa* alters plant growth and terpene biosynthesis. *Planta, 244*(1), 215-230.

Xie, Y., Wu, G., Tang, J., Luo, R., Patterson, J., Liu, S., . . . Li, S. (2014). SOAPdenovo-Trans: *de novo* transcriptome assembly with short RNA-Seq reads. *Bioinformatics, 30*(12), 1660-1666.

Xiong, Z., Gaeta, R. T., & Pires, J. C. (2011). Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. *Proc Natl Acad Sci U S A, 108*(19), 7908-7913. doi:10.1073/pnas.1014138108

Xu, Z., & Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res, 35*(suppl_2), W265-W268.

Yadav, S. K. (2010). Cold stress tolerance mechanisms in plants. A review. *Agron Sustain Dev, 30*(3), 515-527.

Yang, T. J., Kim, J. S., Kwon, S. J., Lim, K. B., Choi, B. S., Kim, J. A., . . . Park, B. S. (2006). Sequence-level analysis of the diploidization process in the triplicated FLOWERING LOCUS C region of *Brassica rapa*. *Plant Cell, 18*(6), 1339-1347. doi:10.1105/tpc.105.040535

Yang, L., Jin, Y., Huang, W., Sun, Q., Liu, F., & Huang, X. (2018). Full-length transcriptome sequences of ephemeral plant Arabidopsis pumila provides insight into gene expression dynamics during continuous salt stress. *BMC genomics, 19*(1), 1-14.

Yano, R., Nakamura, M., Yoneyama, T., & Nishida, I. (2005). Starch-related α-glucan/water dikinase is involved in the cold-induced development of freezing tolerance in Arabidopsis. *Plant Physiol, 138*(2), 837-846.

Yatsuhashi, H., Hashimoto, T., & Shimizu, S. (1982). Ultraviolet action spectrum for anthocyanin formation in broom sorghum first internodes. *Plant Physiol, 70*(3), 735-741.

Yi, L., Liu, L., Melsted, P., & Pachter, L. (2018). A direct comparison of genome alignment and transcriptome pseudoalignment. *BioRxiv*, 444620.

Yogeeswaran, K., Voelckel, C., Joly, S., & Heenan, P. B. (2011). Pachycladon. In *Wild crop relatives: genomic and breeding resources* (pp. 227-249). Springer, Berlin, Heidelberg.

Yuan, F., Yang, H., Xue, Y., Kong, D., Ye, R., Li, C., . . . Krichilsky, B. (2014). OSCA1 mediates osmotic-stress-evoked Ca^{2+} increases vital for osmosensing in Arabidopsis. *Nature, 514*(7522), 367-371.

Yuan, G. F., Jia, C. G., Li, Z., Sun, B., Zhang, L. P., Liu, N., & Wang, Q. M. (2010). Effect of brassinosteroids on drought resistance and abscisic acid concentration in tomato under water stress. *Sci Hortic, 126*(2), 103-108.

Zahaf, O., Blanchet, S., de Zélicourt, A., Alunni, B., Plet, J., Laffont, C., . . . Diet, A. (2012). Comparative transcriptomic analysis of salt adaptation in roots of contrasting *Medicago truncatula* genotypes. *Mol Plant, 5*(5), 1068-1081.

Zhang, G., Liu, X., Quan, Z., Cheng, S., Xu, X., Pan, S., . . . Wang, W. (2012). Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. *Nat Biotechnol, 30*(6), 549.

Zhang, H., Han, B., Wang, T., Chen, S., Li, H., Zhang, Y., & Dai, S. (2012). Mechanisms of plant salt response: insights from proteomics. *J Proteome Res, 11*(1), 49-67.

Zhang, J., Jia, W., Yang, J., & Ismail, A. M. (2006). Role of ABA in integrating plant responses to drought and salt stresses. *Field Crops Res, 97*(1), 111-119.

Zhang, S. L., Peng, D., XU, Y. C., Lü, S. W., & WANG, J. J. (2016). Quantification and analysis of anthocyanin and flavonoids compositions, and antioxidant activities in onions with three different colors. *J Integr Agric, 15*(9), 2175-2181.

Zhang, X. L., Jiang, L., Xin, Q., Liu, Y., Tan, J. X., & Chen, Z. Z. (2015). Structural basis and functions of abscisic acid receptors PYLs. *Front Plant Sci, 6*, 88.

Zhao, B., Liu, L., Tan, D., & Wang, J. (2010). Analysis of phylogenetic relationships of Brassicaceae species based on Chs sequences. *Biochem Syst Ecol, 38*(4), 731-739.

Zhao, C., Wang, P., Si, T., Hsu, C.C., Wang, L., Zayed, O., . . . Tao, W. A. (2017). MAP kinase cascades regulate the cold response by modulating ICE1 protein stability. *Dev Cell, 43*(5), 618-629. e615.

Zhao, Q. Y., Wang, Y., Kong, Y. M., Luo, D., Li, X., & Hao, P. (2011). Optimizing *de novo* transcriptome assembly from short-read RNA-Seq data: a comparative study. *BMC Bioinformatics, 12 Suppl 14*, S2. doi:10.1186/1471-2105-12-S14-S2

Zhao, S., Fung-Leung, W. P., Bittner, A., Ngo, K., & Liu, X. (2014). Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS One*, 9(1), e78644. doi:10.1371/journal.pone.0078644

Zerbino, D. R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn. *Genome Res, 18*(5): 821–829.

Zhu, J. K. (2016). Abiotic stress signaling and responses in plants. *Cell, 167*(2), 313-324.

Zhu, J. K. (2007). Plant salt stress. *eLS*.

Appendix 1. Supplementary Tables and Figures

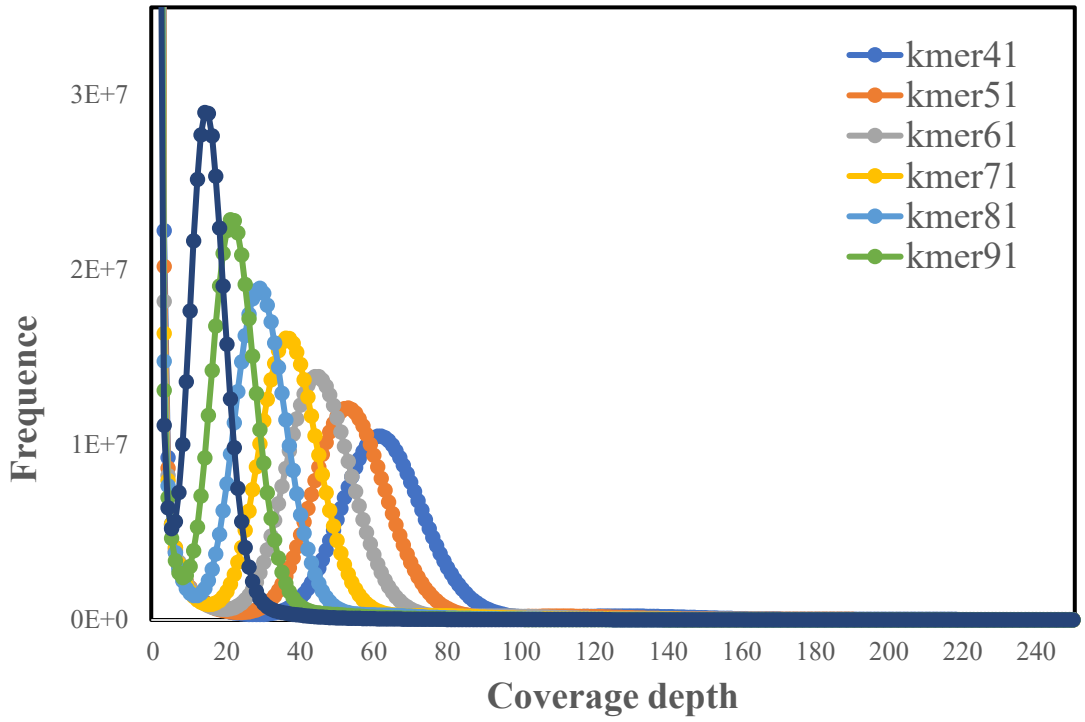

Figure S3-1. K-mer analysis for estimating the genome size of *P. cheesemanii*.

The genome size was estimated by using the formula: G=(N x(L-K+1)-B)/D. G, genome size; N, number of reads; L, length of reads; K, length of k-mer; B, low-frequency k-mers with occurrence less than four times; D, coverage depth corresponding to selected k-mer. 41,51,61,71,81,91, and 101-mer sizes were analyzed, and the coverage depth of 41-mer was selected for genome size estimation.

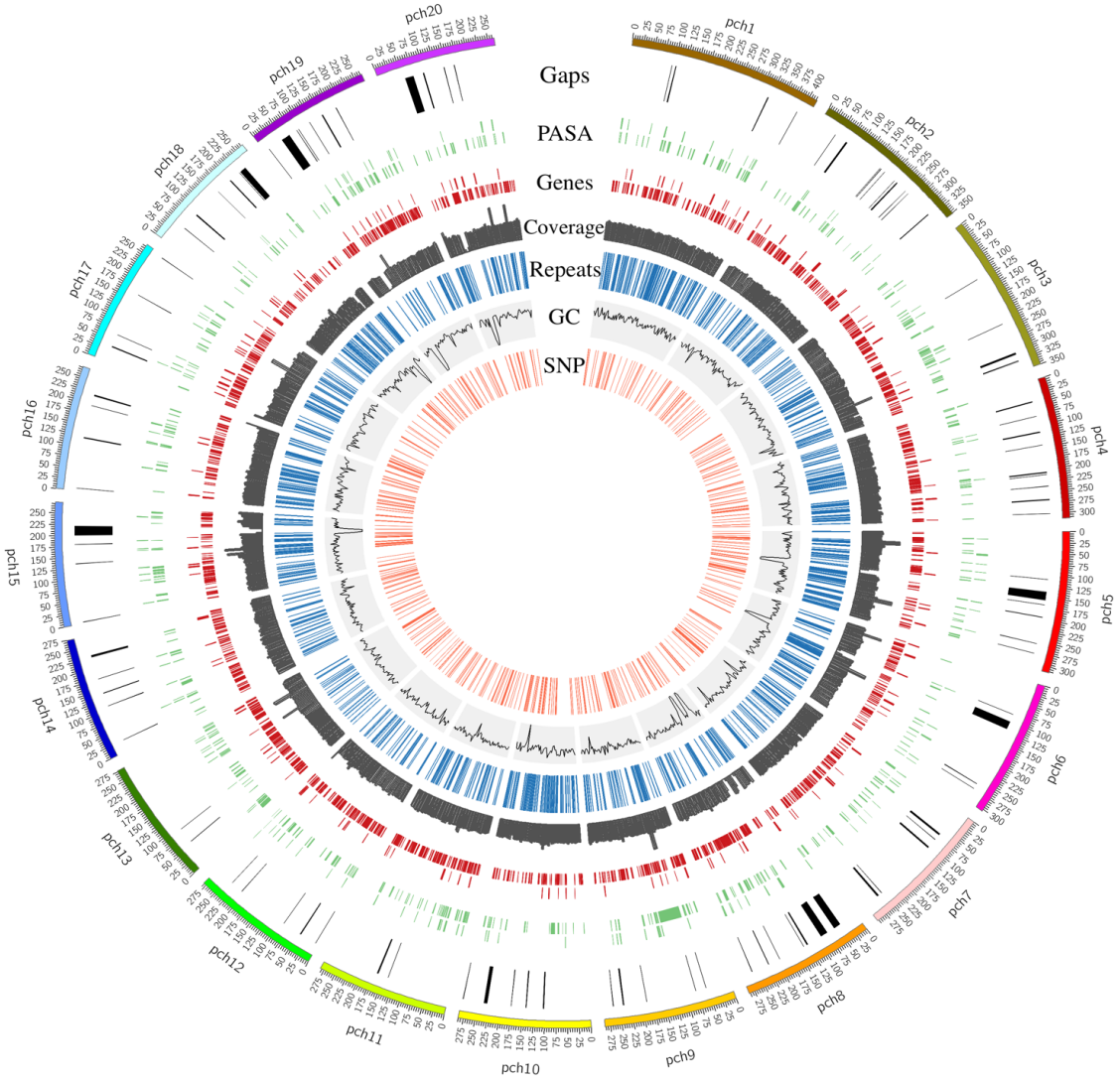

Figure S3-2. Characteristics of the 20 longest scaffolds of the *P. cheesemanii* genome assembly.

The most outer ring represents the 20 longest scaffolds of the *P. cheesemanii* genome assembly, and they are numbered based on the length of the scaffolds from longest to shortest. 1) SNPs, 2) GC content (bin: 5 kb; axis: 0 – 50%), 3) repeats, 4) read coverage (bin: 5 kb), 5) predicted genes, 6) PASA alignments against the leaf transcriptome, 7) distribution of gaps (Ns).

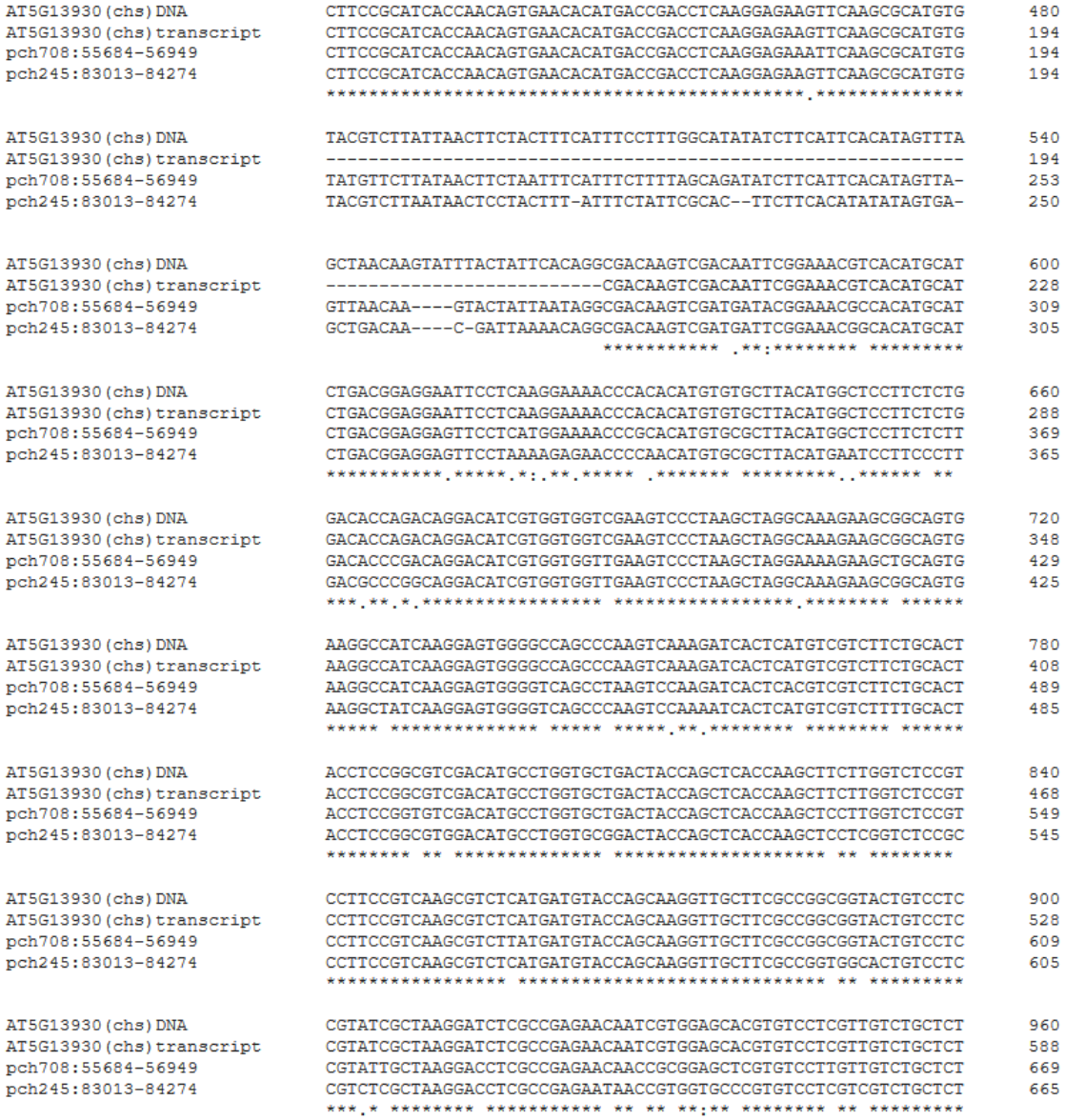

Figure S3-3. Comparison of transcript and genomic DNA sequences between *A. thaliana CHS* and two *P. cheesemanii* homologs.

The two *P. cheesemanii* homologs show obvious sequence differences from *A. thaliana*, and slight differences between them. Full sequences of the *P. cheesemanii* genes are given in Supplementary File 2-5. The indicated sequences were aligned using CLUSTAL OMEGA (1.2.4) (https://www.ebi.ac.uk/Tools/msa/clustalo/) and the conserved nucleotide sequences are indicated by asterisks below the sequences.

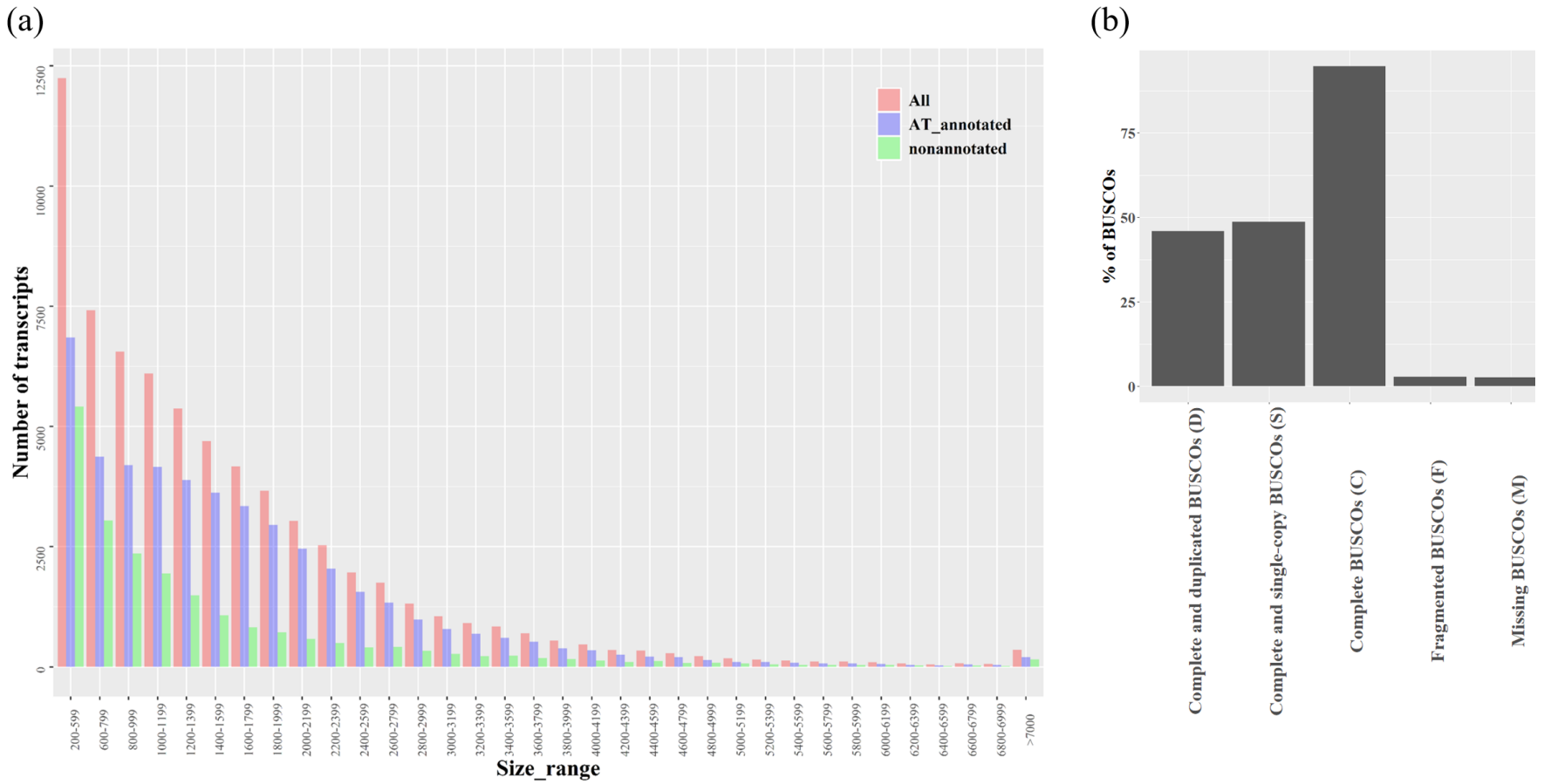

Figure S4-1 Length distribution of the assembled transcripts in *P. cheesemanii* and BUSCOs assessment of assembled transcriptome.

(a) length distribution of the assembled transcripts. Blue bar: the number of the transcripts could be annotated by *A. thaliana*; green bar: the number of the transcripts could not be annotated by *A. thaliana*. (b) BUSCOs assessment.

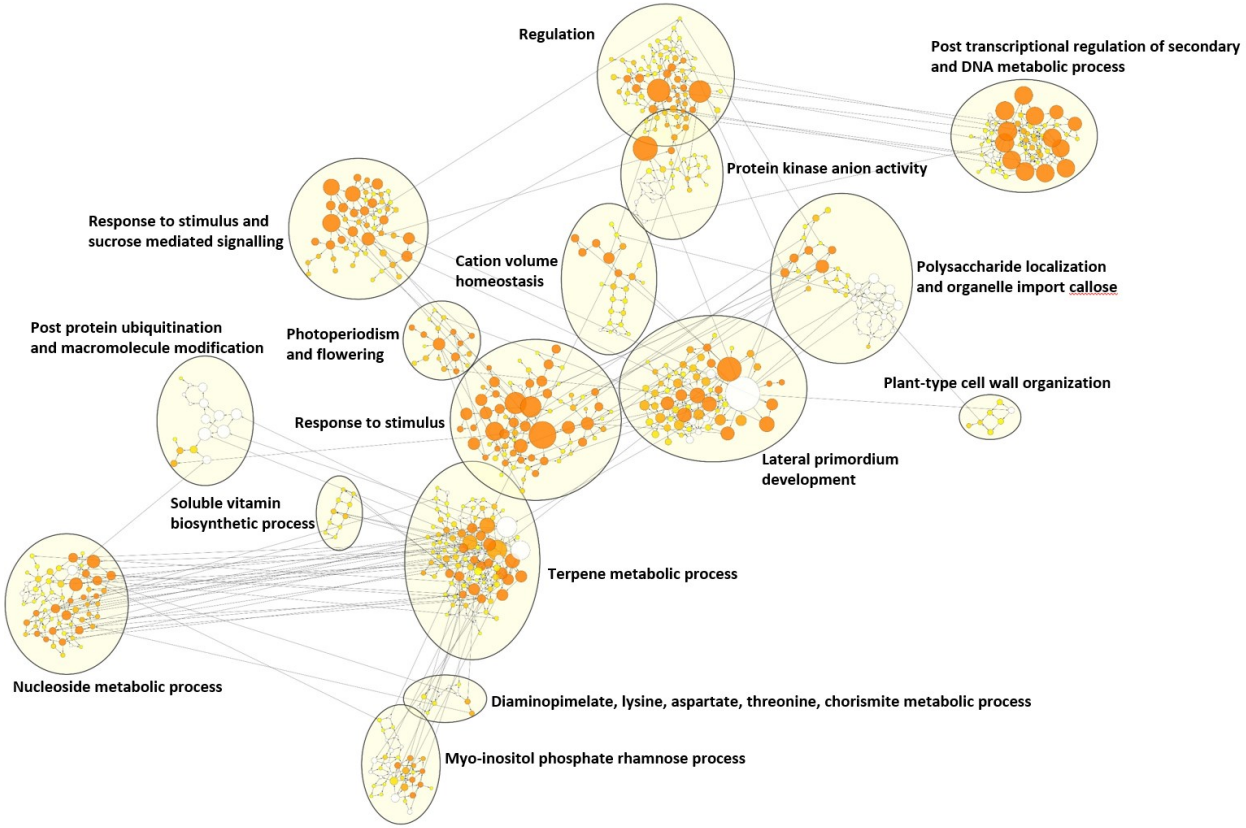

Figure S5-1 Clustering annotation of the gene set of the shared biological processes of *P. cheesemanii* and *A. thaliana* in responding to cold stress.

1139 *A. thaliana* cold responsive genes scattered in 667 nodes, then these nodes were grouped into 16 clusters. Nodes: the terms of GO biological process; colour: significant *p*-value. Higher significant *p*-value, the node colour gets increasingly more orange. Uncoloured nodes are not overrepresented, but they are the parents of overrepresented categories.

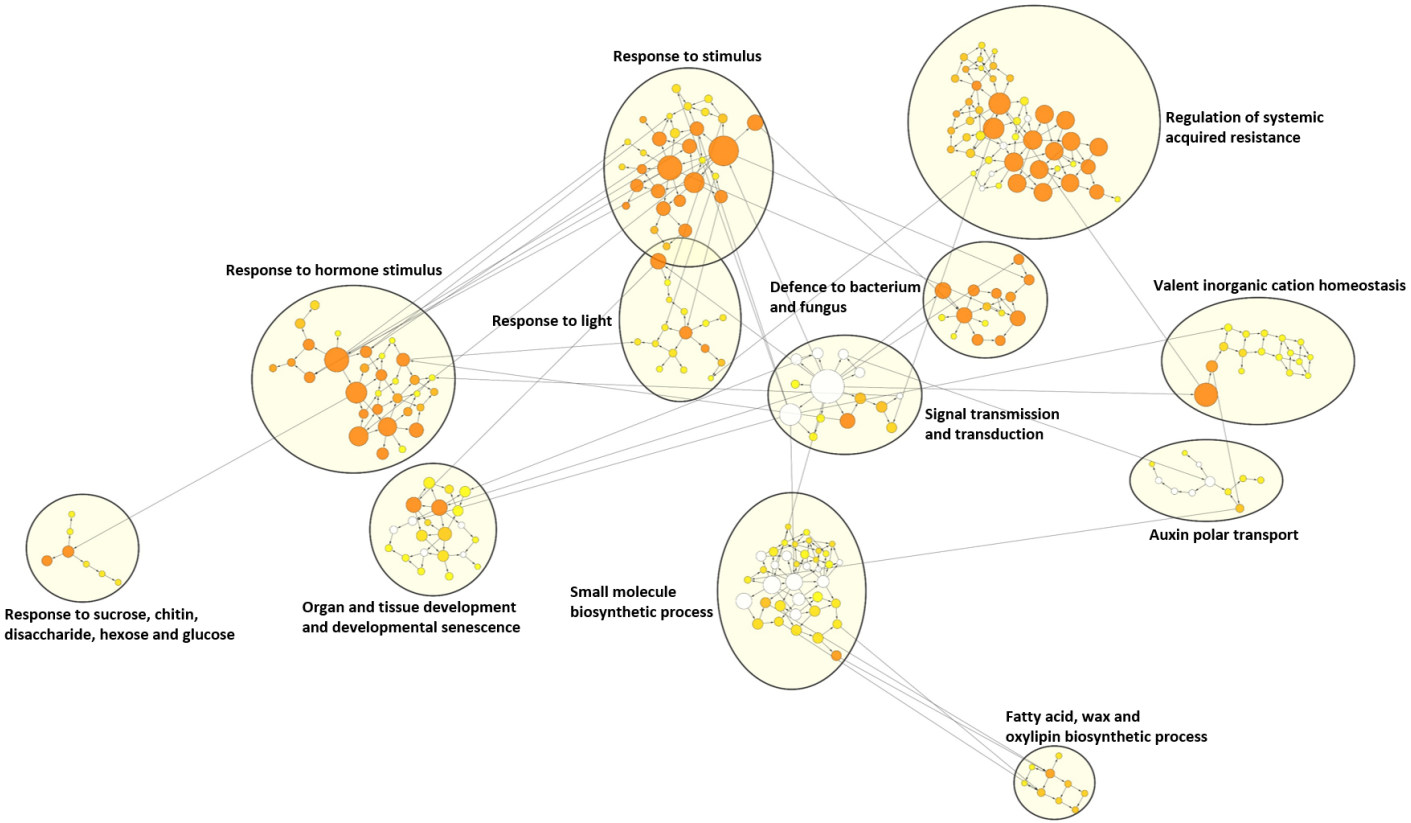

Figure S5-2 Clustering annotation of the gene set of the shared biological processes of *P. cheesemanii* and *A. thaliana* in responding to salt stress.

381 *A. thaliana* salt responsive genes scattered in 257 nodes, then these nodes were grouped into 12 clusters. Nodes: the terms of GO biological process; colour: significant *p*-value. Higher significant *p*-value, the node colour gets increasingly more orange. Uncoloured nodes are not overrepresented, but they are the parents of overrepresented categories.

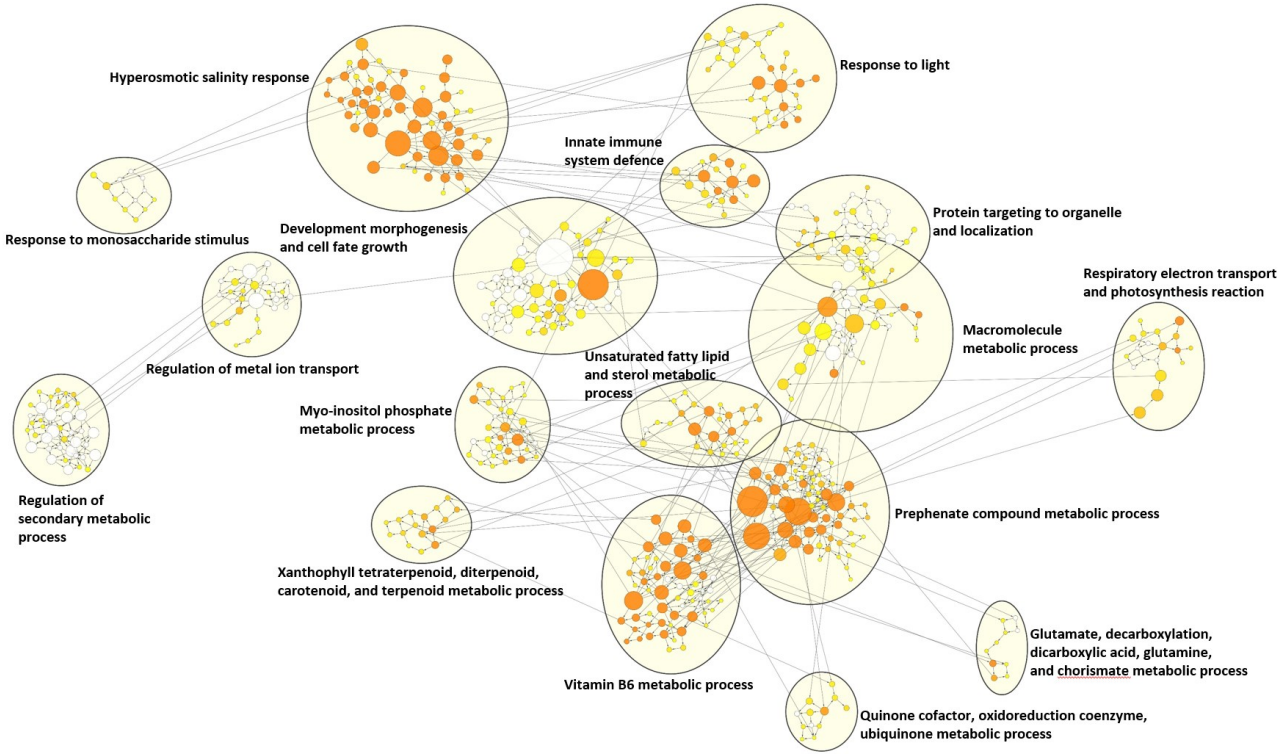

Figure S5-3 Clustering annotation of the gene set of the shared biological processes of *P. cheesemanii* and *A. thaliana* in responding to UV-B irradistion stress.

952 *A. thaliana* UV-B radiation responsive genes scattered in 550 nodes, then these nodes were grouped into 17 clusters. Nodes: the terms of GO biological process; colour: significant *p*-value. Higher significant *p*-value, the node colour gets increasingly more orange. Uncoloured nodes are not overrepresented, but they are the parents of overrepresented categories.

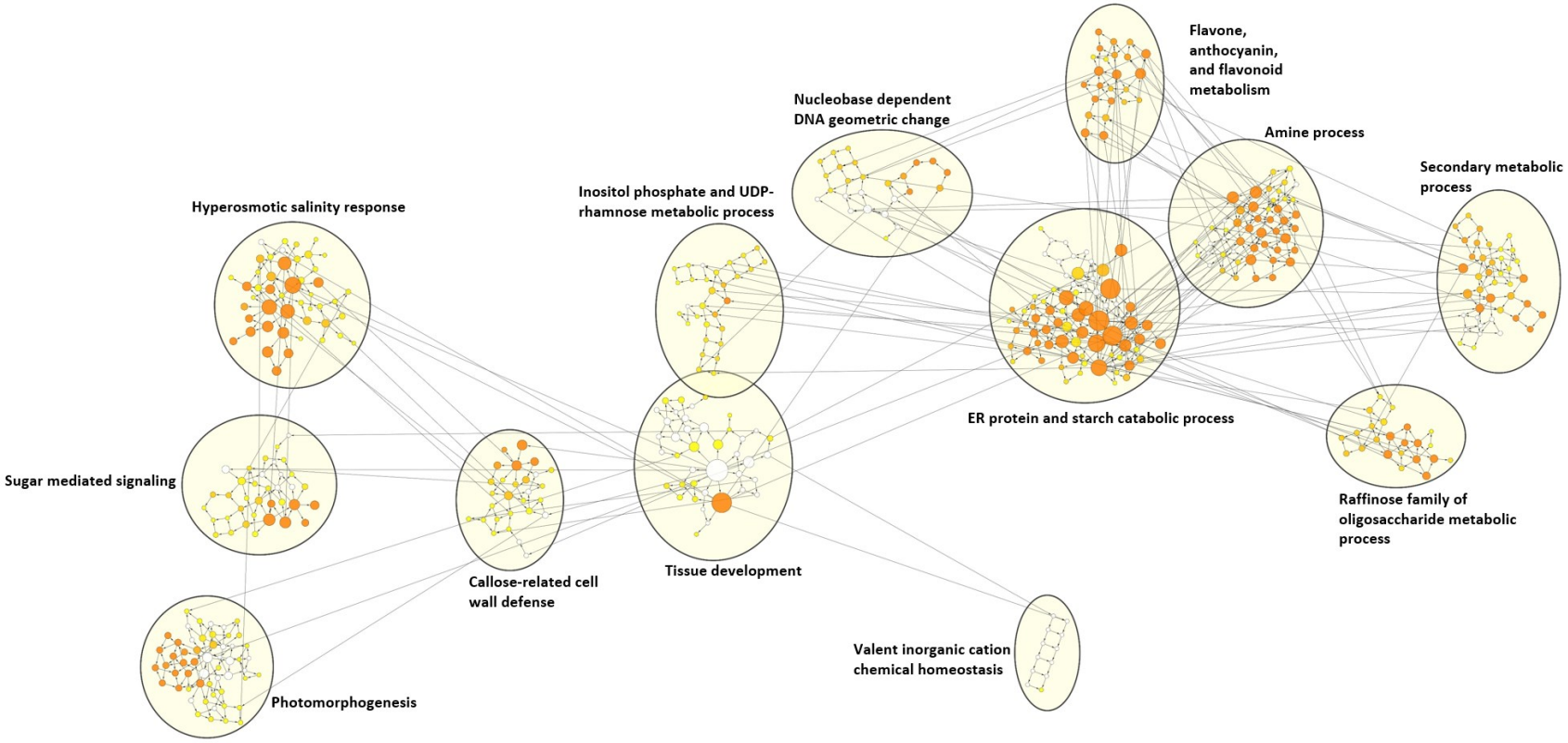

Figure S5-4 Clustering annotation of *A. thaliana* unique gene set in responding to cold stress.

183 genes of *A. thaliana* unique cold response set scattered in 484 nodes, then these nodes were grouped into 13 clusters. Nodes: the terms of GO biological process; colour: significant *p*-value. Higher significant *p*-value, the node colour gets increasingly more orange. Uncoloured nodes are not overrepresented, but they are the parents of overrepresented categories.

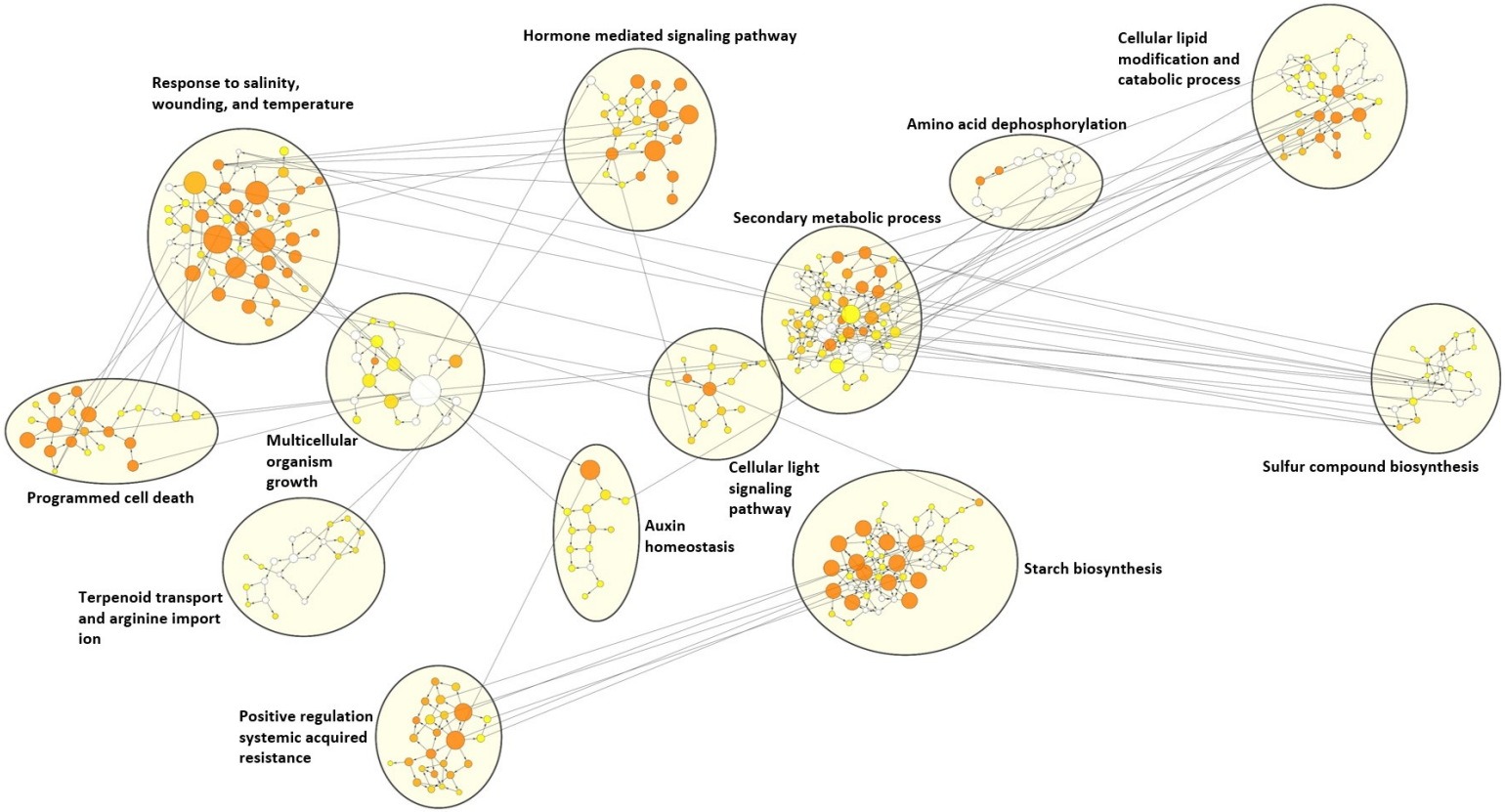

Figure S5-5 Clustering annotation of *A. thaliana* unique gene set in responding to salt stress.

288 genes of *A. thaliana* unique salt response set scattered in 358 nodes, then these nodes were grouped into 13 clusters. Nodes: overrepresented terms of GO biological process; colour: significant *p*-value. Higher significant *p*-value, the node colour gets increasingly more orange. Uncoloured nodes are not overrepresented, but they are the parents of overrepresented categories.

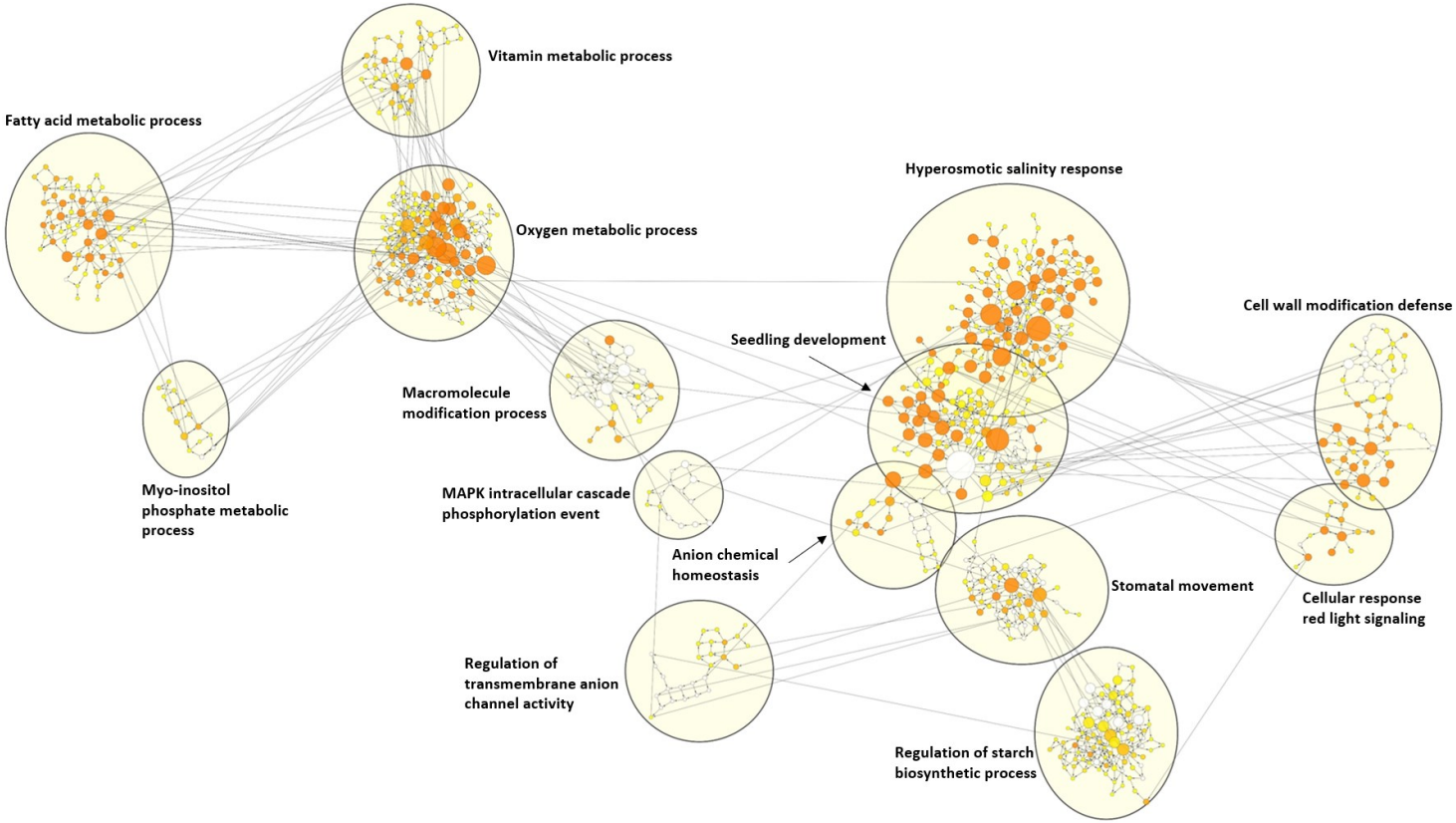

Figure S5-6 Clustering annotation of *A. thaliana* unique gene set in responding to UV-B radiation stress.

464 genes of *A. thaliana* unique UV-B radiation response set scattered in 729 nodes, then these nodes were grouped into 14 clusters. Nodes: overrepresented terms of GO biological process; colour: significant *p*-value. Higher significant *p*-value, the node colour gets increasingly more orange. Uncoloured nodes are not overrepresented, but they are the parents of overrepresented categories.

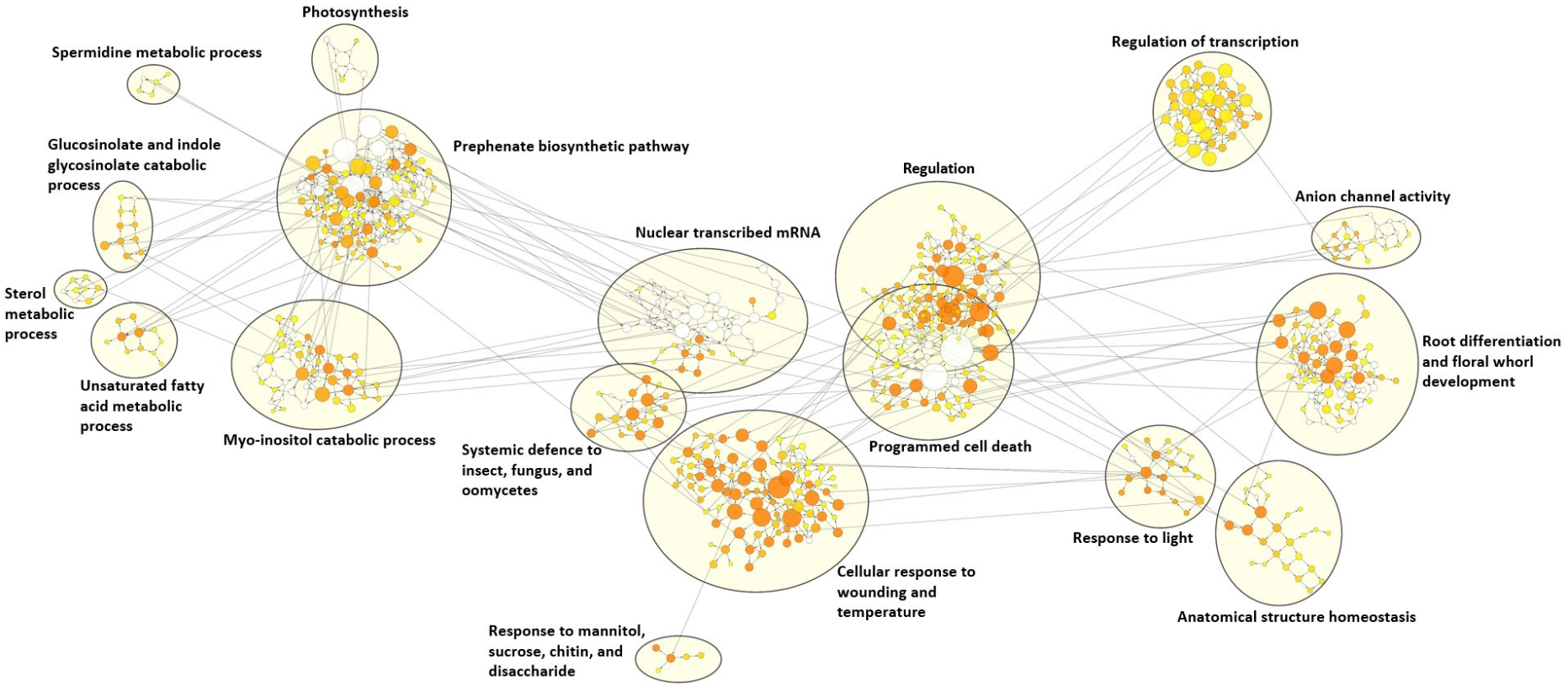

Figure S5-7 Clustering annotation of *P. cheesemanii* unique gene set in responding to cold stress.

464 genes of *P. cheesemanii* unique cold response set scattered in 655 nodes, then these nodes were grouped into 18 clusters. Nodes: overrepresented terms of GO biological process; colour: significant *p*-value. Higher significant *p*-value, the node colour gets increasingly more orange. Uncoloured nodes are not overrepresented, but they are the parents of overrepresented categories.

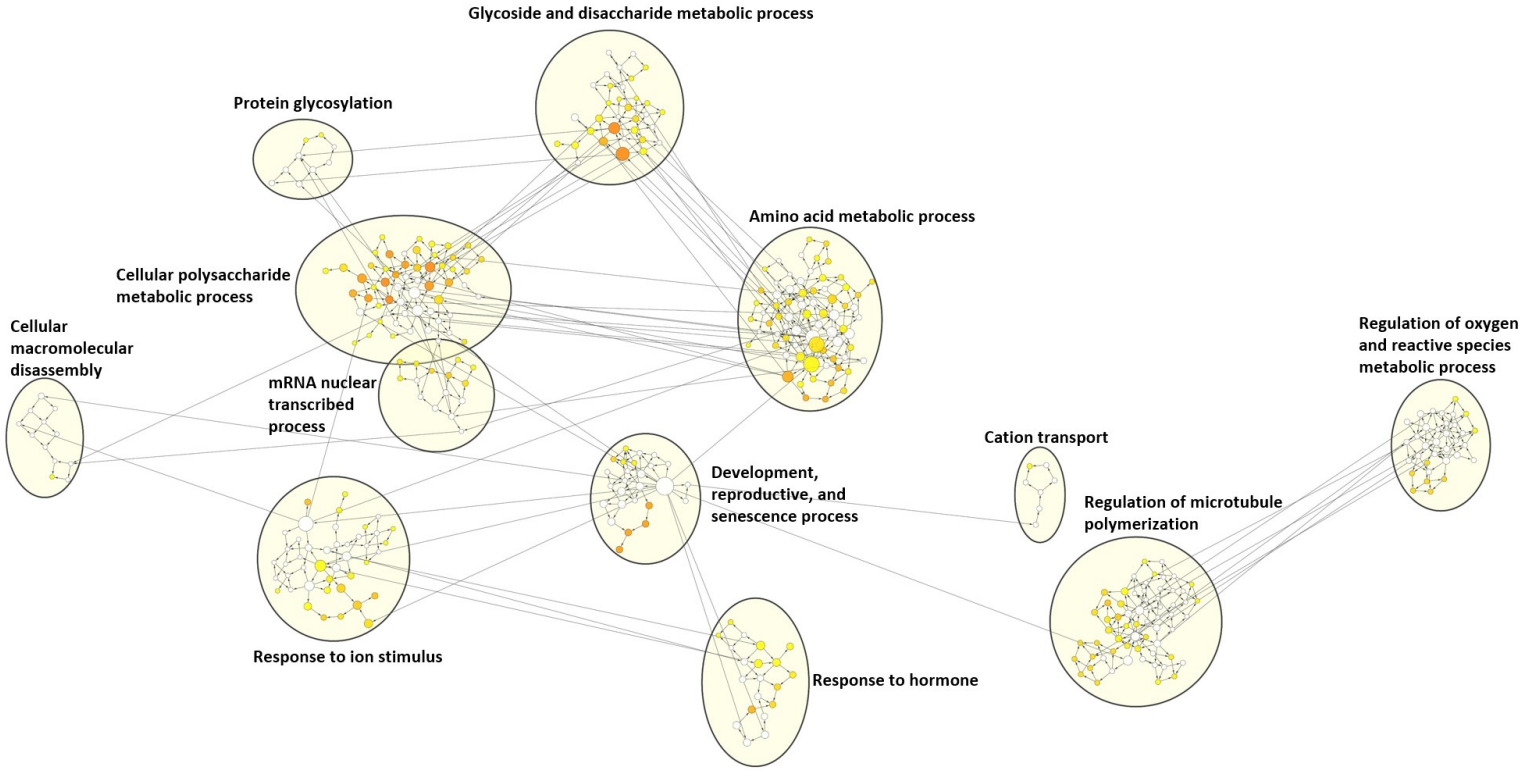

Figure S5-8 Clustering annotation of *P. cheesemanii* unique gene set in responding to salt stress.

90 genes of *P. cheesemanii* unique cold response set scattered in 391 nodes, then these nodes were grouped into 12 clusters. Nodes: overrepresented terms of GO biological process; colour: significant *p*-value. Higher significant *p*-value, the node colour gets increasingly more orange. Uncoloured nodes are not overrepresented, but they are the parents of overrepresented categories.

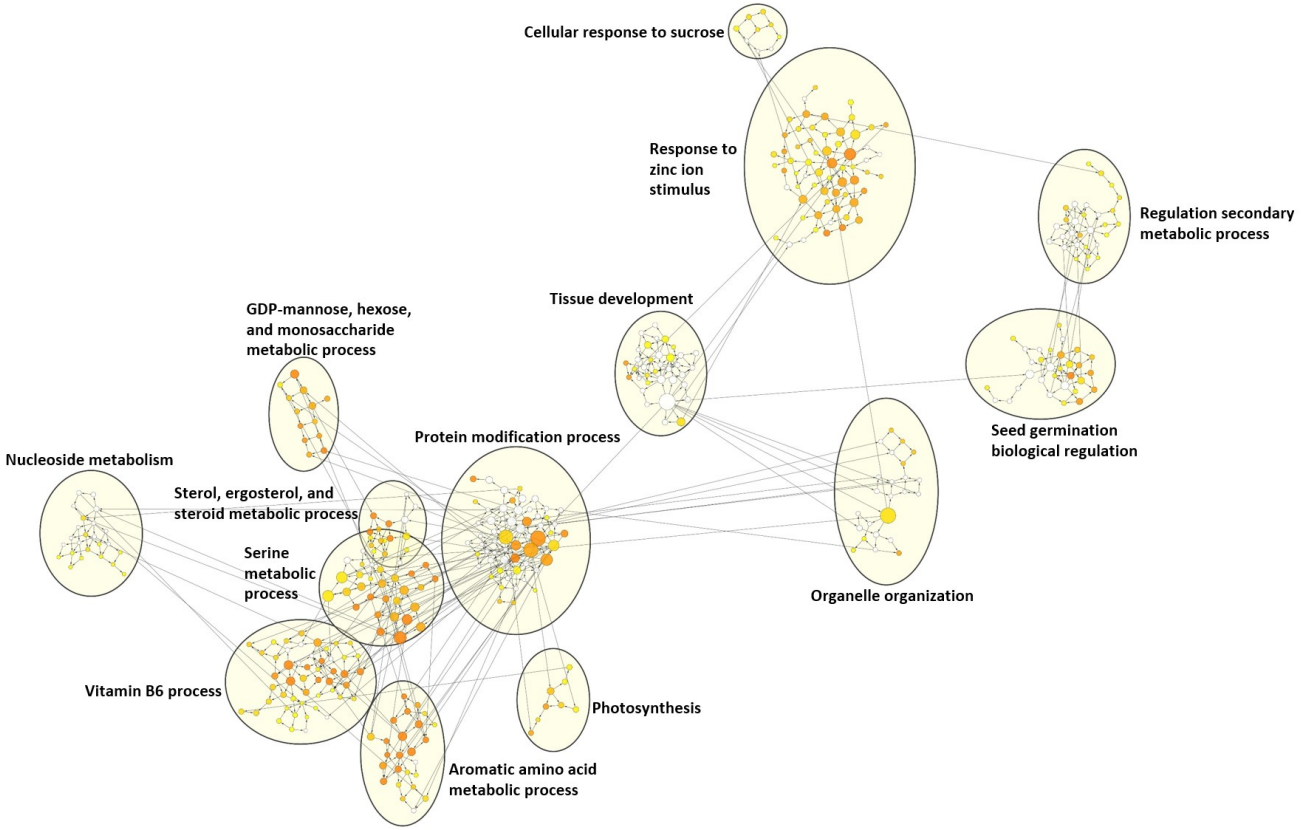

Figure S5-9 Clustering annotation of *P. cheesemanii* unique gene set in responding to UV-B radiation.

109 genes of *P. cheesemanii* unique cold response set scattered in 416 nodes, then these nodes were grouped into 14 clusters. Nodes: overrepresented terms of GO biological process; colour: significant *p*-value. Higher significant *p*-value, the node colour gets increasingly more orange. Uncoloured nodes are not overrepresented, but they are the parents of overrepresented categories.

Appendix 2. Supplementary Files

Supplementary File 2-1 CTAB approach to isolate nuclear enriched genomic DNA for next generation sequencing of Pachycladon.

Supplementary File 2-2 Assembled *Pachycladon cheesemanii* genome

Supplementary File 2-3 Statistics of different assemblies for *Pachycladon cheesemanii*.

Supplementary File 2-4 RNA extraction protocol.

Supplementary File 2-5 Sequences of genes used for RT-qPCR analysis.

Supplementary File 2-6 List of primers used for RT-qPCR analysis.

Supplementary File 2-7 RT-qPCR protocol.

Supplementary File 2-8 Code.

Supplementary File 3-1 Repeats content estimation and classification for *Pachycladon cheesemanii*.

Supplementary File 3-2 Distribution of Simple Sequence Repeats (SSR) markers identified using MISA.

Supplementary File 3-3 KEGG pathway annotation for all genes obtained from KEGG Automated Annotation Server (KAAS).

Supplementary File 3-4 Gene Ontology annotation for all genes obtained from BLAST similarity searches.

Supplementary File 4-1 Assembled *Pachycladon cheesemanii* transcriptome.

Supplementary File 4-2 Gene Ontology annotation of *Pachycladon cheesemanii* transcriptome for all genes obtained from BLAST similarity searches.

Appendix 3. Publications and Presentations

Publications

Dong, Y., Gupta, S., Sievers, R., Wargent, J. J., Wheeler, D., Putterill, J., ... & Dijkwel, P. P. (2019). Genome draft of the Arabidopsis relative *Pachycladon cheesemanii* reveals novel strategies to tolerate New Zealand's high ultraviolet B radiation environment. *BMC genomics, 20*(1), 1-14.

Presentations

2018 Oral presentation, Workshop on Plant Systems Biology, Bulgaria

2017 Poster presentation, Lorne Conference, Australia

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Yanni Dong
Name/title of Primary Supervisor:	Associate Professor Paul Dijkwel
In which chapter is the manuscript /published work:	chapter 3 and 2

Please select one of the following three options:

(●) The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

Dong, Y., Gupta, S., Sievers, R., Wargent, J. J., Wheeler, D., Putterill, J., ... & Dijkwel, P. P. (2019). Genome draft of the Arabidopsis relative Pachycladon cheesemanii reveals novel strategies to tolerate New Zealand's high ultraviolet B radiation environment. BMC genomics, 20(1), 1-14.

( ) The manuscript is currently under review for publication – please indicate:

- The name of the journal:

- The percentage of the manuscript/published work that was contributed by the candidate:

- Describe the contribution that the candidate has made to the manuscript/published work:

( ) It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Yanni dong Digitally signed by Yanni dong Date: 2020.11.01 13:20:50 +13'00'
Date:	01-Nov-2020
Primary Supervisor's Signature:	Paul Dijkwel Digitally signed by Paul Dijkwel DN: cn=Paul Dijkwel, c=NZ, email=p.dijkwel@massey.ac.nz Date: 2020.11.06 16:56:16 +13'00'
Date:	6-Nov-2020

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

GRS Version 5 – 13 December 2019
DRC 19/09/10