Mechanisms regulating inflorescence development and flowering traits in *Arabis alpina*, an alpine perennial

Inaugural-Dissertation

Zur

Erlangung des Doktorgrades

Der Mathematisch-Naturwissenschaftlichen Fakultät

Der Universität zu Köln

Vorgelegt von

Evelyn Obeng-Hinneh

Aus Ghana

Köln, March 2018

Die Vorliegende Arbeit wurde am Max-Planck Institute für Pflanzenzüchtungsforschung in Köln, im Labor von Prof. Dr. Maria Albani angefertigt.

Berichterstatter: Prof. Dr. Maria Albani

Prof. Dr. Ute Hoecker

Prüfungsvorsitz: Prof. Dr. Wolfgang Werr

Tag der mündlichen Prüfung: 14.05.2018

Dedication

There is a time for everything, and a season for every activity under the sky

Ecclesiastes 3

In loving memory of Janet Akosua Gyinaah

Abstract

Flowering and inflorescence development are important plant processes that determine reproductive success, thus survival of many plant species. These processes are controlled by both endogenous and environmental cues. *Arabis alpina*, an alpine perennial and a close relative of the annual species *Arabidopsis thaliana*, initiates flower buds during prolonged cold exposure. Flower emergence then occurs a season later during permissive growth conditions. The Pajares accession of *A. alpina* was used in this study to investigate cold duration effects. Extended vernalization accelerated flower emergence, increased the percentage of flowering inflorescence branches and suppressed floral reversion within the inflorescence. The *A. alpina* gene, *PERPETUAL FLOWERING 1* (*PEP1*) is a floral repressor that contributes to perennial traits and the orthologue of the *A. thaliana* gene, *FLOWERING LOCUS C* (*FLC*). The *pep1* mutant does not require vernalization to flower and exhibits compromised perennial traits. Second site enhancer mutants of *pep1-1, enhancers of perpetual flowering 1* (*eop*), were isolated to identify additional regulators of flowering. Five selected mutants are characterized here for their early flowering and inflorescence phenotypes. One mutant, *eop101*, developed a determinate inflorescence with a terminal flower as it carried lesion in the *A. alpina TERMINAL FLOWER 1*. Genome wide transcriptome analysis on this mutant revealed the up regulation of some flower meristem and organ identity genes and transcription factors associated with regulation of circadian rhythms and flowering. The causal gene of the other *eops* has been identified as a member of an AAA+ ATPase family whose involvement in flowering and inflorescence development has so far not been reported. Characterization of *eop* in a *PEP1* background using introgression lines (ILs), revealed *EOP* to be involved in flowering and inflorescence development in response to vernalization duration. ILs exhibited early saturation of vernalization requirement to accelerate flowering, reduced floral reversion and increased percentage of flowering branches. Furthermore, transcriptome analysis suggests the involvement of *EOP* in stress response, reproductive development and transport of lipids and oligopeptides.

Zusammenfassung

Blüte- und Blütenstandentwicklung sind wichtige Prozesse in der Pflanze, die Reproduktivität, und damit das Überleben vieler Pflanzenarten bestimmen. Diese Prozesse werden sowohl durch endogene Reize als auch durch Umwelteinflüsse kontrolliert. In *Arabis alpina*, einer mehrjährigen Pflanze und nahen Verwandten der einjährigen Pflanze *Arabidopsis thaliana*, wird das Knospen von Blüten durch langzeitige Kälteeinwirkung ausgelöst. Die Blüte findet dann eine Saison später unter günstigen Umweltbedingungen statt. In der vorliegenden Arbeit wurde der *A. alpina* Ökotyp Pajares verwendet, um die Effekte anhaltender Kälte zu untersuchen. Verlängerte Vernalisationszeiten beschleunigten die Blüte, erhöhten die Prozentzahl an blühenden Nebenblüten, und unterdrückten die Blütenreversion innerhalb des Blütenstands. Das *A. alpina* Gen *PERPETUAL FLOWERING 1* (*PEP1*) ist ein Ortholog des *A. thaliana* Gens *FLOWERING LOCUS C* (*FLC*), und stellt einen Inhibitor in der Blüte dar, der zur Ausprägung von Eigenschaften mehrjäriger Pflanzen beiträgt. Die *pep1* Mutante benötigt keine Vernalisation, um die Blüte einzuleiten, und zeigt kompromittierte Mehrjährigen-Eigenschaften. *Second site enhancer* Mutanten von *pep1-1, enhancers of perpetual flowering 1* (*eop*), wurden isoliert, um weitere Blüh-Regulatoren zu identifizieren. Fünf ausgewählte Mutanten wurden in dieser Arbeit aufgrund ihres verfrühten Blühzeitpunktes und ihrer Blütephänotypen charakterisiert. Eine Mutante, *eop101*, entwickelte einen geschlossenen Blütenstand mit einer endständigen Blüte als Resultat einer Deletion im *A. alpina* Gen *TERMINAL FLOWER 1*. Eine Genomweite Transkriptionsanalyse mit dieser Mutante zeigte die erhöhte Expression einiger Blütemeristem- und Organidentitätsgene, und von Transkriptionsfaktoren, die mit der Regulierung des zirkadianen Rhythmus und der Blüte assoziiert sind. Das den anderen *eops* Mutanten zugrunde liegende Gen wurde hier als ein der AAA+ ATPase-Familie zugehöriges Gen identifiziert. Eine Beteiligung dieser Genfamilie bei der Blüte- und Blütenstandentwicklung wurde bisher noch nirgends beschrieben. Die Charakterisierung von *eop* im *PEP1*-Hintergrund mithilfe von introgression lines (ILs) zeigte, dass *EOP* abhängig von der Vernalisationsdauer an der Blüte- und Blütenstandentwicklung beteiligt ist. Die ILs zeigten eine verfrühte Saturierung der Vernalisationsvoraussetzungen, um die Blüte zu beschleunigen, eine verringerte Blütereversion, und eine erhöhte Prozentzahl von Blütezweigen. Weiterhin weist die Transkriptionsanalyse auf eine Beteiliung von *EOP* an der Stressantwort, der Reproduktionsentwicklung, und am Lipid- und Oligopeptidtransport hin.

Table of contents

List of figures

List of tables

List of abbreviations

°C	Degree Celsius
AAA+	ATPases Associated with diverse cellular Activities
ABA	Abscisic acid
ABFs	ABA responsive elements-binding factors
AF	Allele frequency
AGL24	*AGAMOUS-LIKE 24*
Ala (A)	Alanine
AP1	*APETALA1*
ART1	*AERIAL ROSETTE 1*
Asp (D)	Aspartic acid
ATP	Adenosine triphosphate
AUX1	*AUXIN RESISTANT 1*
AXR3	*AUXIN RESISTANT 3*
BAM2	*BARELY ANY MERISTEM 2*
B-CHI or *CHI-B*	*BASIC CHITINASE*
bp	basepair
BSA	Bulk segregant analysis
BT2	*BTB AND TAZ DOMAIN PROTEIN 2*
bZIP	leucine zipper
CAL	*CAULIFLOWER*
CaMV	Cauliflower Mosaic Virus
CAPS	Cleaved amplified polymorphic
CCA1	CIRCADIAN CLOCK ASSOCIATED 1
CCD8	*CAROTENOID CLEAVAGE DIOXYGENASE 8*
CDF1	*CYCLING DOF FACTOR 1*
cDNA	complementary DNA
Chr.	Chromosome
CO	*CONSTANS*
CRC	*CRABS CLAW*
CRY1	CRYPTOCHROME1
CRY2	CRYPTOCHROME2
CYP707A3	*CYTOCHROME P450, FAMILY 707, SUBFAMILY A, POLYPEPTIDE 3*
CYP79B	*CYTOCHROME P450, FAMILY 79, SUBFAMILY B, POLYPEPTIDE*
CYP83B1	*CYTOCHROME P450, FAMILY 83, SUBFAMILY B, POLYPEPTIDE 1*
dCAPS	Derived cleaved amplified polymorphic
DEGs	Differentially expression genes
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleic triphosphate
DREB2A	DRE-BINDING PROTEIN 2A
dsRNAi	doublestrand-RNA-interference
ELF4	EARLY FLOWERING 4
EMB	*EMBRYO-DEFECTIVE*
EMF1	*EMBRYONIC FLOWER 1*
EMS	Ethyl methanesulfonate
eop	*enhancers of perpetual flowering 1*
EPR1	*EARLY-PHYTOCHROME-RESPONSIVE1*

ERF8	*ETHYLENE RESPONSE FACTOR 8*
ERL1	*ERECTA-LIKE 1*
ET	Ethylene
FCA	*
FD	*
FLC	*FLOWERING LOCUS C*
FLD	*FLOWERING LOCUS D*
FLK	*FLOWERING LATE WITH KH MOTIFS*
FPA	*
FRI	*FRIGIDA*
FT	*FLOWERING LOCUS T*
FUL	*FRUITFULL*
FVE	*
FY	*
GA	Gibberellic acid
GAI	GA INSENSITIVE
GFP	Green fluorescent protein
Glu (E)	Glutamic acid
Gly (G)	Glycine
GO	Gene Ontology
H3K27	Histone H3 lysine 27
H3K27me3	H3 trimethyl-K27
H3K9	Histone H3 lysine 9
HB51	*HOMEOBOX 51*
HOS1	*HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENE 1*
hr	Hour
HY5	*ELONGATED HYPOCOTYL 5*
IAA	Indole-3-acetic acid
IAOx	Indole-3-acetaldoxime
ICU2	*INCURVATA2*
ILs	Introgression lines
JA	Jasmonic acid
JAZ	*JASMONATE-ZIM-DOMAIN PROTEINs*
kb	Kilobase
L	Liter
LD	Long day
LD	*LUMINIDEPENDENS*
Lf	*
LFY	*LEAFY*
LHY	*LATE ELONGATED HYPOCOTYL*
Lys (K)	Lysine
M	Mol
MAF1	MADS Affecting Flowering 1
Mb	Megabases
mg	Milligram
miR156	microRNA156
miR172	microRNA172
ml	Milliliter
mRNA	messenger RNA

NGS	Next-generation sequencing
PCR	Polymerase chain reaction
PEP1	*PERPETUAL FLOWERING 1*
PHY	PHYTOCHROME A
PI	*PISTILLATA*
PIF5	*PHYTOCHROME-INTERACTING FACTOR 5*
PR	Potence ratio
Pro (P)	Proline
PRR	*PSEUDO-RESPONSE REGULATOR*
PTF1	*PLASTID TRANSCRIPTION FACTOR 1*
RFP	Red fluorescent protein
RGA	REPRESSOR OF GA1-3
RGL	RGA-LIKE
RNA	ribonucleic acid
RNA-seq	RNA sequencing
RT-qPCR	Reverse transcription quantitative PCR
SAM	shoot apical meristem
SBPs	SQUAMOSA-BINDING PROTEINS
SD	Short day
SDD1	*STOMATAL DENSITY AND DISTRIBUTION 1*
SEP2	*SEPALLATA 2*
SEP4	*SEPALLATA 4*
Ser (S)	Serine
SMZ	*SCHLAFMUTZE*
SNP	Single nucleotide polymorphism
SOC1	*SUPRESSOR OF OVEREXPRESSION OF CONSTANS1*
SPA1	*SUPPRESSOR OF PHYA-105 1*
SPLs	SQUAMOSA PROMOTER BINDING LIKE
SUR1	*SUPERROOT 1*
SUR2	*SUPERROOT 2*
SVP	*SHORT VEGETATIVE PHASE*
TEM1	*TEMPRANILLO 1*
TF	Terminal flower
TFL1	*TERMINAL FLOWER1*
Thr (T)	Threonine
TMM	*TOO MANY MOUTHS*
TOC1	TIMING OF CAB1 EXPRESSION
TSF	*TWIN SISTER OF FT*
VRN	*VERRNALIZATION*
WAV	Weeks after vernalization
WLD	Weeks in LD
WV	Weeks in vernalization
µg	Microgram
µM	microMolar

1 Introduction

1.1 Flowering and inflorescence development in higher plants

One of the most complicated and yet fascinating processes in the life cycle of higher plants is flowering. This process marks a major transition in the plants life as a change from vegetative development to reproductive development occurs. The shoot apical meristem (SAM), at this point, switches from the production of vegetative structures, such as leaves and axillary branches, to the production of flowers. Within a plant's architecture, the region that tends to bear clusters of flowers is referred to as the inflorescence [1]. Thus, the SAM changes from being a vegetative meristem to an inflorescence meristem. This developmental switch leads to different morphogenetic changes in a flowering plant including the modification of plant architecture. The inflorescence part of the plant, in most cases, can be distinguished from the vegetative region by means of differences in their foliage which includes changes in leave morphology or the complete lack of leaves [1, 2]. In the model plant species *Arabidopsis thaliana*, the shoot apical meristem undergoes transition from a vegetative phase, where nodes are closely compressed and petiolated rosette leaves are produced, to a reproductive phase which involves the elongation of the main stem to give rise to the inflorescence, a process referred to as bolting [3, 4]. The inflorescence comprises of an early inflorescence zone (I1) and a late inflorescence zone (I2) (Figure 1). The I1 zone contains inflorescence branches subtended by modified leaves and nodes separated by internode elongation whereas the I2 zone consists of solitary flowers with no subtending leaves (Figure 1) [5-7]. The modified leaves within I1 are referred to as cauline leaves and unlike the rosettes, these are smaller non-petiolated leaves that are directly attached to the stem. This change in leaf shape corresponds to early inflorescence and thus can serve as a morphological marker in determining inflorescence zones [8]. The inflorescence form or architecture plays determinant roles in reproductive success since it can affect pollination and fruit set [9, 10]. The forms of inflorescence can be classified based on several criteria, however a main parameter considered for classification is whether or not the shoot apices ends with a terminal flower [1, 2, 11, 12]. An inflorescence that does not terminate is classified as indeterminate. For this type of inflorescence, the apical meristem has the ability to grow indefinitely while producing floral

meristems from its periphery. A common example here, is the typical inflorescence of *A. thaliana* [4, 13, 14]. On the contrary, an inflorescence meristems can be converted to a floral identity resulting in the production of a terminal flower and these are classified as determinate inflorescence. An example here is the inflorescence of *Solanum lycopersicum* (tomato) [15-17]. In several plants species, the proper timing of flowering is of great importance to ensure reproductive success. In *A. thaliana* and many other species, transient exposure of plants to floral inductive signals are usually sufficient to cause stable floral transition [18]. The ability of the plants to continue flowering even after removal of these inductive signals is referred to as floral commitment [19]. In some other species and some *A. thaliana* mutants, the switch of the SAM from vegetative growth to reproductive development is not always an irreversible process, once flowering commences. This switch to reproductive development and the development of initiated primordia at the SAM are controlled by both endogenous and environmental signals [20, 21]. At the point of flowering induction, changes in environmental factors can lead to floral and inflorescence reversion [22, 23]. This phenomenon of inflorescence or floral reversion refers to the return to vegetative development after, or the intercalation of a vegetative development during, inflorescence development [22]. This points out the fact that, apart from flowering initiation, maintenance is also required. Thus, to maintain reproductive development and ensure survival, plants have evolved mechanisms to control the proper timing of flowering and inflorescence development in response to environmental and endogenous cues. This thesis focuses on the traits of flowering and inflorescence development in the perennial species *Arabis alpina*. As the control of these traits have been extensively studied in *A. thaliana,* a close annual relative of *A. alpina*, a brief description of what has so far been discovered is first given.

1.2 Control of flowering

The mechanisms adopted by plants to regulate flowering are considered to be involve in a complex genetic network of signaling pathways. Extensive efforts have been made to describe the flowering time pathways in *A. thaliana* and these have revealed the presence of a photoperiod pathway, vernalization pathway, gibberellins (GA) pathway, autonomous pathway and the age pathway (Figure 1) [24, 25]. Among the flowering time pathways, GA, autonomous and age react to endogenous cues while photoperiod and vernalization respond to environmental cues [25]. The availability of naturally occurring *A. thaliana* ecotypes with variations in flowering time and the broad spectrum of flowering time mutants have largely driven genetic analysis and aided the identification of genes involved in the activities of these pathways [26, 27].

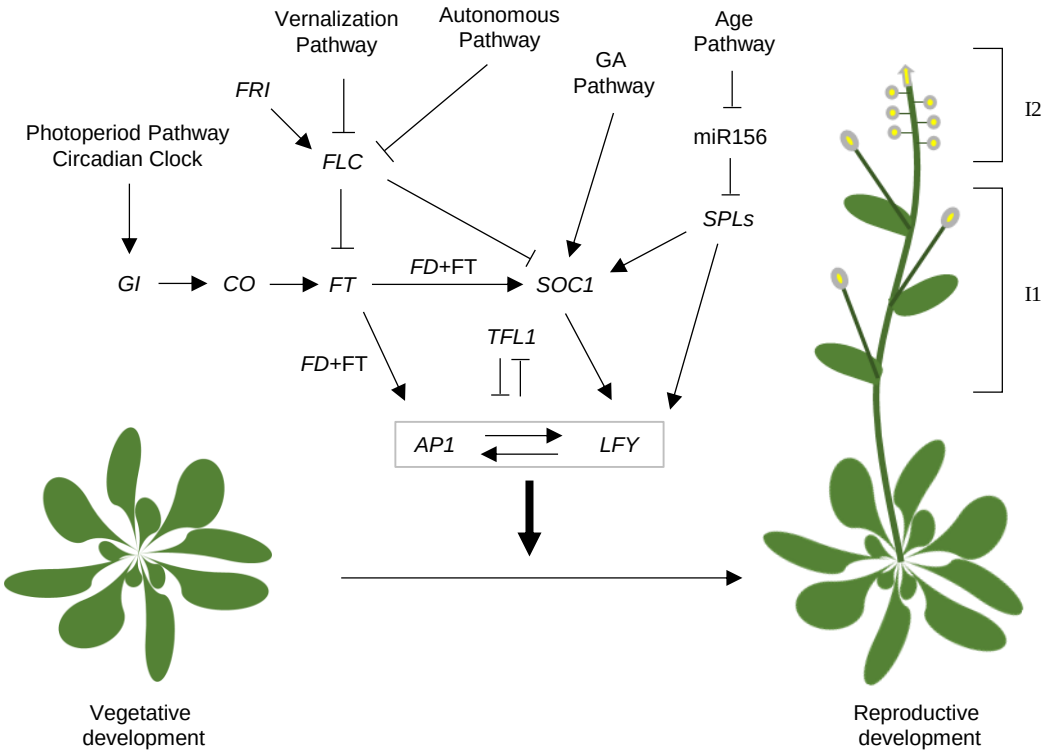


Figure 1. Simplified diagram of the various flowering pathways in Arabidopsis.
Five independent pathways have been identified to control the transition from vegetative to reproductive development. The main stem during reproductive development produces an inflorescence with an I1 zone where inflorescence branches are produced and an I2 zone which consists of solitary flowers.

1.2.1 The age pathway

In several plant species a certain age or size must be attained in order to flower. During vegetative development, the plant shoot passes through a phase referred to as the juvenile phase. At this period the plant is considered incompetent to respond to internal and external signals that normally induce flowering. The tendency of the plant to perceive and react to these floral inducing signals increases with age and the shoot is described as acquiring the competence to flower. The phase at which plants are able to flower in response to these signals is thus referred to as the adult phase [28]. The juvenile phase, depending on the species, can last for a few days to several years. The two phases are distinguishable from each other by unique vegetative traits such as changes in the morphology and physiology of leaves and the orientation of branch growth [28-30]. Recent molecular studies have discovered some underlying mechanisms regulating the vegetative phase changes and competence to flower. The SQUAMOSA PROMOTER BINDING LIKE (SPL) transcription factors are regarded to control a wide range of developmental processes including flowering time and inflorescence development [31, 32]. They were originally isolated from *Antirrhinum majus* and referred to as SQUAMOSA-BINDING PROTEINS (SBPs) [31]. In *A. thaliana*, genes that encoded SBPs related proteins were also isolated and referred to as SPLs [33]. The microRNAs, miR156 and miR172 have become an interesting point of focus in the age pathway as age of the shoot influences their abundance and they also are able to influence flowering time [34]. There are 16 genes in the *Arabidopsis* SPL family, 11 of these have been reported to contain microRNA recognition sites [35]. Out of the 11 SPLs, five members, SPL3, SPL4, SPL5, SPL9 and SPL15 are strongly associated with floral transition and contain miR156 recognition sequences, which inhibits their activities [34]. The mRNA levels of *SPL3*, for example, were demonstrated to increase during floral transition and overexpression caused early flowering [33]. Overexpression of miR156 inhibited the mRNA translation of *SPL3* and the early flowering caused by *SPL3* overexpression was enhanced with the removal of the miR156 recognition sequence [36]. Concentration of SPLs increases as the plant ages and approaches the adult phase, on the other hand, aging of the plant and phase changes results in a decrease in the expression levels of miR156. Overexpression of miR156 caused a delay in juvenile to adult phase transition [36]. In contrast to miR156, the abundance of miR172 increases progressively with plant age and is an activator of flowering and floral development as it targets AP2-like transcription factors that are likely repressors of floral transition [37]. The activities of miR172 are downstream that

of miR156/SPL, in that while high levels of miR156 reduces it expression, the SPLs, especially SPL9 and SPL15, activates its transcription [38-40]. Thus the aging of plants reduces miR156 levels and permits an increase in expression levels of the SPLs, these also activate transcription of *MIR172b* and together they control vegetative phase changes and flowering competence.

1.2.2 The GA pathway

The phytohormone GA is involved in multiple developmental processes in plants including the promotion of flowering in *A. thaliana* [41, 42]. Langridge, 1957, by the exogenous application of GA demonstrated the ability of this hormone to promote flowering [43]. Following this initial research, extensive studies have been conducted using mutations that cause disruptions in GA biosynthesis, signaling or increase the degradation of GA. These mutations not only cause a delay in flowering, which is most evident under non-inductive photoperiods, but also affect other aspects of plant growth and development such as, germination, stem elongation and floral development. *GA1* for instance, encodes the enzyme ent-kaurene synthetase A which catalyses the first committed step in GA biosynthesis [44]. *A. thaliana ga1* mutants flower later than wild type in long day conditions and are unable to flower under short day conditions [41]. The mutants can also exhibit male-sterility, dwarfism and in the case of more severe alleles non-germinating phenotypes [45, 46]. These mutant phenotypes can however be converted to wild type phenotypes by the continuous application of GA [45]. DELLA proteins are considered as key negative regulators of GA signaling and the genome of *A. thaliana* encodes five DELLA proteins, GA-INSENSITIVE (GAI), REPRESSOR OF GA1-3 (RGA) and three RGA-LIKE proteins (RGL1, RGL2 and RGL3), which have distinct but also overlapping functions [47]. These proteins act as repressors of almost all studied GA-dependent processes, however GA is able to relieve the repressive activities of these proteins [48-50]. It was hypothesized that GAI and RGA might function as "GA-derepressible repressors," which was based on a prediction that in the absence of GA, GA responses can still occur once *GAI* and *RGA* genes are inactivate [51]. Accordingly, inactivation of *RGA* and *GAI* in the *ga1* mutant suppressed the nonflowering phenotype of *ga1,* as the triple mutant flowered slightly earlier than the wild type, and rescued the stem growth phenotype of the *ga1-3* mutant [51, 52]. *RGA* and *GAI* are therefore considered as repressors of vegetative growth and floral induction. Not all aspects of the *ga1* phenotype was rescued by the *rga* and *gai* mutations. Seed germination and floral development mutant phenotypes were not restored in the triple mutant. This points

out the possible regulation of these process by the *RGLs*. Using a loss-of-function *rgl1* line, it was observed that the mutants displayed GA-independent activation of stem elongation, seed germination and floral development [53]. Thus, compared to *GAI* and *RGA*, *RGL1* plays greater roles in seed germination and floral development. The expression profile of *RGL1* transcripts was predominantly detected in the inflorescence, more specifically in developing ovules and anthers of flowers as compared to the other DELLAs, indicating that *RGL1* functions as a negative regulator of GA response within the inflorescence [53].

1.2.3 The photoperiodic pathway

Photoperiodism can be defined as the response of living organisms to the lengths of daily light period which enables them adapt to seasonal changes within their environment. Plants are able to detect light by the use of photoreceptors such as cryptochromes for blue light and ultraviolet-A detection, and phytochromes for red (R) and far-red (FR) light detection [54]. There are at least five phytochromes in *A. thaliana*, from PHYA to PHYE, and two crytochromes, CRY1 and CRY2 [54]. The role of photoperiod in flowering begins with the photoreceptors, as they act together to initiate signals that interact with the entrain circadian clock [55]. The circadian clock measures the length of the daily light period in leaves, this is then followed by the convey of a signal to the shoot apex which results in the initiation of floral transition [56]. The circadian rhythms have a period length of about 24 hours and have so far been observed in leaf movement, stomata aperture, gene transcription and many other plant processes [57, 58]. The internal mechanism that generates these rhythms can be organized into three components, which includes, the central oscillators which generate the mechanism of 24 hours, the input pathways involved in the synchronization of clock oscillators to the daily cycles of light and dark and the output pathways which are involved in the regulation of specific processes [59]. There are at least four proteins which function as the core of the circadian oscillator, these are CIRCADIAN CLOCK ASSOCIATED 1 (CCA1), LATE ELONGATED HYPOCOTYL (LHY), TIMING OF CAB1 EXPRESSION (TOC1) and EARLY FLOWERING 4 (ELF4). While the circadian regulated mRNA expressions of *CCA1* and *LHY* peaks at dawn, those of *TOC1* and *ELF4* peaks at dusk [60, 61]. These proteins have been demonstrated to participate in a feedback loop of about 24 hours to regulate their own expressions which also determines the time of day flowering time genes in the long day pathway such as *CONSTANT* (*CO*), *GIGANTEA* (*GI*) and *FLOWERING LOCUS T* (*FT*) are expressed [60-63]. Genetic analysis

have demonstrated that *GI* activates the transcriptional regulator *CO* which in turn activates a key floral integrator gene, *FT* in the leaves during long days [64]. FT protein then moves to the meristems where it activates the program of floral development by activating meristem identity genes such as *LEAFY* (*LFY*) and *APETALA1* (*AP1*). This results in change of the shoot apical meristem fate from vegetative to floral [18]. Based on their photoperiod responses plants can be classified as long day (LD), short day (SD) or day neutral. *A. thaliana* for example is a facultative long day plant, thus it flowers earlier under LD conditions and much later in short days. Other plants like rice, are regarded as short day plants as they flower early in SD [65] and then there are day neutral plants like tomato which flower independently of day length [66].

1.2.4 The vernalization pathway

For many plant species grown in temperate climates, periods of low temperatures, such as during winter, are required to enable them flower. The exposure of plants to prolong periods of cold temperatures in order to promote flowering is known as vernalization. This is a useful adaption especially for plants that flower in spring. The process of vernalization requires active metabolism, therefore dry seeds for instance cannot be vernalized. Imbibed seeds of many cereals, on the other hand, have been shown to be responsive to vernalization and in some species, like rye, the embryo can also be vernalized [67, 68]. In several other plants response to vernalization is age dependent, thus they are not able to respond to vernalization until a critical age or stage of development is reached [67]. *FLOWERING LOCUS C* (*FLC)* encodes a MADS-domain containing transcription factor and is one of the central regulators of vernalization response in *A. thaliana* [69]. The competence of a meristem to respond to promotive floral signals is antagonized by high levels of *FLC* expression. It acts as a repressor of flowering by repressing floral pathway integrators such as *FLOWERING LOCUS T* (*FT*) and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS* (*SOC1*) [69-71]. *FRIGIDA* (*FRI)* is a gene that encodes a protein with two potential coiled-coil domains that is considered to interact with other nucleic acids or proteins [72, 73]. *FRI* confers a vernalization requirement through the uprequlation of *FLC* expression. Vernalization is able to promote flowering by reducing the transcript levels of *FLC*, thus antagonizing the effects of *FRI* [74, 75]. Upon the down regulation of *FLC* transcript levels by vernalization, they remain reduced throughout subsequent development, suggesting vernalization causes a mitotic stability on an epigenetic basis [76].

The mechanisms by which exposure to low temperatures causes reductions in *FLC* mRNA levels has been under investigations in recent years. So far, the stable repression of *FLC* by vernalization has been associated with repressive modifications to *FLC* chromatin, which involves histone modifications such as methylation of histone H3 lysine 9 (H3K9) and lysine 27 (H3K27) [76-78]. Enrichment of H3 trimethyl-K27 (H3K27me3) across the *FLC* locus, for instance, is initiated during cold and further increases after return of plants to ambient temperatures [78]. *VERRNALIZATION* (*VRN*) genes involved in the vernalization induced reduction of *FLC* transcript levels have also been genetically identified through the isolation of mutants. These mutants exhibited reductions in the acceleration of flowering caused by vernalization [79]. In both *vrn1* and *vrn2* mutants, *FLC* transcript levels were normally reduced in response to vernalization, however, once plants were returned to warm conditions *FLC* mRNA levels increased again instead of remaining low [80]. This indicates, *VRN1* and *VRN2* are required to maintain the stable repression of *FLC* after vernalization when plants are returned to warm temperatures to enable follow up developmental processes such as flowering.

1.2.5 The autonomous pathway

The identification of flowering time mutants that flowered late under all photoperiods and were strongly responsive to vernalization lead to the discovery of the autonomous pathway [81]. These mutants, which included, *fca*, *fpa*, *fy*, *fve*, *ld*, *fld*,and *flk*, exhibited similarities in the fact that they all contained high levels of *FLC* mRNA compared to the wild type indicating, the autonomous promotive pathway acts redundantly with the vernalization promotive pathway [82-85]. However, in contrast to the vernalization pathway in which its components repress *FLC* during prolonged low temperatures and also maintain the repression even after return to warm temperatures, the autonomous pathway, at late development, represses *FLC* in response to internal signals and involves a collection of genetic components that target the regulation of *FLC* at different levels [86]. Depending on the naturally occurring *FLC* allele, interaction with the autonomous components may differ. The *FLC* allele present in *Landsberg erecta* is sensitive to loss-of-function alleles in *fca*, *fy* and *fpa* as these mutations cause late flowering in this *A. thaliana* ecotype. On the other hand *ld* mutation causes no late flowering in Landsberg *erecta* but does so in the Columbia ecotype which suggests Columbia is sensitive to *ld* mutations [87, 88]. By controlling FLC, the autonomous pathway indirectly promotes flowering by enabling responsiveness to floral inductive cues.

1.3 Integration of pathways that regulate flowering

The different genetic signaling pathways, previously described, integrate by converging to regulate the expressions of overlapping set of common downstream targets. For instance, within the floral primordia, these pathways lead to the transcriptional regulation of the same group of floral identity genes [89]. Considering the photoperiod and GA pathways, these increase the expression levels of a group of genes identified as the floral integrators, which includes genes such as *FT, TWIN SISTER OF FT* (*TSF*), *SOC1*, *LFY* and *AGAMOUS-LIKE 24* (*AGL24*) [19]. Activities of these genes promote the conversion of the shoot apical meristem from a vegetative to a reproductive fate. Another group of genes have been identified as enablers of floral transition and pathways involved here are considered to regulate expression of floral repressors. The activities of these enabling genes therefore antagonize activities of floral integrators. Genes here include *FLC*, *TERMINAL FLOWER 1* (*TFL1*), and *SHORT VEGETATIVE PHASE* (*SVP*) [19, 90].

The photoperiod and *FLC*-mediated pathways at some point converge to regulate expression of a few targets which includes *FT* and *SOC1*. One of the earliest molecular markers for the transition from vegetative meristem to reproductive meristem is expression of the floral integrator *SOC1*. Upon floral induction the expression of *SOC1* shows a sharp increase in the apex [91]. This early activation is induced by long day conditions and dependent on *FT* which after interacting with the bZIP transcription factor, FD, in the SAM, directly regulate gene expression and activate meristem identity genes such as *APETALA1* (*AP1*) [69, 92]. *FLC*, involved in both the vernalization and autonomous pathways, is able to delay flowering without causing a reduction in the expression of *CO*, in a similar manner, *CO* does not affect the expression of *FLC* [93]. However, genes involved in flowering time have so far been identified that act downstream of *CO* and *FLC*. Overexpression of *CO* causes activation of *SOC1* but this activation was completely blocked by the overexpression of *FLC* [70]. *FT* expression was found to reduce with mutations that disrupted functions of the long-day and autonomous pathway and by the direct binding of FLC protein to first intron of *FT*, while overexpression of *CO* activated its expression [91, 94, 95]. Thus *SOC1* and *FT* are both activated by *CO* and repressed by *FLC*. *SOC1* is also regulated by the GA pathway as the flowering time of GA biosynthetic and signaling mutants correlated well with *SOC1* expression levels [96]. Among the floral identity genes, *LFY* expression has been demonstrated to be responsive to both GA

and long-day. *LFY* promoter activities are reduced in *ga1-3* mutant but increases with exogenous application of GA in both the mutant and wild-type [42]. Within the *LFY* promoter, the deletion of an MYB transcription binding site responsible for the GA induced expression, prevented its activation by GA but its activation by long day was intact. The *lfy* mutant with a *LFY* transgene in which the GA-responsive element was deleted produced wild type looking plants under long day but *lfy* mutant phenotype in short days [97]. Thus, both GA and long-day pathways converge on the promoter of the floral identity gene *LFY*.

1.4 The interacting effects of floral and shoot identity genes on inflorescence architecture

The development of floral meristem and inflorescence shoots needs to be regulated such that both can co-exist in close proximity. Several genes, including *LFY*, *AP1* and *TFL1*, have been demonstrated to play important roles in specifying *Arabidopsis* flower and inflorescence morphogenesis. The role of these three genes are the best analyzed in *Arabidopsis* and many other species as they seem to form the backbone of the flower and inflorescence shoot network. Besides its role as a floral integrator, *LFY* plays a major role in flower initiation and floral development and therefore essential for floral meristem and floral organ patterning determinacy [98, 99]. The occurrence of flower reversion into inflorescence shoot development in *lfy* mutants, demonstrates the importance of *LFY* for maintaining floral meristem identity [100]. Another phenotype observed for the *lfy* mutant was the formation of flowers subtended by leaves, that is bracts [101]. This was absent in the wild types and thus suggested a role of *LFY* in the suppression of bracts. The effects of *lfy* mutation are most obvious in the first positions within the inflorescence whereas subsequent structures formed at more apical positions steadily acquired floral identity [98, 99], this can be attributed to other floral meristem identity genes that might be independently activated such as *AP1*. *AP1* gene is essential for the establishment of organ identity, and required for normal development of sepals and petals [102]. In *ap1* mutants flowers, sepals develop as bracts while petal development and differentiation does not occur [102, 103]. Mutation in *AP1* also causes partial transformation of flowers into inflorescence shoots [103]. In agreements with its role in floral meristem identity, overexpression of *AP1* cause early flowering and the conversion of shoots to flowers [104]. The reversion of flowers into inflorescence shoots is much enhanced

in the *lfy* and *ap1* double mutants, this thus indicates some level of redundancy in their functions [98]. *TFL1* is another key gene that affects flowering time and inflorescence architecture in *A. thaliana*. However, its role is opposite to that of *AP1* and *LFY*. As a repressor of flowering, mutations in *TFL1* not only accelerates flowering time but also converts the indeterminate inflorescence of *A.thaliana* into determinate inflorescence and causes the replacement of inflorescence shoots by flowers [5, 6, 105]. *TFL1* therefore specifies shoot identity.

In *A. thaliana*, *LFY* expression is restricted to floral tissues and can also be detected, even though at low levels in leaf primordia at the vegetative phase which increases gradually as floral transition approaches [106, 107]. The expression level of *LFY* in the apex is actually considered as a marker for determining floral transition [106]. *AP1* expression, similar to that of *LFY*, is restricted to emerging floral meristems shortly after *LFY* expression. *LFY* further has the ability to directly activate *AP1* [108]. The expression of *TFL1*, on the other hand, is restricted to the center of the main and axillary inflorescence meristems. The expressions of *TFL1* to those of *LFY* and *API* are therefore complementary with each confined to a specific region. However, upon *TFL1* mutation, *AP1* and *LFY* become expressed in the center of the inflorescence apical meristem. Additionally, overexpression of *TFL1* causes a delay in the upregulation of *LFY* and *AP1* during the transition from vegetative to reproductive development [4, 98, 109]. Therefore while genes such as *LFY* and *AP1* act as floral meristem identity genes, *TFL1* acts as a shoot identity gene. The antagonistic interaction between these two set of identity genes regulates inflorescence architecture.

Apart from the identity genes, some flowering time genes have also been implicated to affect the inflorescence. In *A. thaliana* Sy-0, a late flowering *Arabidopsis* ecotype, interactions of a flower repressor gene *AERIAL ROSETTE 1 (ART1)* with *FLC* and *FRI* produces inflorescence meristems that display reversion of flowering. This phenotype of Sy-0 is attributed to either the inadequacy of floral-promoting signals or lack of competence to respond to these signals at the shoot apical meristem [110]. Flowering time genes in this case, not only affect the transition to flower but also inflorescence morphology.

1.5 The life history strategies of plants; Annuals, biennials and perennials

Plants have evolved different life strategies to aid coordinate their reproductive development with seasonal changes. A major part of this life history strategy is how long the plant takes to complete its life cycle. Based on this, plants can be categorized as being annuals, biennials or perennials. Annuals undergo their entire life cycle from vegetative to reproductive phase followed by senescence and death within a single growing season. Since they have one attempt to reproduce, annuals commit much of their efforts into producing as many seeds as possible. This includes committing all active SAMs to flower production [111, 112]. *A. thaliana* is an example of a well-studied annual plant and senescence in this plant begins close to the end of reproduction, as seeds mature and a certain number of seeds have been produced. There is then, the arrest of meristems and growth is stopped [113, 114]. It is considered that, senescence occurs as a consequence of diversion of plant nutrients from vegetative to reproductive structures to ensure the production of as many offspring as possible. The plants then gets depleted of energy and can no longer sustain growth [113]. The plant life strategies are of considerable importance in cases where winter or vernalization is involved, then a distinguishing feature here becomes whether the plants has an obligate or facultative vernalization requirement. Many *A. thaliana* ecotypes are summer-annuals which are also referred to as rapid cycling accessions, as they are able to flower rapidly without vernalization. These ecotypes also typically germinate in early spring, flower and set seeds in summer and then go through the cold periods of winter as seeds [115]. The commonly used laboratory ecotypes of *A. thaliana*, Columbia and Landsberg *erecta*, are examples of summer-annuals. There are other ecotypes referred to as winter-annuals, these have facultative vernalization requirements, that is, cold is not required for flowering, but, plants will flower earlier after cold treatment. Winter-annuals usually germinate in fall, go through winter in the vegetative state and initiate reproductive development in spring [115, 116]. Genetic analyses of *A. thaliana* ecotypes have revealed that, winter-annuals usually possess dominant alleles of *FLC* and *FRI* whereas the summer annuals often have mutations in one or both of these genes [87, 115-117]. Biennials generally are plants that remain vegetative the first season, flower and set seeds the second season then senescence and die [118]. This includes crops such as carrot and

sugar beets. Many temperate biennial plants have obligate vernalization requirements and therefore cannot flower without being exposed to cold [116].

Plants following a perennial life strategy, have the ability to live for several growing seasons, and in most cases, repeatedly cycling between vegetative and reproductive development. In contrast to annuals and biennials, majority of perennials are able to maintain vegetative growth after a flowering episode and senescence occurs locally in branches that underwent flowering [111, 112, 119]. The ability of a plant to regenerate photosynthetic organs such as leaves and shoot has been demonstrated to be a determining factor of how long the plant lives [120]. Perennial plants achieve this, by differentially controlling the behavior or fate of meristems within a single plant, whereby while some undergo floral transition, others are kept vegetative [111]. For instance, individual branches in adult poplars, which are perennial trees, produce axillary vegetative buds in addition to reproductive buds to aid the maintenance of vegetative development after flowering [121]. In some cases, in order to maintain vegetative growth, reproductive meristems are reverted back to being vegetative [122]. In perennial species, *Callistemon*, *Metrosideros* and *Cheiranthus cheiri L*., inflorescence reversion were observed and considered to be a mechanism adapted to maintain vegetative growth and support their perennial life history strategy [122-124]. Plants that allocate all meristems to flowering and therefore flower only once during their life cycle, are referred to as monocarpic, while those that maintain some meristems as vegetative and are able to flower multiple times during subsequent seasons are referred to as polycarpic [122, 125, 126]. Based on this, annuals are generally considered as monocarpic. Winter-annuals and biennials can be induced to exhibit perennial habits if flowering is prevented [126]. However once flowering and seed set occurs, the plants die since there are no mechanisms to conserve meristems for subsequent years [125, 126]. Thus, plants following these live strategies are also monocarpic. Many perennials on the other hand are considered to be polycarpic, however there are exceptions. Bamboos after decades of vegetative growth, flower synchronously and die together [125, 127], they are therefore referred to as monocarpic perennials.

1.6 Flower initiation and emergence in annuals and perennials

An important distinguishing factor between annuals and many perennials is the timing of floral initiation and emergence. Under favorable conditions flower emergence quickly follows floral

initiation in *A. thaliana* and many other annuals. In this instance, flower emergence is considered to reflect the time of initiation [126]. However there are other instances where different environmental signals, such as temperature and day length, are required for initiation and emergence, making these two processes uncoupled in some species. In several perennials there is the tendency for a significant delay between the time of flower initiation and emergence whereby floral initiation, which results in the formation of flower buds, can occur a season or more before flower emergence. This phenomenon of organs being initiated a season or more before their maturation and function is referred to as organ preformation [128, 129]. The event of organ preformation is reported to be most common in plants growing in diverse seasonal environments such as in arctic and alpine environments [129-131]. Due to the short growing seasons of spring and summer in these environments, preformation of organs, such as flower buds, is considered as an advantage to enable rapid emergence and the complete development of plant structures, such as flowers, during permissive conditions [130, 132]. In the alpine perennial species, *Acomastylis rossii*, studies on preformation demonstrated their inflorescence is initiated and develop as primordia below ground for about two years [129]. Emergence of the inflorescence occurs in the third year within a week after snowmelt. Similar studies which resulted in similar conclusion have been performed in the alpine perennial species, *Polygonum viviparum L*. [130]. Several patterns of preformation have so far been reported in other temperate perennial trees, shrubs and herbaceous plants [121, 133-135]. The herbaceous perennial, *Dicentra cucullaria*, was identified to preform its leaves, inflorescence and flowers a year before their function [133]. Apple, a woody perennial, initiates floral buds that develop over two consecutive seasons followed by a period of winter dormancy after which development continues and flowers bloom in spring [136]. In many temperate perennials, decreases in temperature and day length during autumn results in growth cessation and the gradual entry of floral buds into a state of dormancy. Prolonged exposure to winter chilling conditions is then required to break the dormancy and enable flower emergence once permissive warm temperatures commence [121, 125, 137, 138]. For instance, in the perennial plant, Junebearing cultivar of strawberry, flower initiation is induced by a combination of cool temperatures and shortening photoperiods [139]. The continuous decrease in temperatures as winter approaches, leads to the cessation of flower initiation and the plant is considered to gradually enter a period of dormancy. The so-called dormant plants still retain the capability for growth although at a much reduced rate. Winter chilling in this

species not only removes dormancy but also switches off flower initiation. Flowers then emerge from the autumn initiated buds during warmer temperature in spring.

1.7 *Arabis alpina,* as a perennial model species

A. thaliana has served as a great model plant for uncovering molecular basis for reproductive development and has allowed the transfer of knowledge gained to several other species. However the study of some perennial traits such as, the delay in flower emergence after initiation, mechanisms involved in the prevention of flowering in some meristems, mechanisms regulating flowering duration and localized senescence, possess as a challenge as they do not exist in this annual model species. In recent years, another member of the *Brassicaceae*, *Arabis alpina* (Alpine rock-cress), has been adopted as a perennial model species for the study of flowering processes and various perennial traits [119, 140-143]. *A. alpina* is a perennial arctic-alpine plant which has a wider distribution compared to most plants in this habitat. Its range of distribution extends from the northern amphi-Atlantic area and the European mountain systems, to the Arabian Peninsula, mountain ranges of Central Asia, North Africa and the high mountains of East Africa [144]. *A. alpina* is a diploid organism with a genome size of about 375 Mb and harbors eight chromosomes, it is self-compatible and also susceptible to *Agrobacterium tumefaciens* transformation [119, 145]. There are over 140 *A. alpina* accessions, which includes plants with and without vernalization requirements [142, 143]. The most commonly used laboratory accession, Pajares, was collected from the Cardillera Cantábrica mountain region of Spain and it exhibits an obligate vernalization requirement [119]. Similar to other polycarpic perennials, some meristems are kept vegetative for subsequent flowering episodes and flowering duration is restricted in Pajares [119]. In recent years the molecular basis for the polycarpic life history strategy of *A. alpina* has been of great interest and being investigated.

Molecular studies in *A. alpina* have led to the identification of some orthologues of genes regulating flowering in *A. thaliana*. *PERPETUAL FLOWERING 1* (*PEP1*) was identified as the orthologue of *A. thaliana FLC* gene and likewise a floral repressor that prevents flowering before vernalization. The onset of vernalization results in the silencing of *PEP1* which is accompanied by increase accumulation of H3K27me3 at the *PEP1* locus [119]. However in contrast to *FLC,* which is stably repressed even after vernalization, *PEP1*, as observed in 12

weeks vernalized Pajares plants, is only transiently repressed during vernalization, and upregulated again after plants are returned to warm conditions [119]. While vernalization reduces *PEP1* expression levels, it activates *AaSOC1* and *AaLFY*, orthologues of *A. thaliana* genes *SOC1* and *LFY*. The expression of *AaLFY*, for instance, was detected at the shoot apical meristem after 5 weeks in vernalization indicating the onset of floral transition [140]. Thus in contrast to *A. thaliana* where floral initiation occurs after vernalization and the stable repression of *FLC*, in *A. alpina* it occurs during vernalization following the transient repression of *PEP1*. With its reactivation after vernalization, *PEP1* is able to contribute to the perennial life cycle of *A. alpina*, by restricting the duration of flowering and preventing flowering of some axillary shoots, thus maintaining vegetative branches for next flowering season [119]. This was supported by phenotypes of *pep1* mutants as they did not require vernalization to flower, flowered perpetually and show compromised polycarpic behavior, whereby more side shoots were induced to flower [119].

Flowering in many perennials occurs only in the adult phase, and so far studies in *A. alpina* have also demonstrated the important role of the age pathway in regulating flowering. Pajares plants must be at least five weeks old to exhibit a flowering response when exposed to prolonged cold [140]. The abundance of miR156 was found to decline in the shoot apex as plants aged and the minimum levels occurred at about 5 weeks which is also the age at which plants are sensitive to vernalization [141]. The decline in miR156 levels was accompanied by an increase in the expression levels of several genes encoding SPL transcription factors in the apices of older plants, which likely are essential to promote flowering in response to vernalization [141]. Thus miR156 in *A. alpina* is considered to act as a timer involved in the control of competence to flower in response to vernalization. Additionally, *PEP2*, the orthologue of *A. thaliana* AP2 transcription factor, was identified as a target of miR172 that prevented flowering before vernalization in *A. alpina*. The ability of *PEP2* to confer vernalization requirement is regarded to be partly through increasing transcription of *PEP1* [141]. Taken together, Bergonzi *et al*., 2013 [141] demonstrated that the parallel repressive activities of miR156 and PEP2/PEP1 ensures meristems of *A. alpina* achieve a certain age before becoming competent to flower in response to cold temperatures. The production of axillary branches with different developmental phases, such as juvenile and adult vegetative and reproductive phases, is very crucial for the maintenance of a perennial life strategy. Park *et al*., 2017 [143] used *A. alpina* accession Pajares in a comparative study with *A. thaliana* Sy-

0, a winter-annual, to understand the molecular mechanisms behind the differential behavior of axillary branches in perennials. They discovered there were varying sensitives of axillary branches to flower in response to vernalization and this can be attributed to highly variable miR156 levels among these branches. Whereas Pajares axillary branches with high levels of miR156 remained juvenile even after vernalization, those with low levels of miR156 flowered after vernalization. On the contrary, in Sy-0, similar levels of miR156 were expressed in all axillary branches which caused flowering in these branches in a synchronous response to vernalization.

A. alpina, *AaTFL1,* the orthologue of *A. thaliana TFL1* gene, has been shown to influence the age dependent response to vernalization. *AaTFL1* prevented the expression of *AaLFY* in young vernalized plants and increased the duration of vernalization required in older plants for *AaLFY* expression and flowering. *TFL1* as previously mentioned as a major gene that affects flowering time and inflorescence architecture in *A. thaliana* and these roles seem to be conserved also in *A. alpina*. Reduction in *AaTFL1* expression caused transgenic plants to flower slightly earlier than wildtype and produce a determinate inflorescence with a terminal flower. These transgenic plants flowered with a shorter vernalization duration of 5 weeks. *AaTFL1*, in addition, acts additively with *PEP1* to prevent flowering in vegetative axillary shoots. A reduction in the activities of both genes caused almost all branches to flower [140]. Therefore, *AaTFL1* by contributing to the maintenance of some axillary shoots in a vegetative state supports the perennial lifecycle of *A. alpina*. The fact that plants with reduced expression of *AaTFL1*, an inflorescence meristem identity gene, are able to response to shorter vernalization durations indicates a link between vernalization durations and reproductive development.

1.8 Impact of cold duration on flowering and inflorescence development

To avoid flowering during periods of temperature fluctuations, such as in autumn, vernalization-requiring plants achieve a vernalized state that promotes flowering only after exposure to sufficient cold durations. The response to cold is also quantitative, in that, by increasing the duration of vernalization, flowering becomes progressively accelerated once plants are returned to warm conditions [67]. In a study involving several *Arabidopsis* ecotypes,

vernalization of 8 weeks, promoted flowering in Swedish accessions, Ull5 and Lov-1, which do not flower in long days (LD) [146]. A follow up study extended the period of cold to 20 weeks and revealed even though 8 weeks vernalization promoted flowering in these accessions, 14 weeks was the required period to fully saturate their vernalization requirement for flowering acceleration [147]. Further studies in these *Arabidopsis* accessions, also revealed, increasing the duration of vernalization quantitatively enhances the stable repression of *FLC* after cold and this correlated with the flowering time acceleration [74, 147]. Accessions, such as Lov-1, which required prolonged cold exposure, exhibited reactivation of *FLC* expression after a non-saturating vernalization and this was associated with the rate of H3K27me3 accumulation at the *FLC* locus [147]. There are additional evidence in different species, such as wheat, cauliflower and raddish, showing that an increase in the exposure of plants to cold, decreases the periods of permissive temperatures required for flowering to occur [147-152]. Furthermore, effects of vernalization requirements not been satisfied on other aspects of reproductive development have been reported. For instance in wheat, the rate of spikelet initiation decreased with insufficient vernalization [149]. In various temperate perennial species, failure to receive sufficient chilling results in erratic bud break, extended time to anthesis and yield reduction [148, 150, 153]. These therefore illustrate the importance of a plants vernalization requirements being met in order to ensure maximum reproductive development.

Following the repression of *PEP1,* floral initiation occurs and this is supported by the detection of *AaLFY* expression at the shoot apical meristem after 5 weeks in vernalization [140]. Even though *AaLFY* expression is detected early in shoot apices during vernalization, 12 weeks is the minimum required for all vernalized plants to flower and develop an inflorescence. Vernalization for shorter periods results in partial flowering plants with reverting inflorescence [119]. This suggest, the additional vernalization period might be important for the regulation of downstream process after floral initiation. An extension of venalization period beyond the previously suggested minimum of 12 weeks, can provide information on the period at which vernalization requirement is saturated in *A. alpina* accession Pajares and how this can influence reproductive development.

1.9 Strategies for gene discovery

Studies in *A. alpina* have so far provided great insights into certain molecular mechanisms involved in the regulation of the polycarpic life strategy of a perennial plant. However, there are more to be discovered in order to fully understand important differences between the life history strategies of plants. Furthermore, even though the *pep1* mutant shows altered perennial traits, it does not die after flowering and setting seeds as observed for monocarpic plants. It is able to maintain the continuous supply of meristems and undergoes numerous episodes of flowering, supporting the fact that, there are additional mechanisms besides *PEP1* regulating the perennial growth habit. Another round of mutagenesis in the background of *pep1* mutation can lead to the discovery of additional genes involved in plant life history strategies.

Isolation of mutants with intriguing phenotypic alterations after mutagenesis requires follow up steps to identify the underlying causal mutations. Identification of such mutations involves the location of chromosomal region containing the causal gene, that is, the mapping interval. Most experimental strategies typically employed for this, includes the generation of a recombinant mapping population using well defined crossing schemes [154]. The recombinant population is screened to isolate individuals with the phenotype of interest, as they are expected to carry the casual mutation. This results in an over-representation of alleles from the mutant background which aids the direction of analyses to the chromosomal region harboring the causal mutation [154]. The process of genetic mapping through the use of allele frequencies within pooled recombinant individuals is generally referred to as bulk segregant analysis (BSA) [155]. With the advent of next-generation sequencing (NGS), a combination of BSA with whole genome resequencing, a process referred to as mapping-by-sequencing [156, 157], allows for the quantification of allele frequency and simultaneous identification of genetic markers, mostly in the form of SNPs. This requires the use of pipelines such as SHOREmap, a computational tool that enables analysis of mapping-by-sequencing data [156, 158]. Following the original application in F2 population of *A. thaliana* [156], mapping-by-sequencing has been successfully implemented in other plant and animal species, and for the isolation of different forms of mutations including large deletions [159-164].

Once a mapping interval is established, a target search of the region enables the isolation of candidate genes. In cases of broad mapping intervals, narrowing down on the putative casual locus through co-segregation analysis of SNPs segregating with the mutant phenotype, permits the reduction of the mapped interval. This was implemented in barley, where mapping-by-sequencing was used to identify a mapping interval harboring a mutation that increased the rate of leaf initiation [160]. Further genotyping of individuals in the mapping population using markers developed from SNPs, assisted the confirmation and refinement of the target interval. This approach exploits the ability of recombination to break up the genome into fragments and enable the correlation with phenotypic variation. However, the influence of recombination is limited in circumstances where certain non-casual mutations are closely linked to the casual one. To circumvent such situations, one can consider resequencing pooled recombinants from two or more allelic or phenotypically identical mutants. Each sequenced bulk will feature several mutations, however, the overlap of mutated genes between independently mutagenized individuals are expected to be extremely small thus, only genes with lesions in all mutants need to be considered. A similar approach was previously proposed for mutation identification through direct sequencing of two or more allelic mutants [165]. As a proof-of-concept, it was applied to two independently isolated mutants *pep1-1* and *pep1-2* which both flowered without vernalization in comparison to wild type [166]. Out of 165 genes from *pep1-1* and 94 from *pep1-2*, only three were common for both mutants. Further functional analysis revealed one of the three genes, *PEP1*, was indeed the causal gene. This in addition proves the necessity of functional analysis to support the discovery of a casual gene. Eventual validation of candidate genes involves the use of transgenic wild-type allele of the gene to complement the mutated allele in order to rescue the mutant phenotype. Once the casual gene is determined, further studies can be carried out to characterize its molecular role or roles.

The development of high-throughput sequencing has further provided a means for quantifying transcriptomes. The transcriptome, which refers to the complete range of messenger RNA (mRNA) molecules expressed by an organism, a particular tissue or a given cell, is not stable and actively changes depending on several factors including the developmental stage of the organism and environmental conditions. The advent of RNA sequencing (RNA-seq) has permitted access to a great deal of information regarding expression changes and relative abundance of a gene's transcript during specific developmental stages or under defined

treatment conditions. To conduct RNA-seq, mRNA is first converted to cDNA, which is then used as input for NGS library preparation. Analysis of acquired NGS dataset involves the use of some powerful computational programs. The raw data obtained first undergoes preprocessing to obtain high-quality reads followed by mapping of those reads to a reference genome with the aid of bioinformatics programs such as Bowtie and TopHat [167, 168]. The number of reads that map to a specific gene serves as the measurement for that gene's expression level. The mapped reads are further assessed with statistical models, such as Cuffdiff from the Cufflinks package [168], to call out differentially expression genes (DEGs). These are genes whose abundance or transcript levels differ across two or more compared samples. The availability of gene annotation data have additionally improved to a great extent functional genomic studies where by DEGs can be grouped into functional categories through Gene Ontology (GO) analysis. GO uses gene annotations to describe the biological roles of genomic products such as genes and proteins and classify them using ontology terms [169, 170]. The classification is performed by comparing a given set of genes with the whole genome to find over-represented GO terms in the set, that is, whether a GO term appears more frequently than as expected by chance. This then helps to narrow in on mechanisms associated with changes in an organism. Overall understanding the changes in the transcriptome of an organism is essential for defining functional components of the genome and provides important insights into biological mechanisms regulating various developmental changes.

1.10 Research aims

As the duration of cold exposure has been demonstrated in some perennials and other species to influence reproductive development, this thesis first aims at providing detailed descriptions on the effects of vernalization duration on flowering and inflorescence development of *A. alpina* accession Pajares. Secondly, the thesis aims at identifying and characterizing the effects of additional genes involved in the regulation of flowering and inflorescence development using second site enhancer mutants of *pep1-1* and the mapping-by-sequencing approach. The last aim is to determine the role of the identified genes in the perennial life cycle of *A. alpina* and the use of genome wide transcriptome analysis to determine possible regulatory mechanisms.

2 Results

2.1 Effects of vernalization duration on flowering and inflorescence development in *A. alpina*

The main shoot of a flowering Pajares plant, after vernalization, can be divided into four main zones and axillary shoots produced differ in their flowering behavior depending on their position on the main shoot (Figure 2). At the early vegetative stage of development in Pajares, the main shoot produces leaves with long petioles and internode elongations, this part of the shoot is referred to here as V1. Axillary branches formed within some axils of leaves in V1 before vernalization, flower alongside the main shoot apices after plant are returned to warm temperatures. The next parts, V2 and V3, above V1 also comprise of leaves with long petioles. The axils of leaves in V2 contain axillary meristems that do not outgrow whereas the leaf axils in V3 are filled with vegetative axillary branches with compact internodes. The onset of flowering during reproductive development, results in the production of the inflorescence lateral branches in the axils of cauline-like leaves, in the part of the inflorescence referred to as I1. These leaves have no petioles and are separated by elongated internodes. I2 represents the part of the inflorescence in which solitary flowers, without bracts, are produced. Studies in this section are centered on the inflorescence I1 and I2 zones and how exposure of Pajares plants to different vernalization durations affects flowering and inflorescence development.

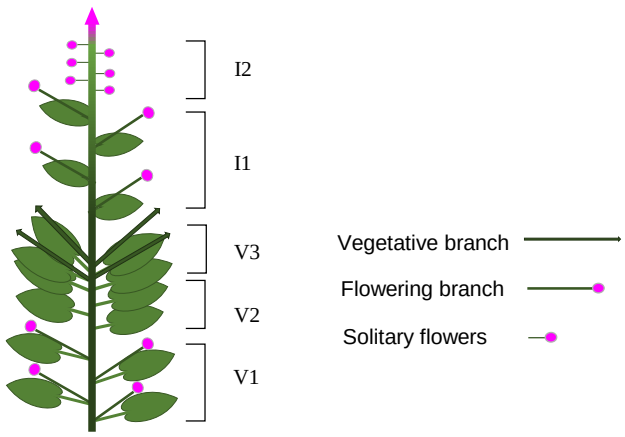


Figure 2. Diagrammatic representation of *A. alpina* accession Pajares flowering plant.
Axillary branches formed within V1 leaf axils flower alongside the main shoot apices. The axils of leaves in V2 contain axillary meristems that do not outgrow whereas V3 are filled with vegetative axillary branches. The inflorescence zone is divided into I1 and I2. At I1 inflorescence branches are produced within cauline-like leaves. I2 consists of solitary flowers without bracts.

2.1.1 Flower emergence and inflorescence outgrowth after different vernalization durations

To examine the effects of vernalization duration on *A. alpina*, Pajares plants were grown for 8 weeks in long days (LD) then vernalized for 8, 12, 15, 18, 21 and 24 weeks. The experiment was designed to ensure all treatments returned to LD conditions at the same time (as shown in Figure 3A). In support of previous results from Wang et al., 2009, vernalization of 12 weeks and above resulted in 100% of plants flowering. However extended duration of vernalization reduced the time to flower emergence, which is indicated as number of days to the first opened flower, after plants were returned to LD conditions (Figure 3B). While plants which experienced 24 weeks of cold flowered within an average of 14 days, 12 weeks vernalized plants required 24 days on average to flower. The total number of leaves after flower emergence were scored (Figure 3C) and although days to flower emergence decreased with longer vernalization, plants had similar number of leaves regardless of the vernalization duration.

The influence of extended vernalization on inflorescence outgrowth or bolting was also studied. As a measure of inflorescence outgrowth (bolting) the length of the inflorescence shoot was scored at 2, 3, 8 and 14 weeks after vernalization. Plants vernalized for 15 weeks and above, showed an outgrowth of their inflorescence shoot 2 weeks after vernalization while those vernalized for less than 12 weeks bolted after 3 weeks in LD (Figure 3D). The development of the inflorescence was observed earlier in plants vernalized for 15 weeks and above (Figure 3E).

This demonstrates, the number of days required for flower emergence reduces with extended vernalization and inflorescence outgrowth is earlier in *A. alpina* accession Pajares with longer vernalization.

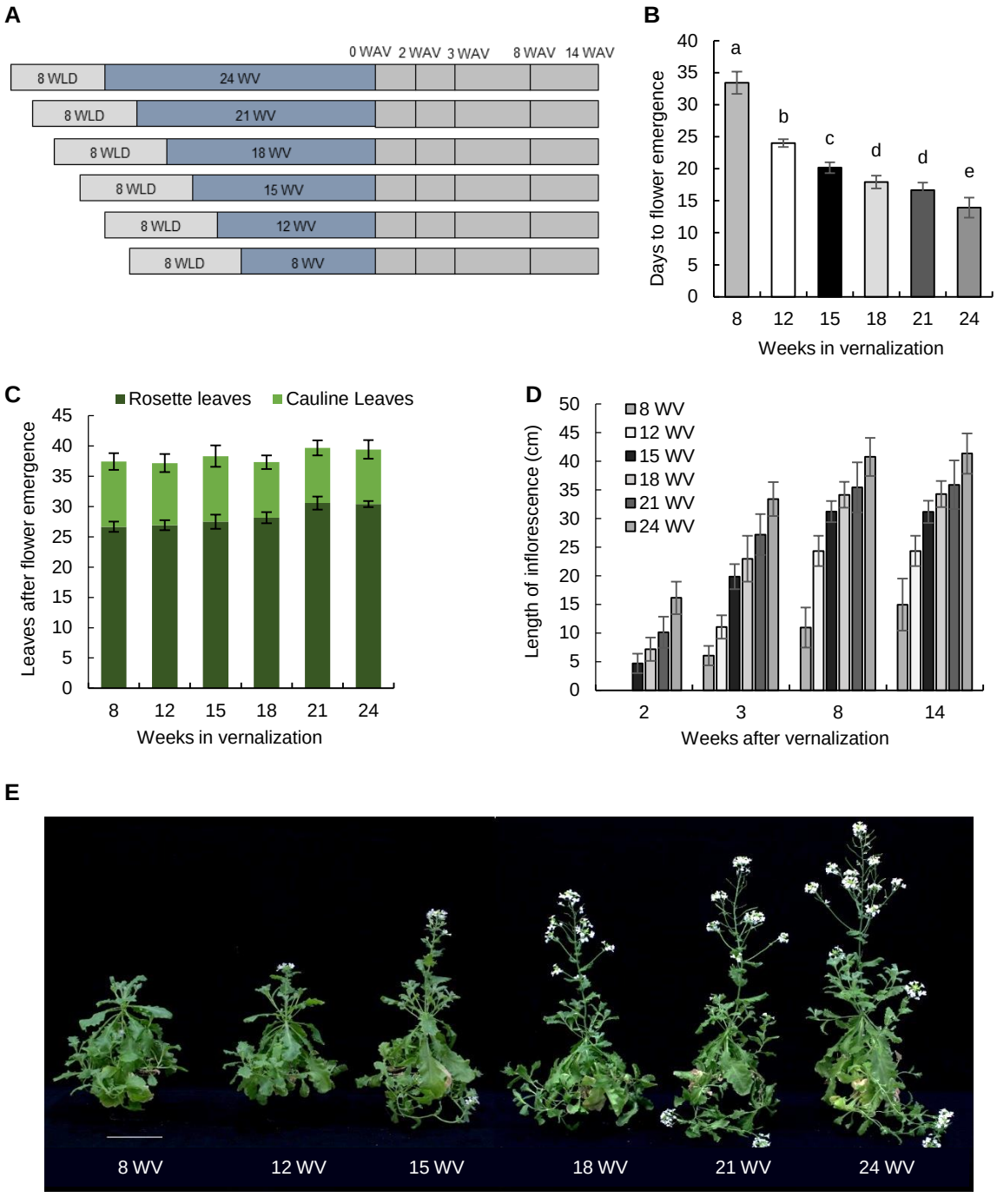


Figure 3. Prolonged vernalization affects flower emergence and inflorescence outgrowth.
(A) Schematic representation of synchronized experimental design. Plants were grown for 8 weeks in LD (WLD) and vernalized (weeks in vernalization, WV) at different times to ensure the return of all plants to warm temperature (weeks after vernalization, WAV) at the same time. **(B)** Number of days to flower emergence after extended vernalization. **(C)** Total number of leaves scored at after flower emergence. **(D)** Inflorescence outgrowth (bolting) after different vernalized durations. This was scored as a measure of the inflorescence length after 2, 3, 8 and 14 weeks in LD. **(E)** *A. alpina* Pajares exposed to different durations of vernalization. Plants were grown for 8 weeks in LD prior to vernalization and returned to LD for 3 weeks. Bar = 10 cm. Data are presented as means ± SD; n = 9 - 12.

2.1.2 Floral reversion and vegetative features within the inflorescence are suppressed with extended vernalization

The presence of bracts on the main inflorescence stem were scored as a measure of floral reversion in Pajares plants exposed to different periods of cold, since bracts are considered as phenotypes associated with floral reversion [23]. Plants that flowered with 8 weeks of vernalization flowered partially (Figure 4A), exhibited floral reversion within the inflorescence in the form of numerous bracts. Pajares plants vernalized for 12 to 18 weeks also exhibited the intercalation of bracts within their inflorescences (Figure 4B). Bracts were however absent in 21 and 24 weeks vernalized plants (Figure 4C). The scoring of number of bracts present within the inflorescence therefore revealed a suppression of bracts with an increase in vernalization (Figure 4D).

During reproductive development in *A. alpina*, Pajares, flower emergence is preceded by the inward curving of leaves closest to the shoot apex (Figure 4E). The shoot apex at this point ceases the production of leaves to favor the production of flowers (Figure 4F). Vegetative lateral branches within inflorescence I1 zone were identified as branches which displayed the inward curving of leaves but then was followed by the production of more leaves (Figure 4G). The influence of vernalization duration on the fate of lateral branches within the inflorescence was investigated here. I1 lateral branches were scored for being either vegetative or flowering at 8 weeks in LD after exposure to different vernalization durations. The number of I1 lateral branches slightly reduced with longer vernalization (Figure 4H), however an increase in the duration of vernalization was observed to cause an increase in the percentage of flowering branches while reducing the percentage of vegetative ones (Figure 4I). To further examine the effect of vernalization duration on the fate of I1 lateral branches, the point with less frequent flowering branches along the main inflorescence axis was followed. To achieve this, the percentages of flowering branches at I1 nodes containing lateral branches in the inflorescence were determined. The starting point was considered as the first node closest to the base of the inflorescence. Inflorescence nodes with lowest frequency of flowering branches ranged from positions 1 to 6 in plants vernalized for 12 and 15 weeks. This however reduced to node 1 in plants with 18 and 21 vernalization (Figure 4J). The overall trend of this analysis suggests, the point with lowest frequency of flowering branches along the inflorescence axis is

influenced by the duration of vernalization. Shorter vernalization durations causes a reduction in the percentage of flowering branches at nodes closest to the base of the inflorescence.

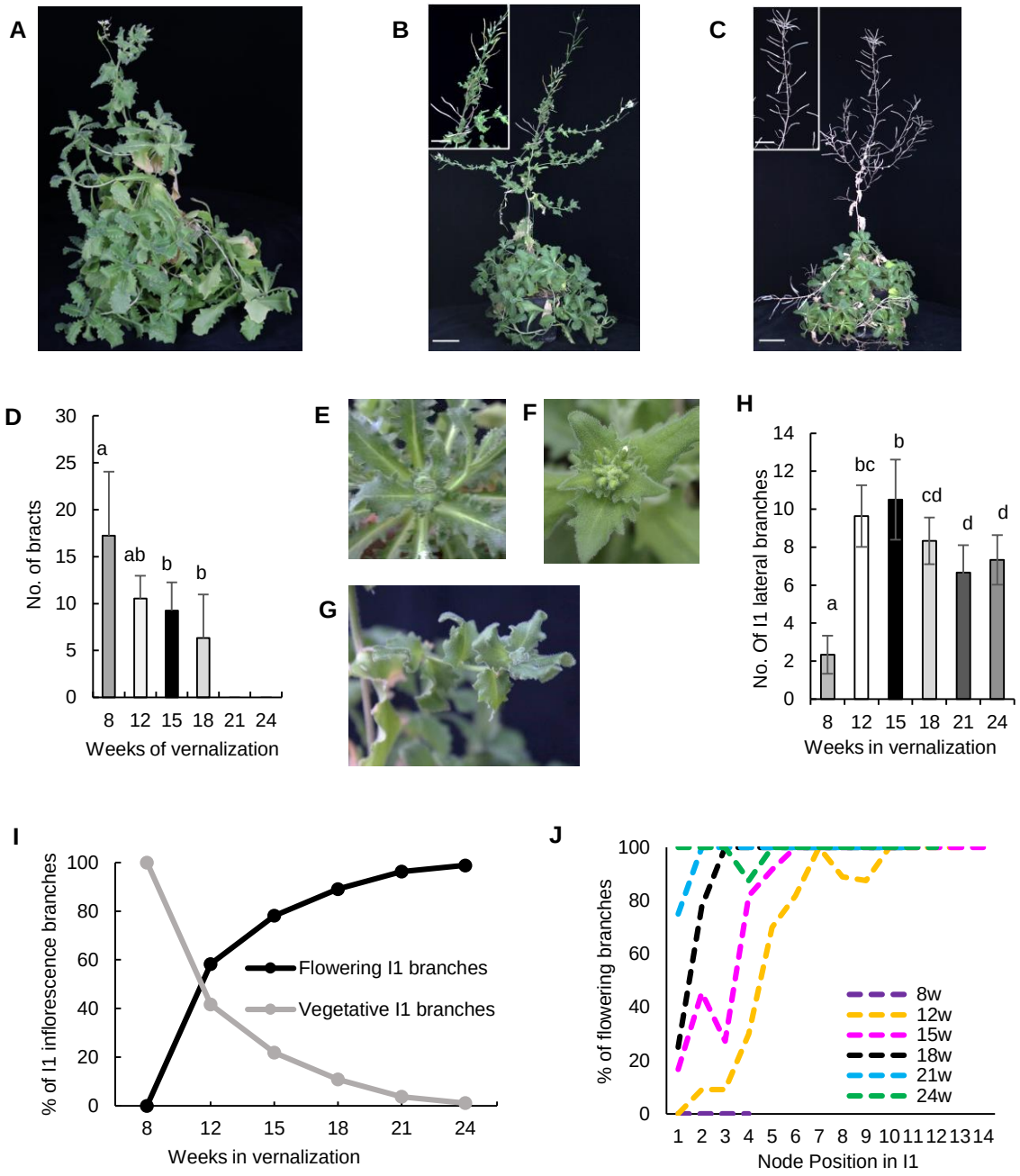


Figure 4. Increase in vernalization duration suppresses floral reversion and increases the number of flowering I1 lateral branches.

(A) 8 weeks vernalized Pajares plant exhibiting partial flowering and floral reversion. **(B)** Intercalation of bracts in inflorescences of 12 weeks vernalized plants. **(C)** Absence of bracts in 24 weeks vernalized plants. The browning part on the inflorescence is due to senescence. Bar = 5cm, insert bar = 3cm. **(D)** Number of bracts with respect to vernalization duration. **(E)** Inward curve of leaves closest to the shoot apex of Pajares. **(F)** Flower bud emergence at the shoot apex. **(G)** A vegetative I1 lateral branch exhibiting inward curving of leaves. **(H)** Number of I1 lateral branches. **(I)** Percentage of vegetative and flowering inflorescence I1 branches with respect to vernalization duration. **(J)** Percentage of flowering branches at various inflorescence I1 node positions. Data is presented as means ± SD; n = 9 - 12. Letters above columns represent significant differences determined by omnibus Kruskal-Wallis test and pairwise multiple comparison using Mann-Whitney U test (α-value of 0.05)

2.1.3 *A. alpina* accession Pajares demonstrates a basipetal differentiation of inflorescence branches

In *Arabidopsis*, inflorescence branches differentiate from activated meristems in axils of pre-exiting leaf primordia. The initiation and development of the inflorescence branches arise basipetally with upper-most inflorescence branches primordium being the biggest and most developed [171]. The fate of inflorescence branches was scored after Pajares plants were exposed to different durations of vernalization and returned to LD. The scoring was conducted 8 and 14 weeks after vernalization (Figure 5A-B). Flowering pattern of inflorescence branches in plants vernalized for different durations occurred basipetally with upper-most branches flowering first as scored 8 weeks after vernalization (Figure 5A). A second scoring conducted after 14 weeks in LD (Figure 5B) revealed some vegetative branches within the inflorescence I1 zone had begun to flower (Figure 5C).

Interestingly, extended vernalization of 24 weeks caused flowering of a few branches in the V3 zone (Figure 5A-B). To further investigate the basipetal development of inflorescence branches, the lengths of branches at each inflorescence I1 node were scored in 12 and 24 weeks vernalized plants after 8 weeks in LD. Branches at the upper part of I1 were longer than those present at the lower part (Figure 5D). The follow up flowering of basal I1 lateral branches, in 12 to 18 weeks vernalized plants and the reduced lengths of lower I1 branches demonstrates that inflorescence branches differentiation occurs basipetally in *A. alpina*, Pajares.

2.1.4 Prolonged vernalization marginally increases silique production

The effect of vernalization duration was also examined on the productivity of the inflorescence. The number of siliques was scored after plants had been 8 weeks in LD. There was a significant increase in total number of siliques produced by plants vernalized for 15 weeks and above (Figure 6A). A separation of the I1 (Figure 6B) and I2 (Figure 6C) also showed similar increases in the number of siliques. The fact that some vegetative inflorescence branches were able to flower after 14 weeks in LD means if kept longer, these plants could to some extent increase the silique number in I1.

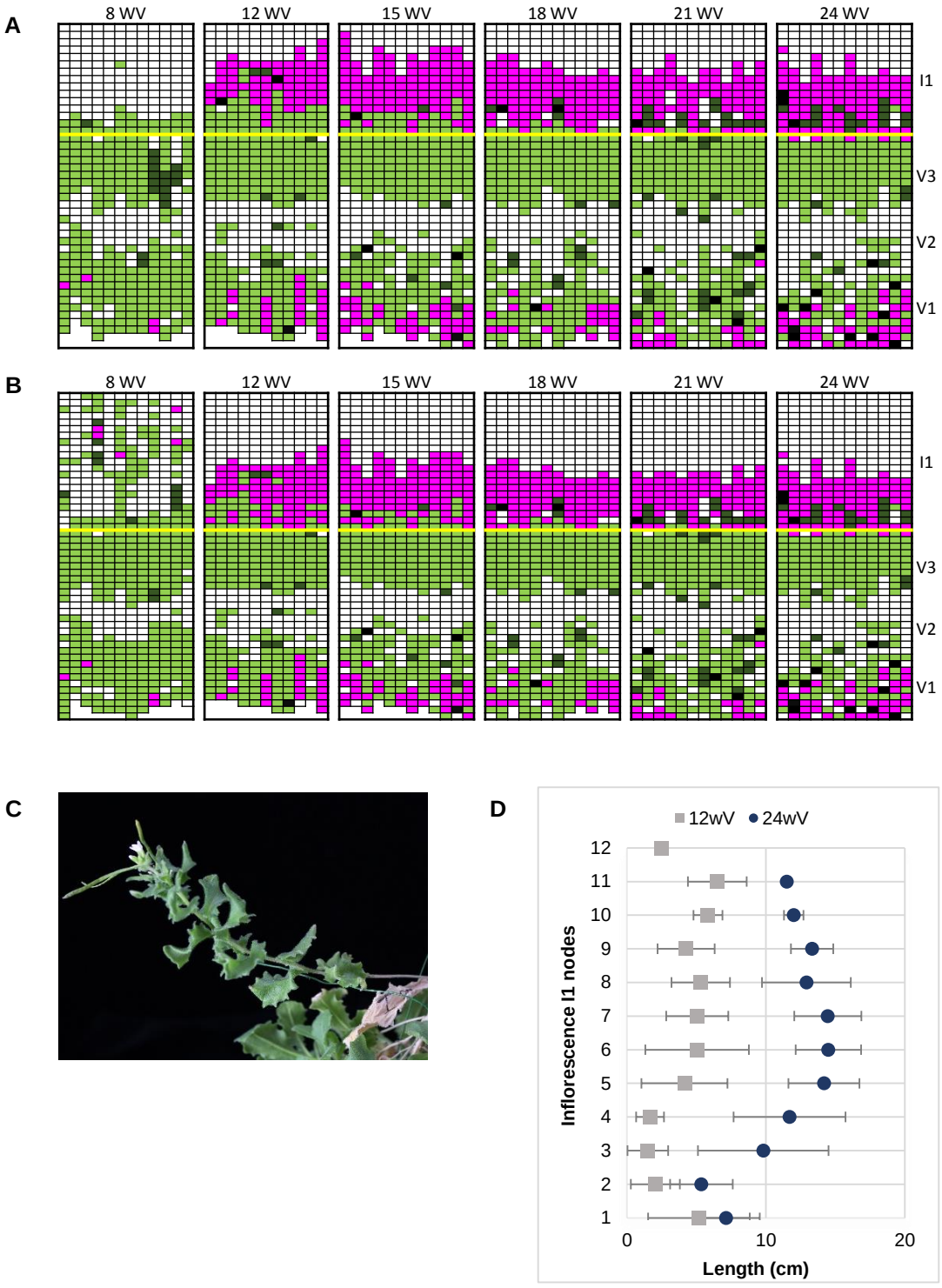


Figure 5. Inflorescence branches of *A. alpina*, Pajares develop basipetally.
(A) Schematic representation of flowering Pajares plants vernalized for different durations after 8 weeks in LD and **(B)** 14 weeks in LD. **(C)** A vegetative I1 lateral branch flowering after 14 weeks in LD. **(D)** The length of axillary branches at different node position in I1 of 12 weeks and 24 weeks vernalized plants. Data is presented as means ± SD; n = 9 – 12. Each column in figure 4A and 4B represents a single plant and each row is a leaf axil. Magenta boxes indicate flowering branches, green boxes below the yellow line depict vegetative branches in V3 and above the yellow line indicate vegetative branches in I1.

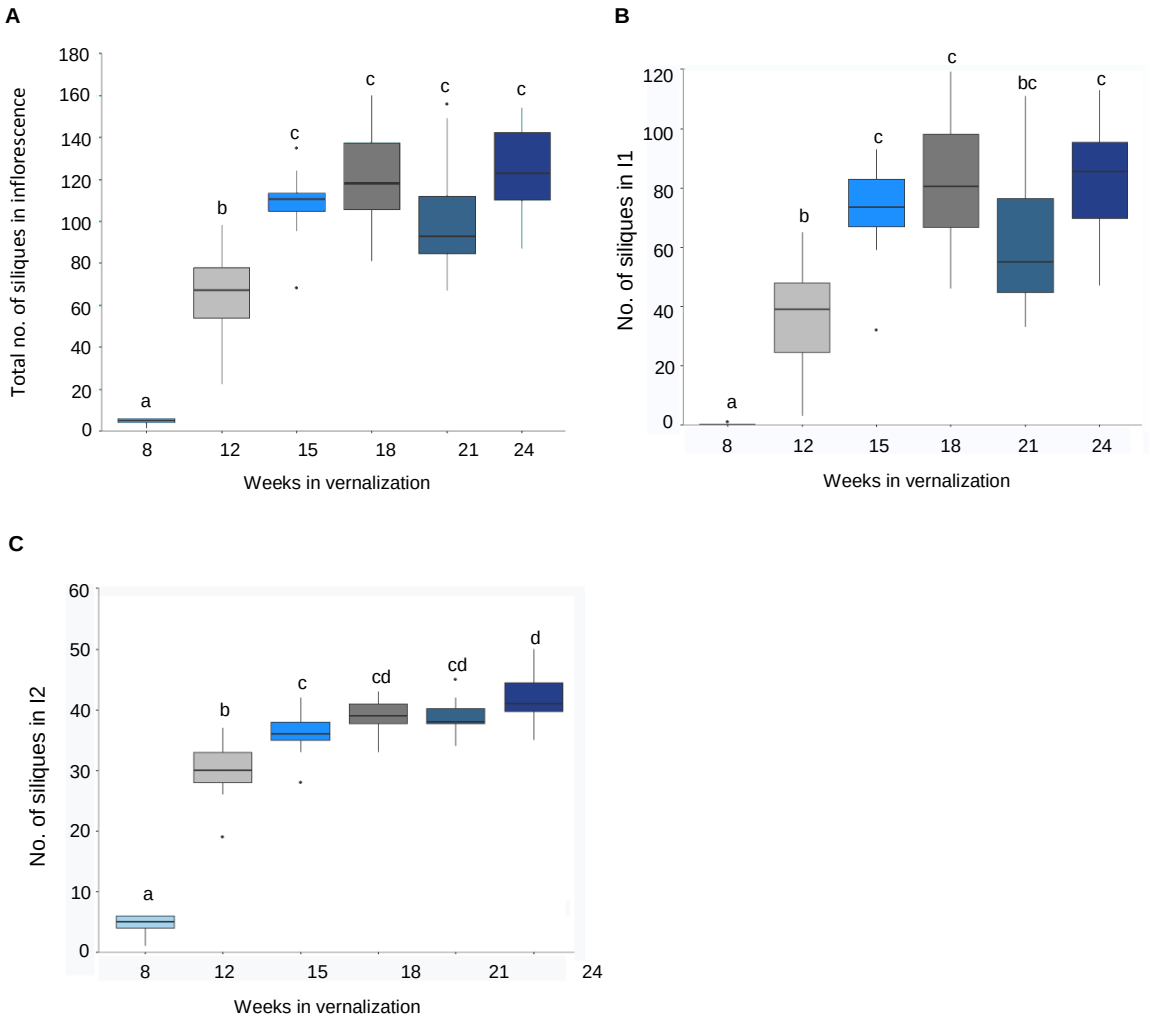


Figure 6. Extended vernalization increases the number of siliques in the inflorescence.
(A) Total number of siliques as counted in *A. alpina* Pajares vernalized for 8, 12, 15, 18, 21 and 24 weeks and subsequently grown for 8 weeks in LD. **(B)** Number of siliques on the inflorescence branches, I1. **(C)** Number of siliques on main inflorescence stem, I2. Data are presented as means ± SD; n = 9 -12. Letters above columns represent significant differences determined by omnibus Kruskal-Wallis test and pairwise multiple comparison using Mann-Whitney U test (α-value of 0.05)

2.2 Analysis of flowering time and inflorescence structure of *pep1-1* and *enhancers of pep1* (*eops*)

In order to identify other factors regulating flowering besides *PEP1*, the *pep1-1* mutant was previously mutagenized using ethyl methanesulfonate (EMS) by Prof Maria Albani. Second site mutants that enhanced the early flowering phenotype of *pep1-1* were isolated. These mutants are referred to as *enhancers of perpetual flowering 1* (*eop*). Five *eop* mutants were isolated, four of which, *eop002*, *eop085*, *eop088* and *eop091,* were early flowering with simpler inflorescence as they produced reduced number of inflorescence branches. The fifth mutant, *eop101,* apart from flowering earlier than *pep1-1* developed a determinate inflorescence by the formation of a terminal flower. To properly compare the inflorescence of the *eops* to that of *pep1-1*, detailed characterizations of the inflorescences are required. The plant architecture and flowering behavior of axillary branches of the *pep1-1* mutant however creates difficulties in determining its inflorescence zone, specifically I1. Upon reproductive development in Pajares, elongation of the main stem (bolting) occurs above the V3 zone to give rise to the inflorescence (Figure 7A). Axillary branches that develop within leaf axils in the V3 zone are usually vegetative with compact internodes making it easier to distinguish them from I1 branches which are mainly flowering branches with separated internodes (Figure 7B). In *pep1-1* mutant, on the other hand, bolting of the main stem at the onset of reproductive development is absent (Figure 7C), there is no zone of compact internodes and all axillary branches flower (Figure 7D).

This section first focuses on the characterization of the *pep1-1* mutant inflorescence as well as that of the *eop* mutants using morphological markers such as leave shape and senescence behavior of the inflorescence. Senescence can be considered as a morphological marker as, to maintain its perenniality, only the main inflorescence of Pajares and that of axillary shoots that flowered, cease growth and senesce as seeds mature (Figure 4C, Section 2.1.2). The fact that the *pep1-1* mutant still remains perennial and is able to maintain the supply of axillary meristems, suggests senescence might also be restricted in this mutant. Phenotypic experiments were conducted three times (Supplementary Figure S1-S4). Results from one experiment are presented here.

Secondly, transcriptome analysis was carried out for the *eop101* mutant to gain insights into possible regulatory mechanisms adopted by this mutant for its reproductive development. For the remaining four *eop* mutants, experiments were performed to determine their level of dominance and to determine whether they are allelic.

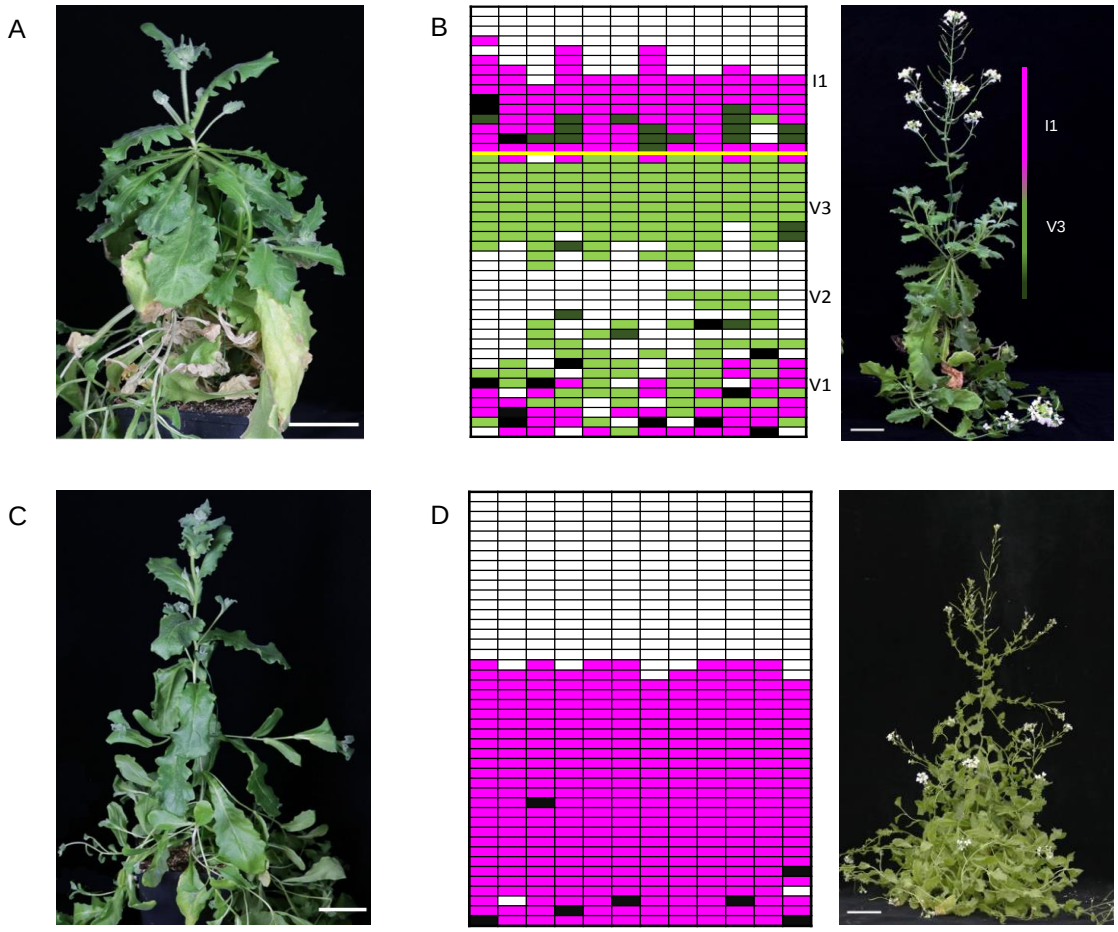


Figure 7. Bolting of main stem is absent in *pep1-1* and all axillary branches flower during reproductive development.
(A) Elongation or bolting of Pajares main shoot to give rise to the inflorescence. **(B)** Flowering behavior of Pajares V3 and I1 branches. **(C)** Absence of main shoot bolting in *pep1-1*. **(D)** Flowering behavior of pep1-1 axillary branches. Scale bar = 5 cm. Each column in **(B)** and **(D)** represents a single plant and each row is a leaf axil. Magenta boxes indicate flowering branches, green boxes below the yellow line in **(B)** depict vegetative branches in V3 and above the yellow line indicate vegetative inflorescence branches in I1.

2.2.1 Inflorescence zone detection in *pep1-1* mutant

Changes in leaf shape and senescence behavior of the inflorescence were studied in both *pep1-1* and Pajares for the determination of inflorescence zone. To avoid the loss of earlier formed leaves in Pajares, the first set were sampled at eight weeks in long day before plants were vernalized (Figure 8A, upper panel). The second set were sampled after vernalization and at the time of flowering (Figure 8A, lower panel). *pep1-1* lacked well defined petiolated leaves as compared to Pajares (Figure 8B). Nevertheless leaves became smaller, more rounded at the base and directly attached to the stem at the inflorescence of *pep1-1* just as observed in Pajares (Figure 8B, I1 zone shaded in magenta). Senescence in Pajares is restricted to the main inflorescence and axially branches that flowered in a given season (Figure 8C). This behavior serves as a good indicator for determining inflorescence zone. Study of senescence behavior in *pep1-1* main shoot revealed, just as observed in Pajares, senescence was restricted to a specific zone in the *pep1-1* shoot (Figure 8D). The inflorescence of *pep1-1*, based on leaf-shape, was identified to have a height of 34.37 ± 3.38 cm and to contain 7.3 ± 0.79 inflorescence branches subtended by cauline-like leaves in I1 (Figure 8E and F). This however changed after scoring with senescence, inflorescence height increased to 39.31 ± 2.98 and number of inflorescence branches to 9.16 ± 0.98 (Figure 8E and F). The number range of *pep1-1* inflorescence branches was found to be similar to that scored in Pajares plants (Figure 4H, Section 2.1.2). Scoring of siliques in I2 is not dependent on these two criteria, as this zone is more distinguishable. The I2 zone of *pep1-1* was scored to contain 23.7 ± 5.62 siliques. In comparison to the number of siliques in I2 of Pajares plants (Figure 6C, Section 1.1.5) with 15 weeks of vernalization and above, *pep1-1* showed a significant reduction in I2 siliques (Table 1).

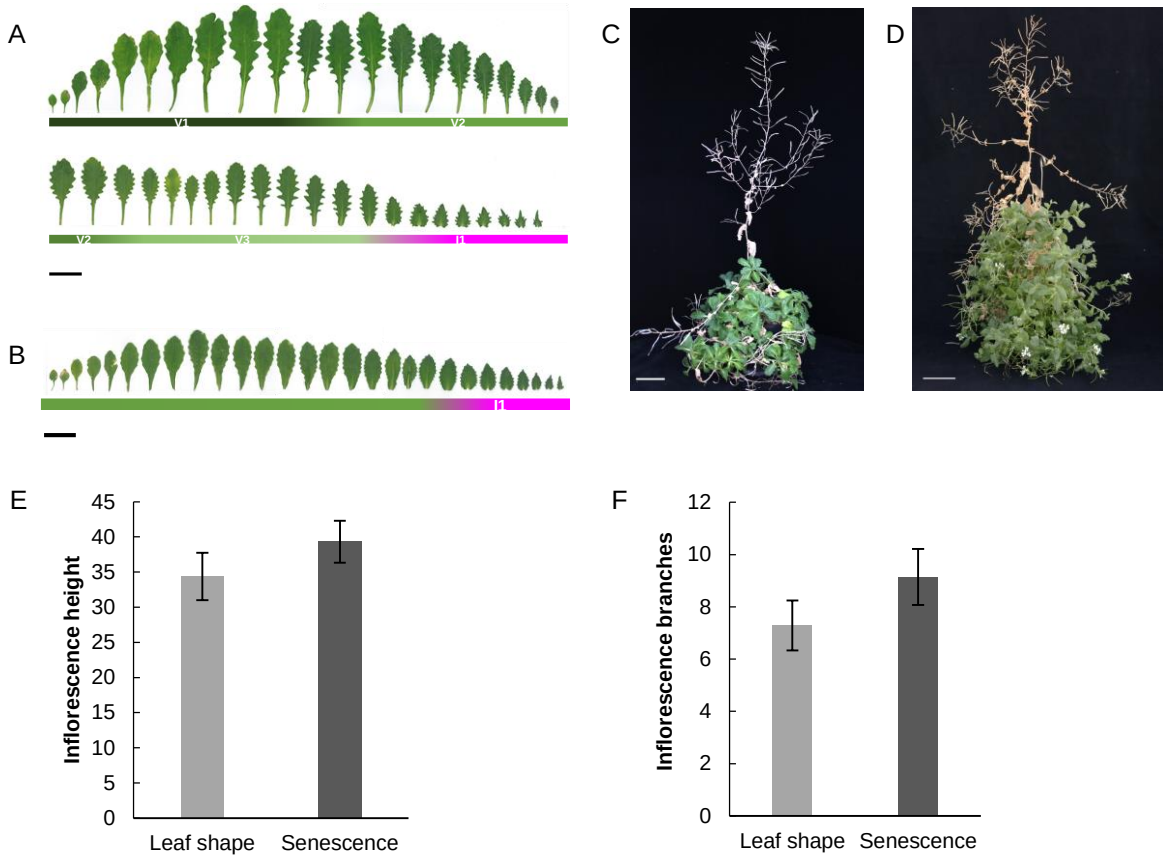


Figure 8. *pep1-1* inflorescence zone is detectable with leaf shape and senescence.
(A) Changes in leaf shape within Pajares vegetative and inflorescence zones. Upper panel are leaves of an 8 weeks old plant before vernalization and lower panel are leaves of a flowering plant after vernalization. **(B**) Leaf shape changes in *pep1-1* non-inflorescence and inflorescence zones. **(C)** Senescence of the inflorescence in Pajares. **(D)** Senescence of *pep1-1* inflorescence. **(E)** Height of *pep1-1* based on leaf shape and senescence. **(F)** Number of inflorescence branches in *pep1-1* scored with leaf shape and senescence. Data are presented as means ± SD, n = 12. Scale bars = 5 cm.

Table 1. Siliques number comparison in *pep1-1* and Pajares plants

	Paj 8w (8.22±5.78)	Paj 12w (27.72±4.22)	Paj 15w (33.5±3.26)	Paj 18w (36.75±2.98)	Paj 21w (38.25±3.93)	Paj 24w (40.41±3.96)
	p value					
pep1-1 (23.7±5.62)	0.0053	0.0699	0.0022	0.0018	0.0017	0.0018

a. Numbers in bracket represent the means and SD of siliques
b. The mean difference is significant at *p* value < 0.05

2.2.2 Flowering time and inflorescence characterization of *eop101*

The *eop101* mutant, through sequencing, was identified to carry lesions in the *A. alpina TERMINAL FLOWER 1* (*AaTFL1*) gene that caused an amino acid substitution of proline (Pro) for serine (Ser) (Figure 9A, Supplementary Dataset S1). *eop101* flowered earlier than *pep1-1* and it first flowered with basal lateral branches (Figure 9B and C). Cauline-like leaves were barely visible at this point therefore rosette leaves were scored for number of leaves at flowering (Figure 9D and E, upper panel). The cauline-like leaves were later observed and scored at flowering of the main shoot apical meristem (Figure 9D). The cauline-like leaves were distinguishable based on leaf shape, as like in *pep1-1*, these were more rounded at the base and directly attached to the stem (Figure 9E, lower panel).

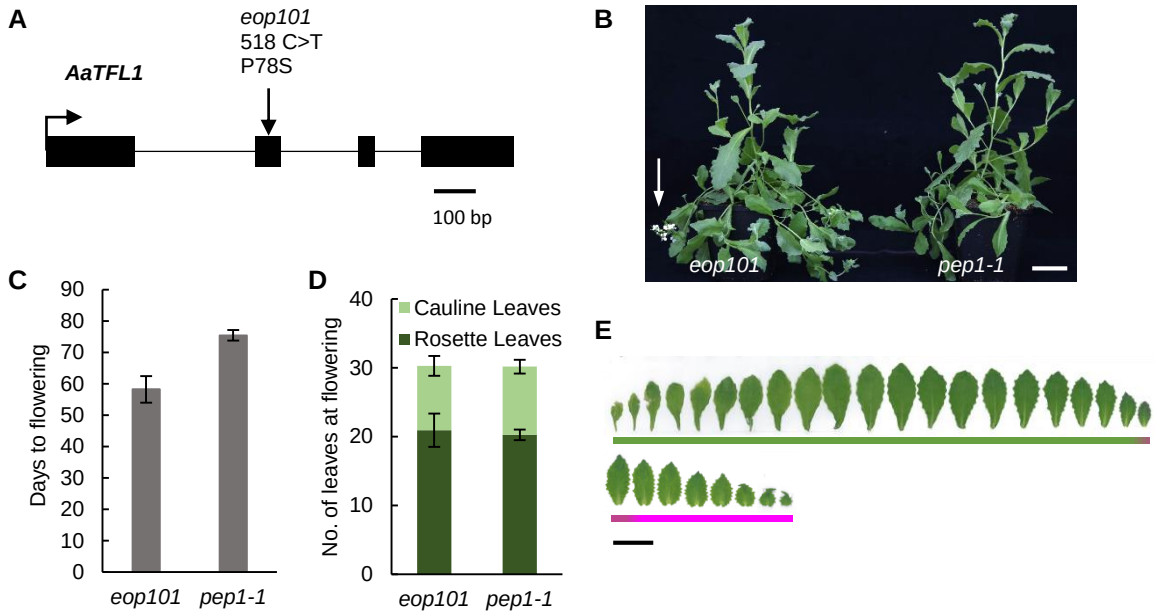


Figure 9. Lesion in *AaTFL1* causes early flowering in *eop101*.
(A) Mutation in the second exon of *AaTFL1* causes a Pro to Ser amino substitution in *eop101*. **(B)** *eop101* flowers earlier than *pep1-1* by the flowering of basal lateral branches (shown with arrow) **(c)** Number of leaves present at flowering of basal lateral branches and main shoot apical meristem (SAM). **(D)** Changes in leaf shape along *eop101* stem. Data is presented as means ± SD, n = 12. Scale bar = 5 cm

Apart from flowering earlier than *pep1-1, eop101* also developed an inflorescence that terminated by the formation of a terminal flower as compared to the indeterminate inflorescence of *pep1-1* (Figure 10A and B). *eop101* and *pep1-1* produced similar number of cauline-like leaves, however the cauline-like leave axils in *eop101,* closest to the main SAM were sometimes occupied by single pedicellate flowers (Figure 10A, insert) while those furthest away were occupied by lateral branches that terminated by forming a single flower. Some leaf axils also possessed just two flowers on a pedicel, these were included in the scoring

of inflorescence branches. The presence of either a branch or solitary pedicellate flower was therefore scored (Figure 10C). Due to the formation of a terminal flower, the inflorescence height and number of siliques were greatly reduced at I2 in *eop101* (Figure 10D and E). Senescence along the stem of *eop101* extended much further than the inflorescence zone scored with leaf shape (Figure 10F). Scoring of inflorescence features, mainly in I1, based on senescence in this mutant revealed significant differences from the number of inflorescence branches and total inflorescence height scored with leaf-shape (Figure 10G and H).

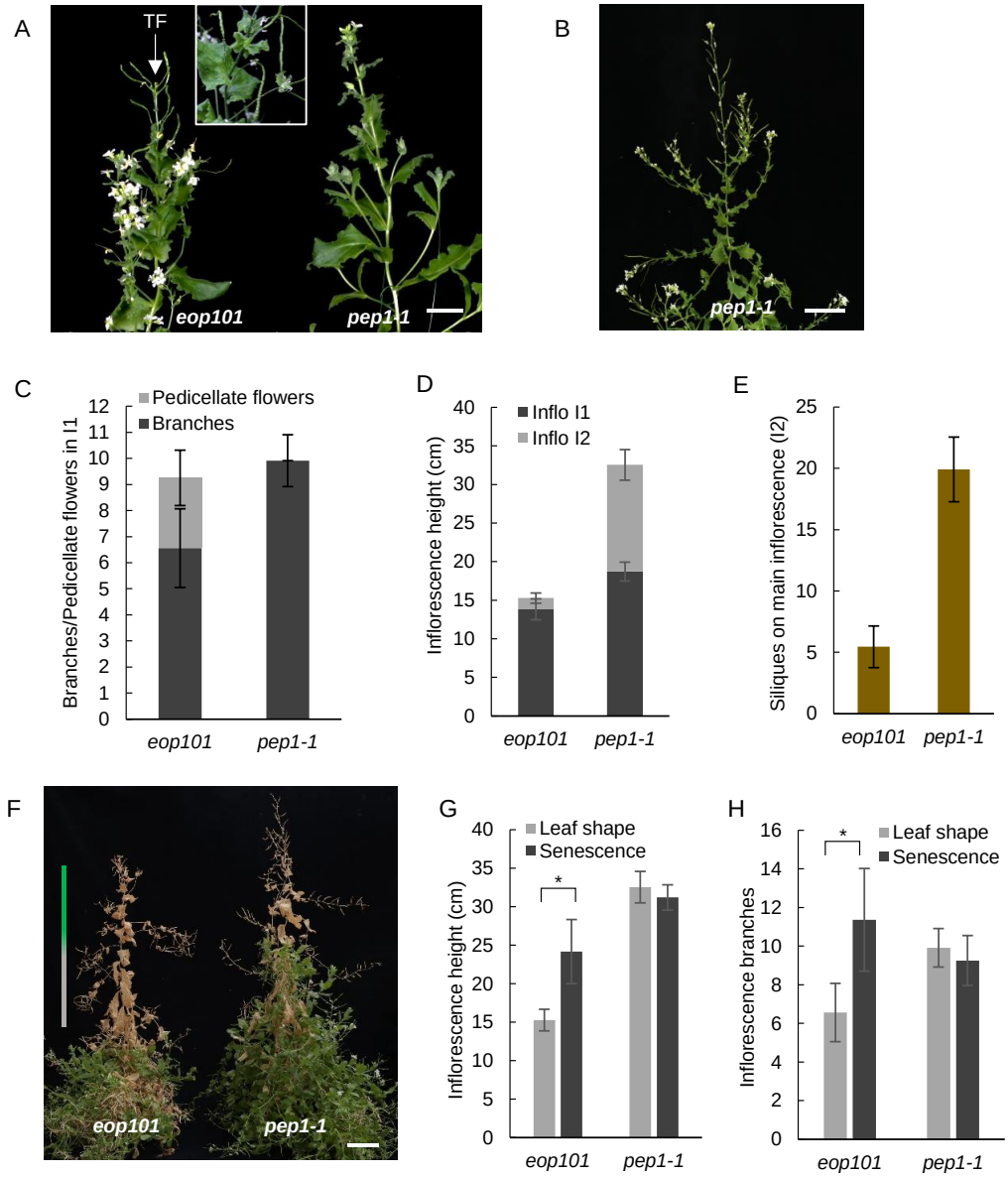


Figure 10. *eop101* possesses a determinate inflorescence by the formation of a terminal flower (TF). **(A)** Determinate inflorescence and pedicellate flowers (insert) in *eop101*. **(B)** Indeterminate inflorescence of *pep1-1*. **(C)** Number of inflorescence branches and pedicellate flowers in *eop101* and *pep1-1*. **(D)** Height of *eop101* and *pep1-1* inflorescence I1 and I2. **(E)** Number of siliques in I2 of *eop101* and *pep1-1*. **(F)** Senescence along *eop101* stem, the bar depicts the extent of senescence, the green part corresponds to inflorescence scored with leaf shape. **(G)** Inflorescence height and **(H)** number of inflorescence branches scored with leaf shape and senescence. Data are presented as means ± SD, n = 12. Scale bar = 5 cm

2.2.3 RNA-seq of *eop101* for genome wide expression analysis

RNA sequencing was performed on samples from leaves and apices of *eop101* and *pep1-1* 3 weeks old seedlings to study transcription changes that might lead to early floral transition. Differential expressed genes (DEGs) were determined by comparing gene expression levels in *eop101* to that in *pep1-1*. Within the leaves, *eop101* contained 247 genes down-regulated and 575 up-regulated. In apices there were 455 genes down-regulated and 312 up-regulated, indicating that in *eop101* the overall number of differentially regulated genes in leaves and apices is relatively similar (Figure 11). However there were more up regulated genes in leaves and more down-regulated genes in apices in comparison to *pep1-1*.

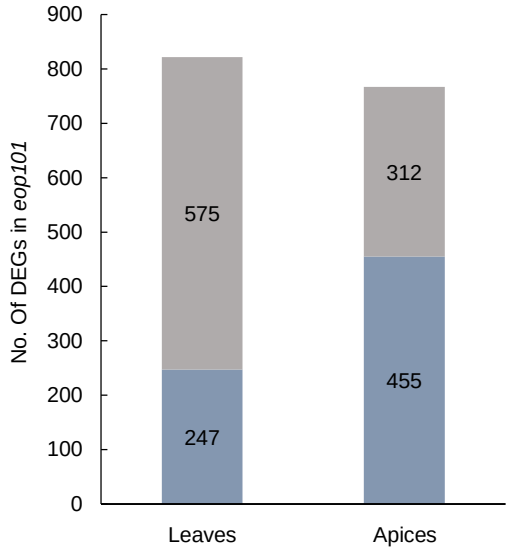


Figure 11. Differentially expressed genes (DEGs) in *eop101* apices and leaves.
Number of DEGs in *eop101* leaves and apices as compared their expressions in *pe1-1*.

2.2.4 Gene Ontology (GO) enrichment analysis for functional classification of DEGs from *eop101*

GO analysis was performed to help group DEGs into functional categories. VirtualPlant software platform [172] was employed as a tool for the identification of overrepresented GO terms in the data set. To enable the use of this tool, *A. thaliana* orthologues had to be assigned to the *A. alpina* genes. The initial assignment of *A. thaliana* orthologues was performed using an orthologue list called out with more stringent parameters (strict-list, obtained from AG Schneeberger, MPIPZ Cologne) to reduce the rate of false positives calls. The drawback of this approach was the fact that, approximately 35% of genes in *A. alpina* were not assigned *A. thaliana* orthologues. Thus, to make up for this, a list with *A. thaliana* orthologues based on less stringent parameters (relaxed-list, obtained from Dr. Julieta L. Mateos) was also used. This helped obtain 86% of *A. alpina* genes with orthologues however it has to be bared in mind that there is relatively higher probabilities of false orthologue calls. Table 2 below summaries the number of *A. thaliana* orthologous genes used for the prediction of enriched GO terms.

Table 2. Number of *A. thaliana* orthologues used for GO-analysis

		Leaves		Apices	
		Strict-list	Relaxed-list	Strict-list	Relaxed-list
eop101	UP	390	464	215	252
	Down	174	196	272	344

To determine enriched GO terms, up and down-regulated genes in apices and leaves were compared to the whole genome of *A. thaliana*. Majority of enriched DEGs were found in the leaves and mainly consisted of up regulated genes. Based on the GO analysis of biological process, DEGs were grouped into 17 main functional categories (Table.3). Among the categories was, response to stress and biological process regulation, these contained the highest numbers of DEGs, followed by overrepresentation of genes involved in carbohydrate, nitrogen, sulfur, transport and developmental processes. The hormones, which included response to abscisic acid, jasmonic acid, ethylene, auxin and salicylic acid, together also possessed a high number of enriched genes.

Table.3 Biological functional categories of *eop101* DEGs

Category		*eop101* DEGs	
		Apices	Leaves
Developmental process	UP	36	4
	Down	0	0
Sulfur process	UP	0	24
	Down	8	0
Carbohydrate process	UP	0	39
	Down	22	0
Nitrogen process	UP	28	19
	Down	0	0
Abscisic acid (ABA)	UP	0	15
	Down	9	0
Jasmonic acid (JA)	UP	0	14
	Down	0	0
Ethylene (ET)	UP	0	7
	Down	0	0
Auxin	UP	0	3
	Down	6	0
Salicylic acid (SA)	UP	0	14
	Down	0	0
Cation homeostasis	UP	0	0
	Down	8	0
Response to stress	UP	40	92
	Down	55	0
Transport	UP	0	12
	Down	38	0
Response to water deprivation	UP	0	18
	Down	0	0
Response to temperature stimulus	UP	0	19
	Down	0	0
Response to light stimulus	UP	15	33
	Down	0	0
Rhythmic process	UP	0	5
	Down	0	0
Biological process regulation	UP	0	99
	Down	18	9

a. DEGs, Differentially Expressed Genes
b. Up-regulated genes are shaded in red and down-regulated in blue

2.2.5 Developmental genes misregulated in *eop101*

In the *eop101* mutant, meristem determinacy, leaf morphogenesis and root development, were the specific terms associated with development process. There were three genes, associated with meristem determinacy and four genes with leaf morphogenesis all up-regulated in apices. On the other hand four genes were involved in root development and up-regulated in leaves (Table 4). Overrepresented root development gene included, *CYTOCHROME P450, FAMILY 83, SUBFAMILY B, POLYPEPTIDE 1* (*CYP83B1*), *SUPERROOT 1* (*SUR1*), *LJRHL1-LIKE 3* (*LRL3*) and *ROP* (*RHO OF PLANTS*) *GUANINE NUCLEOTIDE EXCHANGE FACTOR 11* (*ROPGEF11*). Out of the three genes associated with meristem determinacy, two were specific to meristem determinacy while one was also involved in leaf morphogenesis. *CRABS CLAW* (*CRC*) and *EMBRYONIC FLOWER 1* (*EMF1*) were the two genes involved in meristem determinacy while *CAROTENOID CLEAVAGE DIOXYGENASE 8* (*CCD8*), *INCURVATA2* (*ICU2*) and *PLASTID TRANSCRIPTION FACTOR 1* (*PTF1*) were involved in leaf morphogenesis. *HOMEOBOX 51* (*HB51*), was involved in both processes. *CRC*, among the genes in meristem determinacy, is involved in specifying abaxial cell fate in the carpel and contains four putative binding sites of the floral meristem identity gene *LFY* [173]. *EMF1,* like *TFL1,* acts as a floral repressor [174]. An interesting gene associated with development process, but not placed under a specific term was *SEPALLATA 4* (*SEP4*), it plays an important role in the maintenance of floral meristem identity [175]. Another interesting gene not placed in this category but also up regulated in the apices was *FT*, which together with LFY, promotes flowering [18, 92]. *HB51* was called out as a gene associated with both meristem determinacy and leaf morphogenesis. This is mainly because, *HB51* is a meristem identity regulator that acts together with *LFY* to induce the expression of meristem identity gene, *CAULIFLOWER* (*CAL*). *HB51* also independently of *LFY*, plays roles in leaf morphogenesis [176]. Among the other leaf morphogenesis genes, *CCD8* is involved in the synthesis of an apocarotenoid phytohormone that suppresses lateral shoot growth [177], *PTF1* is involved in heterochronic regulation of leaf development [178] and mutations in *ICU2* causes curled leaves, early flowering and the up regulation of genes responsible for floral organ identity [179]. Taking these together gives indication that, genes acting on shoots identity and genes that interact with floral meristem identity genes are up-regulated in *eop101,* together with genes involved in leaf morphogenesis, at the seedling stage as compared to *pep1-1*.

Table 4. Genes in *eop101* associated with developmental process							
Arabidopsis gene	**Primary Gene Symbol**	**Gene Model Description**	**FPKM_ *eop101***	**FPKM_*pep1-1***	**Root develop-ment**	**Meristem determinacy**	**leaf morpho-genesis**
Genes in *eop101* leaves associated with developmental process							
AT4G31500	*CYTOCHROME P450, FAMILY 83, SUBFAMILY B, POLYPEPTIDE 1 (CYP83B1)*	Required for phytochrome signal transduction in red light. Mutation confers auxin overproduction.	8,20	4,99	X		
AT2G20610	*SUPERROOT 1 (SUR1)*	Confers auxin overproduction. Mutants have an over-proliferation of lateral roots. Encodes a C-S lyase involved in converting S-alkylthiohydroximate to thiohydroximate in glucosinolate biosynthesis.	93,20	58,44	X		
AT5G58010	*LJRHL1-LIKE 3 (LRL3)*	Encodes a basic helix-loop-helix (bHLH) protein that regulates root hair development.	3,73	1,50	X		
AT1G52240	*ROP (RHO OF PLANTS) GUANINE NUCLEOTIDE EXCHANGE FACTOR 11 (ROPGEF11)*	Encodes a member of KPP-like gene family, homolog of KPP (kinase partner protein) gene in tomato.	47,28	31,61	X		
Genes in *eop101* apices associated with developmental process							
AT1G69180	*CRABS CLAW (CRC)*	Involved in specifying abaxial cell fate in the carpel. Contains four putative LFY binding sites (CCANTG) and two potential binding sites for MADS box proteins known as CArG boxes (CC(A/T)6GG)	5,73	3,03		X	
AT5G11530	*EMBRYONIC FLOWER 1 (EMF1)*	Involved in regulating reproductive development	6,20	4,04		X	
AT5G03790	*HOMEOBOX 51 (HB51)*	Encodes a homeodomain leucine zipper class I (HD-Zip I) meristem identity regulator that acts together with LFY to induce CAL expression. The gene also had additional LFY-independent roles in leaf morphogenesis and bract formation.	15,40	10,61		X	X
AT4G32810	*CAROTENOID CLEAVAGE DIOXYGENASE 8 (CCD8)*	Involved in the production of a graft transmissable signal to suppress axillary branching.	1,40	0,92			X
AT5G67100	*INCURVATA2 (ICU2)*	A number of regulatory genes were derepressed in the icu2-1 mutant, including genes associated with flowering time, floral meristem, and floral organ identity. Mutant has curled, involute leaves and causes early flowering.	19,96	16,04			X
AT3G02150	*PLASTID TRANSCRIPTION FACTOR 1 (PTF1)*	TCP gene involved in heterochronic control of leaf differentiation.	6,72	4,74			X

a. FPKM: Fragments Per Kilobase Million
b. X : Gene is associated with the GO Term

2.2.6 Carbohydrate process and auxin biosynthetic genes in *eop101*

The category of auxin biosynthetic process in *eop101* leaves DEGs consisted of genes such as *CYP83B1 (SUR2)* and *SUR1*, which were previously mentioned as part of enriched genes involved in root development in *eop101* (Section 2.4.3). The *sur1* and *sur2* mutants have been demonstrated to accumulate high levels of free auxin due to overproduction [180, 181]. *CYP83B1 (SUR2)* has further been shown to catalyze the conversion of indole-3-acetaldoxime (IAOx), a precursor for auxin biosynthesis, to 1-aci-nitro-2-indolyl-ethane, which is used for the production of S-alkyl-thiohydroximates, the precursors of glucosinolates [182]. Thus, resulting in reduced levels of auxin. Interestingly, in accordance with this, there was an overrepresentation of up-regulated genes involved in carbohydrate process, more specifically glucosinolate biosynthetic process, in *eop101* leaves. *eop101* also possessed some down-regulated genes associated with auxin response in the apices, such as *INDOLEACETIC ACID-INDUCED PROTEIN 10 (IAA10)* and *AUXIN RESISTANT 3* (*AXR3/IAA17*), both AUX/IAA genes whose expression and degradation are induced by auxin. *NAC1*, a transcriptional activator induced by auxin to activate the expression of other auxin-responsive genes [183], happen to be part of the down-regulated genes together with other auxin inducible genes. Taken these together, the presence of *SUR1* and *SUR2* up-regulated in apices and several auxin inducible genes down-regulated in leaves, might mean a reduction of free auxin in *eop101* due to the diversion of auxin precursors into the production of glucosinolates.

2.2.7 *eop101* DEGs involved in biological process regulation

Biological process regulation was the category with the highest set of genes. This consisted of 99 up-regulated genes in the leaves of *eop101*. To obtained a general overview of genes present within the 99 enriched genes, a smaller set of GO was performed on these genes in 'The Arabidopsis Information Resource' (TAIR) [184]. In this analysis, the GO overrepresented terms were based on comparison to the input dataset and not the entire *A. thaliana* genome. This revealed, among the 99 genes ~78% were located in the nucleus (Figure 12A). GO of molecular functions with the highest percentage of genes involved were, transcription factor activities, protein binding and DNA or RNA binding (Figure 12B), additionally, in terms of biological processes, a high percentage of these genes were involved in DNA dependent transcription (Figure 12C). These indicated that most of the up regulated genes in the leaves

of *eop101* are transcription factors. Among these were genes involved in circadian clock-controlled flowering pathway such as *CYCLING DOF FACTOR 1* (*CDF1*), *GI* and *LHY*. Whereas *GI* and LHY are considered to promote flowering through the photoperiodic response pathway, *CDF1* has been shown to repress photoperiodic flowering response by the transcriptional repression of *CO* [64, 185-187]. Another gene *SCHLAFMUTZE* (*SMZ*), which encodes an AP2 domain transcription factor and can repress flowering was also among these set of genes together with *SUPPRESSOR OF PHYA-105 1* (*SPA1*), a gene involved in the regulation of circadian rhythms and flowering time [188, 189]. *TEMPRANILLO 1* (*TEM1*), an ethylene inducible gene, that belongs to the RAV subfamily of transcription factors and is involved in the repression of flowering under long days, was also part of the group here [89, 90]. Transcription factors involved in light response such as, *PHYTOCHROME-INTERACTING FACTOR 5* (*PIF5*), *LONG HYPOCOTYL IN FAR-RED* (*HFR1*), *CRY1* [190-192], were also present. In addition to their roles in light response, these genes are also involved in hormone response, *PIF5* is involved in auxin signaling and ethylene biosynthesis [193, 194], *HFR1* in ABA signaling [195] and *CRY1*, auxin transport [196]. Other genes present included those involved in stress response such as*, BASIC LEUCINE-ZIPPER 1* (*bZIP1*) which positively regulates salt tolerance and drought stress [197], *DREB AND EAR MOTIF PROTEIN 1 (DEAR1)* which is involved in defense and freezing stress responses [198], DRE-BINDING PROTEIN 2A (DREB2A) a gene involved in response to drought and low-temperature stress [199], and WRKY DNA-BINDING PROTEIN 33 (WRKY33) which is involved in response to various abiotic stresses including salt stress, cold and water deprivation [200]. The up regulation of several transcription factors involved in various processes in the *eop101* mutant, provides a possible explanation for the observation of mainly up-regulation enriched genes involved in the various GO categories in the leaves.

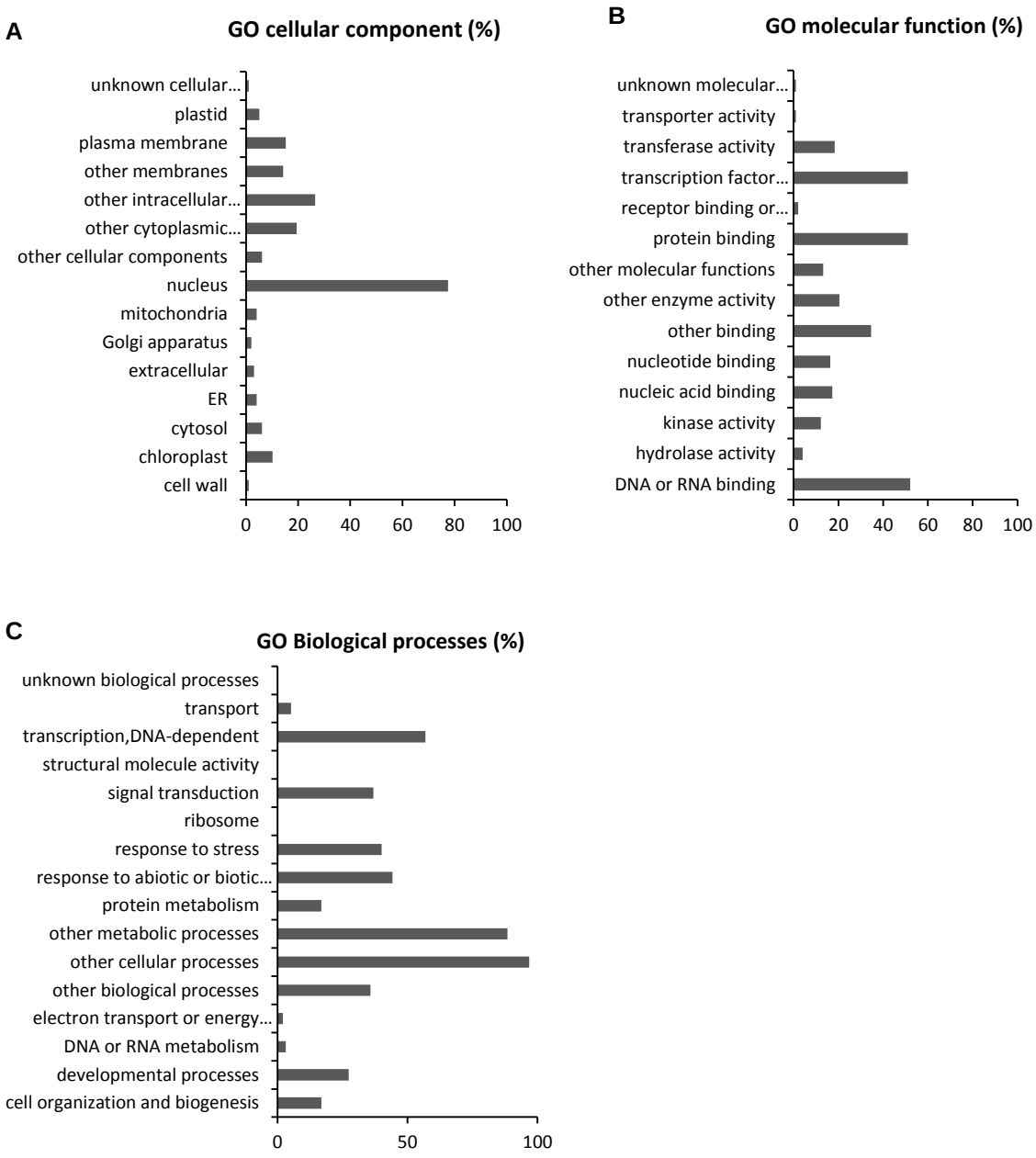


Figure 12. GO analysis of *eop101* DEGs involved in Biological process regulation.
(A) Percentage of genes located in different cellular components **(B)** Percentage of genes involved in various molecular functions **(C)** Percentage of genes involved in different biological processes

2.2.8 Flowering time and inflorescence characterization of *eop002*, *eop085*, *eop088* and *eop091* compared to *pep1-1*

The other *eop* mutants isolated from the mutagenesis of *pep1-1,* unlike *eop101,* produced an indeterminate inflorescence. These mutants, *eop002*, *eop085*, *eop088* and *eop091,* flowered earlier than *pep1-1* with 15 or less leaves, made up of 3-4 cauline-like leaves and 11-12 rosette-like leaves, as compared to 30 ± 0.5 leaves in *pep1-1*, consisting of 10 cauline-like leaves and 20 rosette-like leaves (Figure 13A and B). This was also consistent in terms of days to flower (Supplementary Figure S1), *eops* flowered 7 weeks after sowing whereas *pep1-1* flowered after 10 weeks. For comparison to *pep1-1* inflorescence, the inflorescence zone of the *eops* was identified by using leaf shape and senescence as morphological marker just as conducted for *pep1-1* (Figure 13C and D).

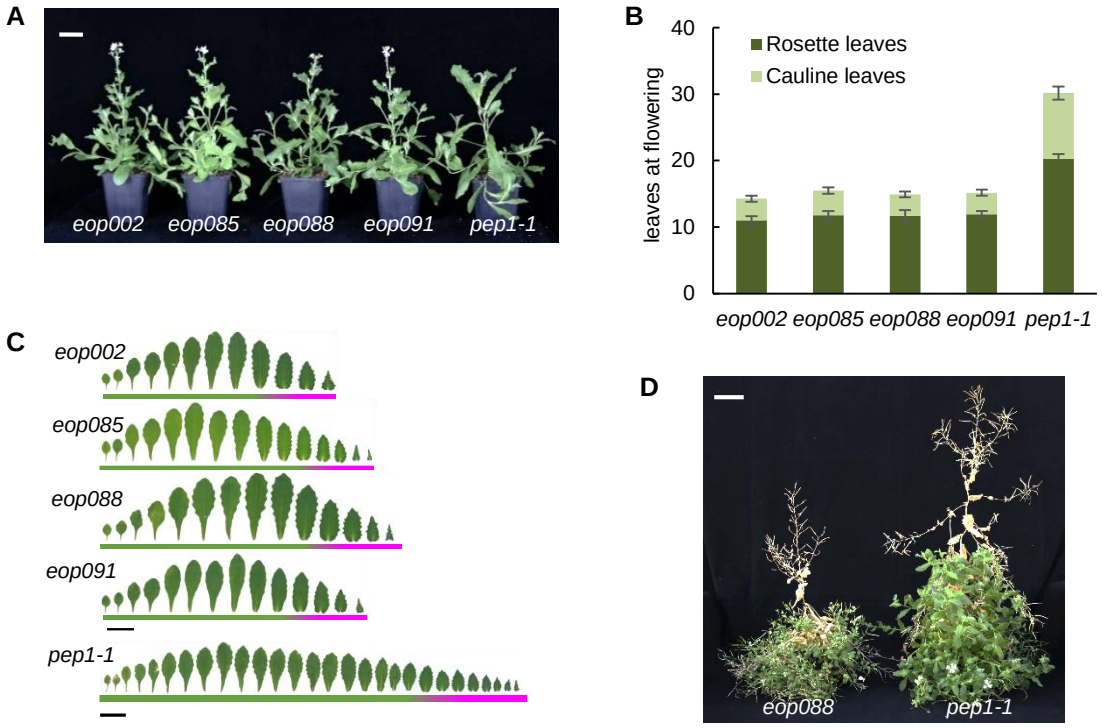


Figure 13. eop mutants flower earlier than *pep1-1* and their inflorescence zone can be detected with leaf shape and senescence
(A) Early flowering of *eop* mutants compared to *pep1-1* **(B)** Number of leaves at flowering in *eop* mutants and *pep1-1* **(C)** Leaf shape changes in *eops*. Non-inflorescence zone is shaded in green and inflorescence zone in magenta **(D)** Inflorscence senescence *eop08*8 compared to *pep1-1*. Data are presented as means ± SD, n = 12. Scale bars = 5 cm

The number of inflorescence lateral branches in I1, in correlation with the number of cauline-like leaves, were much reduced in the *eops*. While *pep1-1* contained on average 9 inflorescence branches, the *eops* possessed 3 to 4 inflorescence branches (Figure 14A). However the scoring of solitary flowers, as siliques present, showed no apparent differences in the number of siliques in the I2 zone of *eops* and *pep1-1*, as all individuals had on average 19 to 20 siliques (Figure 14B). Measurement of the inflorescence height with respect to the I1 and I2 zones, further confirmed a reduction in the I1 zone of the *eops* compared to *pep1-1* but no significant difference in I2 zone (Figure 14C). This reduction in I1 overall affected the total number of siliques produced in the *eops* inflorescence (Figure 14D). This observation of no difference in number of I2 siliques was also true for the other replicates, however *pep1-1* in replicate 3 exhibited a slight reduction in number of I2 siliques compared to the other replicates (Supplementary Figure S3F), the inflorescence height, mainly I2, of *pep1-1* in this replicate was also much reduced compared to previous replicates (Supplementary Figure S2F). Nevertheless, in comparison to *pep1-1* inflorescence, all four isolated *eops* possess reduced I1 zones with less number of inflorescence branches but no apparent changes in their I2 zones.

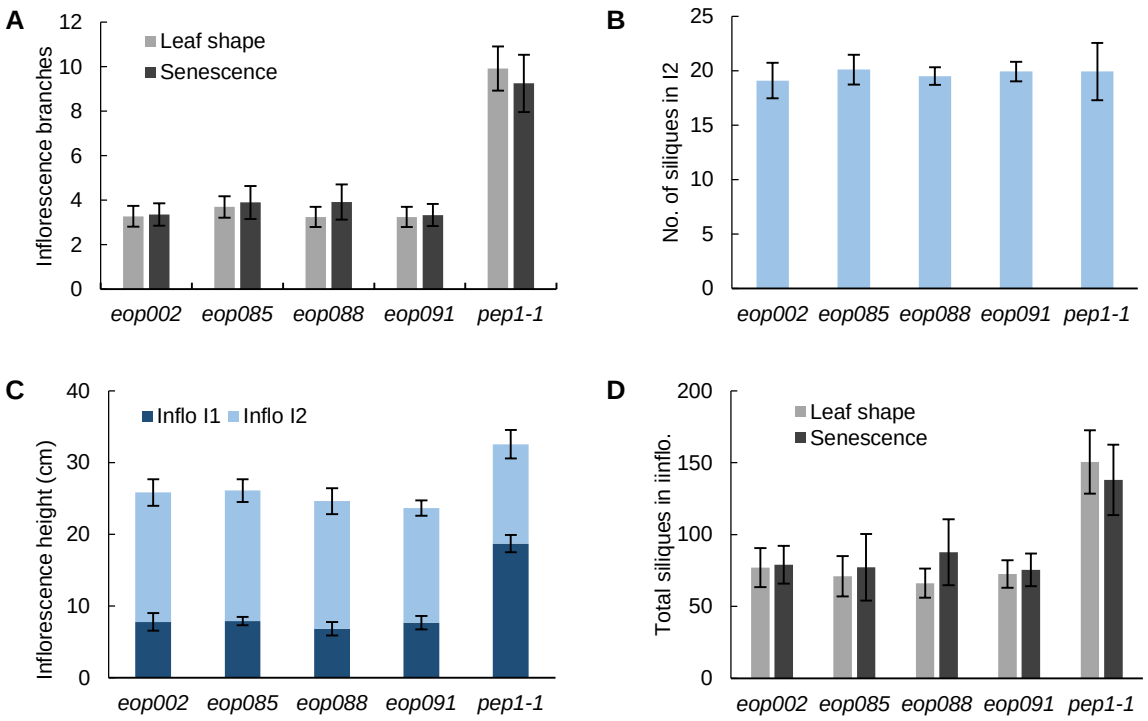


Figure 14. *eop* mutants produce simpler inflorescence with less inflorescence branches
(A) Number of inflorescence branches in I1 of *eops* and *pep1-1* scored based on leaf shape and senescence of the inflorescence **(B)** Number of siliques in I2 of *eop* mutants compared to *pep1-1*. **(C)** Height of inflorescence I1 and I2. **(D)** Total number of siliques in the inflorescence of *eop* mutants and *pep1-1*, scored based on leaf shape and senescence of the inflorescence. Data are presented as means ± SD, n = 12. Scale bar = 10 cm

2.2.9 Level of dominance and allelism testing

The attainment of the four *eop* mutants from the mutagenesis screen with similar phenotypes, lead to the question of whether these mutants are allelic. An allelism test was therefore conducted to assess the combined effect of the mutations on flowering time. In parallel, a dominance test was conducted to study the effect of mutant allele, on flowering time, in the presence of wild type allele. F1 individuals were generated for dominance and allelism testing by crossing the *eop* mutants, *eop002*, *eop085, eop088* and *eop091* with each other and to *pep1-1.* Flowering time was analyzed in terms of leaf number, and among the F1 progeny developed for allelsim test, no wild type, that is *pep1-1*, phenotype was observed indicating that the mutants might be allelic (Figure 15A). These mutants however showed variations in their degrees of dominance (Figure 15A). *eop002* F1s, from a cross to *pep1-1,* displayed intermediate phenotype taking into account the number of leaves *eop002* and *pep1-1* had at flowering. *eop088* F1s also showed such intermediate phenotype but to a lesser extent than *eop002*. The F1s of *eop085* and *eop091* flowered with number of leaves relatively closer to that of *pep1-1*. Potence ratios (PR) involves the use of population means to calculate the degree of dominance [201-203]. PR = 0 is indicative of no dominance, PR = 1 means complete dominance, 0 < PR < 1 means incomplete dominance, -1 < PR < 0 means incomplete recessivity and PR = -1 indicates complete recessivity. PR was calculated from the F1 data sets (according to formula provided by Sharma, J. R., 2006 [202]) to determine the degree of dominance as it seemed to vary among the mutants. The degrees of dominance with respect to the wild type allele obtained ranged from -0.21 in *eop002*, 0 in *eop088*, 0.25 in *eop085* to 0.77 in *eop091*. This indicated *eop091* and *eop085* alleles are recessive as there is a dominance of wild type allele over mutant allele with the strongest effect exhibited in *eop091*, however the dominance is incomplete. On the contrary, in *eop002* the mutant allele rather showed dominance over the wild type and this was also incomplete. *eop088* showed equal levels of dominance from both wild type and mutant allele. To further expand on this analysis, 48 F2 progeny from crosses used for the dominance test were phenotyped for their leaves at flowering. The number of leaves at flowering in the controls ranged from 11-13 in *eop002*, 15-17 in *eop085*, 12-14 in *eop088*, 12-15 in *eop091* and 27-31 in *pep1-1*. Majority of F2s from *eop085* and *eop091* showed a skewness towards *pep1-1* by flowering closer to the range in *pep1-1*. 38 *eop085* F2 individuals flowered with 21-30 leaves while 10 flowered with 15-19 leaves (Figure 15B). Similarly, 37 *eop091* F2 individuals flowered with 22-31 leaves, while 11

flowered with 13-16 (Figure 15C). These segregations fitted a 3:1 ratio (*eop085*: $X_1^2 = 0.444$, $P = 0.51$, *eop091*: $X_1^2 = 0.389$, $P = 0.53$) indicating that these two alleles are recessive to that of *pep1-1*. The opposite was observed for *eop002* as the skewness was towards *eop002* flowering range, 39 individuals flowered with 11-19 and 9 with 22-28 leaves (Figure 15D). This fitted the 3:1 ratio ($X_1^2 = 1$, $P = 0.32$) but suggested the *eop002* allele is dominant over that of *pep1-1*. In *eop088* F2s, there was a distribution of 12 individuals with 12-14 leaves, an intermediate group of 24 with 18-21 leaves and a late group of 12 individuals with 22-32 leaves (Figure 15E). This fitted in good accord with a 1:2:1 ratio, indicating semi-dominance of the *eop088* allele.

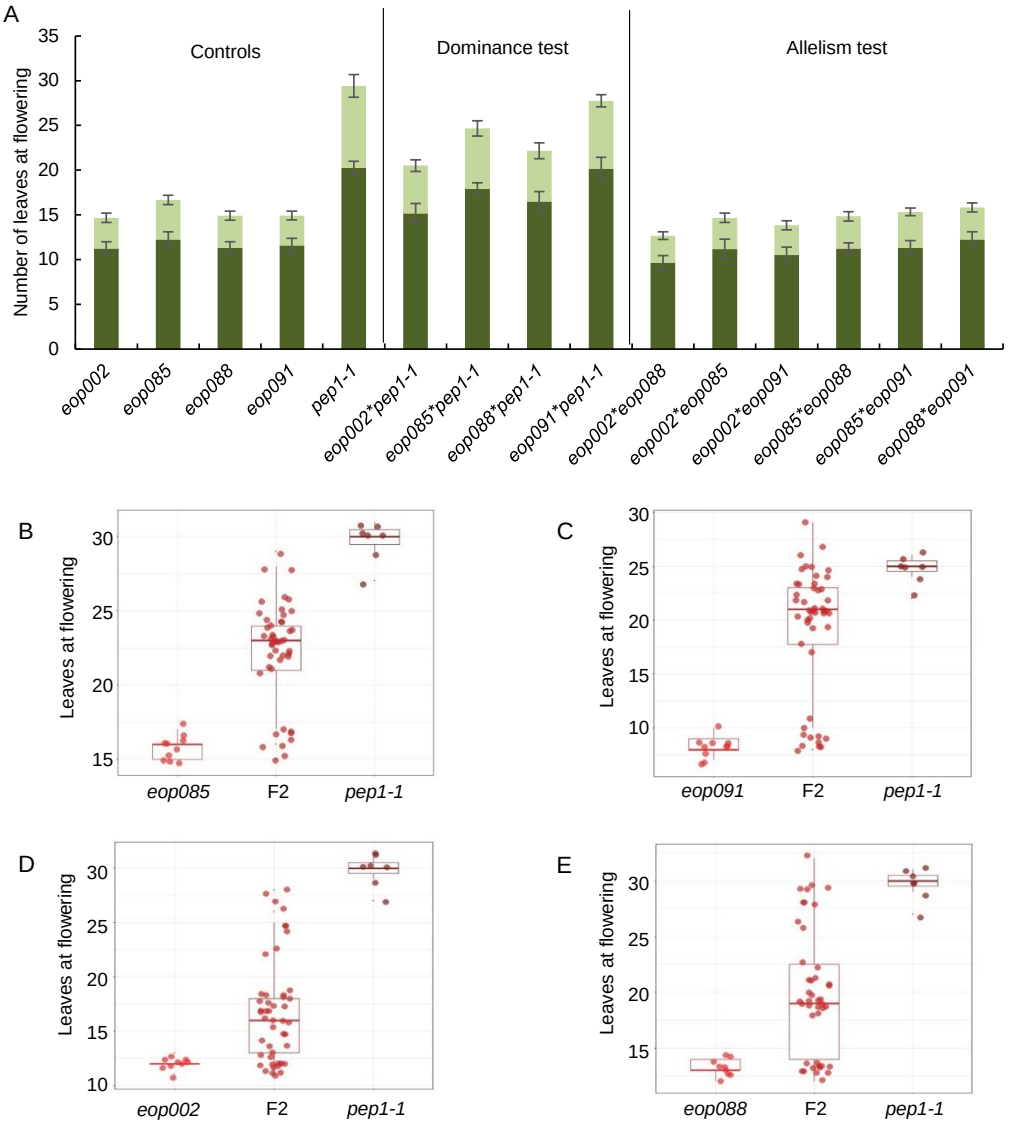


Figure 15. The *eop* mutants display varying degrees of dominance
(A) Number of leaves at flowering of F1s from *eop002*, *eop085*, *eop088* and *eop091* crossed to *pep1-1* for dominance test and crossed with each other for allelism test. **(B)** Flowering time segregation of F2 progeny from *eop* mutants crossed to *pep1-1*.

2.3 Molecular investigations into additional mechanisms affecting flowering and inflorescence in *A. alpina*

2.3.1 Mapping-by-Sequencing and SHOREmap for causal mutation identification

Three enhancer mutants of *pep1-1*, *eop002*, *eop085* and *eop091,* were previously backcrossed to *pep1-1* to generate F2 (BC1F2) mapping populations by Prof. Dr. Maria Albani. To ensure the careful selection of individuals that flowered as early as the mutants for sequencing, the F2s were planted alongside F1 individuals and parental lines (Figure 16A-C). DNA was extracted from a pool of flower bud materials collected from F2 individuals of each *eop* mutant and sent for NGS. Out of the sequences received, 94.86% of *eop002* reads, 96.12% of *eop085*, 93.44% of *eop091* and 92.67% of *pep1-1* reads were successfully aligned to the reference sequence. The average nucleotide genome coverage ranged from 145.46 for *pep1-1*, 102.9 for *eop091*, 89.2 for *eop002* to 20.83 for *eop085.* SAMtools was used for independent detection of differences between reference sequence and the mutant sequences including *pep1-1*. These results were further analyzed using SHOREmap.

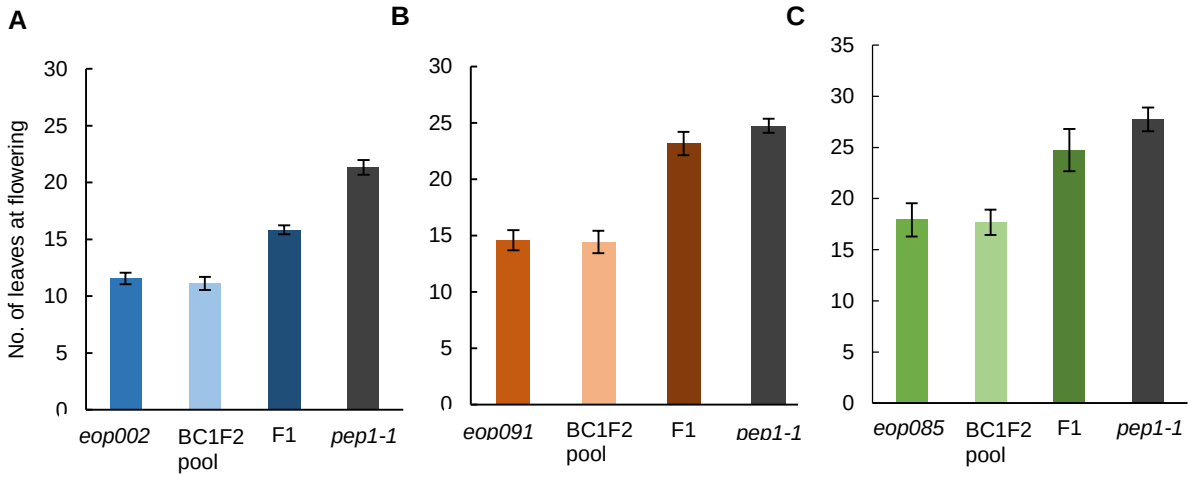


Figure 16. BC1F2 individuals selected for sequencing flowered as early as their *eop* parental lines. Average number of leaves at flowering of, **(A)** 92 *eop002* BC1F2 individuals, **(B)** 136 *eop091* BC1F2s and **(C)** 96 *eop085* BC1F2s compared to the flowering time their respective parental lines and F1s. Data are presented as means ± SD

As the *eop* mutants are second site mutants derived from EMS mutagenesis of *pep1-1*, sequence differences with their origin in *pep1-1* were filtered out together with non-EMS induced changes. SHOREmap provided allele frequency (AF) of EMS induced changes at particular loci on each chromosome (Chr.). The AF is calculated as the ratio of the mutant alleles divided by all reads at that particular locus. In a pool of bulked segregants, the causative change is expected to occur with the highest frequency, thus all EMS induced mutations with allele frequency higher than 0.7 were considered. 301 EMS changes were obtained for *eop002*, 144 for *eop085* and 147 for *eop091* (Supplementary Tables 1-3). Visualization of AF on the *A. alpina* chromosomes revealed frequency distortion on the Chr.8 of all sequenced *eop* mutants (Figure 17A-C). In addition, a distortion of allele frequency was also observed on the Chr.2 of *eop002* (Figure 17A).

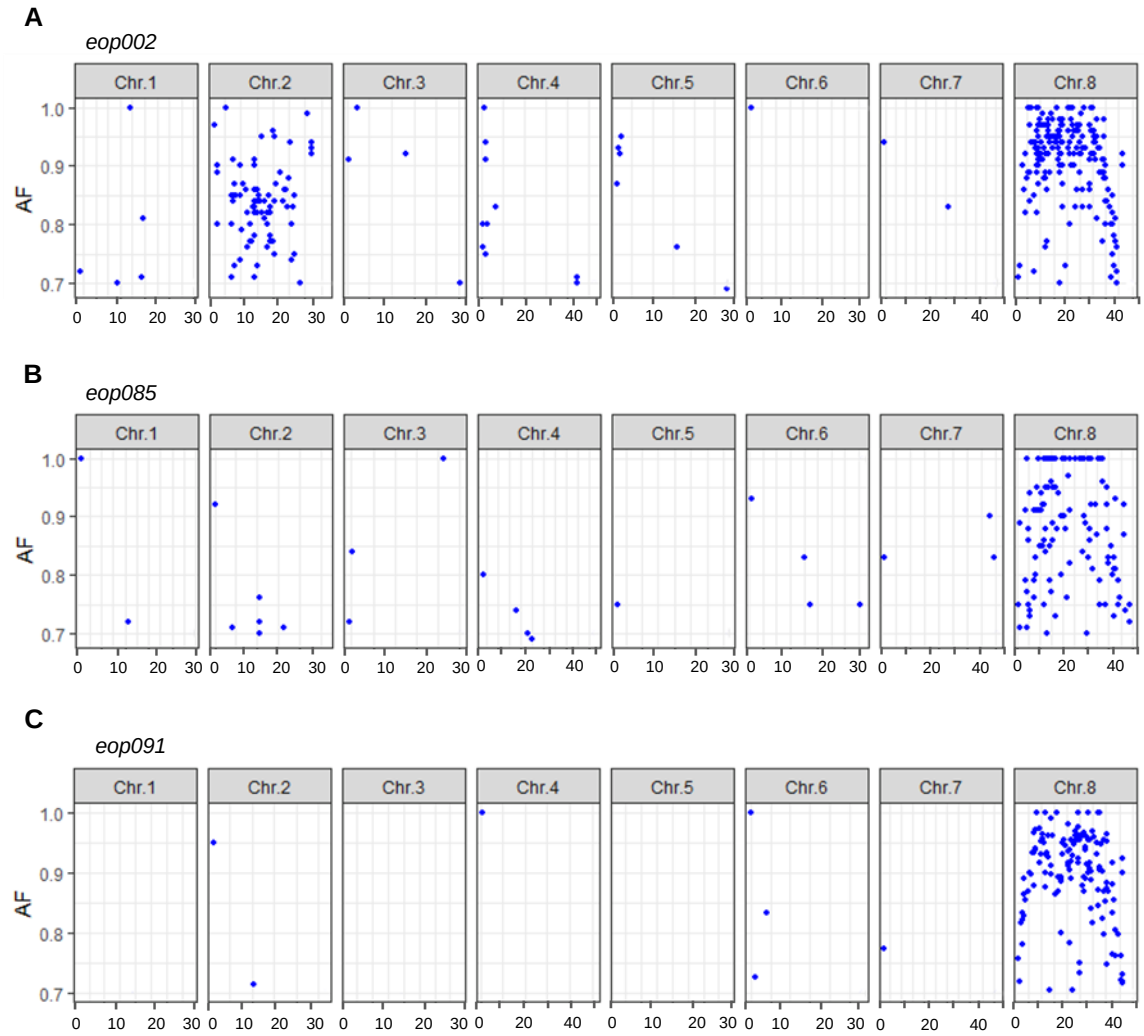


Figure 17. Presence of allele frequency (AF) distortion on Chr8 of the eo*ps* and on Chr2 of just *eop002*. AF estimation of EMS induced mutations from **(A)** *eop002*, **(B)** *eop085* and **(C)** *eop091*, across all 8 *A. alpina* chromosomes, represented in Mb on the x axis.

A mapping interval of 6Mb to 42Mb was defined on the Chr.8 of all the mutants. This region was unexpectedly large and covered 68% of the chromosome. In the case of *eop002*, a second region of 7Mb to 28Mb on Chr.2 was also identified which covered 57% of the chromosome. EMS changes located in splice-sites, coding and intronic regions of genes within the mapped regions were further investigated, Table 5 below summarizes the number of EMS changes obtained. Detailed list of the *A. alpina* genes containing these changes and their *A. thaliana* orthologues are provided in Supplementary Tables 4-6.

Table 5. SNPs within mapping intervals

	Chr 8			Chr 2		
	Coding	Intronic	Splite-site	Coding	Intronic	Splite-site
eop002	20	16	0	4	2	0
eop085	16	7	0	-	-	-
eop091	16	12	1	-	-	-

A candidate region harboring the casual mutation, in a segregating population, is expected to co-segregate with the phenotype of interest. Based on this assumption EMS changes, which served as SNPs, were used for the design of cleaved amplified polymorphic (CAPS) and derived cleaved amplified polymorphic (dCAPS) markers on Chr.2 and Chr.8 of *eop002* to examine their co-segregation with the flowering phenotypes of F2 individuals previously used in the analysis of dominance (Section 2.2.9). Three markers were designed for both chromosomes, one at the beginning of the mapped interval, the second in the middle and the last at the end (Figure 18A and C), however only two markers developed for Chr.2 worked (Information on markers available in Supplementary Table 7). Genotyping results revealed the absence of segregation for markers on Chr.2, all F2 individuals were homozygote for the mutant allele (Figure 18B). On the other hand, a close genotype-phenotype correlation was observed for the middle marker on Chr.8. In the case of the flanking markers, even though segregation was observed they did not completely correlate with the flowering phenotype (Figure 18D). The correlation between markers on Chr.8 and flowering time, and the mapping of candidate region in the other two *eop* mutants to the same chromosome indicated the region on Chr.8 is the candidate region harboring the casual mutation. Thus, subsequent analysis were focused on Chr.8.

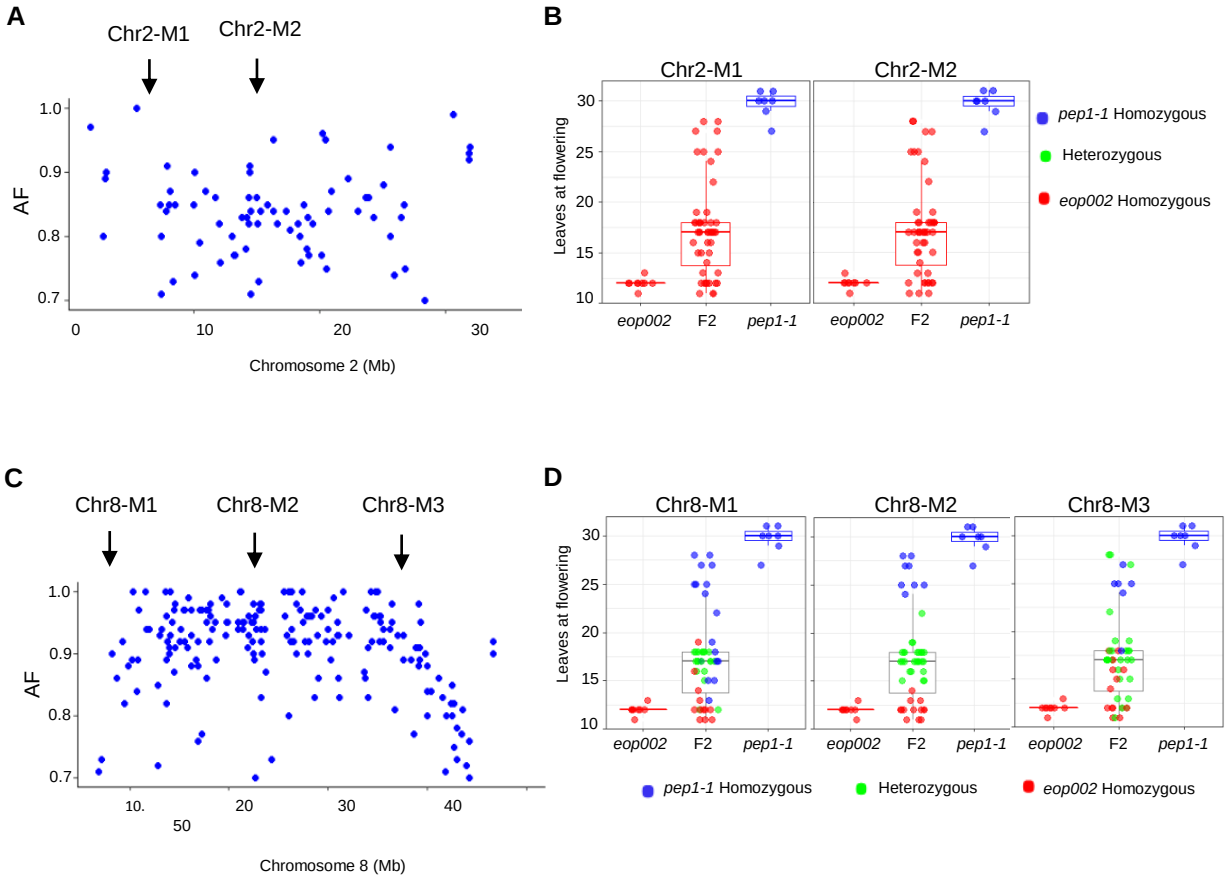


Figure 18. Markers on Chr8 but not Chr2 co-segregates with flowering time in *eop002* F2 segregating population.

(A) Position of markers on Chr.2. **(B)** Segregation of markers on Chr.2 with the flowering time of *eop002* F2 population. **(C)** Position of Chr.8 markers. **(D)** Segregation of markers with the flowering time of *eop002* F2 population,M2 exhibits the strongest genotype-phenotype correlation. N = 48 for F2, n = 12 for controls.

2.3.2 Genotype-phenotype co-segregation analysis in *eop002* and *eop091*

The mapped region of 34Mb on Chr.8 of the *eop* mutants was atypically large, however the discovery of recombination between markers used for co-segregation analysis in *eop002* F2s, provided a possibility to reduce this region. A larger population of 400 F2 individuals for *eop002* and 200 for *eop091* were genotyped to select plants that showed recombination among two extreme-most markers, independently designed for each mutant (Figure 19A and Figure 20A). Additional markers were further added to identify recombination breakpoints among the selected individuals (Information on markers available in Supplementary Table 7). In *eop002*, the region between markers M1 and M4 tightly co-segregated with the phenotypes of early flowering and reduced number of inflorescence branches (Figure 19B). This helped reduce the mapping interval from 34Mb to 8.8Mb and limited the number of candidate genes to seven genes with EMS changes (Figure 19C).

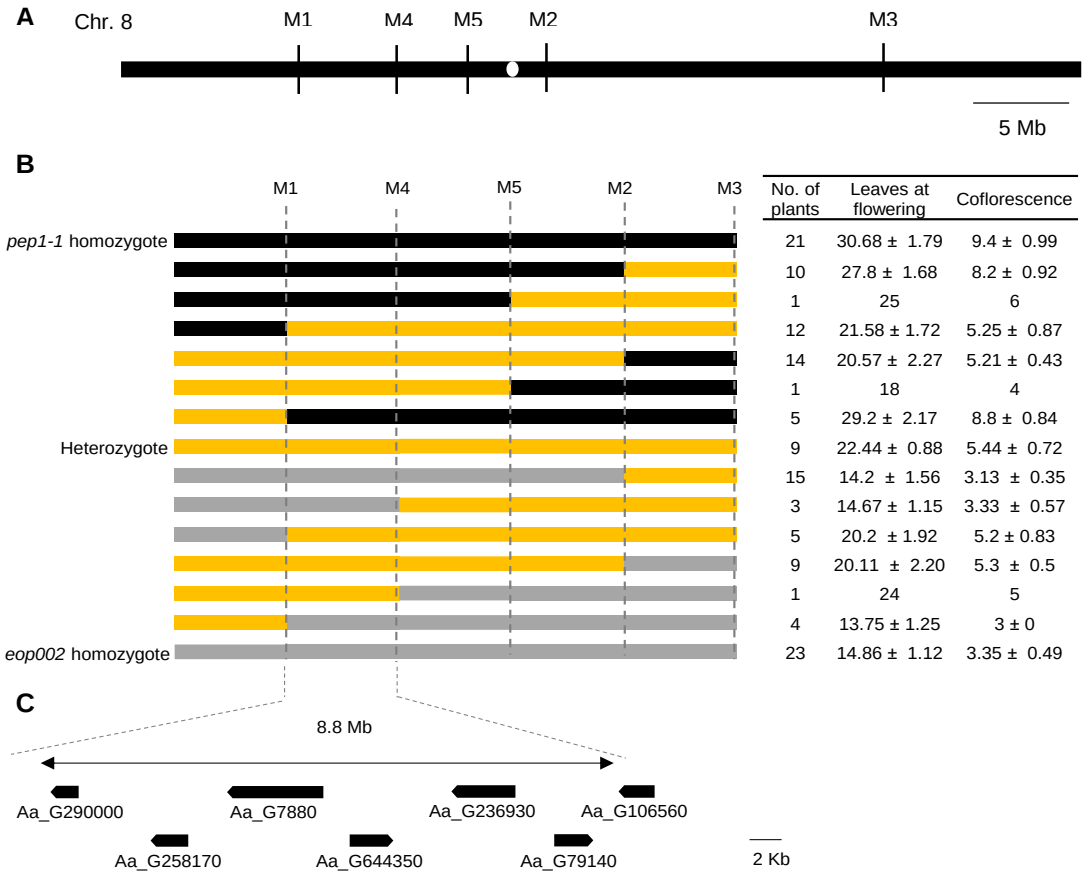


Figure 19. Co-segregation analysis in *eop002* F2 population enables reduction in mapping interval.
(A) Position of SNP markers for genotyping on Chr. 8. The white dot represents the centromere **(B)** Correlation of markers segregation with flowering and inflorescence phenotypes. Black bars represent chromosomal segment of wild type (*pep1-1*) homozygous, yellow bars for heterozygous and grey bars for mutant homozygous. **(C)** Genes with EMS changes within the reduced mapping interval.

In *eop091*, while markers A1, A4 and A5 did not show complete co-segregation with the early flowering and reduced inflorescence branches number, markers A2 and A3 were tightly linked to these phenotypes (Figure 20B). This resulted in the reduction of the mapping interval to 14.5 Mb and number of candidate genes to nine (Figure 20C). To further identify all common candidate genes among the three sequenced *eop* mutants, genes with EMS changes within the entire mapped regions of 34 Mb were compared (Figure 20D). Two genes were found common among the three mutants, these were *Aa_G106560* and *Aa_G70060*. *Aa_G70060* was however located outside the reduced mapping intervals and close to marker M3 in *eop002* and A5 in *eop091*. This therefore excluded lesions in *Aa_G70060* as causative for the observed *eop* phenotypes.

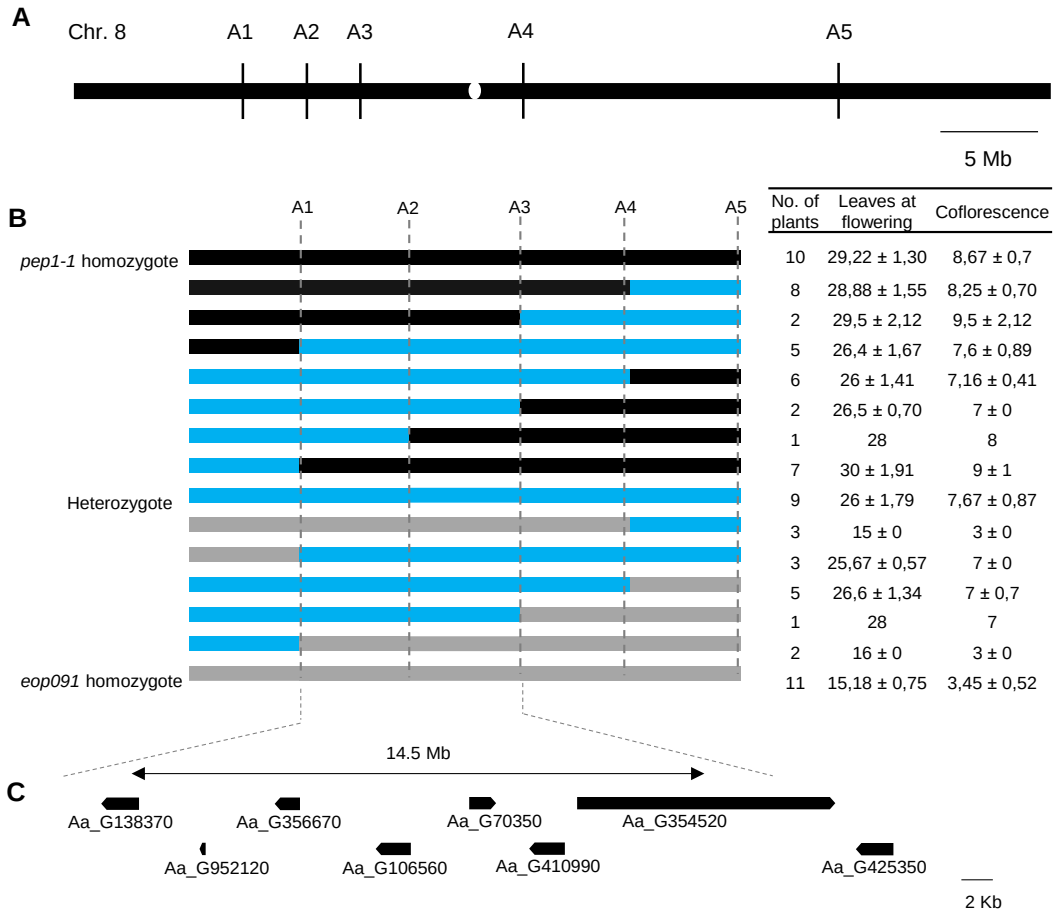


Figure 20. Co-segregation analysis in *eop091* F2 population enables reduction of mapping interval.
(A) Position of SNP markers for genotyping on Chr. 8. The white dot represents the centromere **(B)** Correlation of markers' segregation with flowering and inflorescence phenotypes. Black bars represent chromosomal segment of wild type (*pep1-1*) homozygous, blue bars for heterozygous and grey bars for mutant homozygous. **(C)** Genes with EMS changes within the reduced mapping interval **(D)** Comparison of genes with EMS changes in *eop002*, *eop085* and *eop091* within the mapped region of 34Mb on Chr. 8.

2.3.3 Isolated candidate gene, *AaG106560*, is a member of AAA+ superfamily of ATPases

Comparison of candidate genes within mapping intervals in the sequenced *eop* mutants together with results from genotype-phenotype co-segregation analysis, identified *AaG106560* to be strictly linked to their early flowering and simple inflorescence phenotypes. *AaG106560* is predicted to be a member of AAA+ (ATPases Associated with diverse cellular Activities) superfamily of ringshaped P-loop NTPases. The orthologue of *AaG106560* in *A. thaliana*, *AT5G17760*, is located in a syntenic region on *A. thaliana* chromosome 5 (Figure 21A) and shows 88.7% amino acid identity to *AaG106560* (Table 6). *AaG106560* was found to be one of two tandemly duplicated genes within the syntenic region. The second gene, *AaG106570*, similar to *AaG106560*, is a member of the AAA+ superfamily of ATPases. This tandem duplication event was also observed in *A. thaliana*, however, in this case *AT5G17760* was part of a cluster of four tandemly duplicated genes, all of which are members of the AAA+ superfamily of ATPases. The three additional members include, *AT5G17730*, *AT5G17740* and *AT5G17750*. *AaG106570* showed the closest identity to *AT5G17730* from the cluster (Table 6).

Table 6. Pairwise identity comparison using amino acid sequences

	AaG106560	AaG106570	AT5G17730	AT5G17740	AT5G17750	AT5G17760
AaG106560	100%					
AaG106570	59.31%	100%				
AT5G17730	60.85%	76.13%	100%			
AT5G17740	58.06%	65.45%	64.68%	100%		
AT5G17750	61.22%	61.47%	64.79%	74.23%	100%	
AT5G17760	88.7%	59.54%	60%	58.81%	61.98%	100%

A key defining feature of the AAA+ superfamily of ATPases is a structurally conserved central ATPase domain referred to as AAA+ module, which is responsible for ATP binding and hydrolysis [204-206]. The domain consists of distinctive signature motifs, comprising of the Walker-A motif, with the consensus GxxGxGK[S/T], where x is any amino acid, and the Walker–B motif, with consensus of hhhhDE, where h represents hydrophobic amino acids. Other characteristics features of the AAA+ module include sensor 1, sensor 2 and the arginine finger.

An input of AaG106560 protein sequence, consisting of 496 amino acids, into the NCBI conserved domain database [207], recognized the Walker-A and -B motifs, the arginine finger and additionally, a domain associated with the AAA family (Figure 21B). Sanger sequencing of *AaG106560* from the four isolated *eop* mutants and genotyping with CAP/dCAP markers, confirmed independent EMS mutations which caused amino acid substitutions in all cases (Figure 21C). The position of the mutations were however located outside any well conserved motif.

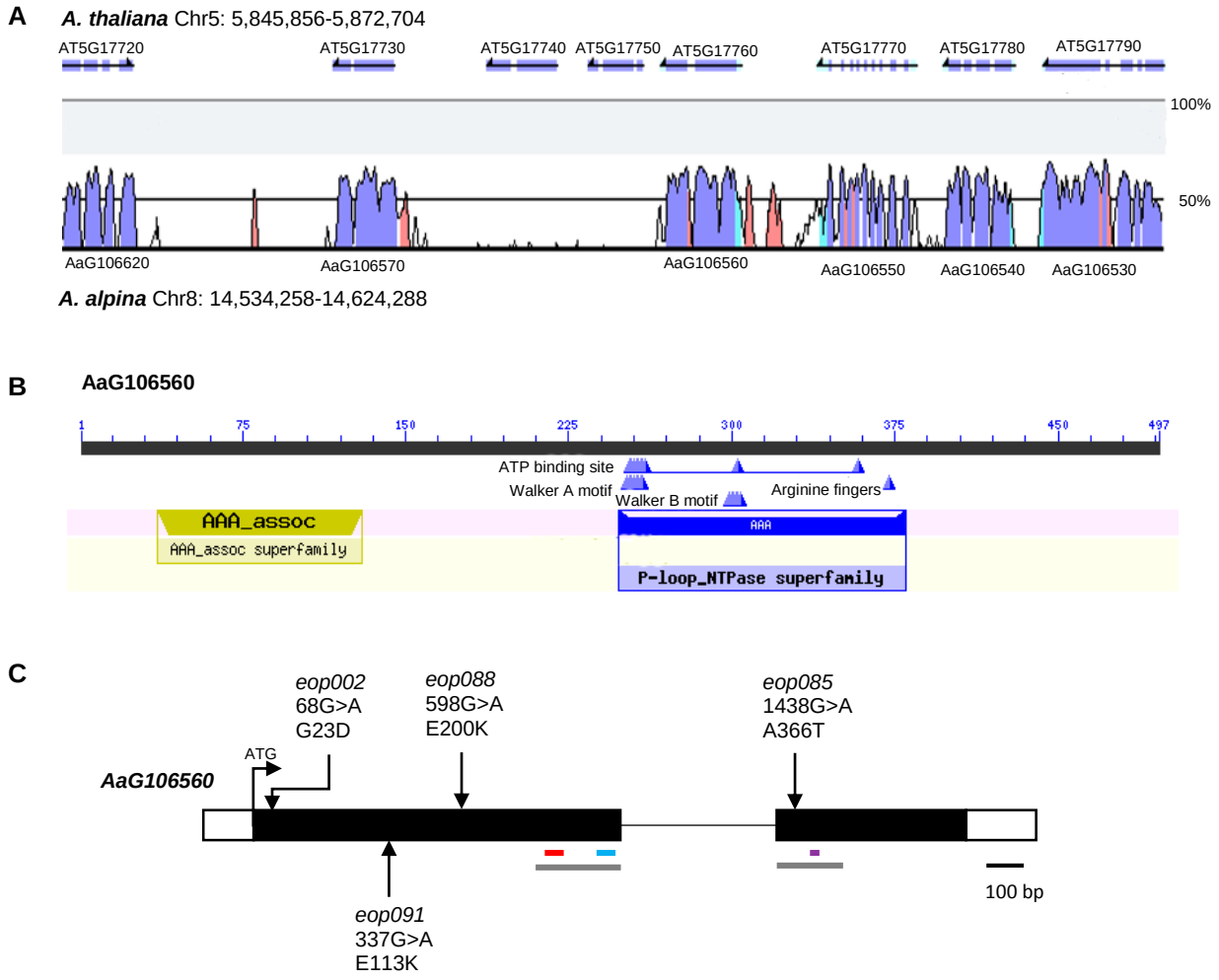


Figure 21. *AaG106560* encodes an AAA+ ATPase.
(A) Synteny mapping between *A. alpina* region surrounding *AaG106560* and *A. thaliana* region containing the orthologue, *AT5G17760*. The region in *A. alpina* was blasted against the entire *A. thaliana* genome using the web-based platform, VISTA (http://www-gsd.lbl.gov/vista/). Percentage lines depicts the percentage identity of genes. **(B)** AAA+ ATPase domain and motifs identified in AaG106560 using NCBI conserved domain database. **(C)** Structure of *AaG106560*, arrows indicate the positions of the *eop* mutations. Exons are represented by black boxes, untranslated regions by white boxes and introns by black line. The AAA+ ATPase domain is indicated by a grey line, the Walker A motif by red line, Walker B motif by blue line and the arginine finger by a purple line.

2.3.4 *AaG106560* Subcellular localization and expression in different tissues

AT5G17760 in *A. thaliana* is predicted to be localized in the plasma membrane. To determine the subcellular localization of *AaG106560*, the *AaG106560* coding sequence was fused with green fluorescent protein (GFP) under the transcriptional control of Cauliflower Mosaic Virus (CaMV) 35S promoter. The final construct of 35S::GFP-*AaG106560* was expressed transiently in epidermal cells of *N. benthamiana*. The mCherry-tagged plasma membrane marker, CD3-1007 [8], was employed. The GFP-*AaG106560* fusion was confined to the plasma membrane, just as predicted for the *A. thaliana* orthologue, and this co-localized with the red fluorescent protein (RFP) fluorescence from CD3-1007 (Figure 22A).

Gene expression patterns can provide essential clues when investigating the function of a gene. The expression analysis of *AaG106560* in different tissues was performed by reverse transcription quantitative PCR (RT-qPCR). Samples for RNA isolation, from roots, stem, leaves, vegetative (Veg.) apices (in V3 zone), reproductive (Rep.) apices, open flowers and siliques, were collected from 12 weeks vernalized Pajares plants after their return to warm conditions. Expression of *AaG106560* was observed in all organs examined, but it was most highly expressed in leaves followed by vegetative apices (Figure 22B). This indicates *AaG106560* might be highly expressed during vegetative stages of development.

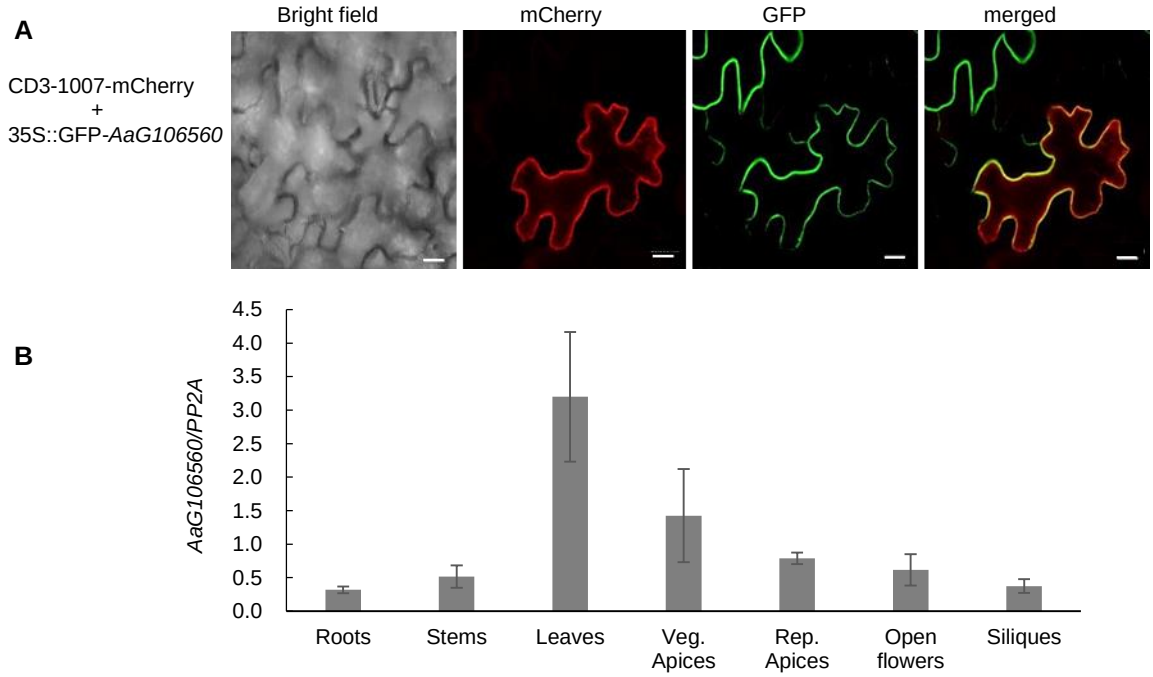


Figure 22. *AaG106560* localizes to plasma membrane and highly expressed in leaves.
(A) Expression of 35S::GFP-*AaG106560* and mCherry-tagged plasma membrane marker, CD3-1007, in tobacco leaf cells **(B)** Expression pattern of *AaG106560* in different Pajares tissues using qRT-PCR. Scale bar = 20 µm and error bars represent the s.d.m.

2.3.5 Complementation of the *eop* mutants phenotype

Although *AaG106560* seems to co-segregate and show tight correlation with the early flowering phenotype of the *eop* mutants, the possibility remains that the observed phenotype could result from a nearby linked mutation. Therefore a 6.8 kb genomic DNA fragment, containing the *AaG106560* gene, a region of 3,654 bp upstream of the 5' UTR considered to contain the native promoter, and a 923 bp region downstream of the 3' UTR, was transformed into the *eop* mutants for complementation analyses. Acquisition of transgenic plants for *eop002* was so far unsuccessful, on the other hand 4 independent T1 lines were obtained for *eop085*, 1 for *eop088* and 1 for *eop091*. From preliminary results, the main shoot apex of the *eop091* T1 plant initially displayed the inward curving of leaves, which normally precedes flower emergence and indicates floral transition (Figure 23), however this was followed by the production of several leaves. The expression of *pAaG106560*::*AaG106560* in *eop085* and *eop088*, on the other hand, rescued the mutant flowering phenotypes. These T1 plants flowered between 10 to 12 weeks after sowing, with 23 to 30 leaves (Table 7), this is similar to what is observed for *pep1-1* (Section 2.2). These data support the notion that, the observed early flowering phenotypes of the mutants are due to disruptions in *AaG106560*, henceforth referred to as *EOP*.

Figure 23. Expression of *pAaG106560::AaG106560* caused late flowering in *eop* T1s.
Image of *eop085*, *eop088* and *eop091* T1s after 12 weeks in LD. Insert shows the reversion of *eop091* T1's shoots apex. Scale bar = 10cm

Table 7. Flowering time of *pAaG106560::AaG106560 eop* T1 lines

	DTF	Leaves at flowering
eop085-1	70	23
eop085-2	78	27
eop085-3	81	30
eop085-4	75	27
eop088	78	26

a. DTF: Days to first opened flower

2.3.6 Overexpression of *EOP* in *pep1-1*

The phenotype of *eop* mutants and that of *pAaG106560*::*AaG106560* transgenic lines suggested a role for *EOP* in maintaining vegetative development. To explore this idea, *EOP* genomic DNA was constitutively expressed under the CaMV 35S promoter in *pep1-1*. Interestingly the transgenic lines obtained, displayed varying flowering time phenotypes (Figure 24A and B). Out of 15 T1 lines, 9 flowered unexpectedly early with number of leaves ranging from 12 to 22. Two of the lines, flowered with number of leaves similar to that of *pep1-1*, that is, 25 and 27 leaves. Another plant, line B3, flowered at 27 leaves by a lower axillary branch, the main shoot, on the other hand, remained vegetative and produced several leaves, as observed previously in *eop091* transgenic line (Figure 24B). The remaining 3 lines flowered extremely late with 36 to 45 leaves, and only by lower axillary branches. Their main shoots apices remained vegetative and continuously produced numerous cauline-like leaves and occasionally vegetative axillary branches. Therefore, in most instances, overexpression of *EOP* in *pep1-1* promoted flowering and in minority of the cases it delayed flowering and promoted vegetative development.

The late lines produced very few seeds, thus to further explore the discrepancies in flowering, T2 progeny of line B3 and an early flowering line, line A1, were examined. The number of independent insertion points were assessed via basta resistance segregation in a subset of the T2 populations. T2 progeny from line B3 showed a segregation of 25 resistant and 5 sensitive individuals, while line A1 progeny showed a ratio of 28 resistant to 9 sensitive individuals. These fitted a 3:1 segregation (line A1: $X_1^2 = 0.009$, $P = 0.92$, line B3: $X_1^2 = 1.111$, $P = 0.29$) supporting single insertions. A second set of T2 lines were screened by PCR, using a combination of transgene-specific primers, to select transformed individuals for further analysis. All 13 T2 plants obtained from line A1 were uniformly early flowering compared to *pep1-1* (Figure 24C). Interestingly, in the 18 T2 progeny from line B3, 6 plants flowered as early as individuals from line A1 (Figure 24D and E). 9 plants flowered by their lower axillary branches with total leaf numbers of 25-32, the main shoot of 5 of these plants eventually flowered while the main shoot of the rest stayed vegetative. The remaining 3 T2 progeny from line B3 were much later flowering than *pep1-1* with 38-42 leaves, and flowered only by their lower side branches as their main shoot remained vegetative (Figure 24E and F).

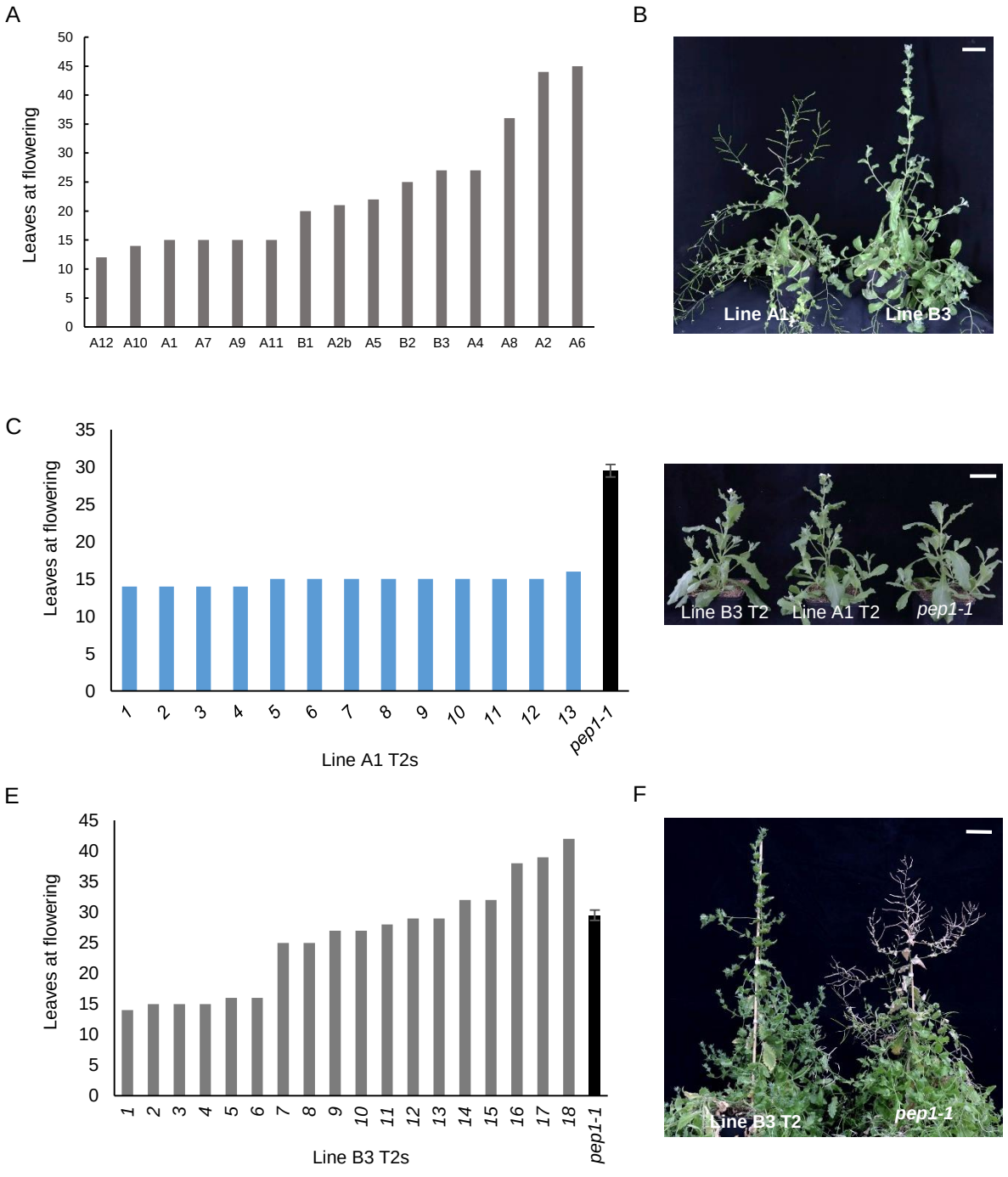


Figure 24. Overexpression of *AaG106560* in *pep1-1* produces varying flowering time phenotypes.
(A) Flowering time of 35S::*AaG106560* *pep1-1* T1 transformants. **(B)** Early flowering, Line A1, and vegetative, Line B3, T1 plants **(C)** Flowering time of T2 progeny of Line A1 compared to *pep1-1* (black bar) **(D)** Early flowering T2 lines from both Line B3 and A1 **(E)** Flowering time of T2 progeny of Line B3 compared to *pep1-1* (black bar) **(F)** A T2 line from Line B3 that exhibited late flowering and continuous production of leaves by the SAM compared with *pep1-1*. Flowering times were scored as leaves at first opened flower. Scale bar = 5 cm

It is unclear why line B3 produced varying flowering individuals. It was speculated that the phenotypes observed might be as a result of differences in quantitative expression levels of the transgene. To determine whether expression levels correlated with flowering time, RNA was extracted from leaves of individual T2 adult plants (in Figure 24, section 2.3.6), collected during flowering of the early flowering lines. Expression levels were examined in 5 T2 individuals from line A1, 3 early flowering lines from B3 and 2 lines that flowered first by lower branches and later by the main shoot, and finally the 3 late flowering T2 lines obtained from line B3. *EOP* expression was predominantly detected in the 3 late flowering lines with vegetative main shoots (Figure 25). The rest of the lines tested displayed low levels of or no overexpression compared to levels observed in the control, *pep1-1*.

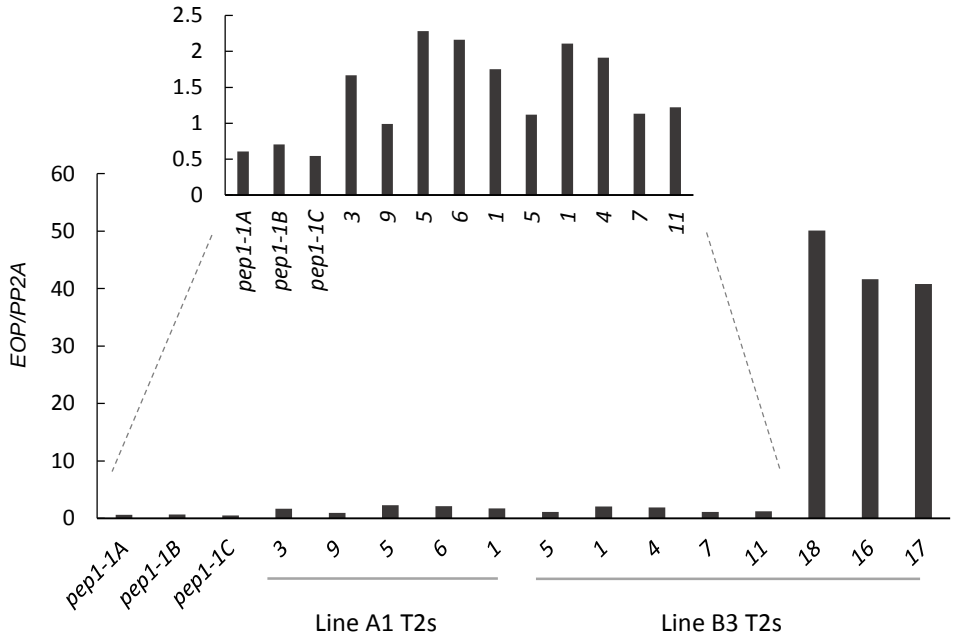


Figure 25. *EOP* is highly expressed in late flowering *pep1-1* T2 transformants.
EOP expression levels in early and late flowering T2 plants from lines A1 and B3. Lines L18, L19 and L24 showed the highest levels of expression and these were latest to flower.

These results together demonstrate that *EOP* might affect flowering time in a quantitative manner. Low levels of overexpression appear to either cause plants to flower much earlier than controls or flower similar as controls but initially by lower side branches whereas a larger increase causes a delay in flowering and vegetativeness of the main shoot.

2.3.7 Physiological analysis of *eop* and *Aatfl1* in a functional *PEP1* background

A. alpina Pajares, formerly described in Section 2.1, requires prolonged exposure to cold to ensure flowering and development of the inflorescence. *AaTFL1* has been shown to contribute to this prolonged requirement as 8 weeks old transgenic Pajares lines, with a 35S:*AaTFL1dsRNAi* construct, flowered with shortened periods of vernalization due to reductions in *AaTFL1* expression levels [140]. This reduction in *AaTFL1* expression also enabled flowering in young vernalized Pajares plants, which usually do not flower in respond to vernalization. A study of *eop* mutation in a functional *PEP1* background is essential to understand its role in the perennial growth habit of *A. alpina*. Thus, to aid determine whether *EOP* also contributed to either the prevention of flowering in young plants exposed to vernalization or the duration of vernalization required for reproductive development in older plants, introgression lines (ILs) were produced with a functional *PEP1*. *PEP1*, *AaTFL1* and *EOP* are all located on *A. alpina* chromosome 8. Backcrossed F1 individuals (BC1F1) for *eop002*, *eop091* and *eop101* (crossed to Pajares) were generated by Prof. Dr. Maria Albani. Through marker assisted selection (Figure 26A), the mapped region of 34 Mb was introgressed into a *PEP1* background in BC1F3 populations of *eop002* and *eop091* (Figure 26B, right lane). Similarly the locus containing *Aatfl1* mutation in *eop101* was introgressed into a *PEP1* background in BC1F3 populations (Figure 26B, right lane). Interestingly, the marker T2, used for genotyping of *eop101* backcrossed individuals, was also identified as a SNP within *AaG106560*, meaning that the *tfl1* mutant also contains a mutation in *EOP*. The ILs obtained for *eop101* were however within the wild type *EOP* background.

Introgression-lines were grown for 8 weeks in long days (LD) and vernalized for 6, 8, 12 and 15 weeks to test their response to shortened vernalization. Just as previous vernalization duration experiment, this was designed to ensure the return of all treatments to LD conditions at the same time. For testing the response of young ILs to vernalization, plants were grown for 3 weeks in LD and vernalized for 15 weeks. As expected, 3 weeks old vernalized ILs from *eop101* flowered after return to LD conditions, on the contrary, lines from *eop002* and *eop091* did not flower (Figure 27), and together with Pajares, these plants grew vegetatively till experiment was terminated. Thus, in contrast to Aa*TFL1*, *EOP* does not contribute to the prevention of flowering in young vernalized *A. alpina* Pajares plants.

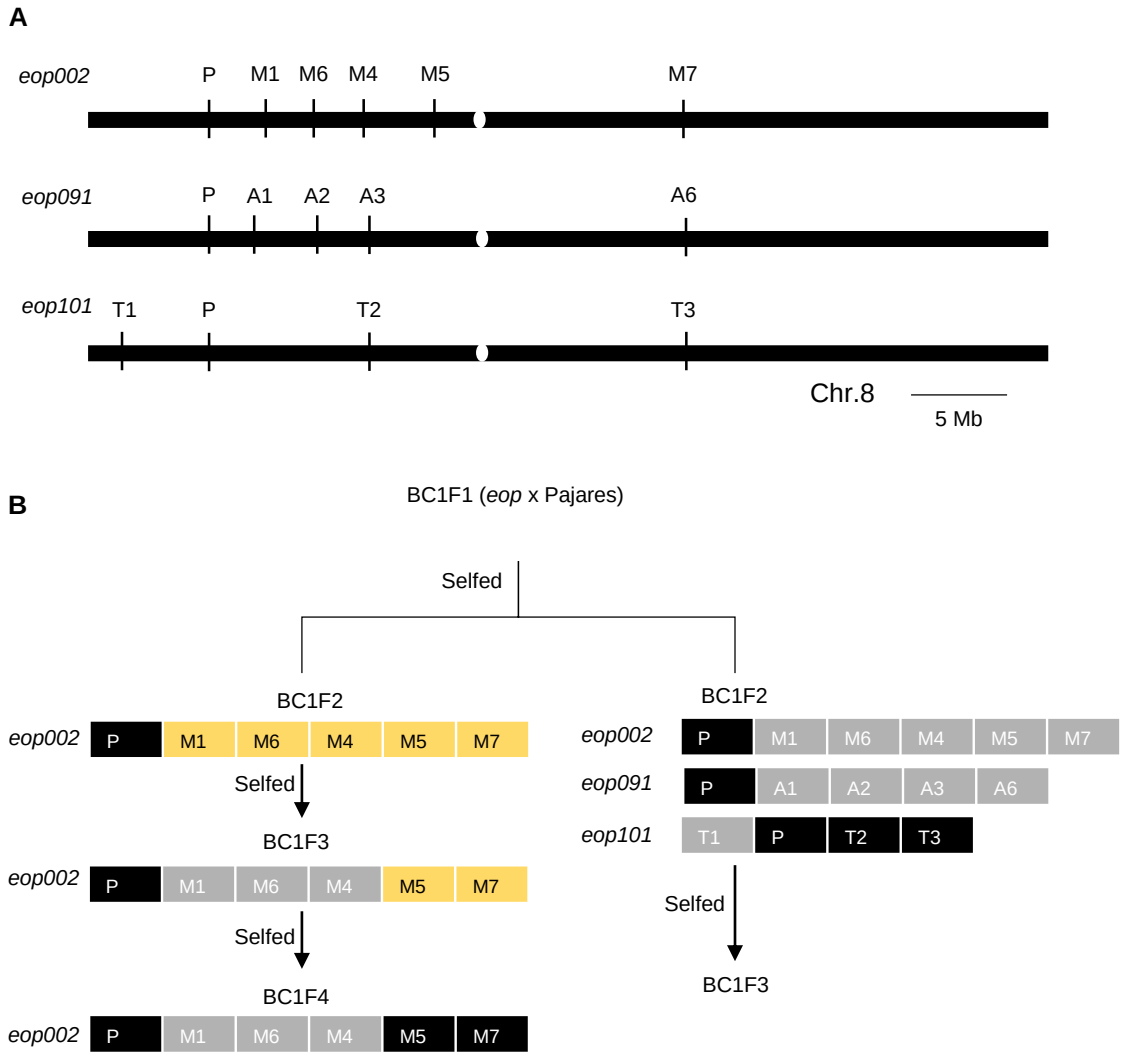


Figure 26. Schematic illustration of introgression lines development.
(A) Position of SNP markers on chromosome 8 for selection of introgressions lines. The white dot represents the centromere. **(B)** Genotypes of introgression lines. Black bars represent chromosomal segment of wild type (Pajares) homozygous, yellow bars for heterozygous and grey bars for mutant homozygous. Marker P, represents SNP marker in *PEP1*, markers M4 in *eop002*, A3 in *eop091* and T2 in *eop101* represents SNPs in *EOP*, while T1 represents SNP in *AaTFL1*.

Figure 27. Young *eop002* and *eop091* introgression lines (ILs) do not flower in response to vernalization.
Image of 3 weeks old flowering *eop101*BC1F3 IL and non-flowering *eop002* and *eop091* ILs, and Pajares plants after 15 weeks vernalization. Picture was taken 5 weeks after return to LD. Scale bar = 5cm

2.3.8 Effects of vernalization duration on flowering and inflorescence development in *eop002*, *eop091* and *eop101* introgression lines

Exposure of eight weeks old Pajares plants to six weeks vernalization failed to induce flowering, on the other hand, this short period of vernalization was sufficient to induce flowering in 90% of ILs from *eop002*, 60% of *eop091* ILs and 100% of *eop101* ILs (Figure 28A). Pajares was able to produce only 30% of flowering plants after eight weeks vernalization. In accordance with previously presented data on vernalization duration effects in Pajares (Section 2.1), prolonged vernalization reduced the number of days required for flower emergence and caused early inflorescence outgrowth in the ILs. The differences in days to flower emergence were however very minute in *eop101* introgression lines and also between 12 and 15 weeks vernalized *eop002* and *eop091* ILs (Figure 28B). Nevertheless, regardless of the vernalization duration, flower emergence was observed earlier in the ILs compared to Pajares plants (Figure 28B).

Inflorescence outgrowth or bolting, measured in terms of inflorescence length, also occurred earlier in ILs after vernalization than in Pajares (Figure 28C - E). Six weeks of vernalization was sufficient for ILs to exhibit inflorescence outgrowth after three weeks in LD (Figure 28C and D). The measure of final inflorescence length at 14 weeks in LD revealed, ILs achieved similar inflorescence height irrespective of the vernalization duration, whereas Pajares plants displayed an increase in inflorescence length with respect to the vernalization duration (Figure 28F). An additional observation was the fact that, Pajares plants vernalized for 12 and 15 weeks, after 14 weeks in LD, possessed higher inflorescence lengths compared to ILs (Figure 28F). Lines from *eop101* exhibited the shortest inflorescence length due to the determinate nature of their inflorescence (Figure 28E). Taken together, these results suggest that, *EOP* contributes to the prevention of flowering in response to short vernalization durations. Also, disruptions in *EOP* can lead to a decrease in the number of days required for flower emergence and inflorescence outgrowth after plants are returned to warm conditions.

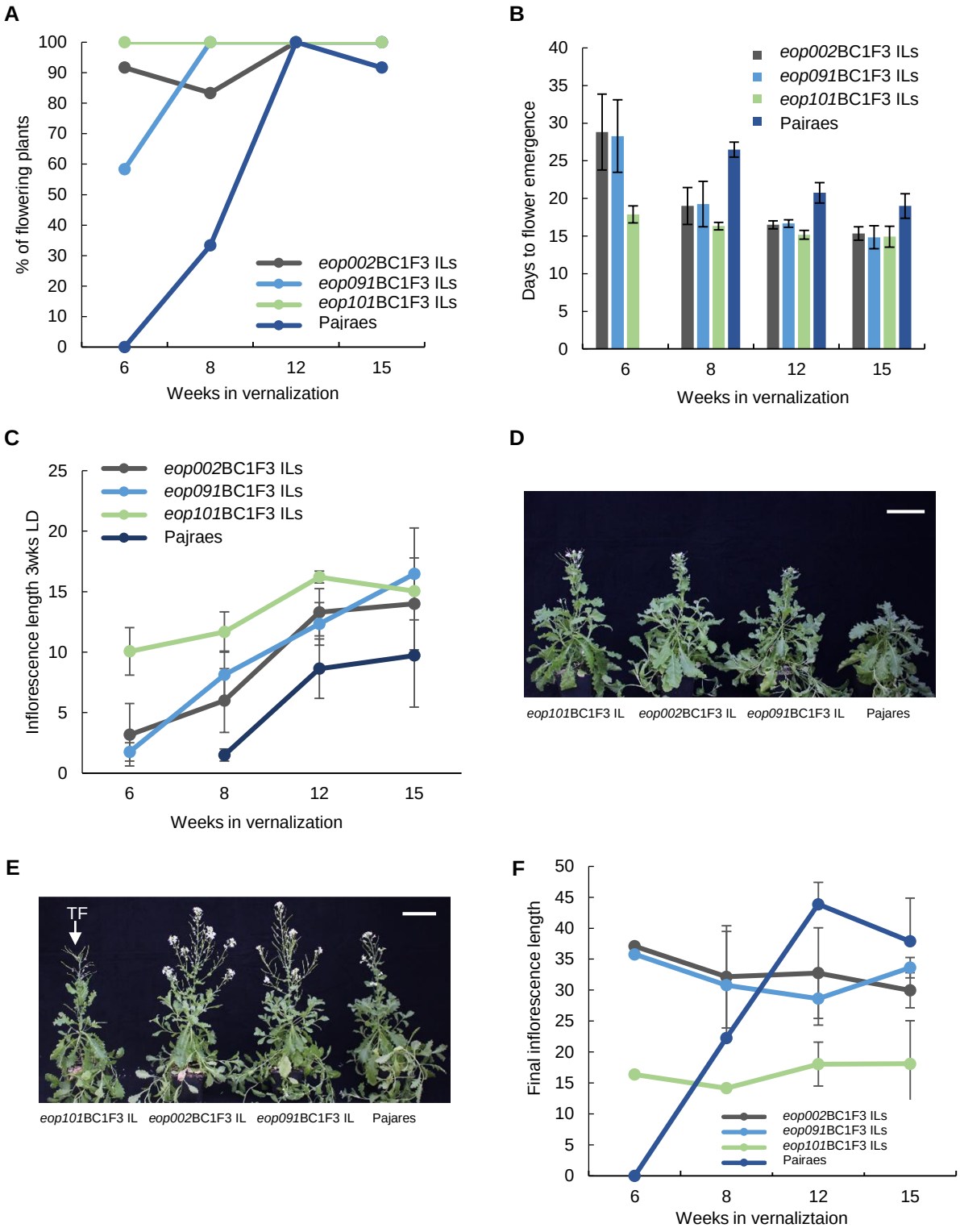


Figure 28. *eop002, eop091* and *eop101* introgression lines (ILs) exhibit early flower emergence and inflorescence outgrowth after different durations of vernalization.
(A) Percentage of ILs that flowered compared to Pajares after different vernalization durations **(B)** Days to flower emergence in ILs and Pajares **(C)** Inflorescence outgrowth (bolting) at 3 weeks in Long day (LD) **(D)** *eop002*, *eop091* and *eop101* ILs, and Pajares vernalized for 6 weeks followed by 3 weeks in LD. *eop101* IL produced determinate inflorescence by formatting a terminal flower (TF). **(E)** ILs and Pajares vernalized for 15 weeks vernalization followed by 3 weeks in LD **(F)** Final length of the inflorescence after 14 weeks in LD. Data are presented as means ± SD, n = 7-12. Scale bar = 10 cm

2.3.9 Study of floral reversion and vegetative phenotypes within the inflorescence of *eop002* and *eop091* introgression lines

Pajares plants that receive insufficient vernalization have been previously demonstrated to display vegetative features and floral reversion in their inflorescence such as branches lacking commitment to flowering in I1 and production of bracts in I2. The fact that ILs require less vernalization to flower suggested their inflorescences might exhibit less of these features. Hence, the number of flowering and vegetative I1 lateral branches where scored in the *eop002* and *eop091* ILs in comparison to Pajares at each vernalization time point. ILs did not show any significant reduction in the total number of I1 lateral branches as compared to Pajares (Figure 29A). However, there were more vegetative branches scored in Pajares compared to the ILs. The large extent to which extended vernalization reduced the number of I1 vegetative branches in Pajares was not equally observed in the ILs. Eight weeks of vernalization seemed sufficient to ensure majority of I1 lateral branches flowered in the ILs, and further vernalization showed minimal or no effect (Figure 29A). A similar trend was observed for the presence of bracts, the number of bracts was reduced in the ILs compared to Pajares, however extended vernalization caused no apparent change in the numbers scored (Figure 29B). Additional interesting observation, was the fact that a proportion of V3 branches, which normally remain vegetative, were committed to flowering in the ILs. This was observed in 12 weeks vernalized *eop002* ILs (Figure 29C) and more pronounced in 15 weeks vernalized plants (Figure 29D). *eop091* ILs exposed to 15 weeks of vernalization also exhibited flowering of some V3 branches (Figure 29D). These results indicate that ILs, in comparison to Pajares, require less vernalization to suppress floral reversion and ensure flowering of I1 branches. Furthermore exposure of ILs to extended vernalization, in certain cases, causes flowering of branches in the V3 zone.

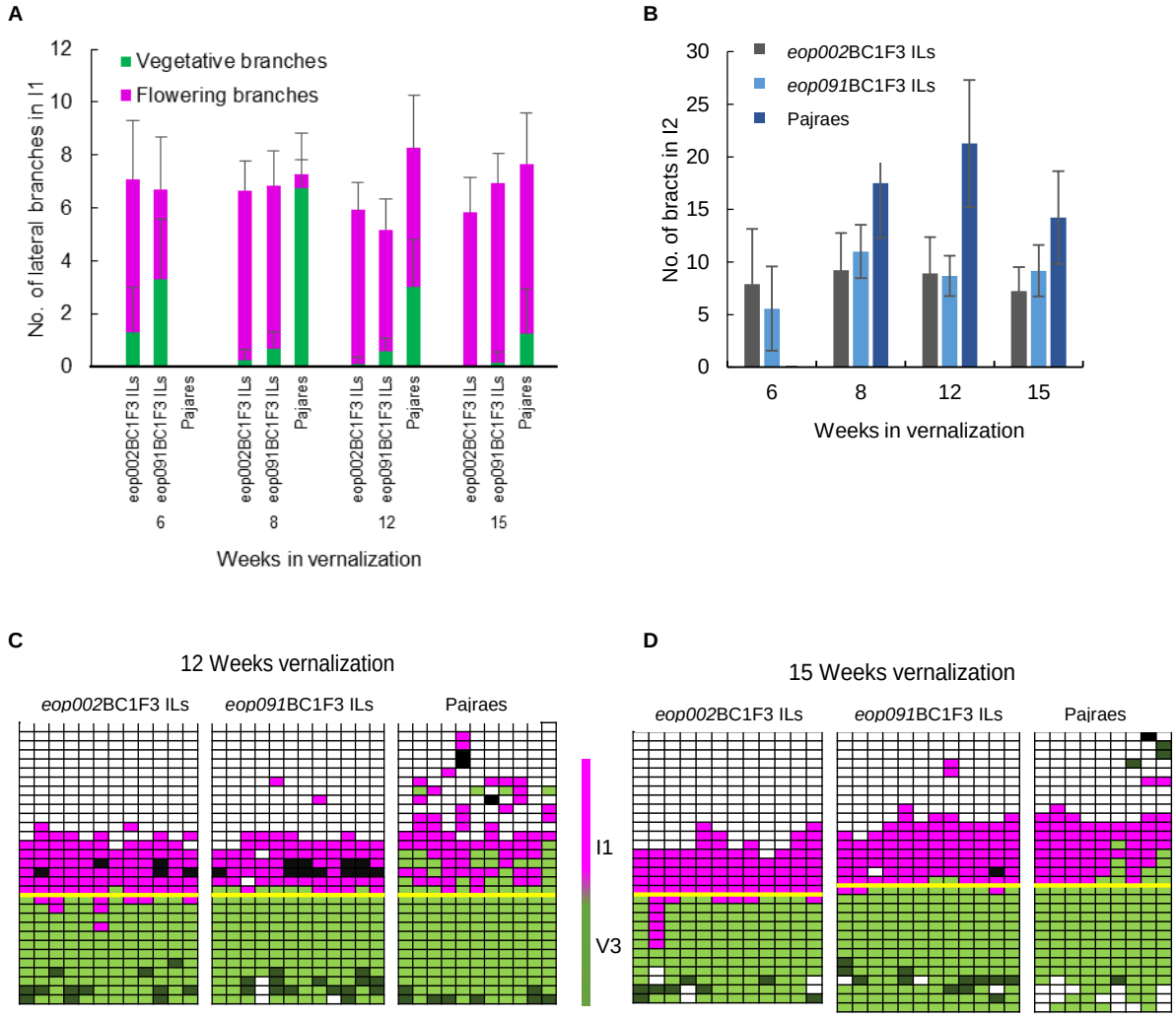


Figure 29. Floral reversion and vegetative branches are reduced in *eop002* and *eop091* ILs inflorescence.
(A) Number of flowering and vegetative branches in the inflorescence of ILs compared to Pajares. **(B)** Number of bracts in I2 of ILs and Pajares. **(C)** and **(D)** Scheme of axillary branches within leaf axils in the V3 zone and I1 zone of *eop002* ILs, *eop091* ILs and Pajares after 12 and 15 weeks of vernalization, respectively. Each column represents an individual plant and each row a leaf axil. Magenta boxes indicate flowering branches, green boxes below the yellow line represent vegetative branches in V3 and above the yellow line indicate vegetative branches in I1. Black and dark green boxes indicate damaged branches and branches less than 0.5 cm, respectively. Data are presented as means ± SD, n = 7-12.

2.3.10 Reduction of introgression region in *eop002* in a functional *PEP1* background

Introgression of the 34Mb mapped region has so far provided information on *eop* mutation effect on vernalization duration requirement and inflorescence development. With further marker assisted selection in *eop002*BC1F4 population, this region was reduced to 12 Mb to narrow in on the *EOP* locus on chromosome 8 (Figure 26B, left lane). To determine whether *EOP* contributed to the obligate vernalization requirement, *eop002*BC1F4 ILs together with Pajares were grown under long day conditions. In addition, ILs were grown for 8 weeks in LD and vernalized for 21 weeks for comparison to Pajares plants after saturation of their vernalization requirement. *eop002*BC1F4 ILs partially overcame the obligate vernalization requirement as 60% of these plants flowered after 128.33 ± 32.57 days in LD, Pajares plants never flowered in these conditions (Figure 30A). Flowering of the ILs was however only by a small portion of lower axillary branches (Figure 30B), the main shoot on the other hand remained vegetative as it continuously produced cauline-like leaves even after an initial displayed of inward curving of leaves closest to the shoot apex.

After extended vernalization of 21 weeks and return of plants to LD conditions, no difference was observed in the number of days to flower emergence between ILs and Pajares (Figure 30C). However, ILs produced more axillary branches compared to Pajares. This included the outgrowth of branches from axillary meristems within leaf axils in V2 zone, where there is usually the arrest of axillary shoot outgrowth in Pajares (Figure 30D). Interestingly, a large proportion of axillary branches formed in the ILs also flowered including branches in the vegetative V3 zone (Figure 30D and E). The ILs exhibited a reduction in their inflorescence height compared to Pajares (Figure 30F). A measure of this reduction revealed, in contrast to the *eop* mutants in which the I1 zone was the prominently reduced part of the inflorescence, the I2 zone in ILs was the greatest contributor to the reduction of inflorescence height. This was also evident in the number of axillary branches in I1 (Figure 30G) and the number of siliques in I2 (Figure 30H) in comparison to those in Pajares. These results indicate that, *EOP* contributes to the prevention of flowering in older *A. alpina* Pajares in response to vernalization requirements and also the prevention of axillary branches outgrowth within some leaf axils. Additionally, in the absence of *EOP*, saturation of vernalization requirement

causes flowering in a greater proportion of axillary branches and a reduction in inflorescence height.

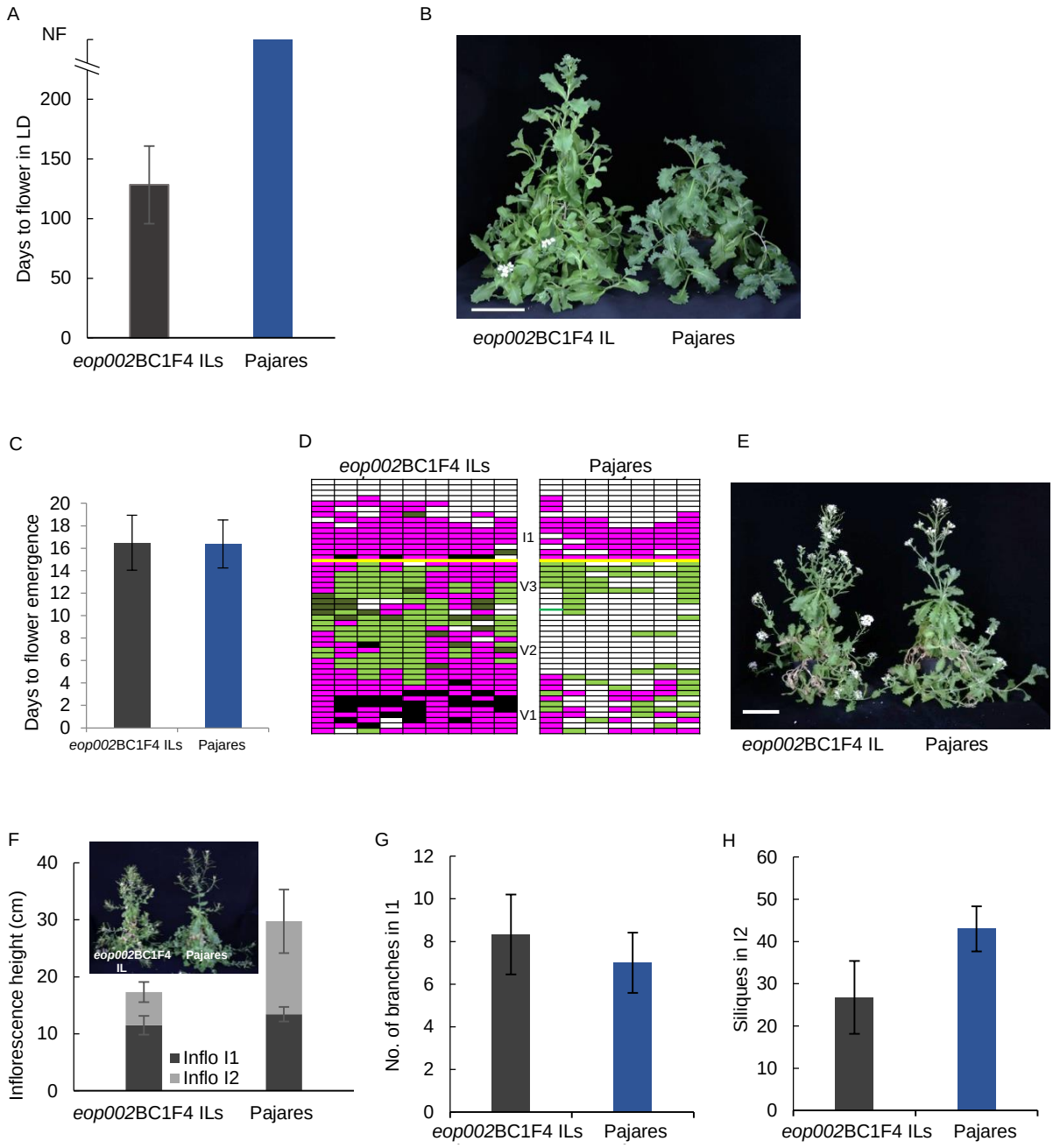


Figure 30. *eop002* ILs with reduced introgression region partially flower in LD and contain increased proportion of flowering branches.

(A) Number of days to flowering in *eop002*BC1F4 ILs and Pajares grown under LD conditions. NF = No flowering **(B)** Flowering *eop002*BC1F4 IL in LD. **(C)** Days to flower emergence in *eop002*BC1F4 ILs and Pajares after 21 weeks of vernalization. **(D)** Scheme of axillary branches in *eop002*BC1F4 ILs and Pajares. Magenta boxes indicate flowering branches, green boxes represent vegetative branches. Black and dark green boxes indicate damaged branches and branches less than 0.5 cm, respectively. Yellow line indicates the starting point of the inflorescence. **(E)** Flowering *eop002*BC1F4 IL and Pajares after 21 weeks vernalization. **(F)** Height of inflorescence, I1 and I2, in *eop002*BC1F4 ILs and Pajares. Insert is a picture taken 8 weeks in LD after vernalization **(G)** Number of branches in I1 of *eop002*BC1F4 ILs and Pajares. **(H)** Number of siliques in I2. Data are presented as means ± SD, n = 6-12. Scale bar = 10 cm

2.4 Genome-wide gene expression analysis in *eop002*

2.4.1 Summary of RNAseq data

To gain insights into possible regulatory mechanisms that might lead to early floral transition in the absence of a functional *EOP,* RNAseq was carried out on the *eop002* mutant. As the *eops* flower much earlier than *pep1-1*, samples for sequencing were collected from leaves and apices of 3 weeks old seedlings since at this point no apparent morphological differences could be observed. Analysis of DEGs in *eop002* compared to *pep1-1* revealed, *eop002* contained 217 DEGs down-regulated and 196 DEGs up-regulated in leaves while in apices, there was the down-regulation of 558 genes and the up-regulation 234 genes (Figure 31). Indicating that, in *eop002*, more genes were differentially regulated in the apices than in the leaves and most of the genes in apices were down-regulated as compared to *pep1-1*.

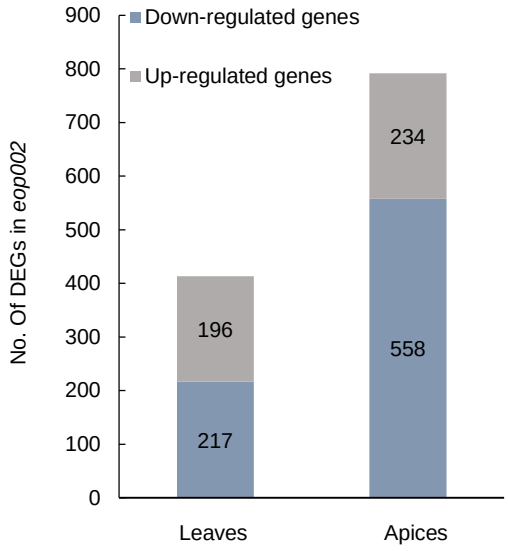


Figure 31. Differentially expressed genes (DEGs) in *eop002* in apices and leaves.
Number of DEGs in leaves and apices of *eop002* as compared their expressions in *pe1-1*.

2.4.1 Gene Ontology (GO) enrichment analysis for functional classification of DEGs from *eop002*

Similar to previous analysis in *eop101*, the VirtualPlant software platform was employed for the identification of overrepresented GO terms in *eop002* DEGs using the *A. thaliana* orthologues of the *A. alpina* genes. Table 8 below gives a summary of the number of *A. thaliana* orthologous genes used for the prediction of enriched GO terms.

Table 8. Number of *A. thaliana* orthologues used for GO-analysis

		Leaves		Apices	
		Strict-list	Relaxed-list	Strict-list	Relaxed-list
eop002	UP	120	142	167	197
	Down	120	164	324	412

Enriched GO terms call out after comparison of DEGs to the whole genome of *A. thaliana* included the overrepresentation of genes involved in developmental, carbohydrate, nitrogen, photosynthesis, biological regulation and rhythmic processes (Table.9). Genes involved in response to stimulus, more specifically, response to stress, water deprivation, hormone, light stimulus and temperature were also enriched in the sets of DEGs. Majority of the enriched DEGs in *eop002* were located in the apices with most being down regulated genes.

The categories of response to stress and transport contained the highest numbers of DEGs, mainly down-regulated (Table.9). This was followed by overrepresentation of genes involved in developmental and nitrogen processes, genes here were all up-regulated in *eop002* apices. The hormones, abscisic acid (ABA) and jasmonic acid (JA), also possessed high numbers of down-regulated enriched genes. Response to water deprivation and to temperature stimulus also consisted of a decent number of down-regulated DEGs in *eop002* apices.

Table.9 Biological functional catergories of *eop002* DEGs			
Category		*eop002* DEGs	
		Apices	Leaves
Developmental process	UP	29	0
	Down	0	0
Sulfur process	UP	7	6
	Down	4	0
Carbohydrate process	UP	3	2
	Down	0	0
Nitrogen process	UP	28	0
	Down	0	0
Abscisic acid (ABA)	UP	0	0
	Down	22	0
Jasmonic acid (JA)	UP	0	0
	Down	14	6
Ethylene (ET)	UP	0	0
	Down	3	0
Auxin	UP	0	0
	Down	4	0
Response to stress	UP	0	0
	Down	77	7
Transport	UP	0	0
	Down	45	0
Response to water deprivation	UP	0	0
	Down	19	0
Response to temperature stimulus	UP	0	0
	Down	17	0
Response to light stimulus	UP	0	11
	Down	6	3
Rhythmic process	UP	0	3
	Down	0	0
Photosynthesis	UP	0	4
	Down	0	0
Biological process regulation	UP	3	0
	Down	12	3

a. DEGs, Differentially Expressed Genes
b. Up-regulated genes are shaded in red and down-regulated in blue

2.4.2 Genes involved in developmental process in *eop002*

Developmental process in *eop002* was narrowed down to more specific terms, embryo development, developmental process involved in reproduction and stomatal morphogenesis. There were 10 genes associated with embryo development, 14 with developmental process involved in reproduction and just 3 genes with stomatal morphogenesis. A close look at genes involved in these terms showed the presence of some common shared genes. 9 genes were found common between embryo development and developmental process involved in reproduction, this meant 90% of embryo development genes and 64% of genes involved in reproduction (Table 10). 6 of the 9 genes were *EMBRYO-DEFECTIVE* (*EMB)* genes which are required for normal development of the embryo [208]. Among the remaining genes involved in reproduction were 2 interesting genes, *GI* and *ERECTA-LIKE 1* (*ERL1*). *GI* is a flowering time gene, in the circadian clock-controlled flowering pathway. *GI* together with *CO* and *FT* promotes flowering under long days [64]. *ERL1* is a gene associated with both developmental process involved in reproduction and stomatal morphogenesis. It is a member of the Arabidopsis ERECTA (ER)-family receptor kinases that regulates stomatal morphogenesis, that is, epidermal stomatal patterning, and also inflorescence architecture and organ shape [209]. Other genes associated with stomatal morphogenesis included *STOMATAL DENSITY AND DISTRIBUTION 1* (*SDD1*) and *TOO MANY MOUTHS* (*TMM*), *EPIDERMAL PATTERNING FACTOR 2* (*EPF2*). These analysis indicate that, within the *eop002* mutant, there is the up regulation of genes involved in embryogenesis, stomatal morphogenesis and a gene involved in photoperiodic flowering and regulation of circadian rhythm compared to *pep1-1* at an early developmental stage.

Table 10. Genes in *eop002* apices associated with developmental process

Arabidopsis gene	Primary Gene Symbol	Gene Model Description	FPKM_ *eop002*	FPKM_ *pep1-1*	stomatal complex morpho-genesis	embryo develop-ment	developmental process involved in reproduction
AT1G04110	*STOMATAL DENSITY AND DISTRIBUTION 1 (SDD1)*	Initially identified as a mutation affecting stomatal development and distribution.	2,02	0,98	X		
AT1G80080	*TOO MANY MOUTHS (TMM)*	Recessive mutation leads to disruption of asymmetric cell division during stomata development.	21,84	17,3	X		
AT5G62230	*ERECTA-LIKE 1 (ERL1)*	It is important for maintaining stomatal stem cell activity and preventing terminal differentiation of the meristemoid into the guard mother cell	8,68	5,73	X		X
AT3G14900	*EMBRYO DEFECTIVE 3120 (EMB3120)*	hypothetical protein;(source:Araport11)	48,83	38,49		X	
AT1G19850	*MONOPTEROS (MP)*	Encodes a transcription factor (IAA24) mediating embryo axis formation and vascular development.	60,16	48,06		X	X
AT1G22770	*GIGANTEA (GI)*	Involved in promotion of flowering under long days in a circadian clock-controlled flowering pathway.	8,8	6,31			X
AT1G47530	*DTX33*	MATE efflux family protein	56,52	44,5			X
AT1G70070	*EMBRYO DEFECTIVE 25 (EMB25)*	Allelic to ISE2(increased size exclusion limit of plasmodesmata 2).	26,21	21,14		X	X
AT2G04030	*CR88*	Involved in red-light mediated deetiolation response	181,7	139,27		X	X
AT3G10110	*MATERNAL EFFECT EMBRYO ARREST 67 (MEE67)*	Mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23 family protein.	46,46	37,02		X	X
AT3G20440	*EMBRYO DEFECTIVE 2729 (EMB2729)*	Involved in organogenesis and somatic embryogenesis by regulating carbohydrate metabolism.	19,33	14,32		X	X
AT3G49240	*EMBRYO DEFECTIVE 1796 (emb1796)*	Encodes NUWA, an imprinted gene that controls mitochondrial function in early seed development.	47,02	35,66		X	X
AT3G49500	*RNA-DEPENDENT RNA POLYMERASE 6 (RDR6)*	Involved in trans-acting siRNA and other siRNA biogenesis.	9,93	7,46			X
AT4G05410	*YAOZHE (YAO)*	Encodes a nucleolar protein with seven WD40-repeats that plays a role in embryo sac development	39,19	31,61		X	X
AT4G29060	*EMBRYO DEFECTIVE 2726 (emb2726)*	Elongation factor Ts family protein.	118,04	93,33		X	X
AT5G49030	*OVULE ABORTION 2 (OVA2)*	tRNA synthetase class I (I, L, M and V) family protein.	47,63	38,3			X
AT5G63420	*EMBRYO DEFECTIVE 2746 (emb2746)*	Plays a vital role in embryo morphogenesis and apical-basal pattern formation by regulating chloroplast development.	35,38	27,94		X	X

a. FPKM: Fragments Per Kilobase Million
b. X : Gene is associated with the GO Term

2.4.3 *eop002* DEGs involved in stress response and transport

Stress and transport were the categories with the highest number of enriched genes. Genes involved in transport were mainly group under specific terms lipid transport, oligopeptide transport and hydrogen peroxide transmembrane transport. Participating genes in lipid transport included *END1-LIKE 1* (*END1*) and several members of the LIPID TRANSFER PROTEIN family, including *LP1, LP4* and *LP6*. Oligopeptide transport consisted of 3 NRT1/ PTR FAMILY genes, *NPF2.10, NPF2.13* and *NPF5.13*, and also *OLIGOPEPTIDE TRANSPORTER* (*OPT3*). *NPF2.10* participates in the distribution of glucosinolates within leaves while *NPF2.13* is involved in nitrate remobilization from older leaves to other tissues [210, 211]. Only two genes were grouped under hydrogen peroxide transmembrane transport, these were *GAMMA TONOPLAST INTRINSIC PROTEIN* (*GAMMA-TIP*) and *PLASMA MEMBRANE INTRINSIC PROTEIN 2A* (*PIP2A*). *GAMMA-TIP* and *PIP2A* are required for the transport of water, *GAMMA-TIP* for example facilitates the transport from the vacuolar compartment to the cytoplasm [212, 213]. The down-regulation of these transporters implies an effect of *eop002* mutation on lipid, glucosinolates, nitrate, and water transport which might be part of the reason for enrichment of genes involved in these processes in the GO analysis.

Genes involved in stress response were mainly down regulated in the apices. The stress specific terms in apices, included response to salt stress, oxidative stress, wounding and defense. 28 genes were associated with defense response, 20 with salt stress, 17 with oxidative stress and 11 with response to wounding. These categories shared some genes in common with response to the hormones, ABA, JA ET and auxin in *eop002* apices, therefore a comparison was made between these categories. As suspected there was an overlap of 21 down regulated genes, which meant 50% of genes associated with hormone response and 34% of genes associated with stress response. These genes were further grouped to identify the specific hormone and stress responses they were involved in (Table 11).

Table 11. Number of *eop002*-apices down-regulated genes common in hormone and stress responses

	JA	ABA	ET	Auxin
Defense response	4	0	2	2
Response to wounding	6	2	0	1
Salt stress	4	11	0	0
Oxidative stress	3	3	0	0

Defense response contained genes responsive to JA, ET and auxin while response to wounding was made up of JA, ABA and auxin inducible genes. In both cases the majority of genes were responsive to JA, with response to wounding having the slightly more number of genes. Salt and oxidative stresses consisted of JA and ABA responsive genes. While oxidative stress had equal amount of genes from both JA and ABA, majority of salt stress genes were associated with ABA response. Genes involved in salt stress associated with ABA included 2 genes encoding abscisic acid responsive elements-binding factors (ABFs), *ABF2* and *ABF3* which are a class of ABA/stress-inducible bZIP transcription factors. *ABF2* and *ABF3* are mainly induced by high salt thus function in osmotic stress signaling [214]. *PHOSPHOLIPASE C1* (*PLC1*), *ANNEXIN 1* (*ANNAT1*), *ANNAT4* and *BETA GLUCOSIDASE 18* (*BGLU18*), among this group, are genes required for ABA signaling and several environmental stresses including salinity [214-219]. For DEGs from *eop002* leaves, the stress category was just the specific term, response to wounding. In a similar manner as in apices, some genes from this category, were associated with the hormone response, more specifically JA, which is also the only hormone response enriched in *eop002* leaves. These results indicate that in *eop002* there is a down regulation of genes associated with stress response which are made up of several hormone inducible genes. JA genes are the main representatives in response to wounding and defense while ABA responsive genes are the most predominant genes affiliated to salt stress.

2.4.4 Genes involved in hormone synthesis and signaling in *eop002*

Genes associated with the hormones where generally grouped under response to stimulus, mainly endogenous and chemical stimulus. Ethylene and auxin associated genes were placed under the term hormone-mediated signaling together with several other genes. Genes involved with ABA, JA, ET and auxin were mainly down-regulated in *eop002* apices and in the case of JA in leaves as well. ABA had the most down-regulated genes in *eop002,* with 22 down regulated genes in the apices, this included several genes that can be induced by ABA stimulus. In addition to this, genes involved in ABA biosynthesis such as *NINE-CIS-EPOXYCAROTENOID DIOXYGENASE 3* (*NCED3*), which is a key enzyme in the biosynthesis of abscisic acid [220] and *REVEILLE 2* (*RVE2*), a gene involved in circadian regulation in *Arabidopsis* but also in the positive regulation of the ABA biosynthesis [221], were among the down-regulated genes enriched for ABA response. *ABF2* and *ABF3* which were also down-regulated are involved in ABA signal transduction during vegetative growth [214]. *ELONGATED HYPOCOTYL 5* (*HY5*), a

bZIP transcription factor, was among enriched genes in the ABA response. *HY5* is suggested to be involved in mediating ABA responses by binding to the promoter of *ABI5*, another bZIP transcription factor involved in ABA signaling, to regulate its expression [222, 223]. Therefore, ABA biosynthesis and signaling seems to be negatively affected in the apices of the *eop002* mutant.

The down regulation of JA-response genes could be due to affected JA synthesis and signaling, as one of the down-regulated genes, *OPR3*, encodes a 12-oxophytodienoate reductase, which is the most relevant isoenzyme for jasmonate biosynthesis [224]. *MYC2*, among the down-regulated genes, is considered to be the key transcriptional activator of JA response genes and therefore plays a very important role in JA signaling [225]. Other JA-related down-regulated genes were, *JASMONATE-ZIM-DOMAIN PROTEINs 6* and *9* (*JAZ6* and *JAZ9*), members of the JAZ protein family. MYC2 and JAZ proteins are involved in a negative feed-back regulatory loop in which MYC2 seems to mediate the induced expression of the *JAZs* in response to jasmonate stimulus [226]. JA was the only hormone with down-regulated genes in *eop002* leaves. This included several genes that were also down-regulated in *eop002* apices. An interesting gene placed in this category was *GIBBERELLIC ACID INSENSITIVE* (*GAI*), which encodes a transcription factor that is a member of the DELLA proteins. JA has been demonstrated to antagonize the GA-mediated degradation of DELLA proteins [227]. Taking these together suggests, like ABA, JA biosynthesis and signaling might be negatively affected within the *eop002* mutant.

Ethylene and auxin, among the enriched hormone responses, contained the least number of enriched genes. Three ethylene associated genes, *BASIC CHITINASE* (*B-CHI* or *CHI-B*), *ETHYLENE RESPONSE FACTOR 8* (*ERF8*) and *ETHYLENE-RESPONSIVE ELEMENT BINDING PROTEIN* (*EBP*), involved in ethylene signaling pathway were down-regulated in *eop002* apices. As transcriptional factors, *EBP* acts as an activator by binding to the GCC-box of some ethylene-responsive genes to regulate their expression while *ERF8* is predicted to be a transcriptional repressor [228, 229]. Therefore mutations in *EOP* appear to negatively affect ethylene signaling and expression of ethylene inducible genes. In terms of auxin, *CYTOCHROME P450, FAMILY 79, SUBFAMILY B, POLYPEPTIDE 2* (*CYP79B2*) and *3* (*CYP79B3*), genes that encode enzymes critical for auxin biosynthesis by converting tryptophan (Trp) to IAOx which is important for auxin biosynthesis [230], were part of the down regulated genes.

An auxin influx transporter gene, *AUXIN RESISTANT 1* (*AUX1*) was also included in the down regulated genes. AUX1 functions in the uptake of Indole-3-acetic acid (IAA), a major form of auxin, by mediating its transport in plant cells [231, 232]. These therefore suggest a negative influence of *eop* mutation on auxin synthesis and transport.

2.4.5 *eop002* DEGs involved in rhythmic process and light response

Rhythmic process was one of the categories with very few genes, and these genes were all up-regulated in leaves. There were only three genes overrepresented in *eop002*, which were *PSEUDO-RESPONSE REGULATOR 7 (PRR7)*, *EARLY-PHYTOCHROME-RESPONSIVE1 (EPR1)* and *BTB AND TAZ DOMAIN PROTEIN 2* (*BT2*). *PRR7* belongs to ARABIDOPSIS PSEUDO-RESPONSE REGULATOR protein family, of which the circadian central oscillator *PRR1/TOC1* also belongs. Mutations in any of the PRR family genes causes disruptions of normal circadian function which includes flowering time control. *PRR7* has been demonstrated to play important roles in regulating certain clock-controlled genes and plants with mutations in *PRR7* and *PRR5* show late flowering phenotypes under long day conditions [233]. *PRR7* together with *PRR9* have been shown to act as transcription repressors of *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1)* and *LATE ELONGATED HYPOCOTYL* (*LHY*), which are other central oscillators in *A. thaliana*, in response to ambient temperature changes [234]. *EPR1*, part of the up-regulated genes, contains a single MYB domain and is considered to be highly similar to *CCA1* and *LHY*. It contributes to the circadian output pathway by mediating the oscillatory behavior of clock-controlled genes [235]. The third gene, *BT2*, is a gene whose expression can be modulated by the circadian clock and light. It plays a role in enhancing some auxin responses while repressing ABA and sugar responses [236]. *GI* as part of this category is up-regulated in *eop002* apices. *GI* plays significant roles in the regulation of circadian rhythms and control of flowering time. It is able to regulate flowering by acting between *CCA1/LHY* and *CO* to promote the expression of *CO* and FT [64]. The up regulation of genes involved in rhythmic processes in *eop002* mutants indicates a possible positive effect of *eop* mutations on some genes involved in these processes.

In terms of light responses, there were 11 up-regulated genes in *eop002* leaves. Seven of these were grouped under more specific terms response to red or far red light. This included, *CYTOCHROME P450, FAMILY 707, SUBFAMILY A, POLYPEPTIDE 3* (*CYP707A3*), *PIF5*,

GIBBERELLIN 2-OXIDASE (*GA2OX2*), *EARLY LIGHT-INDUCIBLE PROTEIN 2* (*ELIP2*), *LIGHT HARVESTING COMPLEX PHOTOSYSTEM II* (*LHCB4.3*) and *PHOTOSYSTEM II LIGHT HARVESTING COMPLEX GENE 2.3* (*LHCB2.3*). *CYP707A3* is involved in ABA catabolism, when overexpressed, it results in increased transpiration rate and therefore reduced drought tolerance during early seedling growth [237]. Its up-regulation in *eop002* implies an increase in water loss.

2.4.6 Genes involved in water deprivation and temperature response in *eop002*

Genes involved in water deprivation and temperature were down-regulated in *eop002* apices. Water deprivation in *eop002* was made up of 19 down-regulated genes, 13 of which were ABA responsive, meaning 68% of the genes, and just two JA inducible genes. There were other genes such as the transcription factor *DREB2A* and *DREB2A-INTERACTING PROTEIN 2* (*DRIP2*). *DREB2A* plays important roles in tolerance to drought by inducing water stress response genes independent of ABA [199]. *DRIP2* on the other hand interacts with *DREB2A* to probably cause the degradation of DREB2A protein thus acting as a negative regulator of drought-responsive gene expression [238]. Both genes, however seem to be down-regulated in *eop002* apices.

Temperature response consisted of 17 down-regulated genes in *eop002* apices. Participating genes in this category were mainly cold or low temperature response genes with just five heat response genes present. The heat genes included *HSP70*, *Hop3* and *MYO-INOSITOL-1-PHOSPHATE SYNTHASE 2* (*MIPS2*), these were part of down-regulated oxidative stress response genes in *eop002* apices. Six genes in this category were also responsive to ABA and two to JA. The rest of cold response genes independent of hormone response included, *COLD REGULATED GENE 27 (COR27)*, *LATE EMBRYOGENESIS ABUNDANT 4-5 (LEA4-5)* and *RARE COLD INDUCIBLE GENE 3* (*RCI3*). *COR27* is an interesting gene whose transcription is regulated by the circadian clock and can also be induced by cold, it then plays significant roles in cold response by negatively regulating freeze tolerance [239]. These results indicate a possible negative impact of *eop* mutation on drought and temperature, mainly cold, response genes.

2.4.7 Analysis of *eop002* and *eop101* overlapping DEGs

Considering the fact that *eop101* was discovered to also contain a mutation in the *EOP* gene, differentially expressed genes in the both *eop002* and *eop101* were compared to identify the possibility of similar regulatory mechanisms. A Venn-diagram was created to evaluate the overlaps between DEGs of *eop002* and *eop101* (***Figure 32***). Up and down-regulated DEGs from *eop002* and *eop101* leaves were grouped in one diagram (Figure 32A) and those from apices were placed in another (Figure 32B). In the leaves, *eop002* and *eop101* had more up-regulated DEGs in common, that is 116, than in down-regulated, with 60 genes. On the contrary, in apices, the two mutants had more down-regulated genes in common, 182, than up-regulated, 64 genes. GO was performed on the common DEGs to aid group them into functional categories (Table 12).

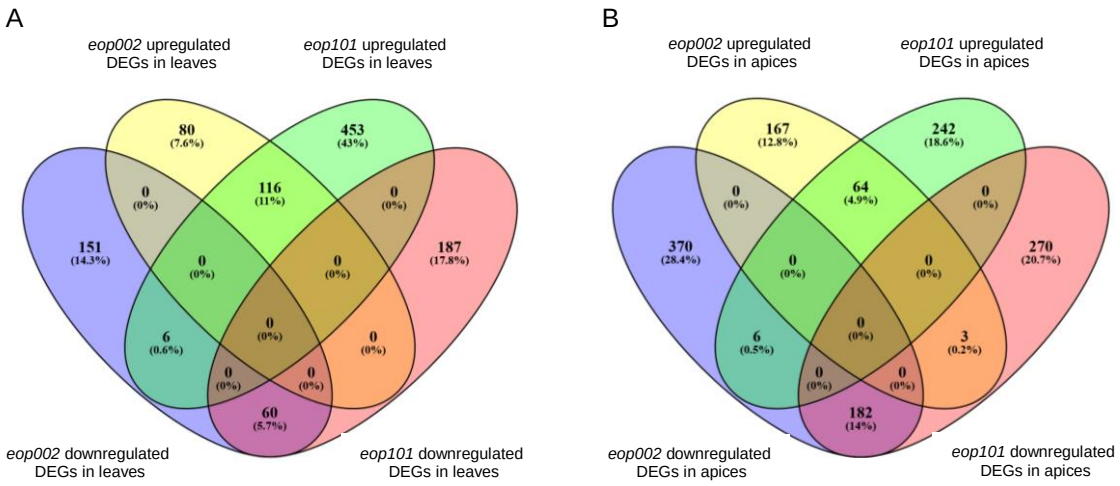


Figure 32. Overlapping DEGs in *eop002* and *eop101* in apices and leaves.
(A) Number of overlapping DEGs in leaves. **(B)** Number of overlapping DEGs apices.

Table 12. Number of genes involved in different functional categories

Category		*eop002-eop101* DEGs overlap	
		Apices	Leaves
Developmental process	UP	2	0
	Down	0	10
Sulfur process	UP	4	6
	Down	0	0
Carbohydrate process	UP	2	3
	Down	0	0
Abscisic acid (ABA)	UP	0	0
	Down	6	0
Auxin	UP	0	0
	Down	2	0
Response to stress	UP	0	0
	Down	24	0
Transport	UP	0	0
	Down	20	1
Response to light stimulus	UP	0	10
	Down	0	
Rhythmic process	UP	0	3
	Down	0	0
Biological process regulation	UP	1	0
	Down	0	5

a. DEGs, Differentially Expressed Genes
b. Up-regulated genes are shaded in red and down-regulated in blue

2.4.8 GO analysis of *eop002* and *eop101* overlapping DEGs

The categories with the highest enriched genes were response to stress and transport which mainly consisted of down-regulated genes in the apices. A closer look was taken at genes participating in these categories. Response to stress was particularly the specific terms, oxidative stress, response to defense and wounding. There were 4 genes in both defense and wounding response and 7 genes associated with oxidative stress. The gene *VEGETATIVE STORAGE PROTEIN 2* (*VSP2*) was found to be involved in all three stress responses and in addition it is considered to be inducible in response to abscisic acid, jasmonic acid, salt and water deficiency, which are all processes previously mentioned as affected in either one or both of the mutants. Other genes in this category included previously mentioned genes, *ARGININE DECARBOXYLASE 2* (*ADC2*), *CORONATINE INDUCED 1* (*CORI3*), *CYP79B2* and

CYP79B3. *ADC2* and *COR13* are also involved in ABA response and part of the overlapping genes enriched in the ABA category. Genes associated with the transport category were group in specific terms, lipid and oligopeptide transport. As obtained for *eop002* (section 2.4.3) genes involved in lipid transport consisted of mostly, members of the LIPID TRANSFER PROTEIN family and *END1* while oligopeptide transport was made up of NRT1/ PTR FAMILY genes, *NPF2.10, NPF2.13* and *NPF5.13*, and *OPT3*.

The categories of developmental process and response to light also contained relatively, high numbers of overrepresented genes. The genes involved in developmental process were mostly down-regulated in the leaves with just a few up-regulated in the apices. On the other hand those associated with light were up-regulated in the leaves. The 2 genes in apices were specifically involved in stomatal complex morphogenesis, these were *ERL1* and *TMM* which were also part of genes involved in developmental process in *eop101* but were not placed under any specific term. Out of the 10 genes in the leaves, 4 were placed under specific terms, 1 gene, *PERMEABLE LEAVES3* (*PEL3*), was associated with tissue morphogenesis and 3, *BARELY ANY MERISTEM 2* (*BAM2*), *PASTICCINO 2* (*PAS2*) and *ZINC FINGER PROTEIN 8* (*ZFP8*), with cell differentiation. *BAM2* encodes a CLAVATA1-related receptor kinase-like protein and is also required for shoot and flower meristem functions. The remaining genes were also found to be involved in reproductive development [240]. *PACLOBUTRAZOL RESISTANCE1* (*PRE1*) which is involved in cell elongation is also directly and negatively regulated by *APETALA3* (*AP3*) and *PISTILLATA* (*PI*), both floral homeotic genes, in petals and required for the regulation of flowering time [241] was among the genes here. *PROMOTION OF CELL SURVIVAL 1* (*PCS1*) plays important roles in cell fate determination during embryonic development and also in reproduction processes [242]. *SEPALLATA 2* (*SEP2*) is a transcription factor that functions with other *SEP* genes to ensure the proper development of floral organ identity [243], another gene also involved in floral organogenesis, *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 8* (*SPL8*), was also part of these genes [244].

In terms of light responses, the overlapping genes were group under more specific terms response to red or far red light. Most of genes in these category have been mentioned in other categories such as, *CYP707A3*, *PRR7*, *PIF5* and *EPR1*.

Taken together, the fact that stress response, transport, and developmental process, specifically reproductive development, are affected in both mutants, implicates the involvement of *EOP* in these processes.

3 Discussion

3.1 Extended vernalization decreases the days required for flower emergence

In many temperate-zone plants, prolonged exposure to cold temperatures, such as during winter, is a requirement to promote flowering. Response to cold temperature exposure (vernalization) can be facultative or obligate. Depending on whether or not annual plants require vernalization to accelerate flowering, they can be referred to as winter or summer annuals. Winter annuals exhibit a facultative response to vernalization while summer annuals do not require vernalization to accelerate flowering. In the winter annuals, floral initiation which is immediately followed by flower emergence occurs after vernalization when plants are returned to permissive temperatures. Several temperate and apline perennials, on the other hand, have obligate requirements for exposure to winter chilling temperatures and show considerable delay between floral initiation and flower emergence. Floral initiation, which results in the formation of flower buds, occurs a season or more before flower emergence [125, 126, 245]. Winter chilling in several cases are crucial to enable bud break and hence flowering in spring, while preformation of flower buds enables rapid flower emergence during the short growing seasons in these environments [125, 130]. In the perennial species, *A. alpina,* flower initiation and flower buds formation occurs during periods of cold exposure [119]. Flower emergence from the preformed buds then occurs once permissive conditions commence. Increasing the exposure time of plants to cold, in several researches, decreased the periods of permissive temperatures required for flowering or budburst followed by anthesis to occur [147-151]. The Pajares accession of *A. alpina,* in this experiment, exhibited a quantitative response to different vernalization durations in terms of days to flower emergence. Plants vernalized for 12 weeks required twice as much days as 24 weeks vernalized plants to enable flower emergence, demonstrating that the reduction in days required for flower emergence depends on the extent of vernalization.

Exceeding the vernalization requirement of a cold requiring plant, does not alter any further the period of warm temperatures required to ensure flowering. Vernalization or chilling requirement are specific for different species and cultivars. Hardy woody perennials species

with late budburst required larger thermal time to enable budburst which decreased with an increase in chilling days. On the contrary, species with early budburst required less thermal time and this barely decreased with increasing chilling days as these species achieved early saturation of chilling requirement [148]. Exposure of different *Arabidopsis* accessions to different periods of vernalization, caused an almost complete acceleration to flowering in some accessions after just 4 weeks vernalization whereas other accessions, such as Lov-1, required 14 weeks to fully saturate their vernalization requirement [147]. Exceeding the vernalization requirement of Lov-1, caused no apparent change in the number of leaves at flowering. After exposing *A. alpina* accession Pajares to different vernalization durations, plants vernalized for 18, 21 and 24 weeks showed marginal decrease in days to flower emergence. This suggests, the vernalization requirement of Pajares to accelerate flowering is close to complete after 18 weeks of vernalization.

3.2 Prolonged vernalization periods are crucial for the development of the inflorescence

Floral commitment ensures that flowering is not only established, as in initiation, but also maintained to enable full development of the inflorescence. Changes in some environmental and endogenous factors during flowering induction can lead to a reversion in the decision to flower when meristems are not yet fully committed [22, 23]. In *A. thaliana,* genes such as *LFY*, *AP1* and *FUL* are important floral promotive signals that confer floral identity, and consequently determinacy, to emerging meristems [19, 246]. Disruptions or reductions in the activity of these genes causes plants to display reversion phenotypes [101, 103]. This indicates the importance of plants to accumulate sufficient floral promotive signals in order to ensure proper development of the entire inflorescence.

Even though days to flower emergence decreased with extended vernalization durations, Pajares plants possessed similar number of leaves at flowering. As previously shown by the up regulation of *AaLFY* in shoot apical meristems of Pajares, flower transition occurs around 5 weeks in vernalization [140], which shows that even the minimum 8 weeks vernalization applied here can cause a transition to flower initiation. However prolonging vernalization, after flower initiation has already occurred, is necessary to ensure the commitment to

flowering and development of the inflorescence. This requirement is evident in the fact that Pajares plants with less than 21 weeks of vernalization exhibited floral reversion and vegetative phenotypes within their inflorescence. These phenotypes included the production of bracts in the I2 zone and the presence of vegetative axillary branches in the I1 zone of the inflorescence. Occurrence of vegetative branches along the inflorescence stem is not random. According to data from this experiment, development of branches in *A. alpina* occur basipetally, with higher branches being bigger than lower ones. In addition, the fate of branches in the inflorescence also proved flowering of inflorescence branches occurred basipetally. The frequency of flowering branches scored were lowest at the basal nodes in the inflorescence I1 zone. The points of lowest percentages ranged from nodes 1 to 6 in 12 and 15 weeks vernalized plants to just node 1 in 18 and 21 weeks vernalized plants. This demonstrates, with extended vernalization the inflorescence meristem accumulates sufficient promotive floral signals to ensure the flowering of its axillary branches in a basipetal manner. In support of this, the percentage of flowering branches increased with longer vernalization. Excessive accumulations of such signals, however, seems to cause flowering of branches just below the inflorescence (in the V3 zone on the plant) which are intended to remain vegetative, as observed in 24 weeks vernalized plants. On the other hand, insufficient vernalization results in the inability of axillary branches at the base of the main inflorescence to pass the required threshold to ensure commitment to flowering. This causes a delay in floral determinacy in those branches thus increasing their likelihood to remain vegetative.

In the perennial sense, floral reversion and non-commitment of some I1 lateral branches to flowering might serve as an advantage to switch between vegetative and reproductive development. An early and rapid response to warm conditions after long cold winters serves as an advantage to make maximum use of the favorable growth conditions for seed set in arctic-alpine environments. As demonstrated here, Pajares plants with vernalization above 15 weeks show early inflorescence outgrowth once in warm conditions and number of siliques increases marginally, but significant, in these plants. Since floral reversion occurs and some inflorescence branches stay vegetative after insufficient vernalization, it can be a potential mechanism adapted by these plants to avoid senescence at a time when flowering might not result in successful seed set before the next winter arrives.

As previously mentioned, non-saturation of Lov-1 vernalization requirement resulted in the reactivation of *FLC* expression after return of plants to warm conditions (Section 1.8). This expression dynamics of *FLC* observed in Lov-1 is quite similar to what has been observed for *PEP1* so far. In addition, the *FLC* levels decreased over four weeks period in cold however 14 weeks was required to accelerate flowering. This prolonged vernalization requirement accelerated flowering by ensuring the stable silencing of *FLC* in Lov-1. Acceleration of flower emergence in Pajares after extended vernalization implicates the involvement of molecular mechanisms such as stable repression of *PEP1* after prolonged vernalization durations. In accordance to this, further experiments conducted by Dr. Ana Lazaro (AG. Albani, University of Cologne, Germany) demonstrates that extended vernalization causes the stable repression of *PEP1* in the inflorescence whereas short durations results in the reactivation of *PEP1*. This repression was specific to flowering inflorescence branches whereas vegetative branches within the inflorescence I1 zone and in V3 zone exhibited high levels of *PEP1*. In addition, exposure of the *pep1-1* mutant to varying vernalization durations revealed a reduction in floral reversion phenotypes in *pep1-1* compared to Pajares. These therefore suggest the involvement of *PEP1* in the regulation of inflorescence development after prolonged exposure to cold. Furthermore within flowering apices, floral meristem identity genes *AaFUL*, *AaLFY* and *AaAP1* showed higher levels of transcript accumulation whereas transcript levels of *AaTFL1*, which as a shoot identity gene in *A. thaliana* has been shown to antagonize activities of floral identity genes, were gradually reduced after extended vernalization. Together these results demonstrates that extended vernalization is essential to promote floral commitment and inflorescence development, and that mechanisms identified to regulate these processes in *A. thaliana* are partly conserved in *A. alpina*. Prolonged vernalization enables the stable repression of *PEP1* which antagonizes floral commitment and increases the quantitative accumulation of flowering identity genes which can be associated with flower initiation [247].

3.3 Adaptation of leaf shape and senescence as morphological markers for inflorescence zone detection in *pep1-1* and *pep1-1* enhancer mutants

The main shoot of flowering plants undergoes different phase changes during its development, which can be characterized by some set of unique morphological attributes. These attributes include leaf shape and size, internode elongation and production of flowers [30]. The reproductive phase in *A. thaliana* causes major changes in the plant structure that are entirely different from those observed at the vegetative phase, such as bolting of the stem and changes in leaf shape. Leaves within the inflorescence are easily identified as they are non-petiolated and directly attached to the stem. These features of leaf shape change and bolting of the inflorescence were as well observed here in the *A. alpina* accession, Pajares, during reproductive phase of development. *A. thaliana* as an annual plant commits all branches to flowering in order to produce as many seeds as possible. This however causes a depletion in the plant's energy and thus ability to support growth, leading to senescence and death of the whole plant. Senescence involves a coordinated process, first, senescence of somatic organs and tissues such as leaves occurs, followed by an arrest of the shoot apical meristem and finally the suppression and prevention of axillary meristems from forming new shoots [248]. At the end of reproductive development, senescence in Pajares, in contrast to *A. thaliana*, was restricted to just the inflorescence zones and axillary branches that flowered. This behavior permits the continuous supply of axillary meristems in support of its perennial life strategy. The *pep1-1* mutant derived from Pajares, contains lesion in the *A. alpina* gene, *PEP1*, which enables it to overcome the obligate vernalization requirement to flower and flowers perpetually [119]. Mutagenesis of *pep1-1* mutant gave rise to second site enhancer mutants to aid the study of additional factors contributing to flowering in *A. alpina*. The attributes of leaf shape and restriction of senescence to the inflorescence zone were adopted as markers for the detection and characterization of inflorescence zone in the *pep1-1* and *enhancers of perpetual flowering 1* (*eop*) mutants. Even though both *pep1-1* and *eop* mutants possess less defined petiolated leaves, inflorescence leaves were distinguishable as they were smaller, more rounded at the base and directly attached to the stem. Senescence in the mutants was also restricted as observed in Pajares which provides the means for these

mutants to remain perennial. In all, these two features of leaf shape and senescence served as good markers for the determination of inflorescence zone in this study.

3.4 Flowering and inflorescence of the enhancer mutant, *eop101*.

The mutant, *eop101*, unlike the other mutants, developed a determinate inflorescence with a terminal flower and was identified to contain a mutation in the *AaTFL1* gene. Mutations in *A. thaliana TFL1* gene have so far been extensively studied. Key features of *tfl1* mutants include, accelerated flowering and the formation of a terminal flower [4-6, 105, 109, 249]. *A. thaliana tfl1* mutants have been reported to flower with less number of rosette leaves compared to wild [6, 109, 249], however there are instances in which, depending on the *tfl1* mutant allele, very minimal differences are observed in rosette leaf number of mutants compared to wild type [5, 105]. In addition, Shannon *et al*., 1991, after studying the vegetative growth rate and morphology of *tfl1* mutant compared to wild type, concluded that these features were indistinguishable, thus the mutant had a normal vegetative development. *TFL1* as a shoot identity gene antagonizes the expression of floral meristem identity genes such as *LFY* and *AP1* [109]. These antagonistic interactions regulate the inflorescence branching pattern. Mutations in *TFL1* leads to ectopic expression of *AP1* and *LFY* in the inflorescence meristems which results in accelerated flowering and commitment of lateral meristems to floral fate instead of inflorescence shoots, with pedicellate flowers usually occurring in place of branches [4, 250]. Branches that do occurred were found to occupy in the lower cauline leaves [5]. In accordance with what has been observed in *A. thaliana*, the *eop101* mutant, showed no apparent deviation from the *pep1-1* vegetative phase as it produced similar number of leaves as *pep1-1*. The inflorescence of this mutant consisted of cauline-like leaves that subtended either inflorescence branches, which also terminated, or pedicellate flowers. The pedicellate flowers mainly occupied the axils of leaves closest the main shoot apex while the inflorescence branches occupied the lower axils. As the number of cauline-like leaves remains similar in *pep1-1* and *eop101*, the presence of pedicellate flowers and reduction of inflorescence branches infers, in the absence of *AaTFL1*, there is the expression of floral meristem identity genes which result in the commitment of some lateral meristems into floral fate just as in *A. thaliana*.

Scoring of inflorescence based on senescence criteria in *eop101* showed significant differences with traits scored with leaf shape. This was not observed for the other *eop* mutants and *pep1-1*. A gibberellin inducible gene *PPF1* has been shown to delay flowering and significantly increase the lifespan of *A. thaliana* plants when overexpressed [251]. *TFL1* was screened out as one of the downstream targets of *PPF1* in the senescence-signaling pathway and its expression was activated by *PPF1*, suggesting that *PPF1* might negatively regulate senescence by stimulating *TFL1* expression [252]. Further, an increase in Malondiadehyde (MDA) content and decrease in chlorophyll content served as indicators of plant senescence, the content of MDA was found to be significantly high and chlorophyll significantly low in the *tfl1-1* mutant compared to wild type [252]. This suggest that the use of senescence as a marker for inflorescence zone detection in *eop101* might be compromised.

3.5 Several floral genes and transcription factors are misregulated in *eop101*

Transcriptome analysis of the *eop101* mutants identified the overrepresentation of genes involved in 17 biological functional categories. The majority of these genes were up regulated and categories such as biological process regulation, response to stress, carbohydrate process, developmental process and response to light stimulus had the highest numbers of genes. This implies *AaTFL1* activities exerts repressive roles on these biological functions.

Important flowering time genes in the photoperiod pathway were up-regulation in *eop101*. *GI* in this pathway activates *CO* which then activates *FT* in leaves [64]. It is interesting that *GI* was up-regulated in *eop101* however an increase of *CO* was not detected. Possibly the activation of *CO* by *GI* is prevented by presence of some negative regulators. Genes such as *CDF1* and *SPA1* which have been identified as negative regulators of *CO* [187, 189] were indeed up-regulated in *eop101*. Thus, in *eop101* increased expression of *CDF1* and *SPA1* may be antagonizing the activation of *CO* by *GI*. In contrast to *CO*, *FT* transcript levels were up regulation in *eop101*. This increase in *FT* without an increase in *CO* transcript levels might be attributed to reduced levels of *FT* repressors, however in *eop101* there is the up regulation of *TEM1*. *TEM1* is a member of the RAV transcription factor family and has been shown to directly bind to the promoter of *FT* to repress its transcription therefore delaying floral induction [253].

GI has been demonstrated to activate *FT* expression in a *CO*-independent manner under short day conditions [254]. It does so by binding directly to the promoter region of *FT* which also turns out to be the binding site of some *FT* repressors including *TEM1*, thereby blocking their access to *FT* promoter region [254]. Thus, *AaTFL1* mutation in *eop101* perhaps elevates *GI* expression levels which in turn causes an increase in *FT* transcript levels independent of *CO*.

TFL1 and *FT* encode a pair of flowering regulators that share approximately 60% amino acid sequence identity however they function in an opposite manner[95, 255]. Whereas *FT* promotes flowering transition, *TFL1* represses this transition. In the shoot apical meristem FT protein binds to the bZIP transcription factor FD to form a complex that positively regulates meristem identity genes, such as *AP1* and *FUL*, which results in flowering induction [92]. TFL1 on the contrary represses several genes downstream of FT, which includes *AP1* and *LFY,* also through interactions with FD, it therefore negatively modulates the FD-dependent transcription of these target genes [105]. The elevation of *FT* transcript levels and mutation of *AaTFL1* in the *eop101* mutant, presents a plausible mechanism by which early flowering transition is achieved in this mutant since the antagonistic effect of *AaTFL1* might be disrupted.

As discussed previously, *TFL1* in *A. thaliana* is considered as a negative regulator of flowering time and a positive regulator of shoot development. It does so by preventing the expression of floral meristem genes such as *AP1* and *LFY* [109]. Analysis of RNAseq data from *eop101* revealed, within the apices of *eop101* there is the up regulation of flower meristem determinacy and organ identity genes *CRC* (which contains *LFY* binding sites) and *SEP4*. *HB51* is another meristem identity regulator that acts together with *LFY* which is also up regulated in *eop101*. This demonstrates that, similar to *A. thaliana TFL1*, *AaTFL1* might be preventing the expression of genes involved in floral meristem identity to favor shoot development. *EMF1* and *TFL1* are considered to regulate similar effects on maintenance of inflorescence growth. The *emf* mutant also produces a reduced inflorescence with a terminal flower [174]. Within the *eop101* mutant there was the up regulation of *EMF1* and other flower repressors, *TEM1* and *ICU2*. Mutations in *ICU2* results in early flowering and increased expression of floral organ identity genes [179]. The TEM transcription factors have further been shown to contribute to the age at which plants are able to sense environmental cues that induce flowering. They act mainly in young *A. thaliana* plants to repress their response to long day conditions and

reduction in *TEM* activities enables young plants to acquire the competence to flower in response to photoperiod [34, 253]. In a similar manner *AaTFL1* increases the duration of the phase at which plants are insensitive to floral inductive signals by preventing the activation of *AaLFY* when young *A. alpina* Pajares plants are vernalized [140]. The up regulation of floral repressors with similar functions as *AaTFL1* in the *eop101* mutant presents a possible redundancy in some of their functions in the absence of *AaTFL1*.

A. thaliana TFL1, considered as a floral repressor, does not possess any form of sequence domains characteristic to transcription factors [105, 256]. Similarly, *FT* which shows high amino acid similarity to *TFL1* also contains no apparent transcription factor domains but it interacts with the transcription factor FD to convey transcriptional activities [256]. *TFL1* has also been proven to act in transcriptional repression as it was able to fuse to a repressor domain, SRDX and rescued the *tfl1* mutant phenotype. This demonstrated *TFL1* is able to function as a transcriptional mediator molecule and is involved in a transcriptional regulator complex whereby it interacts with corepressors [105]. Within *eop101*, majority of the up regulated genes in leaves were transcription factors. Many of these transcription factors were associated with the regulation of circadian rhythms and flowering time. This indicates that *AaTFL1*, similar to *Arabidopsis TFL1*, might be part of a transcriptional complex and interacts with other corepressors, thus its absence interferes with the functions of the repressive transcriptional complex. This might be the reason for up regulation of several flowering and circadian genes.

3.6 *eop002*, *eop085*, *eop088* and *eop091* possess varying degrees of dominance

Characterization of flowering time and inflorescence of *eop002*, *eop085*, *eop088* and *eop091*, revealed very similar phenotypes. They flowered with close to identical number of leaves and possessed similar reduced number of inflorescence branches compared to *pep1-1*. Through phenotyping for number of leaves at flowering, allelism was inferred from F1 individuals, from crosses of the mutants, as they failed to complement the early flowering phenotypes. The F1s in the dominance test of *eops* crossed to *pep1-1*, on the other hand showed varying degrees of dominance. This ranged from the strong recessiveness of *eop091* allele to the weak

recessiveness of *eop085* allele as the dominance of the wild type allele was not complete. The *eop088* allele was semi-dominant whereas *eop002* allele was dominant, but not complete, over the wild type allele. The segregation of F2 individuals were consistent with Mendelian segregation for monogenic basis of the mutations. In several cases allelic dominance depends on the influence of each allele of a gene on specific phenotypes, the genetic background and environmental conditions. There have been reported cases where different alleles of a gene exhibit varying degrees of dominance. In pea plant, *Pisum sativum*, four alleles of the flowering time gene, *Lf*, were described to exert different dominance, which resulted in the dominance order proposal of $Lf^{d} > Lf > lf > lf^{a}$ [257]. The ability of the different alleles to delay flowering correlated with their dominance level. The allele *lf* exhibited dominance of 0.69 over lf^{a} and plants homozygote for *lf* flowered later than lf^{a} plants. In the case of the two furthers alleles, lf^{a} and Lf^{d}, plants homozygous for lf^{a} were the earliest to flower while those homozygous for Lf^{d} flowered the latest. In some cases their dominance ability was influenced by the genetic background and environmental factors. The *Lf* allele showed strong dominance over lf^{a} in long day but very weak dominance in short day. These multiple alleles of a single gene that represses flowering, displayed allelic series in which each allele exhibited a different level of dominance compared to the next. In the case of the *eops*, concluding that they are allelic is rather complicated because even though allelism can be inferred from *eop085* and *eop091,* as they are recessive, the dominance of *eop002* and semi-dominance of *eop088* creates difficulties in the interpretation of their results. In a circumstance like this, the mapping of the casual mutation of these *eops* to the same chromosomal region serve as a good indication of them being allelic.

3.7 SHOREmapping and co-segregation analysis identifies the *EOP* gene

Advances in sequencing technology have to a great extent improved the identification of casual mutations. By implementing mapping-by-sequencing (SHOREmapping) in this study, a candidate region was revealed on *A. alpina* chromosome 8 to harbor the casual mutation of the *eop* mutants. In *eop002* a second region was identified on chromosome 2 with distortions in allele frequency distribution. This region was eliminated as causal after co-segregation analysis as markers used did not segregate but rather were homozygote for the mutant allele

in all F2 individuals. Similar circumstance was reported in a study to identify the causal gene of an *Arabidopsis* clock mutant *early bird* (*ebi-1*) phenotype [258]. They obtained two candidate regions, one on chromosome 1 and the other chromosome 5, however rough mapping placed the location of *ebi-1* on chromosome 5 and the SNPs on chromosome 1 did not segregate with the *ebi-1* phenotype. They therefore focused only on SNPs in chromosome 5. In these instances a possible explanation could be that those SNPs are from the background used for the initial generation of the mutant. However due to their absence in background genome sequence used for filtering analysis, these mutations cannot be filtered out.

The region obtained for all mutants on chromosome 8 was atypically large. *A. alpina* has been implied to possess a greater increase in the extent of non-recombining DNA as compared to *A. thaliana* [145]. This was reported to be due to the fact that chromosomes of *A. alpina* contain regions with suppressed crossovers which co-localize with large parts of the pericentromeres [145]. In regions of low recombination, there are higher chances for 'hitch-hiking' to occur. This is when an allele gets a lift in frequency due to changes in frequency of a nearby allele under selection [259]. The large mapping interval obtained on chromosome 8 can therefore be attributed to the low recombination rate in *A. alpina* and nearby mutations 'hitch-hiking' with the *eop* mutation during selection of the mapping population. Nevertheless, co-segregation analysis with marker assisted selection in *eop002* and *eop091* F2 populations refined the target interval and together with comparisons of mutation within the *eop* mutants, the *EOP* locus was corresponded to a gene annotated as *AaG106560*.

3.8 *EOP* encodes a member of the AAA+ proteins

The isolated *eop* mutants contained independent mutations that caused amino acid substitutions in *AaG106560*. Complementation analysis with the native promoter and genomic sequence of *AaG106560* rescued the mutant phenotypes in *eop088* and *eop085*, supporting the discovery of the *EOP* gene. *EOP* encodes a member of an AAA+ ATPase superfamily whose orthologue in *A. thaliana* has so far not been reported to play a role in flowering or inflorescence development.

AAA+ proteins are present in all organisms. *A. thaliana* happens to have the most, with ~140 AAA+ proteins [260]. The AAA+ proteins display remarkable diversity in their functions ranging

from transport across membranes, chromatin maintenance, proteolysis, DNA recombination, replication and repair. The center of their activities involve ATP binding and hydrolysis which are mediated by different motifs in the AAA+ module [205]. The Walker A motif contains the P-loop which directly interacts with the ATP phosphates while the aspartate and glutamate residues of the Walker B motif can coordinate with Mg2+ and activate water molecule, required for ATP hydrolysis, respectively [261]. The energy acquired from ATP hydrolysis is then used to induce conformational changes or remodeling in target proteins or protein complexes. These changes, in many cases, are sufficient to promote protein unfolding and degradation as they perturb protein structure [260]. Due to their diverse functions, AAA+ proteins can be found in most subcellular compartments [205]. *EOP* as a member of the AAA+ family, appears to localize in the plasma membrane. The plasma membrane, since it serves as an interface between a cell and its surrounding, is important for a wide range of functions including communication with extracellular environment, transport of compounds and processing of signals in response to biotic and abiotic stress [262]. Involvement of plasma membrane localized members of the AAA+ family in flowering time or inflorescence development have so far not been reported. The tandem duplication of the *EOP* orthologue in *A. thaliana* provides a possible explanation as to why it has not yet been discovered in this model species. Gene duplication sometimes results in functional redundancy where mutations disrupting the structure and function of one of the duplicates have no significant effect on the host organism as the necessary basic functions can be provided by the other copy [263]. *A. thaliana* genome is also considered to exhibit pronounced redundancy as it significantly possess a great extent of tandem gene duplications [264]. There have however been reports of other AAA+ family members that affect the flowering time and inflorescence development in plants. *A. thaliana* plants lacking the mitochondria AtFTSH4 protease, a member AAA+ proteins, exhibit delays in their inflorescence emergence as well as shortening of their stems due to premature termination of the inflorescence shoot apices [265]. In a study involving the extension of phenotypic characterization of strong mutants, *RPT2a*, also a member of this family, was discovered to be involved in flowering time and inflorescence stem development in *A. thaliana* [266]. The *rpt2a* mutant flowered late in long day and displayed strong fasciation of the inflorescence stem in short day. These indicate that, members of this family can play significant roles in flowering and inflorescence development.

3.9 *EOP* acts as a repressor of flowering

The ability of *EOP* to repress flowering is supported by the early flowering phenotype of the *eop* mutants. This might be due to *EOP*'s involvement in the maintenance of vegetative development. Overexpression of *EOP* in *pep1-1,* however, caused early flowering in majority of cases. In few instances, late flowering was observed and 35S::*EOP* seemed to disrupt the main shoot apices ability to commit to floral transition. Flowering in the late plants occurred only in basal axillary branches. The main shoot exhibited continuous production of leaves and vegetative branches. Overexpression of floral repressors causing the opposite phenotype of early flowering have been reported. Whereas 35S::*FLC* in Lansberg delayed flowering and in some instances even caused extreme lateness to flower [74, 93, 267], majority of 35S::*FLC* Columbia lines were early flowering [93, 267]. In most instances, more than 50-75% of 35S::*FLC* Columbia T1 lines flowered much earlier than wild type. Among the remaining lines, some flowered similar to wild type or remarkably later than wild type. *MAF1* encodes a protein closely related to FLC and also a repressor of flowering. Overexpression of *MAF1* in Columbia, as observed for *FLC*, caused early flowering in ~77% of T1 plants and flowering in the rest of T1 plants were either later than or same as control [267]. In the case of *MAF1*, it was discovered, the highest levels of overexpression correlated with late flowering and vice versa. The levels of *EOP* overexpression in this study also appears to correlate with flowering time of *pep1-1* T2 plants. Extremely late flowering T2 plants exhibited the highest expression levels of 35S::*EOP*, whereas early flowering lines or lines that flowered similar to *pep1-1* showed very minimal levels of overexpression. Thus, the effect of *EOP* overexpression are comparable to those of other floral repressors. A possible explanation for the observed discrepancies, is that, the overexpression of *EOP* interferes with activities of other floral factors. Minimal overexpression possibly disrupts or cancels out the effects of some factors leading to early flowering, whereas excessively high levels of *EOP* might be able to cause repression even in the presence of these factors. In addition, the AAA+ proteins are known to assemble into oligomers, more commonly, hexamers [204, 260]. This is considered as their biologically active form. The protein encoded by *EOP* may therefore exist as part of a protein complex and changes in its expression levels could compromise protein interaction networks.

3.10 In *A. alpina* Pajares, *EOP* affects flowering in response to vernalization.

The introgression of *eop* in a *PEP1* background enabled plants to partially overcome the obligate vernalization requirement for flowering. In contrast to *pep1* mutants, in which the main shoots and almost all axillary shoots flower under long day conditions [119], disruptions in *EOP* was insufficient to induce full flowering response. Lines which flowered in long day conditions, only did so by a small proportion of basal axillary branches as the *eop* mutation failed to induce flowering in the main shoot apex. This points out the fact that, *PEP1* is a more potent repressor of flowering than *EOP*. Nevertheless, both genes are required to ensure the proper response to vernalization in terms of flowering, which includes the absolute prevention of flowering in 8 weeks old Pajares until plants are exposure to cold treatments. In comparison to another *A. alpina* floral repressor, *AaTFL1* [140], *EOP* appears not to contribute to the prevention of flowering in young vernalized Pajares plants. However, similar to *AaTFL1*, 8 weeks old plants containing *eop* mutation were able to completely flower with a reduced vernalization period of 6 weeks. Pajares plants under such conditions never flowered. This demonstrates that *EOP* contributes to vernalization requirement by increasing the duration of vernalization required for a full flowering response in older plants. It also indicates that, even though plants are able to flower without vernalization in the absence of *EOP*, as shown with *eop002BC1F4* lines, they still require the exposure to ensure full flowering response and commitment of the main shoot apex to floral induction.

After vernalization and return of plants to warmer conditions, disruptions in *EOP* appears to reduce the number of days required for flower emergence in ILs. In addition, whereas Pajares plants require a minimum of 18 weeks to almost complete their acceleration to flowering (Section 3.1), 12 weeks of cold treatment seems sufficient to cause close to complete flowering acceleration in *eop002* and *eop091* ILs. The minimal difference observed in days to flower emergence after 12 weeks vernalization, indicates exceeding this requirement leads to no obvious change in the number of days required for flower emergence. In support of this, *eop002* ILs and Pajares plants vernalized for 21 weeks exhibited similar number of days to flower emergence in LD, since both plants had reached a point of saturation. Thus, a mutation in *EOP* enables plants to achieve an early saturation of vernalization requirement.

3.11 *EOP* acts as a regulator of inflorescence development and plays a role in the perennial life cycle of *A. alpina*

Plants with disruptions in *EOP* not only require minimum vernalization to ensure full flowering response, but also proper development of their inflorescence. Inflorescence emergence was observed earlier in ILs regardless of the vernalization duration, however, after 14 weeks in LD, ILs possessed a reduced inflorescence height in comparison to Pajares. Further investigations into *eop002* ILs revealed this reduction was mainly due to the I2 zone. The ILs contained similar number of inflorescence branches in I1 as Pajares, but then the number of siliques in I2 was close to twice as less as that in Pajares. *pep1-1* mutant also exhibits a similar phenotype in comparison to Pajares (Section 2.2.1). This therefore implies, both *EOP* and *PEP1* play important roles in the development of the inflorescences. As a contributor to the perennial life of *A. alpina* Pajares, *PEP1*, has been demonstrated to facilitate the return to vegetative development after a flowering episode [119]. Another mechanism previously, proposed in Section 3.2, to ensure the return to vegetative development, was through inflorescence reversion, especially after a poor flowering season due to insufficient vernalization. Prolonged vernalization, on the other hand, enables the inflorescence to accumulated sufficient floral promotive signals to ensure flowering of inflorescence branches, however once a threshold is breached this can lead to flowering of unintended branches as in the case of 24 weeks vernalized Pajares plants. The plant therefore needs to put certain mechanisms in place to ensure, in the advent of extreme winters, some branches are still kept vegetative. The fact that in ILs flowering axillary shoots were much elevated after prolonged vernalization and floral reversion phenotypes were suppressed together with presence of inflorescence vegetative branches, even after minimum vernalization, demonstrates that, activities of *EOP* might be part of these mechanisms require by *A. alpina* Pajares to ensure its continuous supply of vegetative shoots. Thus, *EOP* contributes to the polycarpy life strategy of *A. alpina* Pajares by preventing flowering of some axillary shoots after extended vernalization and repressing flowering in some inflorescence branches in the advent of insufficient vernalization.

3.12 Probable mechanisms mediating early floral transition in *eop002*

EOP is a member of the AAA+ family that localizes in the plasma membrane. Analysis of DEGs in *eop002* compared to *pep1-1* and the comparison of shared DEGs in *eop002* and *eop101*, both containing mutation in *EOP*, suggested *EOP*'s involvement in stress response, reproductive development and transport of lipids and oligopeptides. As previously mentioned, the members of AAA+ family are involved in diverse activities including transport and stress response [206, 268] and plasma member genes are known to be involved in transport of compounds and signally response to biotic and abiotic stress [262]. These therefore support a possible function of *EOP* in stress response and transport. Stress response here mainly included response to salt stress, oxidative stress, wounding and defense. There was also overrepresentation of genes involved in response to water deprivation and cold. An ultimate adaption of plants to stress response is stress-induced flowering. This is because, under severe stress plants have higher chances of survival if they flower and produce seeds than as individual plants [269]. In recent years flowering time has been reported to be affected in several plant species upon application of different types of stress, and both early and late flowering have been observed with stress induction [270-273]. When faced with stress plants have the options to either maintain vegetative growth while ignoring the stress which will eventually cause late flowering, induce mechanisms that allows them to tolerate the stress which might also cause late flowering or to escape by flowering as quickly as possible to set seeds. Water deprivation is one of the stress factors known to greatly affect plant productivity. To cope with drought and ensure survival, plants result to these three main strategies mentioned above. In the escape strategy, slowly-developing water deficit leads to an acceleration of flowering before the onset of severe water loss. The strategy of drought avoidance includes reducing water loss through stomata closure and increasing water uptake by the roots, to prevent dehydration [274-276].

To play its role in maintaining vegetative development, *EOP* might be required for the activation of stress tolerance in the advent of stress and the prevention of an escape strategy. This is evident in the down regulation of several stress inducible genes in apices of *eop002*. As no stress conditions were specifically applied in this experiment, and both *eop* mutants and *pep1-1* were grown under the same conditions, it implies in the absence of *EOP,* plants become extra sensitive to environmental factors and always favor the escape strategy to

induce early flowering. A genetically based trade-off between drought avoidance and drought escape was previously demonstrated, whereby mutations that cause late flowering correlated with increases in drought avoidance [277]. Based on this, the early flowering caused by *eop* mutation implies a decrease in drought avoidance. However, stomatal morphogenesis genes *SDD1*, *TMM* and *ER*-family genes including *ERL1,* were up regulated in the apices of *eop002*. The overexpression of these genes have been shown to decrease stomatal density and transpiration, which means they increase drought avoidance [278, 279]. This suggests the plant employs mechanisms such as those to prevent the complete loss of water. *GI* is considered to be a key regulator in the drought escape strategy. The *gi* mutant showed a complete absence of drought escape response which resulted in a significant delay in flowering during restricted watering [280]. The up regulation of *GI* and other rhythmic process genes at an early developmental stage in *eop002* suggest there is the involvement of circadian processes in mediating early floral transition in *eop002*. It also implies *EOP* may be repressing the expression of these genes during vegetative development.

Excessive salinity is an example of a stress condition demonstrated to cause delays in flowering in *Arabidopsis* and other species [281-284]. Salt application extended the vegetative phase duration and delayed flowering in a DELLA-dependent manner, as a DELLA quadruple mutant flowered earlier than wild type under this condition [285]. High levels of salt exposure also increased the level of stress induced genes in both an ABA dependent and independent manner [286, 287]. The works of Achard *et al*., 2007, went on to demonstrate, treatment of *Arabidopsis* seedlings with salt caused an increase in the transcript levels of ethylene genes *ACS2* and *ACS7* as well as ethylene inducible genes *CHI-B* and *ERF4*. In addition flowering was delayed in plants grown in an ethylene-rich environment [285, 288]. They therefore concluded that salt increases the duration of vegetative phase, thus delaying flowering, through the activation of ABA and ethylene signaling, and the integration of their effects on DELLA functions. Therefore, in the *eop002* mutant, the down-regulation of genes involved in salt stress response, ABA biosynthesis and signaling, the DELLA protein encoding gene *GAI,* ethylene biosynthesis gene *ACS4* and ethylene inducible genes, *CHI-B* and *ERF8,* presents a possible mechanism by which the vegetative phase duration is reduced to enable an early switch to the reproductive phase.

Introgression lines with a functional *PEP1* and *eop* mutation showed an early saturation of vernalization requirement for flowering acceleration and flowered with shorter vernalization compared to Pajares (Section 2.3.8). Exposure of *A. thaliana* to short-term cold or overexpression of cold responsive genes delays flowering through the activation of *FLC* [289-291]. These short-term colds also resulted in the degradation of CO and suppression of *FT* [289]. *HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENE 1* (*HOS1*) for instance encodes an E3 ubiquitin ligase which acts as a component of photoperiodic flowering and normally involved in destabilizing CO in response to daylight cues. Under cold stress, HOS1 interferes with, and inhibits *FLC* repression while it promotes CO degradation to inhibit *FT* activation [289, 290]. Genes involved in response to temperature, more specifically cold inducible genes, were down-regulated in the apices of *eop002.* Additionally, there was the up regulation of some genes involved in circadian and photoperiodic processes. This implies, *EOP* might be involved in the up regulation of cold response genes expression, which could possibly include a prevention of *PEP1* repression during downstream processes, in order to repress flowering processes and maintain vegetative development after short-term colds. This also provides likely explanation for the suppression of floral reversion and vegetative phenotypes within the inflorescence of the *eop* introgression lines. Upon cold stress or exposure of plants to insufficient cold, such as temperature fluctuation during autumn, *EOP* activities may be required to stop the flowering processes which might have started in response to the cold, this means a reversion in the decision to flower due to insufficient vernalization.

4 Conclusions and perspectives

Survival of flowering plants greatly depends on the coordination of their flowering with appropriate seasons. Extended vernalization durations has been demonstrated in this study to influence flowering traits and inflorescence development of *A. alpina* accession Pajares, in terms of enabling rapid flower emergence and early inflorescence outgrowth after return of plants to LD conditions, thus permitting the plants to take advantage of short permissive growing seasons. Prolonged vernalization further ensures suppression of reversion phenotypes and increases silique production. The fact that these traits are affected by vernalization points out the involvement of flowering time, inflorescence identity and vernalization response genes. The understanding of genes involved in flowering and inflorescence development in *A. thaliana* has assisted in recent years, the identification of their orthologues in *A. alpina*. Although some genes possess altered or additional functions, which play important roles in maintaining the perennial life strategy of *A. alpina*, similarities still exists between some *A. thaliana* and *A. alpina* gene functions. The molecular mechanism of stable repression of *FLC* correlating with accelerated flowering in *A. thaliana* has now been demonstrated also for *PEP1* in *A. alpina*. There is therefore the possibility that molecular reasons for stable repression of *FLC*, such as the rate of H3K27me3 accumulation at the *FLC* locus, might be the cause of similar effects observed in Pajares after prolonged exposure to cold. Extended molecular works on these factors could shed more light on further regulatory mechanisms.

The flowering and inflorescence phenotype of previously isolated enhancer mutants of *pep1-1* (*eops*) have been described here in detail together with the inflorescence of *pep1-1* using leaf shape and inflorescence senescence as morphological markers. In one of the three experimental replicates conducted, *pep1-1* exhibited a reduction in its inflorescence height compared to other replicates. Since the first two experiments were conducted in growth cabinets and the third in a greenhouse chamber, it implies *pep1-1* might be responsive to changes in certain environmental factors that can influence its inflorescence development. Further phenotyping of the *pep1-1* mutant under different combinations of environmental conditions might lead to interesting findings regarding the responsiveness of the mutant to changes in environmental factors.

The *eop101* mutant has been characterized here to exhibit similar inflorescence phenotypes as *A. thaliana tfl1* mutants. Senescence in *eop101* significantly extends beyond inflorescence zone scored with leaf shape. Since in *A. thaliana*, *TFL1* has been associated with the senescence-signaling pathway, it is plausible that *AaTFL1* might as well play a role in the lifespan of *A. alpina*. Further studies in this direction could aid the understanding of senescence programming in perennials as a whole.

The method of mapping-by-sequencing and analysis of data using SHOREmap employed in this study permitted the identification of a region on the chromosome 8 of *A. alpina* harboring the causal mutation of four isolated *eop* mutants. Further marker assisted co-segregation analysis and complementation studies has identified the *EOP* gene as a member of an AAA+ ATPase family of proteins. In the *eop002* mutant a second region was present after SHOREmap analysis on chromosome 2. As discussed previously this could be due to unfiltered background mutations, thus a closer look at background materials and reference sequence for filtering could shed more light on the cause on this event.

As the transition from vegetative to reproductive growth is of great importance to flowering plants, they have developed complex mechanisms to regulate the timing of this developmental process. *EOP* has been demonstrated here to be part of likely mechanisms adopted by *A. alpina* to prevent flowering in response to insufficient cold. *EOP* contributes to the inhibition of axillary meristem outgrowth within certain leaf axils and to the continuous supply of vegetative axillary branches thereby supporting the perennial life strategy of *A. alpina* Pajares. Extensive genetic and molecular studies have aided the identification of several genes and regulatory networks affecting the timing of floral transition. Here in this study, analysis of RNAseq data from *eop002* and *eop101* have provided possible molecular mechanisms by which the early floral transition is achieved in these mutants. Further experiments are however required to confirm whether the observed changes in expressions are directly or indirectly regulated by these mutations. The fact that there is a possibility for *EOP* to be involved in the prevention of *PEP1* repression also presents an interesting point of focus for future works.

5 Materials and Method

Standard molecular biology techniques such as Polymerase Chain Reaction (PCR) and agarose gel electrophoresis were conducted as described by Sambrook and Russell [292], on less stated otherwise.

5.1 Plants materials and growth conditions

The *A. alpina* accession Pajares was previously described by Wang et al., 2009 as an accession originally collected from the Coedillera Cantábrica mountains in Spain and *pep1-1* was also previously isolated and characterized as an EMS mutant in the Pajares background. Seeds for the enhancer mutants of *pep1-1*, derived through EMS mutagenesis, seeds of *eop002*, *eop085* and *eop091,* backcrossed F2 mapping populations (BC1F2) generated by crossing to *pep1-1* and backcrossed F1 individual (BC1F1) seeds for *eop002*, *eop091* and *eop101* crossed to Pajares were kindly provided by Prof. Dr. Maria Albani (University of Cologne, Cologne, Germany). For the extended vernalization experiment, Pajares plants and introgression lines were grown for eight weeks under LD conditions (16hr light, 20°C; 8hr dark, 18°C) in the greenhouse. Plants were vernalized for different durations at 4°C in SD conditions (8hr light, 16hr dark) and synchronized to enable the return of all plants to greenhouse LD conditions at the same time. For the phenotyping experiments of *pep1-1* and the *eop* mutants, plants in experiment 1 and 2 were grown in a Percival growth chambers under LD conditions (16hr light, 20°C; 8hr dark, 20°C) while plants in experiment 3 were grown in the greenhouse with LD conditions (16 hr light, 20°C; 8 hr dark, 18°C). Plants for dominance and allelism testing were also grown in greenhouse LD conditions.

5.2 Analysis of flowering time and inflorescence development

Flowering time was recorded as the number of days to the first opened flower. In *pep1-1* and *eops* the number of leaves at flowering where also scored at first opened flower and verified at the end of flowering, whereas in Pajares plants exposed to different vernalization durations, the total number of leaves where scored at the end of flowering. Inflorescence phenotypes where scored in Pajares and introgression lines at eight and/or 14 weeks after vernalization,

unless stated otherwise. Inflorescence phenotypes in *pep1-1* and *eops* were scored at end of flowering and/or at senescence of the inflorescence. All experiments were performed with a range of 9 to 12 plants.

5.3 Mapping by sequencing and SHOREmap analysis

DNA was extracted from a pool of flower bud materials collected from 96 F2 individuals for *eop002* and *eop085*, and 136 for *eop091* using DNeasy Plant kits (Qiagen, Hilden-Germany) according to manufacturer's recommendations. The DNA pools and DNA material from *pep1-1* were sequenced with the Illumina Hiseq2000 which yielded 138.04 million reads for *eop002*, 30.15 million for *eop085*, 161.84 million for *eop091* and 264.15 million reads for *pep1-1*. DNA pooling and sequencing of *eop002* were conducted by Prof. Dr. Maria Albani. Data from *eop002* backcross mapping population was previously analyzed by Dr. Hequan Sun (AG. Schneeberger, MPIPZ Cologne) with *A. alpina* reference genome version 3. The *eop002* data was reanalyzed here together with data from *eop085* and *eop091* using the *A. alpina* reference genome version 5. Bowtie2 [167] was used to perform paired-end read alignment to the *A. alpina* reference genome. SAMtools [293] was used to predict mutations by calling out the differences between the reference sequences (from the wild type Pajares) and *eop002*, *eop085*, *eop091* and *pep1-1,* independently. Subsequent analysis from this point were carried out using SHOREmap [158]. EMS usually causes mutations by converting G to A or C to T and therefore SNPs detected were filtered to consider only these type of nucleotide changes. SNPs with their origin in *pep1-1* were filtered out from SNPs detected in the *eops* to determine mutations specific to the *eop* mutants. To enable visualization of these mutations on the chromosomes, SHOREmap was used to generate a plot of the allele frequency (AF) against the mutant loci on each chromosome. The AF is calculated as the number of reads supporting the mutant allele divided by the total number of reads at that particular locus. The SHOREmap annotation function was used to predict the effects of candidate mutations on the original genes. The gene descriptions for the *A. thaliana* orthologues of candidate genes were downloaded from, "The Arabidopsis Information Resource" (TAIR, www.arabidopsis.org) [184].

5.4 Marker development and analysis for validation of SNPs

SNPs derived from SHOREmap analysis were converted into CAPs/dCAPs markers. PCR primers for CAPS/dCAPS analysis were designed using dCAPS finder 2.0 [294]. These markers were used for co-segregation analysis and introgression of candidate region in a *PEP1* background. Genomic DNA for genotyping were extracted from leaves using the BioSprint 96 kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations. A standard PCR protocol was used to amplify products from *eop* mutants, *pep1-1*, Pajares and introgression lines. Restriction digest on the PCR products were performed according to manufacturer's guidelines then run on agarose gels (concentrations depended on the fragment sizes) for genotyping. Restriction enzymes used were obtained from New England Biolabs (NEB, Frankfurt). The primers and restriction enzymes that were used are provided in Supplementary Table 7. Sequencing of *AaG106560* (*EOP*) was conducted using primers stated in Table 11 below. *AaG106560* gene sequence is provided in Supplementary Dataset 2.

Table 11. AaG106560 sequencing primers

Primer ID	Primer name	Primer sequence
A013	Aa_G106560_1F	GTCACCAAATGAAAACCTTCC
A014	Aa_G106560_1R	CCGGTCAAGATCTGCTTTAAC
A015	Aa_G106560_2F	TGACCGTGGTCACAAGAAAG
A016	Aa_G106560_2R	CTGTCTTCTAAGGGAGAAGCTG
A017	Aa_G106560_3F	TCGATTGTGCAGTGGAGTTG
A018	Aa_G106560_3R	CAGCATCCTCGCTCTTCATC
A019	Aa_G106560_4F	TTATGGTCAAGTTGCGGAGAC
A020	Aa_G106560_4R	GGGTCTTCAATGGTTCACAAC

a. F: Forward primer
b. R: Reverse primer

5.5 Statistical analysis

The R software was used for the testing of significant differences. To check for normality of data distribution the Shapiro-Wilk test was employed. For data that did not follow a normal distribution the Kruskal–Wallis test was used as an omnibus test to detect significant differences followed by a posthoc test for pairwise multiple comparisons using the Mann–Whitney U test. The Type I error rate (α) was set at 0.05 and the Bonferroni p-value adjustments method was used.

5.6 Gene expression analysis

For quantitative analysis of *AaG106560* (*EOP*), total RNA was extracted from plant tissues using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), followed by DNase treatment with the Ambion DNAfree-kit DNase treatment (Invitrogen, Karlsruhe, Germany), all procedures were according to manufacturer's instructions. Reverse transcription for cDNA synthesis was performed with SuperScript II Reverse Transcriptase (Invitrogen, Karlsruhe, Germany) using a total RNA of 2ug and oligo dT (18) as primer. The cDNAs were diluted at 1:7.5 to serve as template for RT-qPCR and quantification was performed with an iQ SYBR Green Supermix detection system and the CFX96 Real-Time System all from Bio-Rad. *AaPP2A* was used for the normalization of expression data.

RT-qPCR mix:

Components	Volume (1x µl)	Final concentration
2X SYBR Green Supermix	6	1x
10 µM Forward primer	0.3	0.25 µM
10 µM Reverse primer	0.3	0.25 µM
cDNA	2	
Water	3.4	
Total volume	12	

Thermal cycling conditions:

Steps	Temperature	Time	Cycle
Polymerase Activation and DNA Denaturation	95	3 min	1
Denaturation	95	15 sec	35
Annealing /Extension	60	45 sec	
Melting curve analysis	According to instrument default setting		

Primers used are provided below.

Gene	Primer ID	Primer sequence
AaG106560	A106	F: TCACTATGGAAGCTCAATTTGC
	A107	R: GGCCGTTGTTGAAACTATTG
AaPP2A	A053	F: CAATGATTACGGCCTTGGAC
	A054	R: TTGAGTATGCGGACCATCG

5.7 Plasmids construction

To construct GFP-*AaG106560* for subcellular localization studies, AaG106560 cDNA was amplified and introduced into the Gateway cloning vector pDONR201 (Invitrogen, Karlsruhe, Germany) with BP reaction following manufacture's instruction of the BP clonase (Invitrogen, Karlsruhe, Germany). LR reaction was conducted with LR clonase (Invitrogen, Karlsruhe, Germany), following manufacture's guidelines, to transfer the cloned fragments from the entry vector into the pGWB406 destination vector which contains a cauliflower mosaic virus (CaMV) 35S promoter and GFP [295]. The construct was then introduced into the *Agrobacterium tumefaciens* strain GV3101pmp90RK (antibiotics: rifampicin, ampicillin, gentamicin and kanamycin).

For overexpression analysis in *pep1-1*, AaG106560 genomic DNA (from ATG to 923bp downstream of 3'UTR) was amplified and introduced into the Gateway cloning vector pDONR201. The entry vector was transferred by LR reaction into the destination binary vector pLEELA which contains a double enhancer cauliflower mosaic virus 35S promoter and a GATEWAY cassette. The construct was then introduced into the *Agrobacterium tumefaciens strain* GV3101pmp90RK.

For *eops* complementation studies, a region of 3,654 bp upstream of the 5' UTR considered to contain the native promoter was introduced into the entry vector already containing AaG106560 genomic DNA (from ATG to 923bp downstream of 3'UTR) through Polymerase Incomplete Primer Extension (PIPE) cloning method [296], with the kind help of Dr. Miriam Giesguth. This was followed by introduction of the entry vector into the destination vector pEarlyGate301 via LR reactions. The construct was further introduced into the Agrobacterium tumefaciens strain GV3101pSoup (antibiotics: rifampicin, tetracycline, gentamicin and kanamycin). Presence of plasmids during cloning were examined by colony PCR and the constructs were also verified by sequencing. Primers used for plasmid constructions are listed in Supplementary Table 8.

5.8 Transformation of Agrobacterium tumefaciens

The desired constructs within the destination vectors were introduced into *Agrobacterium tumefaciens* via electroporation. To achieve this, a mixture of 250ng plasmid DNA was added to a 50µl aliquot of Agrobacterium cells. The mix was then exposed to an electric pulse of 2.20kV followed by immediate addition of 600 µl of LB medium. The mix was incubated for 2 hours at 28°C with shaking, followed by plating on LB agar plates with appropriate anti-biotics which were then incubated at 28°C for 72 hours.

5.9 Infiltration of *Nicotiana benthamiana* and generation of transgenic plants

Selected single colonies from the plate were further incubated at 28°C for 24 hours in 5ml LB containing appropriate antibiotics to be used for Nicotiana plants infiltration. The culture media was centrifuged (400rpm, 5min at room temperature) and resulting pellet was re-suspended in 5ml of infiltration buffer (10mM MgCl2, 10mM MES-K (pH 5.6) and 150µM acetosyringone). The Agrobacterium culture with a final DO600 of 0.3 was infiltrated into the lower part of young Nicotiana benthamiana leaves with a syringe together with the gene silencing inhibitor P19 [297]. The Nicotiana benthamiana leaves were also infiltrated with a mCheery-labelled plasma membrane marker, CD3-1007 [298] to enable verification of localization.

Transformation of *pep1-1* and *eop* mutants, was carried out by first selecting single colonies of Agrobacterium carrying the desired constructs, these were grown overnight at 28°C in 1L LB medium containing appropriate antibiotics. Once at an OD600 range of 0.8 to 1.0, cells were harvested by centrifugation (6000rpm; 10 mins). Obtained pellets after centrifugation were resuspended in 1L transformation buffer (50g sucrose, 500µl silvet-L77, pH5.7). Transgenic plants were generated through the floral dipping method, whereby flower buds were dipped into the transformation buffer for two minutes. These were then covered with plastic trays and incubated for approximately 24 hours. Plants were subsequently grown in LD until seeds were ready for collection. T1 plants carrying the plasmid were selected on soil based on their resistance to BASTA (Bayer, Leverkusen, Germany).

5.10 Subcellular localization of Aa_G106560

Full length cDNA of *Aa_G106560* was amplified and introduced into the GATEWAY pDONR201 vector. Absence of PCR induced mutations in the constructs was confirmed by DNA sequencing. The construct from the pDONR201 vector generated was transferred by LR recombination to the destination vector pGWB406. pGWB406 is a binary vector that carries a cauliflower mosaic virus (CaMV) 35S promoter and an sGFP fluorescent reporter at the N-terminus [295]. The resulting construct, pGWB406:Aa_G106560, placed *Aa_G106560* downstream of the sGFP. Plasmid DNA from the construct was then introduced into *Agrobacterium tumefaciens* strain GV3101pmp90RK containing the helper plasmid pTiC58DT by electroporation. Plants were incubated for 48 hr at room temperature after which fluorescence was observed by laser scanning confocal microscopy (LSM780; Carl Zeiss).

5.11 RNAseq: Sample preparation and RNA extraction for illumina sequencing

Isolation of mRNA was performed with leaves and apices of 3 weeks old *eop002*, *eop101* and *pep1-1* seedlings at which stage no apparent morphological differences could be observed. Four biological replicates were used. Library preparation and sequencing were performed by the MPIPZ genome center, Cologne. Illumina Hiseq2500 sequencing platform was employed for single-end sequencing. Combining replicates, *eop002* leaves yielded 56 million reads, *eop101* and *pep1-1* yielded 55.9 and 54.2 million reads respectively. The apices samples produced, 58.7, 53.8 and 61 million reads for *eop002, eop101* and *pep1-1* respectively. Read alignment to *A. alpina* reference genome and the report of differentially expressed genes (DEGs) were carried out using TopHat and Cufflinks [168]. The overall read alignment rate was 95% for leaf samples and 96% for apices. At a significant thresh-hold of p-value < 0.05, the up or down regulation of the DEGs were identified by comparing their expression levels in *eop002* and *eop101* with that in *pep1-1*, which served as the control. For GO analysis of DEGs were performed using the VirtualPlant software platform [172].

6 References

1. Weberling, F. and R.J. Pankhurst, *Morphology of Flowers and Inflorescences*. 1992: Cambridge University Press.
2. Coen, E.S. and J.M. Nugent, *Evolution of flowers and inflorescences.* Development, 1994. **1994**(Supplement): p. 107-116.
3. Pouteau, S. and C. Albertini, *The significance of bolting and floral transitions as indicators of reproductive phase change in Arabidopsis.* Journal of Experimental Botany, 2009. **60**(12): p. 3367-3377.
4. Bradley, D., et al., *Inflorescence Commitment and Architecture in Arabidopsis.* Science, 1997. **275**(5296): p. 80-83.
5. Alvarez, J., et al., *terminal flower: a gene affecting inflorescence development in Arabidopsis thaliana.* The Plant Journal, 1992. **2**(1): p. 103-116.
6. Shannon, S. and D.R. Meeks-Wagner, *A mutation in the Arabidopsis TFL1 gene affects inflorescence meristem development.* The Plant Cell, 1991. **3**(9): p. 877-892.
7. Ratcliffe, O.J., D.J. Bradley, and E.S. Coen, *Separation of shoot and floral identity in Arabidopsis.* Development, 1999. **126**(6): p. 1109-1120.
8. Martínez-Zapater, J.M., et al., *Arabidopsis late-flowering fve mutants are affected in both vegetative and reproductive development.* The Plant Journal, 1995. **7**(4): p. 543-551.
9. Kirchoff, B.K. and R. Claßen-Bockhoff, *Inflorescences: concepts, function, development and evolution.* Annals of Botany, 2013. **112**(8): p. 1471-1476.
10. Harder, L.D. and P. Prusinkiewicz, *The interplay between inflorescence development and function as the crucible of architectural diversity.* Annals of Botany, 2013. **112**(8): p. 1477-1493.
11. Benlloch, R., et al., *Floral Initiation and Inflorescence Architecture: A Comparative View.* Annals of Botany, 2007. **100**(3): p. 659-676.
12. Prenner, G., F. Vergara-Silva, and P.J. Rudall, *The key role of morphology in modelling inflorescence architecture.* Trends in Plant Science, 2009. **14**(6): p. 302-309.
13. Shannon, S. and D.R. Meeks-Wagner, *Genetic Interactions That Regulate Inflorescence Development in Arabidopsis.* The Plant Cell, 1993. **5**(6): p. 639-655.
14. Gong, D., et al., *Constitutive activation and transgenic evaluation of the function of an Arabidopsis PKS protein kinase.* Journal of Biological Chemistry, 2002. **277**(44): p. 42088-42096.
15. Pnueli, L., et al., *The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1.* Development, 1998. **125**(11): p. 1979-1989.
16. Schmitz, G. and K. Theres, *Genetic control of branching in Arabidopsis and tomato.* Current Opinion in Plant Biology, 1999. **2**(1): p. 51-55.
17. Lippman, Z.B., et al., *The Making of a Compound Inflorescence in Tomato and Related Nightshades.* PLOS Biology, 2008. **6**(11): p. e288.
18. Corbesier, L., et al., *FT Protein Movement Contributes to Long-Distance Signaling in Floral Induction of Arabidopsis.* Science, 2007.
19. Adrian, J., S. Torti, and F. Turck, *From decision to commitment: the molecular memory of flowering.* Molecular plant, 2009. **2**(4): p. 628-642.
20. Bernier, G., *The control of floral evocation and morphogenesis.* Annual Review of Plant Physiology and Plant Molecular Biology, 1988. **39**(1): p. 175-219.
21. McDaniel, C.N., S.R. Singer, and S.M.E. Smith, *Developmental states associated with the floral transition.* Developmental Biology, 1992. **153**(1): p. 59-69.
22. Asbe, A., et al., *Floral Reversion in Arabidopsis suecica Is Correlated with the Onset of Flowering and Meristem Transitioning.* PLoS ONE, 2015. **10**(5): p. e0127897.
23. Muller-Xing, R., D. Schubert, and J. Goodrich, *Non-inductive conditions expose the cryptic bract of flower phytomeres in Arabidopsis thaliana.* Plant Signal Behav, 2015. **10**(4): p. e1010868.
24. Fornara, F., A. de Montaigu, and G. Coupland, *SnapShot: control of flowering in Arabidopsis.* Cell, 2010. **141**(3): p. 550-550. e2.
25. Gordon G. Simpson, a. Anthony R. Gendall, and C. Dean, *When to Switch to Flowering.* Annual Review of Cell and Developmental Biology, 1999. **15**(1): p. 519-550.
26. Levy, Y.Y. and C. Dean, *The transition to flowering.* The Plant Cell, 1998. **10**(12): p. 1973-1989.
27. Maarten Koornneef, et al., *GENETIC CONTROL OF FLOWERING TIME IN ARABIDOPSIS.* Annual Review of Plant Physiology and Plant Molecular Biology, 1998. **49**(1): p. 345-370.
28. Poethig, R.S., *Phase Change and the Regulation of Shoot Morphogenesis in Plants.* Science, 1990. **250**(4983): p. 923-930.
29. Lawson, E.J.R. and R.S. Poethig, *Shoot development in plants: time for a change.* Trends in Genetics, 1995. **11**(7): p. 263-268.
30. Poethig, R.S., *Phase Change and the Regulation of Developmental Timing in Plants.* Science, 2003. **301**(5631): p. 334-336.

31. Klein, J., H. Saedler, and P. Huijser, *A new family of DNA binding proteins includes putative transcriptional regulators of theAntirrhinum majus floral meristem identity geneSQUAMOSA.* Molecular and General Genetics MGG, 1996. **250**(1): p. 7-16.
32. Cardon, G., et al., *Molecular characterisation of the Arabidopsis SBP-box genes.* Gene, 1999. **237**(1): p. 91-104.
33. Cardon, G.H., et al., *Functional analysis of the Arabidopsis thaliana SBP-box gene SPL3: a novel gene involved in the floral transition.* The Plant Journal, 1997. **12**(2): p. 367-377.
34. Hyun, Y., R. Richter, and G. Coupland, *Competence to Flower: Age-Controlled Sensitivity to Environmental Cues.* Plant Physiology, 2017. **173**(1): p. 36-46.
35. Guo, A.-Y., et al., *Genome-wide identification and evolutionary analysis of the plant specific SBP-box transcription factor family.* Gene, 2008. **418**(1): p. 1-8.
36. Wu, G. and R.S. Poethig, *Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3.* Development, 2006. **133**(18): p. 3539-3547.
37. Aukerman, M.J. and H. Sakai, *Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes.* The Plant Cell, 2003. **15**(11): p. 2730-2741.
38. Hyun, Y., et al., *Multi-layered regulation of SPL15 and cooperation with SOC1 integrate endogenous flowering pathways at the Arabidopsis shoot meristem.* Developmental cell, 2016. **37**(3): p. 254-266.
39. Wu, G., et al., *The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis.* Cell, 2009. **138**(4): p. 750-759.
40. Xu, M., et al., *Developmental functions of miR156-regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) genes in Arabidopsis thaliana.* PLoS genetics, 2016. **12**(8): p. e1006263.
41. Wilson, R.N., J.W. Heckman, and C.R. Somerville, *Gibberellin is required for flowering in Arabidopsis thaliana under short days.* Plant Physiology, 1992. **100**(1): p. 403-408.
42. Blázquez, M.A., et al., *Gibberellins promote flowering of Arabidopsis by activating the LEAFY promoter.* The Plant Cell, 1998. **10**(5): p. 791-800.
43. Langridge, J., *Effect of day-length and gibberellic acid on the flowering of Arabidopsis.* Nature, 1957. **180**(4575): p. 36-37.
44. Sun, T.P. and Y. Kamiya, *The Arabidopsis GA1 locus encodes the cyclase ent-kaurene synthetase A of gibberellin biosynthesis.* The Plant Cell, 1994. **6**(10): p. 1509-1518.
45. Koornneef, M. and J.H. van der Veen, *Induction and analysis of gibberellin sensitive mutants in Arabidopsis thaliana (L.) heynh.* Theoretical and Applied Genetics, 1980. **58**(6): p. 257-263.
46. Koornneef, M., et al., *Genetic fine-structure of the GA-1 locus in the higher plant Arabidopsis thaliana (L.) Heynh.* Genetics Research, 1983. **41**(1): p. 57-68.
47. Davière, J.-M. and P. Achard, *Gibberellin signaling in plants.* Development, 2013. **140**(6): p. 1147-1151.
48. Peng, J., et al., *The Arabidopsis GAI gene defines a signaling pathway that negatively regulates gibberellin responses.* Genes & development, 1997. **11**(23): p. 3194-3205.
49. Silverstone, A.L., C.N. Ciampaglio, and T.-p. Sun, *The Arabidopsis RGA gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway.* The Plant Cell, 1998. **10**(2): p. 155-169.
50. Achard, P. and P. Genschik, *Releasing the brakes of plant growth: how GAs shutdown DELLA proteins.* Journal of experimental botany, 2009. **60**(4): p. 1085-1092.
51. King, K.E., T. Moritz, and N.P. Harberd, *Gibberellins are not required for normal stem growth in Arabidopsis thaliana in the absence of GAI and RGA.* Genetics, 2001. **159**(2): p. 767-776.
52. Dill, A. and T.-p. Sun, *Synergistic derepression of gibberellin signaling by removing RGA and GAI function in Arabidopsis thaliana.* Genetics, 2001. **159**(2): p. 777-785.
53. Wen, C.-K. and C. Chang, *Arabidopsis RGL1 Encodes a Negative Regulator of Gibberellin Responses.* The Plant Cell, 2002. **14**(1): p. 87-100.
54. Thomas, B. and D. Vince-Prue, *3 - Photoperiodic Photoreceptors*, in *Photoperiodism in Plants (Second Edition)*. 1997, Academic Press: London. p. 63-84.
55. Somers, D.E., P.F. Devlin, and S.A. Kay, *Phytochromes and Cryptochromes in the Entrainment of the Arabidopsis Circadian Clock.* Science, 1998. **282**(5393): p. 1488-1490.
56. Johansson, M. and D. Staiger, *Time to flower: interplay between photoperiod and the circadian clock.* Journal of Experimental Botany, 2015. **66**(3): p. 719-730.
57. Salomé, P.A. and C.R. McClung, *The Arabidopsis thaliana Clock.* Journal of Biological Rhythms, 2004. **19**(5): p. 425-435.
58. Doherty, C.J. and S.A. Kay, *Circadian Control of Global Gene Expression Patterns.* Annual review of genetics, 2010. **44**: p. 419-444.
59. Strayer, C.A. and S.A. Kay, *The ins and outs of circadian regulated gene expression.* Current Opinion in Plant Biology, 1999. **2**(2): p. 114-120.
60. Schaffer, R., et al., *The late elongated hypocotyl Mutation of Arabidopsis Disrupts Circadian Rhythms and the Photoperiodic Control of Flowering.* Cell. **93**(7): p. 1219-1229.
61. Wang, Z.-Y. and E.M. Tobin, *Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression.* Cell, 1998. **93**(7): p. 1207-1217.
62. Alabadí, D., et al., *Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock.* Science, 2001. **293**(5531): p. 880-883.
63. Doyle, M.R., et al., *The ELF4 gene controls circadian rhythms and flowering time in Arabidopsis thaliana.* Nature, 2002. **419**(6902): p. 74-77.

64. Mizoguchi, T., et al., *Distinct roles of GIGANTEA in promoting flowering and regulating circadian rhythms in Arabidopsis.* The Plant Cell, 2005. **17**(8): p. 2255-2270.
65. Yano, M., et al., *Genetic Control of Flowering Time in Rice, a Short-Day Plant.* Plant Physiology, 2001. **127**(4): p. 1425-1429.
66. Cao, K., et al., *Four Tomato FLOWERING LOCUS T-Like Proteins Act Antagonistically to Regulate Floral Initiation.* Frontiers in Plant Science, 2015. **6**: p. 1213.
67. Chouard, P., *Vernalization and its Relations to Dormancy.* Annual Review of Plant Physiology, 1960. **11**(1): p. 191-238.
68. Purvis, O., *The physiological analysis of vernalization.* Encyclopedia of plant physiology, 1961. **16**: p. 76-122.
69. Searle, I., et al., *The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis.* Genes & development, 2006. **20**(7): p. 898-912.
70. Hepworth, S.R., et al., *Antagonistic regulation of flowering-time gene SOC1 by CONSTANS and FLC via separate promoter motifs.* The EMBO Journal, 2002. **21**(16): p. 4327-4337.
71. Michaels, S.D., et al., *Integration of Flowering Signals in Winter-Annual Arabidopsis.* Plant Physiology, 2005. **137**(1): p. 149-156.
72. Johanson, U., et al., *Molecular Analysis of FRIGIDA, a Major Determinant of Natural Variation in Arabidopsis Flowering Time.* Science, 2000. **290**(5490): p. 344-347.
73. Choi, K., et al., *The FRIGIDA Complex Activates Transcription of FLC, a Strong Flowering Repressor in Arabidopsis, by Recruiting Chromatin Modification Factors.* The Plant Cell, 2011. **23**(1): p. 289-303.
74. Michaels, S.D. and R.M. Amasino, *FLOWERING LOCUS C Encodes a Novel MADS Domain Protein That Acts as a Repressor of Flowering.* The Plant Cell, 1999. **11**(5): p. 949-956.
75. Sheldon, C.C., et al., *The molecular basis of vernalization: The central role of FLOWERING LOCUS C (FLC).* Proceedings of the National Academy of Sciences, 2000. **97**(7): p. 3753-3758.
76. Bastow, R., et al., *Vernalization requires epigenetic silencing of FLC by histone methylation.* Nature, 2004. **427**: p. 164.
77. Sung, S., R.J. Schmitz, and R.M. Amasino, *A PHD finger protein involved in both the vernalization and photoperiod pathways in Arabidopsis.* Genes & development, 2006. **20**(23): p. 3244-3248.
78. Finnegan, E.J. and E.S. Dennis, *Vernalization-Induced Trimethylation of Histone H3 Lysine 27 at FLC Is Not Maintained in Mitotically Quiescent Cells.* Current Biology, 2007. **17**(22): p. 1978-1983.
79. Chandler, J., A. Wilson, and C. Dean, *Arabidopsis mutants showing an altered response to vernalization.* The Plant Journal, 1996. **10**(4): p. 637-644.
80. Levy, Y.Y., et al., *Multiple Roles of Arabidopsis VRN1 in Vernalization and Flowering Time Control.* Science, 2002. **297**(5579): p. 243-246.
81. Koornneef, M., C. Hanhart, and J. Van der Veen, *A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana.* Molecular and General Genetics MGG, 1991. **229**(1): p. 57-66.
82. Michaels, S.D. and R.M. Amasino, *Loss of FLOWERING LOCUS C Activity Eliminates the Late-Flowering Phenotype of FRIGIDA and Autonomous Pathway Mutations but Not Responsiveness to Vernalization.* The Plant Cell, 2001. **13**(4): p. 935-941.
83. Lee, I., et al., *Isolation of LUMINIDEPENDENS: a gene involved in the control of flowering time in Arabidopsis.* Plant cell, 1994. **6**(1): p. 75-84.
84. Sanda, S.L. and R.M. Amasino, *Ecotype-specific expression of a flowering mutant phenotype in Arabidopsis thaliana.* Plant Physiology, 1996. **111**(2): p. 641-644.
85. Lim, M.-H., et al., *A new Arabidopsis gene, FLK, encodes an RNA binding protein with K homology motifs and regulates flowering time via FLOWERING LOCUS C.* The Plant Cell, 2004. **16**(3): p. 731-740.
86. Koornneef, M., et al., *Genetic Interactions Among Late-Flowering Mutants of Arabidopsis.* Genetics, 1998. **148**(2): p. 885-892.
87. Koornneef, M., et al., *The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg erecta wild-type.* The Plant Journal, 1994. **6**(6): p. 911-919.
88. Lee, I., et al., *The late-flowering phenotype of FRIGIDA and mutations in LUMINIDEPENDENS is suppressed in the Landsberg erecta strain of Arabidopsis.* The Plant Journal, 1994. **6**(6): p. 903-909.
89. Piñeiro, M. and G. Coupland, *The Control of Flowering Time and Floral Identity in Arabidopsis.* Plant Physiology, 1998. **117**(1): p. 1-8.
90. Boss, P.K., et al., *Multiple Pathways in the Decision to Flower: Enabling, Promoting, and Resetting.* The Plant Cell, 2004. **16**(suppl 1): p. S18-S31.
91. Samach, A., et al., *Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis.* Science, 2000. **288**(5471): p. 1613-1616.
92. Abe, M., et al., *FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex.* Science, 2005. **309**(5737): p. 1052-1056.
93. Sheldon, C.C., et al., *The FLF MADS Box Gene: A Repressor of Flowering in Arabidopsis Regulated by Vernalization and Methylation.* The Plant Cell, 1999. **11**(3): p. 445-458.
94. Helliwell, C.A., et al., *The Arabidopsis FLC protein interacts directly in vivo with SOC1 and FT chromatin and is part of a high-molecular-weight protein complex.* The Plant Journal, 2006. **46**(2): p. 183-192.
95. Kardailsky, I., et al., *Activation tagging of the floral inducer FT.* Science, 1999. **286**(5446): p. 1962-1965.
96. Moon, J., et al., *The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis.* The Plant Journal, 2003. **35**(5): p. 613-623.

97. Blázquez, M.A. and D. Weigel, *Integration of floral inductive signals in Arabidopsis.* Nature, 2000. **404**(6780): p. 889-892.
98. Weigel, D., et al., *LEAFY controls floral meristem identity in Arabidopsis.* Cell, 1992. **69**(5): p. 843-859.
99. Parcy, F., et al., *A genetic framework for floral patterning.* Nature, 1998. **395**(6702): p. 561-566.
100. Okamuro, J.K., et al., *Flowers into shoots: photo and hormonal control of a meristem identity switch in Arabidopsis.* Proceedings of the National Academy of Sciences, 1996. **93**(24): p. 13831-13836.
101. Schultz, E.A. and G.W. Haughn, *LEAFY, a homeotic gene that regulates inflorescence development in Arabidopsis.* The Plant Cell, 1991. **3**(8): p. 771-781.
102. Irish, V.F. and I.M. Sussex, *Function of the apetala-1 gene during Arabidopsis floral development.* The Plant Cell, 1990. **2**(8): p. 741-753.
103. Bowman, J.L., et al., *Control of flower development in Arabidopsis thaliana by APETALA1 and interacting genes.* Development, 1993. **119**(3): p. 721-743.
104. Mandel, M.A. and M.F. Yanofsky, *A gene triggering flower formation in Arabidopsis.* Nature, 1995. **377**(6549): p. 522-524.
105. Hanano, S. and K. Goto, *Arabidopsis TERMINAL FLOWER1 is involved in the regulation of flowering time and inflorescence development through transcriptional repression.* Plant Cell, 2011. **23**(9): p. 3172-84.
106. Blazquez, M.A., et al., *LEAFY expression and flower initiation in Arabidopsis.* Development, 1997. **124**(19): p. 3835-3844.
107. Hempel, F.D., et al., *Floral determination and expression of floral regulatory genes in Arabidopsis.* Development, 1997. **124**(19): p. 3845-3853.
108. Wagner, D., R.W.M. Sablowski, and E.M. Meyerowitz, *Transcriptional Activation of APETALA1 by LEAFY.* Science, 1999. **285**(5427): p. 582-584.
109. Ratcliffe, O.J., et al., *A common mechanism controls the life cycle and architecture of plants.* Development, 1998. **125**(9): p. 1609-1615.
110. Poduska, B., et al., *The Synergistic Activation of FLOWERING LOCUS C by FRIGIDA and a New Flowering Gene AERIAL ROSETTE 1 Underlies a Novel Morphology in Arabidopsis.* Genetics, 2003. **163**(4): p. 1457-1465.
111. Friedman, J. and M.J. Rubin, *All in good time: understanding annual and perennial strategies in plants.* American journal of botany, 2015. **102**(4): p. 497-499.
112. Amasino, R., *Floral induction and monocarpic versus polycarpic life histories.* Genome biology, 2009. **10**(7): p. 228.
113. Bleecker, A.B. and S.E. Patterson, *Last exit: senescence, abscission, and meristem arrest in Arabidopsis.* The Plant Cell, 1997. **9**(7): p. 1169-1179.
114. Hensel, L.L., et al., *The fate of inflorescence meristems is controlled by developing fruits in Arabidopsis.* Plant Physiology, 1994. **106**(3): p. 863-876.
115. Ausin, I., C. Alonso-Blanco, and J.-M. Martinez-Zapater, *Environmental regulation of flowering.* International Journal of Developmental Biology, 2004. **49**(5-6): p. 689-705.
116. Michaels, S.D. and R.M. Amasino, *Memories of winter: vernalization and the competence to flower.* Plant, Cell & Environment, 2000. **23**(11): p. 1145-1153.
117. Lee, I., A. Bleecker, and R. Amasino, *Analysis of naturally occurring late flowering in Arabidopsis thaliana.* Molecular and General Genetics MGG, 1993. **237**(1): p. 171-176.
118. Hart, R., *Why are biennials so few?* The American Naturalist, 1977. **111**(980): p. 792-799.
119. Wang, R., et al., *PEP1 regulates perennial flowering in Arabis alpina.* Nature, 2009. **459**(7245): p. 423-7.
120. Noodén, L.D. and J.P. Penney, *Correlative controls of senescence and plant death in Arabidopsis thaliana (Brassicaceae).* Journal of Experimental Botany, 2001. **52**(364): p. 2151-2159.
121. Mohamed, R., et al., *Populus CEN/TFL1 regulates first onset of flowering, axillary meristem identity and dormancy release in Populus.* The Plant Journal, 2010. **62**(4): p. 674-688.
122. Tooke, F., et al., *Mechanisms and function of flower and inflorescence reversion.* Journal of Experimental Botany, 2005. **56**(420): p. 2587-2599.
123. Sreekantan, L., et al., *Cycles of floral and vegetative development in Metrosideros excelsa (Myrtaceae).* International Journal of Plant Sciences, 2001. **162**(4): p. 719-727.
124. Diomaiuto, J., *Periodic flowering or continual flowering as a function of temperature in a perennial species: the Ravenelle wallflower (Cheiranthus cheiri L.).* Phytomorphology, 1988.
125. Battey, N.H. and F. Tooke, *Molecular control and variation in the floral transition.* Current Opinion in Plant Biology, 2002. **5**(1): p. 62-68.
126. Battey, N.H., *Aspects of seasonality.* Journal of Experimental Botany, 2000. **51**(352): p. 1769-1780.
127. Janzen, D.H., *Why bamboos wait so long to flower.* Annual Review of Ecology and Systematics, 1976. **7**(1): p. 347-391.
128. Geber, M.A. and M.A. Watson, *Organ Preformation, Development, and Resource.* Plant resource allocation, 1997: p. 113.
129. Meloche, C.G. and P.K. Diggle, *Preformation, architectural complexity, and developmental flexibility in Acomastylis rossii (Rosaceae).* American Journal of Botany, 2001. **88**(6): p. 980-991.
130. Diggle, P., *Extreme preformation in alpine Polygonum viviparum: an architectural and developmental analysis.* American Journal of Botany, 1997. **84**(2): p. 154-154.
131. Bliss, L., *Adaptations of arctic and alpine plants to environmental conditions.* Arctic, 1962. **15**(2): p. 117-144.

132. Billings, W.D. and H.A. Mooney, *The ecology of arctic and alpine plants.* Biological reviews, 1968. **43**(4): p. 481-529.
133. Walton, G.B. and L. Hufford, *Shoot architecture and evolution of Dicentra cucullaria (Papaveraceae, Fumarioideae).* International Journal of Plant Sciences, 1994. **155**(5): p. 553-568.
134. Jones, C.S. and M.A. Watson, *Heteroblasty and preformation in mayapple, Podophyllum peltatum (Berberidaceae): developmental flexibility and morphological constraint.* American Journal of Botany, 2001. **88**(8): p. 1340-1358.
135. Hayes, P., T. Steeves, and B. Neal, *An architectural analysis of Shepherdia canadensis and Shepherdia argentea: patterns of shoot development.* Canadian journal of botany, 1989. **67**(6): p. 1870-1877.
136. Tränkner, C., et al., *Over-expression of an FT-homologous gene of apple induces early flowering in annual and perennial plants.* Planta, 2010. **232**(6): p. 1309-1324.
137. Rohde, A. and R.P. Bhalerao, *Plant dormancy in the perennial context.* Trends in plant science, 2007. **12**(5): p. 217-223.
138. Bergonzi, S. and M.C. Albani, *Reproductive competence from an annual and a perennial perspective.* Journal of experimental botany, 2011. **62**(13): p. 4415-4422.
139. Darnell, R.L., et al., *The physiology of flowering in strawberry.* Hortic. Rev, 2003. **28**: p. 325-349.
140. Wang, R., et al., *Aa TFL1 confers an age-dependent response to vernalization in perennial Arabis alpina.* Plant Cell, 2011. **23**(4): p. 1307-21.
141. Bergonzi, S., et al., *Mechanisms of Age-Dependent Response to Winter Temperature in Perennial Flowering of Arabis alpina.* Science, 2013. **340**(6136): p. 1094-1097.
142. Albani, M.C., et al., *PEP1 of Arabis alpina is encoded by two overlapping genes that contribute to natural genetic variation in perennial flowering.* PLoS genetics, 2012. **8**(12): p. e1003130.
143. Park, J.-Y., H. Kim, and I. Lee, *Comparative analysis of molecular and physiological traits between perennial Arabis alpina Pajares and annual Arabidopsis thaliana Sy-0.* Scientific Reports, 2017. **7**(1): p. 13348.
144. Koch, M.A., et al., *Three times out of Asia Minor: the phylogeography of Arabis alpina L. (Brassicaceae).* Molecular Ecology, 2006. **15**(3): p. 825-839.
145. Willing, E.-M., et al., *Genome expansion of Arabis alpina linked with retrotransposition and reduced symmetric DNA methylation.* Nature Plants, 2015. **1**: p. 14023.
146. Shindo, C., et al., *Role of FRIGIDA and FLOWERING LOCUS C in Determining Variation in Flowering Time of Arabidopsis.* Plant Physiology, 2005. **138**(2): p. 1163-1173.
147. Shindo, C., et al., *Variation in the epigenetic silencing of FLC contributes to natural variation in Arabidopsis vernalization response.* Genes & development, 2006. **20**(22): p. 3079-3083.
148. Murray, M.B., M.G.R. Cannell, and R.I. Smith, *Date of Budburst of Fifteen Tree Species in Britain Following Climatic Warming.* Journal of Applied Ecology, 1989. **26**(2): p. 693-700.
149. González, F.G., G.A. Slafer, and D.J. Miralles, *Vernalization and photoperiod responses in wheat pre-flowering reproductive phases.* Field Crops Research, 2002. **74**(2): p. 183-195.
150. Jones, H.G., et al., *An approach to the determination of winter chill requirements for different Ribes cultivars.* Plant Biology, 2013. **15**: p. 18-27.
151. Engelen-Eigles, G. and J.E. Erwin, *A model plant for vernalization studies.* Scientia Horticulturae, 1997. **70**(2): p. 197-202.
152. Guo, D.-P., et al., *The interaction of plant growth regulators and vernalization on the growth and flowering of cauliflower (Brassica oleracea var. botrytis).* Plant growth regulation, 2004. **43**(2): p. 163-171.
153. Cook, N.C. and G. Jacobs, *Suboptimal winter chilling impedes development of acrotony in apple shoots.* HortScience, 1999. **34**(7): p. 1213-1216.
154. Schneeberger, K., *Using next-generation sequencing to isolate mutant genes from forward genetic screens.* Nat Rev Genet, 2014. **15**(10): p. 662-676.
155. Michelmore, R.W., I. Paran, and R.V. Kesseli, *Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations.* Proceedings of the National Academy of Sciences, 1991. **88**(21): p. 9828-9832.
156. Schneeberger, K., et al., *SHOREmap: simultaneous mapping and mutation identification by deep sequencing.* Nat Methods, 2009. **6**(8): p. 550-1.
157. James, G.V., et al., *User guide for mapping-by-sequencing in Arabidopsis.* Genome Biol, 2013. **14**(6): p. R61.
158. Sun, H. and K. Schneeberger, *SHOREmap v3.0: fast and accurate identification of causal mutations from forward genetic screens.* Methods Mol Biol, 2015. **1284**: p. 381-95.
159. Pankin, A., et al., *Mapping-by-Sequencing Identifies HvPHYTOCHROME C as a Candidate Gene for the early maturity 5 Locus Modulating the Circadian Clock and Photoperiodic Flowering in Barley.* Genetics, 2014. **198**(1): p. 383-396.
160. Mascher, M., et al., *Mapping-by-sequencing accelerates forward genetics in barley.* Genome Biol, 2014. **15**(6): p. R78.
161. Hartwig, B., et al., *Fast isogenic mapping-by-sequencing of ethyl methanesulfonate-induced mutant bulks.* Plant Physiol, 2012. **160**(2): p. 591-600.
162. Abe, A., et al., *Genome sequencing reveals agronomically important loci in rice using MutMap.* 2012. **30**: p. 174.
163. Liu, S., et al., *Gene Mapping via Bulked Segregant RNA-Seq (BSR-Seq).* PLOS ONE, 2012. **7**(5): p. e36406.
164. Leshchiner, I., et al., *Mutation mapping and identification by whole-genome sequencing.* Genome Research, 2012. **22**(8): p. 1541-1548.

165. Schneeberger, K. and D. Weigel, *Fast-forward genetics enabled by new sequencing technologies.* Trends in Plant Science. **16**(5): p. 282-288.
166. Nordstrom, K.J., et al., *Mutation identification by direct comparison of whole-genome sequencing data from mutant and wild-type individuals using k-mers.* Nat Biotechnol, 2013. **31**(4): p. 325-30.
167. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nature methods, 2012. **9**(4): p. 357.
168. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nature protocols, 2012. **7**(3): p. 562.
169. Ashburner, M., et al., *Gene Ontology: tool for the unification of biology.* Nat Genet, 2000. **25**(1): p. 25-29.
170. Consortium, G.O., *Gene ontology consortium: going forward.* Nucleic acids research, 2015. **43**(D1): p. D1049-D1056.
171. Hempel, F.D. and L.J. Feldman, *Bi-directional inflorescence development inArabidopsis thaliana: Acropetal initiation of flowers and basipetal initiation of paraclades.* Planta, 1994. **192**(2): p. 276-286.
172. Katari, M.S., et al., *VirtualPlant: A Software Platform to Support Systems Biology Research.* Plant Physiology, 2010. **152**(2): p. 500-515.
173. Bowman, J.L. and D.R. Smyth, *CRABS CLAW, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix-loop-helix domains.* Development, 1999. **126**(11): p. 2387-2396.
174. Aubert, D., et al., *EMF1, a novel protein involved in the control of shoot architecture and flowering in Arabidopsis.* The Plant Cell, 2001. **13**(8): p. 1865-1875.
175. Ditta, G., et al., *The SEP4 Gene of Arabidopsis thaliana Functions in Floral Organ and Meristem Identity.* Current Biology, 2004. **14**(21): p. 1935-1940.
176. Saddic, L.A., et al., *The LEAFY target LMI1 is a meristem identity regulator and acts together with LEAFY to regulate expression of CAULIFLOWER.* Development, 2006. **133**(9): p. 1673-1682.
177. Auldridge, M.E., et al., *Characterization of three members of the Arabidopsis carotenoid cleavage dioxygenase family demonstrates the divergent roles of this multifunctional enzyme family.* The Plant Journal, 2006. **45**(6): p. 982-993.
178. Efroni, I., et al., *A Protracted and Dynamic Maturation Schedule Underlies Arabidopsis Leaf Development.* The Plant Cell, 2008. **20**(9): p. 2293-2306.
179. Barrero, J.M., et al., *INCURVATA2 Encodes the Catalytic Subunit of DNA Polymerase α and Interacts with Genes Involved in Chromatin-Mediated Cellular Memory in Arabidopsis thaliana.* The Plant Cell, 2007. **19**(9): p. 2822-2838.
180. Boerjan, W., et al., *Superroot, a recessive mutation in Arabidopsis, confers auxin overproduction.* The Plant Cell, 1995. **7**(9): p. 1405-1419.
181. Delarue, M., et al., *Sur2 mutations of Arabidopsis thaliana define a new locus involved in the control of auxin homeostasis.* The Plant Journal, 1998. **14**(5): p. 603-611.
182. Bak, S., et al., *CYP83B1, a Cytochrome P450 at the Metabolic Branch Point in Auxin and Indole Glucosinolate Biosynthesis in Arabidopsis.* The Plant Cell, 2001. **13**(1): p. 101-112.
183. Xie, Q., et al., *Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development.* Genes & development, 2000. **14**(23): p. 3024-3036.
184. (TAIR), T.A.I.R., *http://www.arabidopsis.org/tools/bulk/go/index.jsp, on www.arabidopsis.org*. 2017.
185. Mizoguchi, T., et al., *LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in Arabidopsis.* Developmental cell, 2002. **2**(5): p. 629-641.
186. Fujiwara, S., et al., *Circadian Clock Proteins LHY and CCA1 Regulate SVP Protein Accumulation to Control Flowering in Arabidopsis.* The Plant Cell, 2008. **20**(11): p. 2960-2971.
187. Imaizumi, T., et al., *FKF1 F-Box Protein Mediates Cyclic Degradation of a Repressor of CONSTANS in Arabidopsis.* Science, 2005. **309**(5732): p. 293-297.
188. Schmid, M., et al., *Dissection of floral induction pathways using global expression analysis.* Development (Cambridge, England), 2003. **130**(24): p. 6001-6012.
189. Ishikawa, M., et al., *The Arabidopsis SPA1 gene is required for circadian clock function and photoperiodic flowering.* The Plant journal : for cell and molecular biology, 2006. **46**(5): p. 736-746.
190. Pedmale, U.V., et al., *Cryptochromes Interact Directly with PIFs to Control Plant Growth in Limiting Blue Light.* Cell, 2016. **164**(1-2): p. 233-245.
191. Yang, S.W., et al., *FAR-RED ELONGATED HYPOCOTYL1 and FHY1-LIKE associate with the Arabidopsis transcription factors LAF1 and HFR1 to transmit phytochrome A signals for inhibition of hypocotyl elongation.* The Plant cell, 2009. **21**(5): p. 1341-1359.
192. Yu, X., et al., *The Cryptochrome Blue Light Receptors.* The Arabidopsis book, 2010. **8**: p. e0135.
193. Khanna, R., et al., *The basic helix-loop-helix transcription factor PIF5 acts on ethylene biosynthesis and phytochrome signaling by distinct mechanisms.* The Plant cell, 2007. **19**(12): p. 3915-3929.
194. Hornitschek, P., et al., *Phytochrome interacting factors 4 and 5 control seedling growth in changing light conditions by directly controlling auxin signaling.* The Plant journal : for cell and molecular biology, 2012. **71**(5): p. 699-711.
195. Zhao, R., et al., *The Arabidopsis Ca2+-dependent protein kinase CPK12 negatively regulates abscisic acid signaling in seed germination and post-germination growth.* New Phytologist, 2011. **192**(1): p. 61-73.
196. Zeng, J., et al., *Arabidopsis cryptochrome-1 restrains lateral roots growth by inhibiting auxin transport.* Journal of Plant Physiology, 2010. **167**(8): p. 670-673.

197. Sun, X., et al., *The Arabidopsis AtbZIP1 transcription factor is a positive regulator of plant tolerance to salt, osmotic and drought stresses.* Journal of plant research, 2012. **125**(3): p. 429-438.
198. Tsutsui, T., et al., *DEAR1, a transcriptional repressor of DREB protein that mediates plant defense and freezing stress responses in Arabidopsis.* Journal of plant research, 2009. **122**(6): p. 633.
199. Sakuma, Y., et al., *Functional analysis of an Arabidopsis transcription factor, DREB2A, involved in drought-responsive gene expression.* The Plant Cell, 2006. **18**(5): p. 1292-1309.
200. Jiang, Y. and M.K. Deyholos, *Functional characterization of Arabidopsis NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses.* Plant molecular biology, 2009. **69**(1-2): p. 91-105.
201. Mather, K. and J.L. Jinks, *Components of means: additive and dominance effects*, in *Biometrical Genetics*. 1971, Springer. p. 65-82.
202. Sharma, J.R., *Statistical and biometrical techniques in plant breeding*. 2006: New Age International.
203. Stone, B.F., *A formula for determining degree of dominance in cases of monofactorial inheritance of resistance to chemicals.* Bulletin of the World Health Organization, 1968. **38**(2): p. 325-326.
204. Neuwald, A.F., et al., *AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes.* Genome research, 1999. **9**(1): p. 27-43.
205. Snider, J., G. Thibault, and W.A. Houry, *The AAA+ superfamily of functionally diverse proteins.* Genome Biology, 2008. **9**(4): p. 216.
206. Ogura, T. and A.J. Wilkinson, *AAA+ superfamily ATPases: common structure - diverse function.* Genes Cells, 2001. **6**.
207. Marchler-Bauer, A., et al., *CDD/SPARCLE: functional classification of proteins via subfamily domain architectures.* Nucleic Acids Research, 2017. **45**(Database issue): p. D200-D203.
208. Tzafrir, I., et al., *Identification of Genes Required for Embryo Development in Arabidopsis.* Plant Physiology, 2004. **135**(3): p. 1206-1220.
209. Shpak, E.D., et al., *Synergistic interaction of three ERECTA-family receptor-like kinases controls Arabidopsis organ growth and flower development by promoting cell proliferation.* Development, 2004. **131**(7): p. 1491-1501.
210. Nour-Eldin, H.H., et al., *NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds.* Nature, 2012. **488**(7412): p. 531-534.
211. Fan, S.-C., et al., *The Arabidopsis nitrate transporter NRT1.7, expressed in phloem, is responsible for source-to-sink remobilization of nitrate.* The Plant cell, 2009. **21**(9): p. 2750-2761.
212. Maurel, C., et al., *The vacuolar membrane protein gamma-TIP creates water specific channels in Xenopus oocytes.* The EMBO journal, 1993. **12**(6): p. 2241-2247.
213. Kammerloher, W., et al., *Water channels in the plant plasma membrane cloned by immunoselection from a mammalian expression system.* The Plant Journal, 1994. **6**(2): p. 187-199.
214. Choi, H.-i., et al., *ABFs, a Family of ABA-responsive Element Binding Factors.* Journal of Biological Chemistry, 2000. **275**(3): p. 1723-1730.
215. Hirayama, T., et al., *A gene encoding a phosphatidylinositol-specific phospholipase C is induced by dehydration and salt stress in Arabidopsis thaliana.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(9): p. 3903-3907.
216. Sanchez, J.-P. and N.-H. Chua, *Arabidopsis PLC1 Is Required for Secondary Responses to Abscisic Acid Signals.* The Plant Cell, 2001. **13**(5): p. 1143-1154.
217. Lee, S., et al., *Proteomic Identification of Annexins, Calcium-Dependent Membrane Binding Proteins That Mediate Osmotic Stress and Abscisic Acid Signal Transduction in Arabidopsis.* The Plant Cell, 2004. **16**(6): p. 1378-1391.
218. Cantero, A., et al., *Expression profiling of the Arabidopsis annexin gene family during germination, de-etiolation and abiotic stress.* Plant Physiology and Biochemistry, 2006. **44**(1): p. 13-24.
219. Lee, K.H., et al., *Activation of glucosidase via stress-induced polymerization rapidly increases active pools of abscisic acid.* Cell, 2006. **126**(6): p. 1109-1120.
220. Iuchi, S., et al., *Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis.* The Plant Journal, 2001. **27**(4): p. 325-333.
221. Yanhui, C., et al., *The MYB Transcription Factor Superfamily of Arabidopsis: Expression Analysis and Phylogenetic Comparison with the Rice MYB Family.* Plant Molecular Biology, 2006. **60**(1): p. 107-124.
222. Chen, H., et al., *Integration of light and abscisic acid signaling during seed germination and early seedling development.* Proceedings of the National Academy of Sciences, 2008. **105**(11): p. 4495-4500.
223. Chen, H. and L. Xiong, *Role of HY5 in abscisic acid response in seeds and seedlings.* Plant Signaling & Behavior, 2008. **3**(11): p. 986-988.
224. Schaller, F., et al., *12-Oxophytodienoate reductase 3 (OPR3) is the isoenzyme involved in jasmonate biosynthesis.* Planta, 2000. **210**(6): p. 979-984.
225. Lorenzo, O., et al., *JASMONATE-INSENSITIVE1 Encodes a MYC Transcription Factor Essential to Discriminate between Different Jasmonate-Regulated Defense Responses in Arabidopsis.* The Plant Cell, 2004. **16**(7): p. 1938-1950.
226. Chini, A., et al., *The JAZ family of repressors is the missing link in jasmonate signalling.* Nature, 2007. **448**(7154): p. 666-671.
227. Yang, D.-L., et al., *Plant hormone jasmonate prioritizes defense over growth by interfering with gibberellin signaling cascade.* Proceedings of the National Academy of Sciences, 2012. **109**(19): p. E1192–E1200.

228. Büttner, M. and K.B. Singh, *Arabidopsis thaliana ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an ocs element binding protein.* Proceedings of the National Academy of Sciences, 1997. **94**(11): p. 5961-5966.
229. Yang, Z., et al., *Arabidopsis ERF4 is a transcriptional repressor capable of modulating ethylene and abscisic acid responses.* Plant molecular biology, 2005. **58**(4): p. 585-596.
230. Zhao, Y., et al., *Trp-dependent auxin biosynthesis in Arabidopsis: involvement of cytochrome P450s CYP79B2 and CYP79B3.* Genes & development, 2002. **16**(23): p. 3100-3112.
231. Bennett, M.J., et al., *Arabidopsis AUX1 Gene: A Permease-Like Regulator of Root Gravitropism.* Science, 1996. **273**(5277): p. 948-950.
232. Yang, Y., et al., *High-Affinity Auxin Transport by the AUX1 Influx Carrier Protein.* Current Biology, 2006. **16**(11): p. 1123-1127.
233. Nakamichi, N., et al., *The Arabidopsis pseudo-response regulators, PRR5 and PRR7, coordinately play essential roles for circadian clock function.* Plant and Cell Physiology, 2005. **46**(4): p. 609-619.
234. Salomé, P.A., D. Weigel, and C.R. McClung, *The Role of the Arabidopsis Morning Loop Components CCA1, LHY, PRR7, and PRR9 in Temperature Compensation.* The Plant Cell, 2010. **22**(11): p. 3650-3661.
235. Kuno, N., et al., *The Novel MYB Protein EARLY-PHYTOCHROME-RESPONSIVE1 Is a Component of a Slave Circadian Oscillator in Arabidopsis.* The Plant Cell, 2003. **15**(10): p. 2476-2488.
236. Mandadi, K.K., et al., *BT2, a BTB Protein, Mediates Multiple Responses to Nutrients, Stresses, and Hormones in Arabidopsis.* Plant Physiology, 2009. **150**(4): p. 1930-1939.
237. Umezawa, T., et al., *CYP707A3, a major ABA 8'-hydroxylase involved in dehydration and rehydration response in Arabidopsis thaliana.* The Plant Journal, 2006. **46**(2): p. 171-182.
238. Qin, F., et al., *Arabidopsis DREB2A-interacting proteins function as RING E3 ligases and negatively regulate plant drought stress–responsive gene expression.* The Plant Cell, 2008. **20**(6): p. 1693-1707.
239. Li, X., et al., *Blue Light- and Low Temperature-Regulated COR27 and COR28 Play Roles in the Arabidopsis Circadian Clock.* The Plant Cell, 2016. **28**(11): p. 2755-2769.
240. DeYoung, B.J., et al., *The CLAVATA1-related BAM1, BAM2 and BAM3 receptor kinase-like proteins are required for meristem function in Arabidopsis.* The Plant journal : for cell and molecular biology, 2006. **45**(1): p. 1-16.
241. Mara, C.D., T. Huang, and V.F. Irish, *The Arabidopsis floral homeotic proteins APETALA3 and PISTILLATA negatively regulate the BANQUO genes implicated in light signaling.* The Plant cell, 2010. **22**(3): p. 690-702.
242. Ge, X., et al., *An Arabidopsis aspartic protease functions as an anti-cell-death component in reproduction and embryogenesis.* EMBO reports, 2005. **6**(3): p. 282-288.
243. Favaro, R., et al., *MADS-box protein complexes control carpel and ovule development in Arabidopsis.* The Plant cell, 2003. **15**(11): p. 2603-2611.
244. Zhang, Y., et al., *SPL8, a local regulator in a subset of gibberellin-mediated developmental processes in Arabidopsis.* Plant molecular biology, 2007. **63**(3): p. 429-439.
245. Sedgley, M. and A.R. Griffin, *Sexual reproduction of tree crops*. 2013: Academic press.
246. Sablowski, R., *Flowering and determinacy in Arabidopsis.* Journal of Experimental Botany, 2007. **58**(5): p. 899-907.
247. Lazaro, A., E. Obeng- Hinneh, and M.C. Albani, *Extended vernalization regulates inflorescence fate in Arabis alpina by stably silencing PERPETUAL FLOWERING 1.* Plant Physiology, 2018.
248. Guo, Y. and S. Gan, *AtMYB2 regulates whole plant senescence by inhibiting cytokinin-mediated branching at late stages of development in Arabidopsis.* Plant Physiology, 2011. **156**(3): p. 1612-1619.
249. Schultz, E.A. and G.W. Haughn, *Genetic analysis of the floral initiation process (FLIP) in Arabidopsis.* Development, 1993. **119**(3): p. 745-765.
250. Liu, C., et al., *A conserved genetic pathway determines inflorescence architecture in Arabidopsis and rice.* Dev Cell, 2013. **24**(6): p. 612-22.
251. Wang, D., et al., *Transgenic expression of a putative calcium transporter affects the time of Arabidopsis flowering.* The Plant Journal, 2003. **33**(2): p. 285-292.
252. Wang, D.Y., et al., *PPF1 may suppress plant senescence via activating TFL1 in transgenic Arabidopsis plants.* Journal of integrative plant biology, 2008. **50**(4): p. 475-483.
253. Castillejo, C. and S. Pelaz, *The balance between CONSTANS and TEMPRANILLO activities determines FT expression to trigger flowering.* Current Biology, 2008. **18**(17): p. 1338-1343.
254. Sawa, M. and S.A. Kay, *GIGANTEA directly activates Flowering Locus T in Arabidopsis thaliana.* Proceedings of the National Academy of Sciences, 2011. **108**(28): p. 11698-11703.
255. Hanzawa, Y., T. Money, and D. Bradley, *A single amino acid converts a repressor to an activator of flowering.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(21): p. 7748-7753.
256. Mach, J., *TERMINAL FLOWER1 Acts in Transcriptional Repression.* The Plant Cell, 2011. **23**(9): p. 3084-3084.
257. Murfet, I., *Flowering in Pisum: Multiple alleles at the Lf locus.* Heredity, 1975. **35**(1): p. 85-98.
258. Ashelford, K., et al., *Full genome re-sequencing reveals a novel circadian clock mutation in Arabidopsis.* Genome biology, 2011. **12**(3): p. R28.
259. Smith, J.M. and J. Haigh, *The hitch-hiking effect of a favourable gene.* Genetics Research, 1974. **23**(1): p. 23-35.
260. Hanson, P.I. and S.W. Whiteheart, *AAA+ proteins: have engine, will work.* Nat Rev Mol Cell Biol, 2005. **6**(7): p. 519-29.
261. Iyer, L.M., et al., *Evolutionary history and higher order classification of AAA+ ATPases.* J Struct Biol, 2004. **146**.
262. Larsson, C., et al., *Plant Plasma Membrane*, in *eLS*. 2001, John Wiley & Sons, Ltd.
263. Zhang, J., *Evolution by gene duplication: an update.* Trends in ecology & evolution, 2003. **18**(6): p. 292-298.

264. The Arabidopsis Genome, I., *Analysis of the genome sequence of the flowering plant Arabidopsis thaliana.* Nature, 2000. **408**: p. 796.
265. Dolzblasz, A., et al., *The mitochondrial protease AtFTSH4 safeguards Arabidopsis shoot apical meristem function.* Scientific Reports, 2016. **6**: p. 28315.
266. Lee, K.-H., et al., *The RPT2 Subunit of the 26S Proteasome Directs Complex Assembly, Histone Dynamics, and Gametophyte and Sporophyte Development in Arabidopsis.* The Plant Cell, 2011. **23**(12): p. 4298-4317.
267. Ratcliffe, O.J., et al., *Regulation of Flowering in Arabidopsis by an FLCHomologue.* Plant Physiology, 2001. **126**(1): p. 122-132.
268. Izumi, N., A. Yamashita, and S. Ohno, *Integrated regulation of PIKK-mediated stress responses by AAA+ proteins RUVBL1 and RUVBL2.* Nucleus, 2012. **3**(1): p. 29-43.
269. Takeno, K., *Stress-induced flowering: the third category of flowering response.* Journal of Experimental Botany, 2016. **67**(17): p. 4925-4934.
270. Magome, H., et al., *dwarf and delayed-flowering 1, a novel Arabidopsis mutant deficient in gibberellin biosynthesis because of overexpression of a putative AP2 transcription factor.* The Plant Journal, 2004. **37**(5): p. 720-729.
271. Martínez, C., et al., *Salicylic acid regulates flowering time and links defence responses and reproductive development.* The Plant Journal, 2004. **37**(2): p. 209-217.
272. Yaish, M.W., J. Colasanti, and S.J. Rothstein, *The role of epigenetic processes in controlling flowering time in plants exposed to stress.* Journal of Experimental Botany, 2011. **62**(11): p. 3727-3735.
273. Pieterse, A.H., *Is flowering in Lemnaceae stress-induced? A review.* Aquatic Botany, 2013. **104**(Supplement C): p. 1-4.
274. Verslues, P.E. and T.E. Juenger, *Drought, metabolites, and Arabidopsis natural variation: a promising combination for understanding adaptation to water-limited environments.* Current Opinion in Plant Biology, 2011. **14**(3): p. 240-245.
275. Bohnert, H.J. and R.G. Jensen, *Strategies for engineering water-stress tolerance in plants.* Trends in Biotechnology, 1996. **14**(3): p. 89-97.
276. Chaves, M.M., J.P. Maroco, and J.S. Pereira, *Understanding plant responses to drought—from genes to the whole plant.* Functional plant biology, 2003. **30**(3): p. 239-264.
277. McKay, J.K., J.H. Richards, and T. Mitchell-Olds, *Genetics of drought adaptation in Arabidopsis thaliana: I. Pleiotropy contributes to genetic correlations among ecological traits.* Molecular Ecology, 2003. **12**(5): p. 1137-1151.
278. Yoo, C.Y., et al., *The Arabidopsis GTL1 transcription factor regulates water use efficiency and drought tolerance by modulating stomatal density via transrepression of SDD1.* The Plant Cell, 2010. **22**(12): p. 4128-4141.
279. Shpak, E.D., et al., *Stomatal patterning and differentiation by synergistic interactions of receptor kinases.* Science, 2005. **309**(5732): p. 290-293.
280. Riboni, M., et al., *GIGANTEA enables drought escape response via abscisic acid-dependent activation of the florigens and SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1.* Plant Physiology, 2013. **162**(3): p. 1706-1719.
281. Apse, M.P., et al., *Salt Tolerance Conferred by Overexpression of a Vacuolar Na⁺/H⁺ Antiport in Arabidopsis.* Science, 1999. **285**(5431): p. 1256-1258.
282. Zapryanova, N. and B. Atanassova, *Effects of Salt Stress on Growth and Flowering of Ornamental Annual Species.* Biotechnology & Biotechnological Equipment, 2009. **23**(sup1): p. 177-179.
283. Sulpice, R., et al., *Enhanced formation of flowers in salt-stressed Arabidopsis after genetic engineering of the synthesis of glycine betaine.* The Plant Journal, 2003. **36**(2): p. 165-176.
284. Pushpavalli, R., et al., *Salt stress delayed flowering and reduced reproductive success of chickpea (Cicer arietinum L.), a response associated with Na+ accumulation in leaves.* Journal of agronomy and crop science, 2016. **202**(2): p. 125-138.
285. Achard, P., et al., *Integration of plant responses to environmentally activated phytohormonal signals.* Science, 2006. **311**(5757): p. 91-94.
286. Suzuki, N., et al., *ABA Is Required for Plant Acclimation to a Combination of Salt and Heat Stress.* PLOS ONE, 2016. **11**(1): p. e0147625.
287. Zhu, J.-K., *SALT AND DROUGHT STRESS SIGNAL TRANSDUCTION IN PLANTS.* Annual Review of Plant Biology, 2002. **53**(1): p. 247-273.
288. Achard, P., et al., *The plant stress hormone ethylene controls floral transition via DELLA-dependent regulation of floral meristem-identity genes.* Proceedings of the National Academy of Sciences, 2007. **104**(15): p. 6484-6489.
289. Jung, J.-H., P.J. Seo, and C.-M. Park, *The E3 ubiquitin ligase HOS1 regulates Arabidopsis flowering by mediating CONSTANS degradation under cold stress.* Journal of Biological Chemistry, 2012. **287**(52): p. 43277-43287.
290. Jung, J.-H. and C.-M. Park, *HOS1-mediated activation of FLC via chromatin remodeling under cold stress.* Plant Signaling & Behavior, 2013. **8**(12): p. e27342.
291. Seo, E., et al., *Crosstalk between Cold Response and Flowering in Arabidopsis Is Mediated through the Flowering-Time Gene SOC1 and Its Upstream Negative Regulator FLC.* The Plant Cell, 2009. **21**(10): p. 3185-3197.
292. Sambrook, J. and D.W. Russell, *Molecular cloning: a laboratory manual. 2001*. 2001, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
293. Li, H., et al., *The sequence alignment/map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-2079.
294. Neff, M.M., E. Turk, and M. Kalishman, *Web-based primer design for single nucleotide polymorphism analysis.* TRENDS in Genetics, 2002. **18**(12): p. 613-615.

295. Nakagawa, T., et al., *Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation.* Journal of bioscience and bioengineering, 2007. **104**(1): p. 34-41.
296. Klock, H.E. and S.A. Lesley, *The Polymerase Incomplete Primer Extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis*, in *High Throughput Protein Expression and Purification*. 2009, Springer. p. 91-103.
297. Voinnet, O., et al., *Retracted: an enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus.* The Plant Journal, 2003. **33**(5): p. 949-956.
298. Nelson, B.K., X. Cai, and A. Nebenführ, *A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants.* The Plant Journal, 2007. **51**(6): p. 1126-1136.

7 Supplementary information

7.1 Supplementary Figures

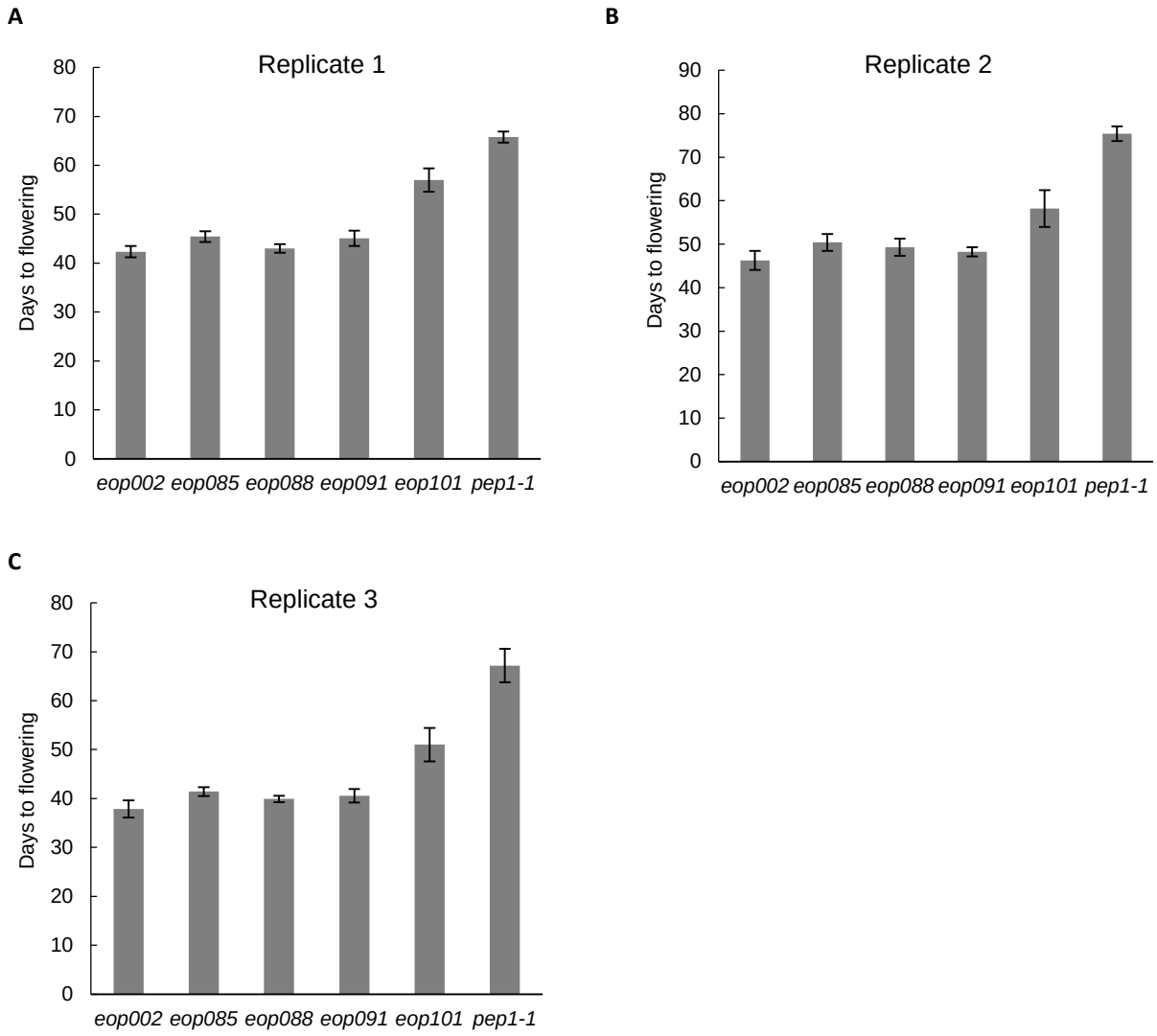


Figure S1. Days to flowering time.
(A), (B) and **(C)** represents experimental replictates of *eops* days to flowering compared to *pep1-1.* Data are presented as means ± SD, n = 12.

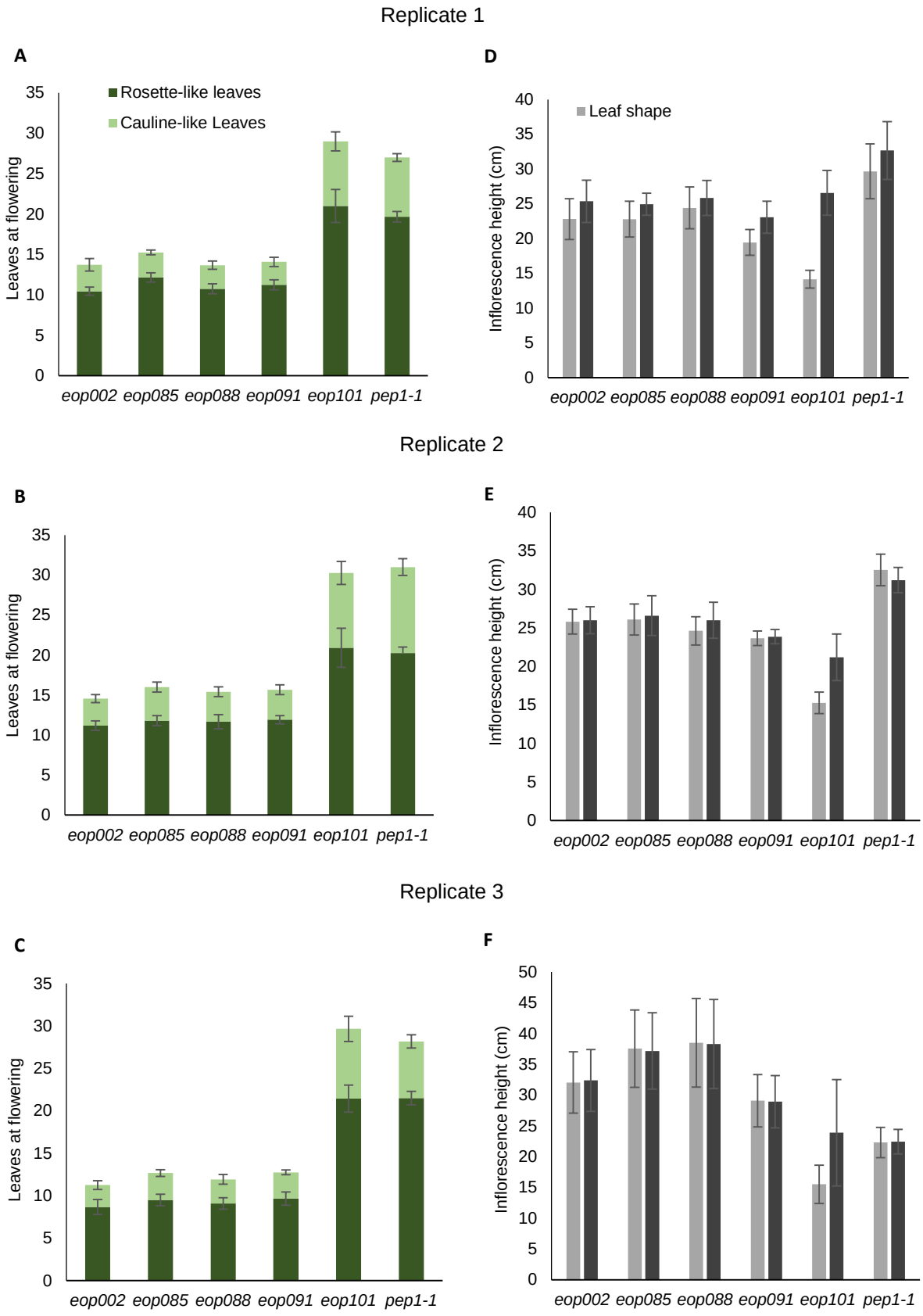


Figure S2. Leaves at flowering and inflorescence height of *eop* mutants.
(A), (B) and **(C)** represents experimental replicates of number of leaves at flowering in *eop* mutants compared to *pep1-1* **(D), (E)** and **(F)** Height of *eop* mutants and *pep1-1* inflorescence with three biological replicates, scored based on leaf shape and senescence of the inflorescence. Data are presented as means ± SD, n = 12.

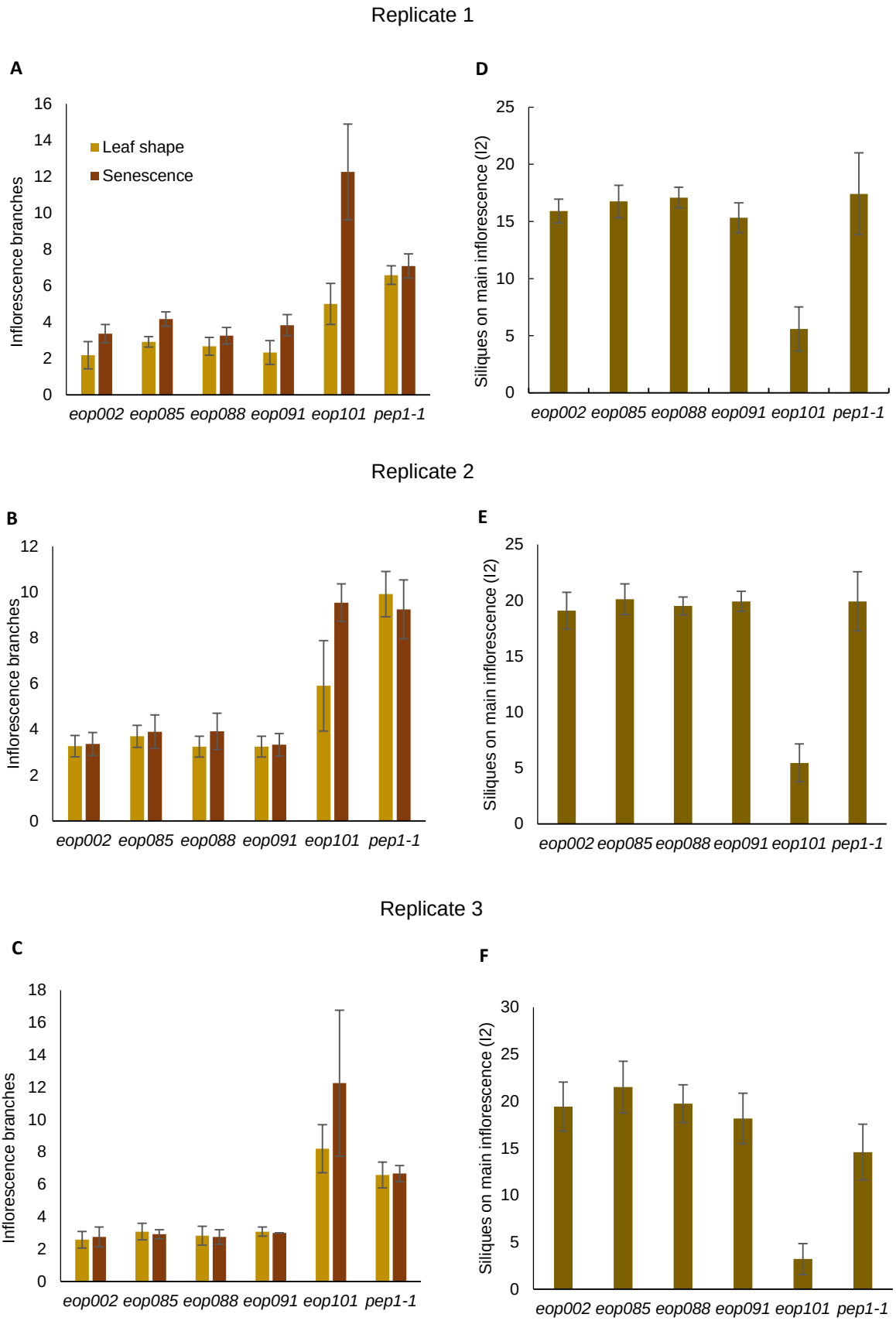
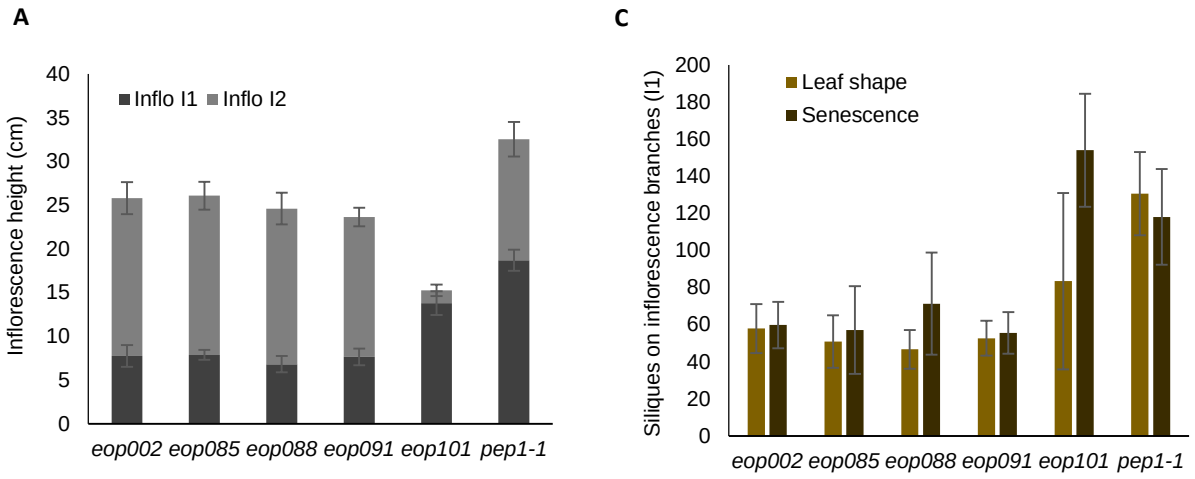


Figure S3. Inflorescence branches in I1 and siliques in I2 of *eop* mutants.
(A), (B) and **(C)** Number of inflorescence branches in I1 of *eops* and *pep1-1* scored based on leaf shape and senescence of the inflorescence with three experimental replicates. **(D), (E)** and **(F)** Number of siliques in I2 of *eop* mutants compared to *pep1-1*. Data are presented as means ± SD, n = 12.

Replicate 1



Replicate 2

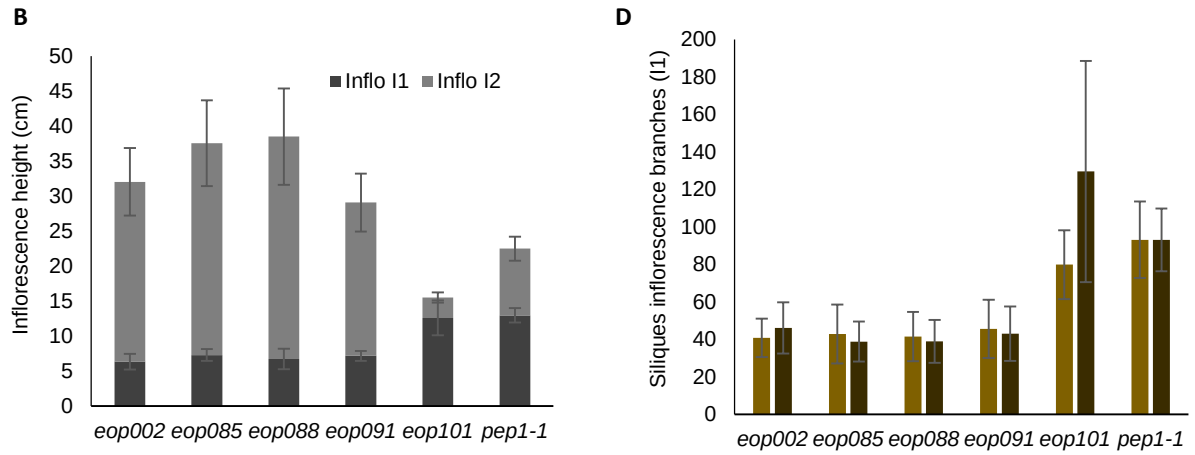


Figure S4. Inflorescence height (I1 and I2) and siliques in I2 of *eop* mutants and *pep1-1*.
(A) and **(B)** Height of inflorescence I1 and I2 of *eops* and *pep1-1* scored in two experimental replicates. **(C)** and **(D)** Number of siliques on inflorescence branches of *eop* mutants compared to *pep1-1,* scored based on leaf shape and senescence of the inflorescence. Data are presented as means ± SD, n = 12.

7.2 Supplementary Data

7.2.1 Dataset S1: *AaTFL1* sequence

ATGGAGAATATTGGAAGTAGACTGATAGAGCCATTGATAGTTGGAAGAGTGGTAGGAGAAGT
TCTTGATTACTTCACTCCAACAATTAAAATGAATGTGAGTTACAACAAGAAGCAAGTcTCCA
ATGGCCATGAGCTTTTCCCTTCCACTGTcTCCTCCAAGCCTAGGGTTGAGATCCATGGTGGT
GATcTCAGATCCTTcTTCACCTTGGTATATACATATATTAATTATAAACTCATCAATAATGT
GGCTTTGAGTAGTATATATTGTAGCAATACATATTCAAAGTcTCATTTTTTGCCACTTACAA
GTTTTCCATTTTcTGACTcTACAGAGAAAAATATAGGGACAGATCCAAAAAATcTATTAAGT
TGTACTTTTTTTTTGTAATATAATAGGATGATATATCTTAAGTTTCTATCATTCCTTTTTTT
CTTTCATGTCATTTATTTATATATTTATTTGTACTGATATGGTAAAACTAAACTGAAAACAG
GTGATGATAGACCCTGATGTTCCAGGTCCTAGTGACCCCTTCCTAAAAGAACACCTGCATTG
GTACGTTCAGTTTTATTTTGTCTTTTCATAATTTTGGGCCCATGGCTCATATGCATTGCATT
TCAAAAAAAACTTATAACCCTAGTAAAGTTATAATTAATACTTATATTTCGTATAAACTTGA
GTTGGTCAAAATCTCTTCCTAATCACATGTCACAACACCATTTTAAATATTTCAACAGGATC
GTTACAAACATCCCCGGTACAACAGATGCTACATTTGGTAAGGTCTTCTCATGAaTTTTTTT
TTTGCTAAACTGAATTTTGTATTTTAAATACTTGTATATTGAAAAACATGTTAAATATAAAT
TATAAATATTTGAATTGAAATATAGGAAAAGAGGTGGTGAGCTATGAATTGCCAAGGCCAAG
CATAGGGATACATAGGTTTGTGTTTGTTCTCTTCAAGCAGAAGCAAAGACGCGTCATCTTCC
CCAATATACCTTCGAGAGATCACTTCAACACTCGTGAATTCGCAGTCGAGTATGATCTTGGT
CTTCCTGTCGCTGCGGTGTTCTTTAACGCTCAAAGGGAAACCGCGGCTCGCAAACGCTAG

7.2.2 Dataset S1: *AaG106560* sequence

>Aa_G106560_t1
ATGTTTTCTTTCTCTAGGGATCTTCCTTCACCCACCTCGTTTTTCACAGCTTATGCATCAAT
GGCTGGTTACATGATGATGATAAGATCAATGGCTCATGAGCTTATCCCAGCTCCAGTCCAAA
ATTTCATTTACTCTTCTCTAAGGTCTTTCTTCGTTCGTTCTACCTCCACAACTTTGACTCTA
ACCATCGATGACGACAACATGGGTATGAACAACGAGATCTACAGAGCAGCTCAGACTTATCT
CTCTACCAAGATCAGTCCAGACGCAGTCAGGTTAAGAATCAGTAAAGGCCATAAAGACAAAC
ATGTCAATCTCTATCTCAGCAATGGAGAAATCGTCAACGATGTTTACCAAGACGTCGAGCTC
ACATGGAGGTTTGTTACTGACCGTGGTCACAAGAAAGGAGGAGGAGGAGGAGGAGGAAGCCC
TGTTGATGACGATGGTAAAAGCGAGTATTTTGAGCTGAGTTTCGAGAAAGAGCATAAAGATT
TGATCTTGAACTCTTACGTTCCTTACATTGAGAGTAAAGCTAAAGTGATTAAGAACGAGAGG
AGGATCTTGATGCTGCATTCTCTCAATAACTCCCGTTGGGAATCTGTTATCCTCGAGCATCC
TTCGACTTTCGAGACAATGGCGATGGAAGATGAGCTCAAGCGTGAGGTTAAAGCAGATCTTG
ACCGGTTTATTAGAAGGAAAGAGTTTTACAAGAGAGTAGGGAAAGCTTGGAAAAGGGGTTAC
TTGTTGTACGGACCGCCTGGAACCGGGAAGTCTAGTTTGGTTGCAGCTATGGCTAATTACCT
CAAATTTGATGTCTATGATCTTCAGCTAGCGAGTGTGATGCGTGACTCTGATCTTAGGAGGC
TTTTACTCGCCACACGTAACCGTTCGATTCTTGTCATTGAAGATATCGATTGTGCAGTGGAG

TTGCCTAATAGAGTGGGGCAGTCTGTTGAAGGCAAGAACCGTGGCGAATCTCAGGTTTGGTT
CTGTCTCTAATGAAATTCTATTAGATATTTTTTTTTACTATAATTGGTATTGGTATAGGTTT
GGGATTTTTAACCAAATCGAACCAAACCAATATTCGTACCAATGTTTTATTTATAACCAAAG
TTTTTTTTATTTTTATTTATAACCAAAGTTTGTCTCCTATACCAAACTGAATTCAGCTTCTC
CCTTAGAAGACAGAATCTTGATTAGTCTAGAGAATTAGTATATTTTAATGATTTTTAACCAC
ATTACTAAATCTAACCAAACCATAATCCGTAGCAAGATTTTATTCGATGAATCTACATTTAG
TATAACAAGACTGTGTCTCATTAGTATATAGAGAATTATTAGTACTATATCTAAATGGTTTC
TTTTTGTTGTTTTCTATAAAATTGTGTGTTTAGGGACCATTGACGTTATCGGGGCTGCTGAA
TTTCATTGACGGATTATGGTCAAGTTGCGGAGACGAGCGTATTATTATATTCACGACGAACC
ACAAAGATAGGCTTGACCCGGCATTGTTACGTCCGGGTCGAATGGATATGCACATTTACATG
GGACATTGCTCGTTTCAAGGATTCAAGACCTTAGCTTCCAACTATTTGTGCTTGAACGACCC
CACAATGCCTCACCGTCTCTACCCCGAGATCGAGCGTTTGATGGAAGGGGACGTGATCACCC
CGGCACAAGTGGCAGAGGAGTTGATGAAGAGCGAGGATGCTGACATGGCGCTTGAGGGCTTG
GTGGATGTTTTAAAGAAGATGAGGTTAAAATCTGAGGAATCGAATCCGGGGATGAAGAAAAA
GGAGAATAGGTTGGAGATGGAGGAGATGAGATTAAAGGGTGTTACAGAGGGTTCTCCGAGGA
AGAACAGCAAAAGAATTAAGAAACTTGTTCTGTTTTGCACCTAA

xxxx = Start codon

xxxx = Exon

xxxx = Intron

xxxx = Stop codon

7.3 Supplementary Tables

Supplementary Table 1. EMS mutations with in *eop002*

Chr [1]	Pos [2]	R[3]	M[4]	N[5]	AF[6]	Mq[7]	Region[8]	Gene ID[9]	Type[10]	AR[11]	AM[12]
chr1	6797317	C	T	13	0.72	211	intergenic				
chr1	14421199	C	T	19	0.70	225	intergenic				
chr1	17023359	G	A	2	1.00	31	intergenic				
chr1	19497517	C	T	5	0.71	100	intergenic				
chr1	19684523	C	T	146	0.81	101	intergenic				
chr2	1763307	G	A	122	0.97	222	intergenic				
chr2	2795828	G	A	40	0.80	201	intergenic				
chr2	2900769	G	A	32	0.89	118	intergenic				
chr2	3031340	G	A	26	0.90	56	intergenic				
chr2	5401756	G	A	53	1.00	144	intergenic				
chr2	7322892	G	A	79	0.85	225	CDS	Aa_G62080.t1	Nonsyn	G	S
chr2	7360332	G	A	97	0.71	40	intronic/noncoding	Aa_G711220.t1			
chr2	7412494	G	A	65	0.80	225	intronic/noncoding	Aa_G711220.t1			
chr2	7769585	G	A	66	0.84	222	intergenic				
chr2	7855399	G	A	96	0.91	222	CDS	Aa_G510360.t1	Nonsyn	V	M
chr2	8015048	G	A	52	0.85	119	intergenic				
chr2	8082334	G	A	135	0.87	65	intergenic				
chr2	8341157	G	A	67	0.73	152	intergenic				
chr2	8518036	G	A	47	0.85	225	intergenic				
chr2	9946910	G	A	35	0.85	178	intergenic				
chr2	10024359	C	T	179	0.90	69	intergenic				
chr2	10024420	C	T	182	0.74	106	intergenic				
chr2	10445966	G	A	67	0.79	61	intergenic				
chr2	10875022	G	A	46	0.87	132	intergenic				
chr2	11657360	G	A	88	0.86	225	intergenic				
chr2	11994491	G	A	78	0.82	215	intergenic				
chr2	12109271	G	A	123	0.76	75	intergenic				
chr2	12979094	C	T	81	0.80	198	intergenic				
chr2	13141687	C	T	219	0.77	158	CDS	Aa_G100860.t1	Nonsyn	R	Q
chr2	13214570	C	T	46	0.77	225	intergenic				
chr2	13811196	C	T	29	0.83	225	intergenic				
chr2	14099612	C	T	168	0.78	94	intergenic				
chr2	14158207	C	T	44	0.83	48	intergenic				
chr2	14333524	C	T	113	0.82	197	intergenic				
chr2	14362712	C	T	31	0.86	116	intergenic				
chr2	14419565	C	T	77	0.91	222	intergenic				
chr2	14436363	C	T	35	0.90	113	intergenic				
chr2	14468752	C	T	81	0.84	37	intergenic				
chr2	14493402	C	T	12	0.71	60	intergenic				
chr2	14986658	G	A	77	0.86	190	intergenic				
chr2	15055997	G	A	55	0.82	225	intergenic				
chr2	15152313	G	A	48	0.73	154	intergenic				

chr2	15254226	G	A	74	0.84	203	intergenic				
chr2	15821900	G	A	143	0.85	222	CDS	Aa_G201310.t1	Nonsyn	T	I
chr2	16287050	C	T	157	0.84	69	intergenic				
chr2	16326607	C	T	21	0.95	222	intergenic				
chr2	16632655	C	T	72	0.82	219	intergenic				
chr2	17308273	G	A	66	0.84	81	intergenic				
chr2	17597241	G	A	89	0.81	122	intergenic				
chr2	18231233	G	A	84	0.82	71	intergenic				
chr2	18420496	G	A	90	0.80	137	intergenic				
chr2	18465340	G	A	39	0.76	77	intergenic				
chr2	18720577	G	A	84	0.85	71	intergenic				
chr2	18981387	G	A	18	0.78	32	intergenic				
chr2	19029086	G	A	54	0.83	86	intergenic				
chr2	19139874	G	A	121	0.77	84	intergenic				
chr2	19401891	G	A	49	0.82	225	intergenic				
chr2	20099821	G	A	30	0.77	202	intergenic				
chr2	20163001	G	A	22	0.96	222	intergenic				
chr2	20450257	G	A	19	0.95	64	intergenic				
chr2	20478766	G	A	24	0.75	77	intergenic				
chr2	20641109	G	A	153	0.84	34	intergenic				
chr2	20882189	G	A	34	0.87	161	intergenic				
chr2	22205786	G	A	42	0.89	222	intergenic				
chr2	23013830	G	A	135	0.84	173	intergenic				
chr2	23629788	G	A	110	0.86	222	intergenic				
chr2	23887355	G	A	54	0.86	139	intergenic				
chr2	24297513	G	A	49	0.83	34	intergenic				
chr2	25075513	G	A	101	0.88	120	intergenic				
chr2	25564239	G	A	29	0.94	222	intronic/noncoding	Aa_G562950.t1			
chr2	25629508	G	A	76	0.80	225	intergenic				
chr2	25868420	G	A	132	0.74	97	intergenic				
chr2	26449223	G	A	50	0.83	62	intergenic				
chr2	26672887	G	A	126	0.85	129	intergenic				
chr2	26783046	G	A	56	0.75	161	intergenic				
chr2	28301731	G	A	46	0.70	66	intergenic				
chr2	30612834	G	A	110	0.99	131	intergenic				
chr2	31821205	G	A	85	0.92	73	intergenic				
chr2	31862213	G	A	55	0.93	222	intergenic				
chr2	31921612	G	A	65	0.94	222	intergenic				
chr3	5262636	C	T	182	0.91	42	intergenic				
chr3	7386563	G	A	2	1.00	33	intergenic				
chr3	18037392	G	A	11	0.92	212	intergenic				
chr4	57980	G	A	69	0.80	89	intergenic				
chr4	208951	G	A	35	0.76	136	intergenic				
chr4	397358	C	T	54	0.76	225	intergenic				
chr4	1024577	G	A	2	1.00	50	intergenic				
chr4	1566552	G	A	49	0.94	80	intergenic				
chr4	1663060	C	T	40	0.91	110	intergenic				

chr4	1683866	C	T	36	0.75	225	intergenic				
chr4	2274621	C	T	33	0.80	187	intergenic				
chr4	6179218	C	T	5	0.83	76	intronic/noncoding	Aa_G95120.t1			
chr4	44831847	C	T	42	0.71	63	CDS	Aa_G69950.t1	Nonsyn	P	L
chr4	44831848	C	T	43	0.70	60	CDS	Aa_G69950.t1	Nonsyn	P	L
chr5	1460913	G	A	84	0.87	92	intergenic				
chr5	2078095	G	A	96	0.93	222	intergenic				
chr5	2161089	G	A	124	0.92	102	intergenic				
chr5	2532130	G	A	95	0.95	222	CDS	Aa_G83240.t1	Syn	I	I
chr5	16121506	G	A	13	0.76	97	intergenic				
chr5	27978593	C	T	61	0.69	96	intergenic				
chr6	12992424	C	T	3	1.00	34	intergenic				
chr7	44932330	C	T	140	0.94	50	intergenic				
chr7	48150555	C	T	68	0.83	109	intergenic				
chr8	6857883	C	T	20	0.71	170	intergenic				
chr8	7174932	C	T	24	0.73	225	intergenic				
chr8	8220357	C	T	9	0.90	118	intergenic				
chr8	8734412	C	T	18	0.86	131	intergenic				
chr8	9313334	C	T	89	0.92	222	CDS	Aa_G526410.t1	Nonsyn	S	F
chr8	9412272	C	T	60	0.82	222	intergenic				
chr8	9880210	C	T	81	0.88	222	CDS	Aa_G290000.t1	Nonsyn	A	T
chr8	10243759	C	T	150	0.89	222	intergenic				
chr8	10331702	C	T	22	1.00	222	intronic/noncoding	Aa_G258170.t1			
chr8	10715716	C	T	51	0.84	222	CDS	Aa_G7880.t1	Nonsyn	G	E
chr8	10776864	C	T	59	0.89	222	intergenic				
chr8	10887335	C	T	56	0.97	222	intronic/noncoding	Aa_G644350.t1			
chr8	11625706	C	T	10	1.00	155	intergenic				
chr8	11634026	C	T	63	0.94	222	intronic/noncoding	Aa_G236930.t1			
chr8	11987102	C	T	33	0.94	222	intergenic				
chr8	12811867	C	T	123	0.72	44	intergenic				
chr8	12834836	C	T	250	0.85	222	intergenic				
chr8	13034187	C	T	70	0.93	222	intergenic				
chr8	13202635	G	A	58	0.94	222	CDS	Aa_G79140.t1	Nonsyn	G	S
chr8	13621177	C	T	13	1.00	222	intergenic				
chr8	13661268	C	T	162	0.96	222	intergenic				
chr8	13664070	C	T	155	0.82	83	intergenic				
chr8	13711234	C	T	17	0.89	103	intergenic				
chr8	13760479	C	T	70	0.92	222	intergenic				
chr8	13955120	C	T	18	0.90	107	intergenic				
chr8	13980325	C	T	139	0.91	222	intergenic				
chr8	14000397	C	T	45	0.90	222	intergenic				
chr8	14041350	C	T	63	0.93	222	intergenic				
chr8	14123831	C	T	20	1.00	222	intergenic				
chr8	14257447	C	T	88	0.95	222	intergenic				
chr8	14424150	C	T	100	0.97	222	intergenic				
chr8	14439225	C	T	156	0.87	222	intergenic				
chr8	14520456	C	T	169	0.91	222	intergenic				

chr8	14606687	C	T	65	0.98	222	CDS	Aa_G106560.t1	Nonsyn	G	D
chr8	15014460	C	T	98	0.92	222	intergenic				
chr8	15181961	C	T	98	0.94	222	intergenic				
chr8	15384427	C	T	145	0.97	222	intergenic				
chr8	15489471	C	T	122	0.97	173	intergenic				
chr8	15511433	C	T	131	0.92	197	intergenic				
chr8	15833454	C	T	202	0.91	222	intergenic				
chr8	15956702	C	T	70	0.99	222	intergenic				
chr8	16164654	C	T	34	0.97	222	intergenic				
chr8	16185062	C	T	54	0.93	222	intergenic				
chr8	16209488	C	T	205	0.88	222	intergenic				
chr8	16713521	C	T	150	0.89	222	intergenic				
chr8	16874023	C	T	111	0.92	222	intergenic				
chr8	16876788	C	T	14	0.88	40	intergenic				
chr8	16889208	C	T	103	0.76	215	intergenic				
chr8	17088532	C	T	152	0.97	222	intergenic				
chr8	17109487	C	T	151	0.97	222	intergenic				
chr8	17237921	C	T	140	0.97	222	intergenic				
chr8	17240921	C	T	67	0.77	222	intergenic				
chr8	17704084	C	T	116	0.95	200	intergenic				
chr8	17715681	C	T	6	0.86	149	intergenic				
chr8	17734340	C	T	116	0.97	222	intergenic				
chr8	17929469	C	T	107	0.97	222	intergenic				
chr8	18006989	C	T	264	0.92	222	intergenic				
chr8	18109904	C	T	48	0.96	88	intergenic				
chr8	18145960	C	T	85	0.98	222	CDS	Aa_G255500.t1	Nonsyn	L	F
chr8	18338213	C	T	30	0.94	222	intergenic				
chr8	18570759	C	T	97	0.95	222	intergenic				
chr8	18684871	C	T	94	0.89	173	intergenic				
chr8	19307861	C	T	98	0.99	222	intergenic				
chr8	19640684	C	T	84	0.93	222	intergenic				
chr8	19737390	C	T	297	0.95	222	intergenic				
chr8	19801428	C	T	48	0.98	221	intergenic				
chr8	19839786	C	T	182	0.95	136	intergenic				
chr8	20875190	C	T	96	0.95	222	CDS	Aa_G429030.t1	Nonsyn	P	L
chr8	20958184	C	T	90	0.94	222	intergenic				
chr8	21127920	C	T	33	1.00	222	intergenic				
chr8	21205314	C	T	11	1.00	177	intergenic				
chr8	21369766	C	T	162	0.94	198	intergenic				
chr8	21373465	C	T	174	0.95	103	intronic/noncoding	Aa_G25950.t1			
chr8	21463312	C	T	20	0.95	126	intergenic				
chr8	21885041	C	T	268	0.91	222	intergenic				
chr8	21957575	C	T	57	0.93	222	intergenic				
chr8	21989344	C	T	90	0.96	222	CDS	Aa_G251110.t1	Syn	L	L
chr8	22290630	C	T	63	0.95	222	intergenic				
chr8	22358235	C	T	195	0.92	222	intergenic				
chr8	22413266	C	T	161	0.95	222	intergenic				

chr8	22425315	C	T	111	0.98	222	intergenic				
chr8	22478822	C	T	91	0.90	222	intergenic				
chr8	22584316	C	T	170	0.89	222	intergenic				
chr8	22620391	C	T	40	0.70	48	intronic/noncoding	Aa_G891490.t1			
chr8	22675214	C	T	82	0.96	222	intergenic				
chr8	22704092	C	T	120	0.94	222	intergenic				
chr8	23010791	C	T	83	0.97	222	intergenic				
chr8	23121084	C	T	78	0.97	185	intergenic				
chr8	23203039	C	T	132	0.92	222	intergenic				
chr8	23265021	C	T	166	0.94	222	intergenic				
chr8	23267373	C	T	111	0.97	222	intergenic				
chr8	23269313	C	T	117	0.97	222	intergenic				
chr8	23274585	C	T	19	0.83	58	intergenic				
chr8	23289571	C	T	132	0.90	222	intergenic				
chr8	23628464	C	T	33	0.94	202	intronic/noncoding	Aa_G1094820.t1			
chr8	23740783	C	T	111	0.87	47	intergenic				
chr8	24324694	C	T	36	0.73	39	intergenic				
chr8	25547597	C	T	19	1.00	222	intergenic				
chr8	25621460	C	T	66	0.96	204	intergenic				
chr8	25818825	C	T	272	0.93	222	intergenic				
chr8	25987592	C	T	4	0.80	46	intergenic				
chr8	26055279	C	T	3	1.00	81	intergenic				
chr8	26208920	C	T	87	0.92	222	intergenic				
chr8	26302685	C	T	18	1.00	222	intergenic				
chr8	26332336	C	T	15	0.94	55	intergenic				
chr8	26477645	C	T	121	0.92	222	intronic/noncoding	Aa_G23190.t1			
chr8	26760582	C	T	46	0.90	187	intergenic				
chr8	26797352	C	T	39	0.97	222	intergenic				
chr8	27205792	C	T	59	0.97	222	intergenic				
chr8	27390844	C	T	5	1.00	103	intergenic				
chr8	27475487	C	T	127	0.95	222	CDS	Aa_G102190.t1	Syn	G	G
chr8	27625936	C	T	108	0.92	222	intergenic				
chr8	27845862	C	T	85	0.96	222	intergenic				
chr8	27954261	C	T	163	0.90	222	intergenic				
chr8	28008646	C	T	57	0.92	198	intergenic				
chr8	28177351	C	T	149	0.96	206	intergenic				
chr8	28377243	C	T	112	0.96	35	intergenic				
chr8	28635028	C	T	140	0.83	120	intergenic				
chr8	28741382	C	T	61	0.97	222	intergenic				
chr8	28891148	C	T	137	0.93	222	intergenic				
chr8	29043168	G	A	92	0.86	72	intergenic				
chr8	29315243	C	T	27	0.96	222	intergenic				
chr8	29814921	C	T	123	0.92	119	intergenic				
chr8	29940015	C	T	64	0.97	222	intronic/noncoding	Aa_G188010.t1			
chr8	29975495	C	T	48	0.94	222	intronic/noncoding	Aa_G424180.t1			
chr8	30234664	C	T	143	0.89	222	intergenic				
chr8	30435031	C	T	118	0.92	222	intergenic				

chr8	30592921	C	T	68	0.99	222	intergenic				
chr8	30907629	C	T	60	0.95	84	intergenic				
chr8	30953322	C	T	74	0.93	222	CDS	Aa_G675250.t1	Syn	N	N
chr8	30992873	C	T	117	0.93	192	intergenic				
chr8	31005143	C	T	57	0.86	222	intergenic				
chr8	31150891	C	T	5	0.83	137	intergenic				
chr8	31539099	C	T	9	1.00	153	intergenic				
chr8	32056783	C	T	74	0.93	222	intergenic				
chr8	32120637	C	T	161	0.93	213	intergenic				
chr8	33619479	C	T	40	0.87	222	CDS	Aa_G636600.t1	Nonsyn	V	I
chr8	33759002	C	T	147	0.86	222	CDS	Aa_G175520.t1	Nonsyn	D	N
chr8	33835619	C	T	86	0.97	222	intergenic				
chr8	33985612	C	T	82	0.83	144	intergenic				
chr8	34022261	C	T	104	0.91	222	intergenic				
chr8	34308396	C	T	10	1.00	208	intergenic				
chr8	34387450	C	T	56	0.92	222	intergenic				
chr8	34549832	C	T	134	0.96	222	intergenic				
chr8	34570985	C	T	55	0.96	222	intergenic				
chr8	34677814	C	T	95	0.96	194	intergenic				
chr8	34763272	C	T	23	1.00	222	intergenic				
chr8	35005804	C	T	8	1.00	105	intronic/noncoding	Aa_G496120.t1			
chr8	35189685	C	T	135	0.96	222	intronic/noncoding	Aa_G496120.t1			
chr8	35194191	C	T	70	0.92	222	intronic/noncoding	Aa_G496120.t1			
chr8	35236211	C	T	58	0.95	222	intronic/noncoding	Aa_G496120.t1			
chr8	35291533	C	T	110	0.96	222	intronic/noncoding	Aa_G496120.t1			
chr8	35291533	C	T	110	0.96	222	CDS	Aa_G88110.t1	Nonsyn	P	S
chr8	35507850	C	T	218	0.94	201	intergenic				
chr8	35511867	C	T	155	0.92	166	intergenic				
chr8	36216713	C	T	115	0.94	222	CDS	Aa_G149070.t1	Syn	N	N
chr8	36252343	C	T	99	0.91	222	CDS	Aa_G542640.t1	Nonsyn	E	K
chr8	36397324	C	T	107	0.95	62	intergenic				
chr8	36449900	C	T	41	0.98	222	intronic/noncoding	Aa_G605090.t1			
chr8	36751839	C	T	44	0.81	222	CDS	Aa_G59080.t1	Nonsyn	S	N
chr8	37039767	C	T	92	0.93	222	CDS	Aa_G571550.t1	Syn	G	G
chr8	37603211	C	T	16	0.89	222	intergenic				
chr8	37606946	C	T	27	0.93	87	intergenic				
chr8	38300011	C	T	85	0.89	222	intronic/noncoding	Aa_G22070.t1			
chr8	38633610	C	T	110	0.77	104	intergenic				
chr8	38873205	C	T	175	0.91	222	intergenic				
chr8	39075205	C	T	86	0.88	222	CDS	Aa_G69080.t1	Syn	K	K
chr8	39089388	C	T	46	0.98	222	intergenic				
chr8	39130314	C	T	69	0.91	222	intronic/noncoding	Aa_G307370.t1			
chr8	39216163	C	T	39	0.89	32	intergenic				
chr8	39217072	C	T	48	0.91	60	intergenic				
chr8	39294663	C	T	52	0.95	222	intergenic				
chr8	39796839	C	T	53	0.90	226	intronic/noncoding	Aa_G31530.t1			
chr8	39953425	C	T	66	0.88	222	intronic/noncoding	Aa_G70060.t1			

chr8	39953425	C	T	66	0.88	222	intronic/noncoding	Aa_G93740.t1			
chr8	39990062	C	T	85	0.80	225	intronic/noncoding	Aa_G70060.t1			
chr8	40010585	C	T	86	0.84	222	intronic/noncoding	Aa_G70060.t1			
chr8	40010585	C	T	86	0.84	222	CDS	Aa_G93880.t1	Nonsyn	T	I
chr8	40391906	C	T	78	0.84	225	CDS	Aa_G345460.t1	Nonsyn	A	V
chr8	41169609	C	T	85	0.86	225	intergenic				
chr8	41500336	C	T	57	0.83	225	intergenic				
chr8	41852076	C	T	15	0.71	225	intergenic				
chr8	41882197	C	T	43	0.80	225	intergenic				
chr8	42358614	C	T	160	0.82	164	CDS	Aa_G447610.t1	Syn	N	N
chr8	42481804	C	T	43	0.80	225	intergenic				
chr8	42492751	C	T	34	0.85	225	intergenic				
chr8	42655518	C	T	76	0.75	225	intergenic				
chr8	42657958	C	T	51	0.82	225	intergenic				
chr8	42936933	C	T	51	0.78	225	intergenic				
chr8	42968078	C	T	8	0.73	175	intergenic				
chr8	43438523	C	T	136	0.77	191	CDS	Aa_G68440.t1	Nonsyn	V	I
chr8	43509454	C	T	50	0.81	225	intronic/noncoding	Aa_G478040.t1			
chr8	43941960	C	T	92	0.72	142	intergenic				
chr8	44186275	C	T	142	0.76	225	CDS	Aa_G70630.t1	Nonsyn	A	V
chr8	44230492	C	T	23	0.70	225	intergenic				
chr8	46586909	G	A	3043	0.92	44	intergenic				
chr8	46586988	G	A	2103	0.92	168	intergenic				
chr8	46587047	G	A	425	0.90	157	intergenic				

1) Chr: Chromosome. **(2)** Pos: Mutated nucleotide position. **(3)** R: Nucleotide in reference genome. **(4)** M: Nucleotide in mutant genome. **(5)** N: Number of reads supporting the mutation. **(6)** AF: Allele frequency. **(7)** Mq: Mapping quality (MAPQ), higher = more unique (min. 30, implies a 0.001 probability of false read alignment). **(8)** Region: Region affected by the mutation. **(9)** Gene ID: Gene identifier. **(10** Type: Mutation type (nonsynonymous or synonymous). **(11)** AR: Amino acid in reference genome. **(12)** AM: Amino acid in mutant genome.

Supplementary Table 2. EMS mutations with in *eop085*

Chr [1]	Pos [2]	R[3]	M[4]	N[5]	AF[6]	Mq[7]	Region[8]	Gene ID[9]	Type[10]	AR[11]	AM[12]
chr1	17830271	G	A	3	1.00	49	intergenic				
chr1	17830349	G	A	2	1.00	37	intergenic				
chr1	22842094	C	T	31	0.72	40	intronic/noncoding	Aa_G367140.t1			
chr2	10024176	G	A	11	0.92	43	intergenic				
chr2	14053905	C	T	17	0.71	38	intergenic				
chr2	20613031	G	A	53	0.70	35	intergenic				
chr2	20613053	G	A	71	0.72	62	intergenic				
chr2	20657436	G	A	89	0.76	63	intergenic				
chr2	26326035	G	A	41	0.71	47	intergenic				
chr3	4399783	C	T	13	0.72	224	intergenic				
chr3	5262636	C	T	26	0.84	86	intergenic				
chr3	26208296	G	A	4	1.00	78	intergenic				
chr4	3296814	G	A	4	0.80	72	intergenic				
chr4	17707723	G	A	14	0.74	225	intergenic				
chr4	22895997	G	A	14	0.70	225	intergenic				
chr4	24707519	G	A	9	0.69	92	intergenic				
chr5	11179979	G	A	3	0.75	56	CDS	Aa_G102620.t1	Nonsyn	R	*
chr6	12827137	C	T	366	0.93	35	intergenic				
chr6	22633758	C	T	5	0.83	111	intronic/noncoding	Aa_G129040.t1			
chr6	23676755	G	A	9	0.75	160	intergenic				
chr6	32682052	C	T	6	0.75	91	intergenic				
chr7	11473356	C	T	5	0.83	46	intergenic				
chr7	48150555	C	T	9	0.90	96	intergenic				
chr7	49346123	G	A	15	0.83	41	intronic/noncoding	Aa_G893800.t1			
chr8	911653	C	T	6	0.75	125	intergenic				
chr8	1003423	C	T	8	0.89	155	CDS	Aa_G85260.t1	Syn	L	L
chr8	1378693	C	T	5	0.71	98	CDS	Aa_G451890.t1	Syn	F	F
chr8	3619514	C	T	22	0.79	225	intronic/noncoding	Aa_G1098230.t1			
chr8	3619514	C	T	22	0.79	225	CDS	Aa_G209840.t1	Nonsyn	S	L
chr8	3769653	C	T	10	0.91	181	intronic/noncoding	Aa_G1098230.t1			
chr8	4254051	C	T	5	1.00	96	intergenic				
chr8	4442794	C	T	5	0.71	123	intergenic				
chr8	4499424	C	T	10	0.77	174	intergenic				
chr8	4717312	C	T	12	0.86	209	intergenic				
chr8	4795355	C	T	18	0.75	73	intergenic				
chr8	5164219	C	T	15	0.88	222	intergenic				
chr8	5426953	C	T	14	0.74	214	intronic/noncoding	Aa_G210860.t1			
chr8	5480047	C	T	15	0.94	222	CDS	Aa_G635950.t1	Nonsyn	E	K
chr8	5605368	C	T	19	0.73	225	CDS	Aa_G130430.t1	Syn	L	L
chr8	7671717	C	T	13	0.76	66	intergenic				
chr8	7697892	C	T	15	0.79	208	intronic/noncoding	Aa_G267490.t1			
chr8	7736172	C	T	10	0.91	190	intronic/noncoding	Aa_G756430.t1			
chr8	7824809	C	T	15	0.83	61	intronic/noncoding	Aa_G756430.t1			
chr8	7913277	C	T	20	0.91	225	CDS	Aa_G278980.t1	Nonsyn	G	D
chr8	8045084	C	T	24	0.80	225	intergenic				
chr8	8536874	C	T	18	0.95	222	intronic/noncoding	Aa_G952120.t1			
chr8	8751951	C	T	18	0.95	67	intergenic				

chr8	9074819	C	T	10	0.91	195	intergenic				
chr8	9438623	C	T	3	1.00	68	intergenic				
chr8	10085022	C	T	22	0.85	222	CDS	Aa_G72830.t1	Syn	R	R
chr8	10185903	C	T	21	0.91	225	intergenic				
chr8	10308489	C	T	17	0.94	222	CDS	Aa_G258210.t1	Nonsyn	R	K
chr8	10590870	C	T	10	0.91	143	intergenic				
chr8	11090942	C	T	23	0.92	164	CDS	Aa_G308210.t1	Syn	L	L
chr8	11236078	C	T	11	0.85	207	intronic/noncoding	Aa_G304340.t1			
chr8	11485342	C	T	6	0.75	102	intergenic				
chr8	11806942	C	T	14	1.00	127	intergenic				
chr8	11847184	C	T	11	0.92	82	intergenic				
chr8	11948923	C	T	18	0.86	40	intergenic				
chr8	12029485	C	T	5	1.00	92	intergenic				
chr8	12252782	C	T	26	0.84	95	intergenic				
chr8	12367437	C	T	18	0.95	222	intergenic				
chr8	12479954	C	T	9	1.00	176	intergenic				
chr8	12717465	C	T	15	0.88	190	intronic/noncoding	Aa_G530170.t1			
chr8	12971632	C	T	14	0.70	38	intergenic				
chr8	13036894	C	T	20	0.95	222	intergenic				
chr8	13706634	C	T	19	1.00	119	intergenic				
chr8	13741890	C	T	18	1.00	222	CDS	Aa_G56670.t1	Syn	K	K
chr8	13765439	C	T	18	0.95	177	intergenic				
chr8	14197298	C	T	17	0.85	120	intergenic				
chr8	14263283	C	T	11	0.79	220	CDS	Aa_G220420.t1	Syn	T	T
chr8	14605246	C	T	10	0.77	199	CDS	Aa_G106560.t1	Nonsyn	A	T
chr8	15080766	C	T	8	1.00	175	intergenic				
chr8	15107402	C	T	22	0.96	222	intronic/noncoding	Aa_G85860.t1			
chr8	15386685	C	T	21	0.95	222	intergenic				
chr8	15506044	C	T	12	0.86	187	intergenic				
chr8	15649583	C	T	16	0.89	131	intergenic				
chr8	16171021	C	T	5	1.00	51	intergenic				
chr8	16471306	C	T	21	1.00	222	intergenic				
chr8	16655981	C	T	15	1.00	222	CDS	Aa_G446950.t1	Nonsyn	R	K
chr8	16776493	C	T	6	1.00	149	intergenic				
chr8	16879065	C	T	18	0.95	177	intergenic				
chr8	16959013	C	T	19	0.73	33	intergenic				
chr8	17131084	C	T	22	0.88	222	intergenic				
chr8	17915436	C	T	16	0.94	138	intergenic				
chr8	19166187	C	T	9	0.90	154	intergenic				
chr8	19482487	C	T	20	0.80	37	intergenic				
chr8	19696634	C	T	12	1.00	209	intergenic				
chr8	20638488	C	T	18	0.90	222	intergenic				
chr8	20831090	C	T	23	0.88	50	intergenic				
chr8	21075537	C	T	5	1.00	98	intergenic				
chr8	21131233	C	T	12	1.00	222	intergenic				
chr8	21753623	C	T	13	0.76	222	intergenic				
chr8	22404296	C	T	29	0.97	222	intergenic				
chr8	22944560	C	T	8	1.00	194	intergenic				
chr8	23035704	C	T	21	0.91	222	intergenic				

chr8	23103505	C	T	14	0.82	124	intergenic				
chr8	25289947	C	T	7	1.00	133	intergenic				
chr8	27312087	C	T	10	1.00	159	intergenic				
chr8	28321474	C	T	21	0.84	222	intergenic				
chr8	28563235	C	T	5	1.00	131	intergenic				
chr8	28915285	G	A	20	1.00	157	intergenic				
chr8	29262216	C	T	18	0.90	222	intergenic				
chr8	29375488	C	T	18	0.90	222	intergenic				
chr8	29636277	C	T	16	0.89	222	intergenic				
chr8	30400344	C	T	19	0.70	202	intergenic				
chr8	30765602	C	T	17	1.00	222	intergenic				
chr8	30896577	C	T	19	0.83	222	CDS	Aa_G84120.t1	Nonsyn	P	L
chr8	31376787	C	T	21	0.88	156	intergenic				
chr8	31566057	C	T	19	0.86	51	intergenic				
chr8	32054886	C	T	23	0.92	191	intergenic				
chr8	32273877	C	T	6	1.00	52	intergenic				
chr8	32839589	C	T	17	0.81	104	intergenic				
chr8	34012166	C	T	11	0.92	222	CDS	Aa_G206860.t1	Syn	V	V
chr8	34585542	C	T	13	0.87	223	intergenic				
chr8	35656431	C	T	18	1.00	222	CDS	Aa_G54540.t1	Nonsyn	P	S
chr8	35943035	C	T	20	1.00	137	intergenic				
chr8	36020416	C	T	18	0.75	222	intergenic				
chr8	36228954	C	T	11	0.79	226	CDS	Aa_G149040.t1	Nonsyn	V	I
chr8	36400342	C	T	7	1.00	158	intergenic				
chr8	37010535	C	T	24	0.96	222	intergenic				
chr8	37041592	C	T	21	1.00	222	intergenic				
chr8	38166222	C	T	3	0.75	81	intergenic				
chr8	38787747	C	T	19	0.95	222	intergenic				
chr8	38942858	C	T	7	0.88	145	intergenic				
chr8	38975713	C	T	11	0.92	206	intergenic				
chr8	39539501	C	T	25	0.83	222	CDS	Aa_G227720.t1	Nonsyn	D	N
chr8	39948203	C	T	18	0.82	225	intronic/noncoding	Aa_G70060.t1			
chr8	39948203	C	T	18	0.82	225	CDS	Aa_G93730.t1	Nonsyn	E	K
chr8	41003710	C	T	11	0.85	168	CDS	Aa_G482910.t1	Nonsyn	Q	*
chr8	41489315	C	T	16	0.80	174	CDS	Aa_G456980.t1	Syn	L	L
chr8	42072451	C	T	19	0.73	121	intergenic				
chr8	42112960	C	T	19	0.83	42	intergenic				
chr8	42372423	C	T	22	0.81	121	intergenic				
chr8	42512363	C	T	13	0.93	195	intergenic				
chr8	42699580	C	T	13	0.81	225	intronic/noncoding	Aa_G9070.t1			
chr8	43994243	C	T	3	0.75	56	intergenic				
chr8	44037145	C	T	15	0.79	112	intergenic				
chr8	44830330	C	T	13	0.76	206	intergenic				
chr8	46586988	G	A	467	0.92	94	intergenic				
chr8	46587047	G	A	165	0.87	36	intergenic				
chr8	46921251	G	A	17	0.74	225	CDS	Aa_G13750.t1	Nonsyn	T	I
chr8	48986330	G	A	18	0.72	77	intergenic				
chr8	49159511	G	A	3	0.75	55	intergenic				

Supplementary Table 3. EMS mutations with in *eop091*

Chr [1]	Pos [2]	R[3]	M[4]	N[5]	AF[6]	Mq[7]	Region[8]	Gene ID[9]	Type[10]	AR[11]	AM[12]
chr2	11159428	C	T	94	0.95	222	intergenic				
chr2	20175781	G	A	5	0.71	89	intergenic				
chr4	40154927	G	A	3	1.00	38	intergenic				
chr6	4911929	G	A	5	1.00	31	intergenic				
chr6	6256569	G	A	8	0.73	37	intergenic				
chr6	8903027	G	A	5	0.83	59	intergenic				
chr7	48150555	C	T	92	0.77	48	intergenic				
chr8	6406645	C	T	53	0.76	225	CDS	Aa_G98160.t1	Nonsyn	G	E
chr8	6803653	C	T	36	0.72	159	intergenic				
chr8	7046265	C	T	71	0.82	225	intronic/noncoding	Aa_G765980.t1			
chr8	7540193	C	T	75	0.78	225	CDS	Aa_G116810.t1	Syn	G	G
chr8	7563151	C	T	80	0.83	226	intronic/noncoding	Aa_G116860.t1			
chr8	7691667	C	T	46	0.82	225	intronic/noncoding	Aa_G267480.t1			
chr8	8272741	C	T	113	0.89	222	CDS	Aa_G256480.t1	Nonsyn	C	Y
chr8	8491434	C	T	101	0.86	222	CDS	Aa_G138250.t1	Nonsyn	A	T
chr8	8565397	C	T	67	0.83	222	intronic/noncoding	Aa_G138370.t1			
chr8	8565397	C	T	67	0.83	222	intronic/noncoding	Aa_G952120.t1			
chr8	8714751	C	T	53	0.85	225	intergenic				
chr8	9860455	C	T	60	0.87	222	intergenic				
chr8	10483249	C	T	63	0.90	222	intergenic				
chr8	11146907	C	T	35	0.90	222	intergenic				
chr8	11891244	C	T	70	0.93	62	intergenic				
chr8	11974246	C	T	150	0.88	222	CDS	Aa_G356670.t1	Syn	E	E
chr8	12130775	C	T	56	0.93	222	intergenic				
chr8	12364202	C	T	87	0.97	133	intergenic				
chr8	12491338	C	T	61	0.94	58	intergenic				
chr8	12783685	C	T	67	0.97	222	intergenic				
chr8	12957212	C	T	31	0.94	222	intergenic				
chr8	13065735	C	T	45	1.00	222	intergenic				
chr8	14286417	C	T	37	0.97	222	intergenic				
chr8	14532392	C	T	22	0.92	178	intergenic				
chr8	14606418	C	T	106	0.97	222	CDS	Aa_G106560.t1	Nonsyn	E	K
chr8	14732900	C	T	67	0.93	222	intergenic				
chr8	15068358	C	T	40	0.93	222	intronic/noncoding	Aa_G70350.t1			
chr8	15073001	C	T	40	0.95	222	intergenic				
chr8	15397035	C	T	43	0.96	222	intergenic				
chr8	15678994	C	T	52	0.96	222	intergenic				
chr8	15702887	C	T	152	0.96	222	intergenic				
chr8	15906089	C	T	114	0.95	222	CDS	Aa_G410990.t1	Nonsyn	R	K
chr8	16340908	C	T	89	1.00	222	intergenic				
chr8	16350326	C	T	92	0.93	222	intergenic				
chr8	16511796	C	T	35	0.88	222	intergenic				
chr8	16646555	C	T	90	0.90	222	intergenic				
chr8	16980761	C	T	42	0.93	222	intronic/noncoding	Aa_G354520.t1			
chr8	17509577	C	T	62	0.93	222	intergenic				

chr8	17658057	C	T	77	0.96	172	intergenic				
chr8	18071503	C	T	50	0.70	158	intergenic				
chr8	18562533	C	T	52	0.90	212	intergenic				
chr8	18680954	C	T	73	0.91	222	intergenic				
chr8	19034042	C	T	88	0.99	190	intergenic				
chr8	19483314	C	T	76	0.96	207	intergenic				
chr8	20514723	C	T	38	0.86	222	intergenic				
chr8	21190629	C	T	5	1.00	80	intergenic				
chr8	21332817	C	T	33	0.87	138	intergenic				
chr8	21694861	C	T	50	0.89	222	intergenic				
chr8	22479025	C	T	113	0.89	222	intergenic				
chr8	22650747	C	T	52	0.80	125	intergenic				
chr8	22745814	C	T	108	0.89	206	intergenic				
chr8	22757391	C	T	54	0.89	222	CDS	Aa_G425350.t1	Nonsyn	G	R
chr8	22993986	C	T	115	0.95	222	intronic/noncoding	Aa_G90320.t1			
chr8	23141386	C	T	68	0.93	222	intergenic				
chr8	23688874	C	T	62	0.95	111	intergenic				
chr8	24263506	C	T	67	0.94	222	intergenic				
chr8	25224117	C	T	94	0.91	73	intergenic				
chr8	25226879	C	T	73	0.94	222	intergenic				
chr8	25338271	C	T	51	0.98	222	intergenic				
chr8	25619560	C	T	44	0.92	222	intergenic				
chr8	25746155	C	T	84	0.90	222	intergenic				
chr8	25823384	C	T	89	0.94	222	intergenic				
chr8	25861568	C	T	40	0.78	222	intergenic				
chr8	26439370	C	T	121	0.89	222	intergenic				
chr8	26452056	C	T	63	0.95	222	CDS	Aa_G382520.t1	Syn	L	L
chr8	26571255	C	T	102	0.92	222	intergenic				
chr8	26993562	C	T	51	0.93	222	intergenic				
chr8	27310384	C	T	69	0.70	191	intergenic				
chr8	27632236	C	T	71	0.95	222	CDS	Aa_G286990.t1	Nonsyn	R	Q
chr8	27922253	C	T	76	0.96	222	intergenic				
chr8	28449335	C	T	31	0.97	222	intergenic				
chr8	28549358	C	T	79	0.96	122	intergenic				
chr8	28921034	C	T	21	0.95	222	intergenic				
chr8	29232127	C	T	60	1.00	222	intergenic				
chr8	29326092	C	T	76	0.92	222	CDS	Aa_G442830.t1	Syn	F	F
chr8	29475586	C	T	40	0.98	222	intergenic				
chr8	29508828	C	T	48	0.92	194	intergenic				
chr8	29772077	C	T	42	0.95	222	intergenic				
chr8	29818950	G	A	3	0.75	56	intergenic				
chr8	29835075	C	T	11	0.73	46	intergenic				
chr8	30449182	C	T	70	0.96	222	intergenic				
chr8	30675652	C	T	105	0.96	154	intergenic				
chr8	31094473	G	A	43	0.88	222	intergenic				
chr8	31274952	C	T	86	0.97	222	splice_site_change	Aa_G101870.t1			
chr8	31339054	C	T	124	0.89	222	intergenic				

chr8	31599424	C	T	60	0.87	222	intergenic				
chr8	31823620	C	T	121	0.94	222	CDS	Aa_G375200.t1	Nonsyn	V	M
chr8	31979459	C	T	49	0.89	170	intergenic				
chr8	32305647	C	T	66	0.94	97	intergenic				
chr8	32604864	C	T	115	0.91	212	intergenic				
chr8	32754835	C	T	93	0.96	222	intergenic				
chr8	33123549	C	T	4	1.00	106	intergenic				
chr8	33138029	C	T	21	0.95	222	intergenic				
chr8	33161409	C	T	71	0.90	222	intergenic				
chr8	33426038	C	T	85	0.96	222	intergenic				
chr8	33649847	C	T	97	0.91	125	intergenic				
chr8	33890849	C	T	55	0.92	222	intergenic				
chr8	34025448	C	T	84	0.84	110	intergenic				
chr8	34138144	C	T	95	0.89	222	intergenic				
chr8	34170071	C	T	65	0.90	222	intergenic				
chr8	34908744	C	T	62	0.97	222	intergenic				
chr8	35032991	C	T	67	0.82	125	intronic/noncoding	Aa_G496120.t1			
chr8	35145947	C	T	70	0.96	222	CDS	Aa_G141180.t1	Nonsyn	P	S
chr8	35145947	C	T	70	0.96	222	intronic/noncoding	Aa_G496120.t1			
chr8	36210157	C	T	95	0.93	222	CDS	Aa_G149090.t1	Nonsyn	Q	*
chr8	36365383	C	T	57	0.95	222	intergenic				
chr8	36762060	C	T	75	0.87	206	intergenic				
chr8	37056748	C	T	54	0.84	79	intronic/noncoding	Aa_G673820.t1			
chr8	37090032	C	T	34	1.00	222	intergenic				
chr8	37476117	C	T	67	0.91	222	intergenic				
chr8	37533206	C	T	32	1.00	222	intergenic				
chr8	37682754	C	T	80	0.91	222	intergenic				
chr8	38292125	C	T	36	0.95	222	intergenic				
chr8	38391087	C	T	73	0.90	222	intergenic				
chr8	38710690	C	T	42	0.82	94	intergenic				
chr8	39132239	G	A	92	0.87	222	CDS	Aa_G307370.t1	Nonsyn	G	R
chr8	39311444	C	T	42	0.95	222	intergenic				
chr8	39413624	C	T	79	0.80	222	intergenic				
chr8	39942706	C	T	52	0.85	222	intronic/noncoding	Aa_G70060.t1			
chr8	40154866	C	T	68	0.88	222	intronic/noncoding	Aa_G148050.t1			
chr8	40161925	C	T	58	0.95	222	intergenic				
chr8	40380808	C	T	26	0.96	222	intergenic				
chr8	40501566	C	T	62	0.87	222	intergenic				
chr8	40502985	C	T	62	0.75	225	intergenic				
chr8	40976101	C	T	59	0.87	222	CDS	Aa_G482940.t1	Nonsyn	L	F
chr8	42353001	C	T	22	0.92	222	CDS	Aa_G447600.t1	Syn	K	K
chr8	42362504	C	T	75	0.83	179	intergenic				
chr8	42469941	C	T	44	0.88	105	intergenic				
chr8	42622712	C	T	78	0.76	225	CDS	Aa_G8930.t1	Syn	Q	Q
chr8	43244010	C	T	65	0.86	225	intronic/noncoding	Aa_G15680.t1			
chr8	43493190	C	T	54	0.76	225	CDS	Aa_G230450.t1	Nonsyn	S	N
chr8	43844783	C	T	37	0.80	225	CDS	Aa_G113030.t1	Nonsyn	A	V

chr8	44841964	C	T	83	0.80	225	CDS	Aa_G45400.t1	Nonsyn	D	N
chr8	45570829	C	T	83	0.72	225	CDS	Aa_G104880.t1	Nonsyn	G	R
chr8	45574139	C	T	35	0.76	225	intergenic				
chr8	46138848	C	T	61	0.72	225	CDS	Aa_G448760.t1	Nonsyn	P	S
chr8	46303198	C	T	28	0.72	113	intergenic				
chr8	46337274	C	T	65	0.73	225	intronic/noncoding	Aa_G401160.t1			
chr8	46586988	G	A	1804	0.90	68	intergenic				
chr8	46587047	G	A	625	0.92	34	intergenic				

1) Chr: Chromosome. **(2)** Pos: Mutated nucleotide position. **(3)** R: Nucleotide in reference genome. **(4)** M: Nucleotide in mutant genome. **(5)** N: Number of reads supporting the mutation. **(6)** AF: Allele frequency. **(7)** Mq: Mapping quality (MAPQ), higher = more unique (min. 30, implies a 0.001 probability of false read alignment). **(8)** Region: Region affected by the mutation. **(9)** Gene ID: Gene identifier. **(10** Type: Mutation type (nonsynonymous or synonymous). **(11)** AR: Amino acid in reference genome. **(12)** AM: Amino acid in mutant genome.

Supplementary Table 4. Genes with EMS changes within *eop002* mapped intervals

Chr [1]	Pos [2]	Gene ID[3]	NCBI ID[4]	Arabidopsis gene[5]	Gene description[6]	R[7]	M[8]	N[9]	AF[10]	Mq[11]	Region[12]	Type[13]	AR[14]	AM[15]
Aa.chr2	7322892	Aa_G62080	AALP_AA4G015900	AT1G54490	ACC INSENSITIVE 1, AIN1, ATXRN4, EIN5, ETHYLENE INSENSITIVE 5, EXORIBONUCLEASE 4, XRN4. Involved in the ethylene response.	G	A	79	0.85	225	CDS	Nonsyn	G	S
Aa.chr2	7360332	Aa_G711220	AALP_AAs43532U000300	NA	NA	G	A	97	0.71	40	intronic/noncoding			
Aa.chr2	7855399	Aa_G510360	AALP_AAs68695U000100	NA	NA	G	A	96	0.91	222	CDS	Nonsyn	V	M
Aa.chr2	13141687	Aa_G100860	AALP_AA2G059200	AT1G51172	Hypothetical protein	C	T	219	0.77	158	CDS	Nonsyn	R	Q
Aa.chr2	15821900	Aa_G201310	AALP_AA2G078300	AT1G66210	Subtilisin-like serine endopeptidase family protein; FUNCTIONS IN: identical protein binding, serine-type endopeptidase activity; INVOLVED IN: proteolysis, negative regulation of catalytic activity.	G	A	143	0.85	222	CDS	Nonsyn	T	I
Aa.chr2	25564239	Aa_G562950	AALP_AA2G073200	AT1G65580	FRAGILE FIBER3 (FRA3); Endonuclease/exonuclease/phosphatase family protein. FUNCTIONS IN: inositol or phosphatidylinositol phosphatase activity.	G	A	29	0.94	222	intronic/noncoding			
Aa.chr8	9313334	Aa_G526410	AALP_AA8G136900	AT5G13600	Phototropic-responsive NPH3 family protein; FUNCTIONS IN: signal transducer activity; INVOLVED IN: response to light stimulus.	C	T	89	0.92	222	CDS	Nonsyn	S	F
Aa.chr8	9880210	Aa_G290000	AALP_AA8G139400	AT5G13920	GRF zinc finger / Zinc knuckle protein; FUNCTIONS IN: zinc ion binding, nucleic acid binding.	C	T	81	0.88	222	CDS	Nonsyn	A	T
Aa.chr8	10331702	Aa_G258170	AALP_AA8G145100	AT5G14390	Alpha/beta-Hydrolases superfamily protein; FUNCTIONS IN: molecular_function unknown.	C	T	22	1	222	intronic/noncoding			
Aa.chr8	10715716	Aa_G7880	AALP_AA8G149400	AT5G14720	Protein kinase superfamily protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation.	C	T	51	0.84	222	CDS	Nonsyn	G	E
Aa.chr8	10887335	Aa_G644350	AALP_AA8G151800	AT5G14930	SAG101, SENESCENCE-ASSOCIATED GENE 101, encodes an acyl hydrolase involved in senescence	C	T	56	0.97	222	intronic/noncoding			

Aa.chr8	11634026	Aa_G236930	AALP_AA8G160600	AT5G15450	ALBINO AND PALE GREEN 6, APG6, ATCLPB3, CASEIN LYTIC PROTEINASE B-P, CASEIN LYTIC PROTEINASE B3, CLPB-P, CLPB3	C	T	63	0.94	222	intronic/noncoding			
Aa.chr8	13202635	Aa_G79140	AALP_AA8G177600	AT5G16700	Glycosyl hydrolase superfamily protein; FUNCTIONS IN: cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process.	G	A	58	0.94	222	CDS	Nonsyn	G	S
Aa.chr8	14606687	Aa_G106560	AALP_AA8G190000	AT5G17760	P-loop containing nucleoside triphosphate hydrolases superfamily protein; FUNCTIONS IN: nucleoside-triphosphatase activity, nucleotide binding, ATP binding.	C	T	65	0.98	222	CDS	Nonsyn	G	D
Aa.chr8	18145960	Aa_G255500	AALP_AA8G211700	AT2G18900	Transducin/WD40 repeat-like superfamily protein.	C	T	85	0.98	222	CDS	Nonsyn	L	F
Aa.chr8	20875190	Aa_G429030	AALP_AA8G257900	AT5G27870	Plant invertase/pectin methylesterase inhibitor superfamily; FUNCTIONS IN: enzyme inhibitor activity, pectinesterase activity; INVOLVED IN: cell wall modification.	C	T	96	0.95	222	CDS	Nonsyn	P	L
Aa.chr8	21373465	Aa_G25950	AALP_AA2G192200	NA	NA	C	T	174	0.95	103	intronic/noncoding			
Aa.chr8	21989344	Aa_G251110	AALP_AA8G253000	AT5G27270	EMBRYO DEFECTIVE 976 (EMB976); INVOLVED IN: biological_process unknown.	C	T	90	0.96	222	CDS	Syn	L	L
Aa.chr8	22620391	Aa_G891490	AALP_AAs39126U000100	ATMG00860	ORF158, DNA/RNA polymerases superfamily protein	C	T	40	0.70	48	intronic/noncoding			
Aa.chr8	23628464	Aa_G1094820	AALP_AAs48402U000100	AT5G16715	EMB2247, EMBRYO DEFECTIVE 2247. Involved in embryo development ending in seed dormancy	C	T	33	0.94	202	intronic/noncoding			
Aa.chr8	26477645	Aa_G23190	AALP_AA6G090400	NA	NA	C	T	121	0.92	222	intronic/noncoding			
Aa.chr8	27475487	Aa_G102190	AALP_AA6G096000	AT5G48270	Plant protein of unknown function (DUF868).	C	T	127	0.95	222	CDS	Syn	G	G
Aa.chr8	29940015	Aa_G188010	AALP_AA8G282300	AT5G50770	ATHSD6, HSD6, HYDROXYSTEROID DEHYDROGENASE 6. Encodes a putative hydroxysteroid dehydrogenase (HSD).	C	T	64	0.97	222	intronic/noncoding			
Aa.chr8	29975495	Aa_G424180	AALP_AA8G282900	AT5G50790	ATSWEET10, SWEET10, Encodes a member of the SWEET sucrose efflux transporter family proteins.	C	T	48	0.94	222	intronic/noncoding			
Aa.chr8	30953322	Aa_G675250	AALP_AA8G288500	AT5G51480	SKU5 similar 2 (SKS2); FUNCTIONS IN: oxidoreductase activity, copper ion binding; INVOLVED IN: oxidation reduction.	C	T	74	0.93	222	CDS	Syn	N	N

Aa.chr8	33619479	Aa_G636600	AALP_AA8G309200	AT5G53040	GRD, GROUNDED, RKD4, RWP-RK DOMAIN-CONTAINING 4, Encodes GROUNDED (GRD), a putative RWP-RK-type transcription factor broadly expressed in early development.	C	T	40	0.87	222	CDS	Nonsyn	V	I
Aa.chr8	33759002	Aa_G175520	AALP_AA8G311000	NA	NA	C	T	147	0.86	222	CDS	Nonsyn	D	N
Aa.chr8	35005804	Aa_G496120	AALP_AA8G325600	NA	NA	C	T	8	1.00	105	intronic/noncoding			
Aa.chr8	35291533	Aa_G88110	AALP_AA8G324300	AT5G54130	Calcium-binding endonuclease/exonuclease/phosphatase family; FUNCTIONS IN: calcium ion binding.	C	T	110	0.96	222	CDS	Nonsyn	P	S
Aa.chr8	36216713	Aa_G149070	AALP_AA8G334300	AT5G54870	unknown protein; FUNCTIONS IN: molecular_function unknown.	C	T	115	0.94	222	CDS	Syn	N	N
Aa.chr8	36252343	Aa_G542640	AALP_AA8G332700	AT1G67120	ATPases;nucleotide binding;ATP binding;nucleoside-triphosphatases;transcription factor binding; FUNCTIONS IN: nucleoside-triphosphatase activity, ATPase activity, transcription factor binding, nucleotide binding, ATP binding; INVOLVED IN: regulation of transcription, DNA-dependent, regulation of protein complex assembly.	C	T	99	0.91	222	CDS	Nonsyn	E	K
Aa.chr8	36449900	Aa_G605090	AALP_AA8G336400	AT5G55050	GDSL-like Lipase/Acylhydrolase superfamily protein; FUNCTIONS IN: hydrolase activity, acting on ester bonds, lipase activity, carboxylesterase activity; INVOLVED IN: lipid metabolic process.	C	T	41	0.98	222	intronic/noncoding			
Aa.chr8	36751839	Aa_G59080	AALP_AA7G036500	AT5G55390	Encodes EDM2 (enhanced downy mildew 2). The predicted protein bears typical features of transcriptional regulators.	C	T	44	0.81	222	CDS	Nonsyn	S	N
Aa.chr8	37039767	Aa_G571550	AALP_AA8G342000	AT5G55840	Pentatricopeptide repeat (PPR) superfamily protein.	C	T	92	0.93	222	CDS	Syn	G	G
Aa.chr8	38300011	Aa_G22070	AALP_AA8G354900	AT5G56680	EMB2755, EMBRYO DEFECTIVE 2755, SYNC1, SYNC1 ARATH, Encodes a putative cytosolic asparaginyl-tRNA synthetase. The mRNA is cell-to-cell mobile	C	T	85	0.89	222	intronic/noncoding			
Aa.chr8	39075205	Aa_G69080	AALP_AA8G361000	AT5G57210	Ypt/Rab-GAP domain of gyp1p superfamily protein; FUNCTIONS IN: RAB GTPase activator activity; INVOLVED IN: regulation of Rab GTPase activity.	C	T	86	0.88	222	CDS	Syn	K	K
Aa.chr8	39130314	Aa_G307370	AALP_AA8G362100	AT5G57330	Galactose mutarotase-like superfamily protein; FUNCTIONS IN: isomerase activity, carbohydrate binding, aldose 1-epimerase activity, catalytic activity; INVOLVED IN: galactose metabolic process, carbohydrate	C	T	69	0.91	222	intronic/noncoding			

Aa.chr8	39796839	Aa_G31530	AALP_AA8G370800	AT5G58110	Chaperone binding;ATPase activators; FUNCTIONS IN: ATPase activator activity, chaperone binding; INVOLVED IN: biological_process unknown.	C	T	53	0.9	226	intronic/noncoding			
Aa.chr8	39953425	Aa_G70060	AALP_AA4G167100	NA	NA	C	T	66	0.88	222	intronic/noncoding			
Aa.chr8	39953425	Aa_G93740	AALP_AA8G374400	AT5G58450	Tetratricopeptide repeat (TPR)-like superfamily protein; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown.	C	T	66	0.88	222	intronic/noncoding			
Aa.chr8	40010585	Aa_G93880	AALP_AA8G375800	AT5G58560	FARNESOL KINASE, FOLK. FOLK is a farnesol kinase that can phosphorylate farnesol using an NTP donor. It can also phosphorylate geraniol, or geranylgeraniol, but it prefers farnesol in experiments performed using yeast membranes.	C	T	86	0.84	222	CDS	Nonsyn	T	I
Aa.chr8	40391906	Aa_G345460	AALP_AA8G381800	AT1G25350	OVA9, OVULE ABORTION 9. INVOLVED IN: Glutaminyl-tRNA aminoacylation, glutamyl-tRNA aminoacylation, plant ovule development, translation.	C	T	78	0.84	225	CDS	Nonsyn	A	V

(1) Chr: Chromosome. **(2)** Pos: Mutated nucleotide position. **(3)** Gene ID: Gene identifier. **(4)** NCBI ID: Gene identifier as on NCBI. **(5)** Arabidopsis gene: Orthologues of the *A. alpina* genes in *A. thaliana*. . **(6)** Gene description description of the *A. thaliana* orthologue as given by The Arabidopsis Information Resource (TAIR, www.arabidopsis.org), NA indicates an orthologue is not available. **(7)** R: Nucleotide in reference genome. **(8)** M: Nucleotide in mutant genome. **(9)** N: Number of reads supporting the mutation. **(10)** AF: Allele frequency. **(11)** Mq: Mapping quality (MAPQ), higher = more unique (min. 30, implies a 0.001 probability of false read alignment). **(12)** Region: Region affected by the mutation. **(13)** Type: Mutation type (nonsynonymous or synonymous). **(14)** AR: Amino acid in reference genome. **(15)** AM: Amino acid in mutant genome.

Supplementary Table 5. Genes with EMS changes within *eop085* mapped intervals

Chr [1]	Pos [2]	Gene ID[3]	NCBI ID[4]	Arabidopsis gene[5]	Gene description[6]	R[7]	M[8]	N[9]	AF[10]	Mq[11]	Region[12]	Type[13]	AR[14]	AM[15]
Aa.chr8	7697892	Aa_G267490	AALP_AA8G113300	AT5G11650	Alpha/beta-Hydrolases superfamily protein; FUNCTIONS IN: hydrolase activity.	C	T	15	0.79	208	intronic/noncoding			
Aa.chr8	7736172	Aa_G756430	AALP_AA8G115700	AT1G66900	Alpha/beta-Hydrolases superfamily protein	C	T	10	0.91	190	intronic/noncoding			
Aa.chr8	7913277	Aa_G278980	AALP_AA8G116400	AT5G11370	FBD / Leucine Rich Repeat domains containing protein	C	T	20	0.91	225	CDS	Nonsyn	G	D
Aa.chr8	8536874	Aa_G952120	AALP_AAs50802U000100	NA	NA	C	T	18	0.95	222	intronic/noncoding			
Aa.chr8	10085022	Aa_G72830	AALP_AA8G141800	AT5G14170	BAF60, CHC1, SWP73B. CHC1 is predicted to encode a protein that belongs to the chromodomain remodeling complex.	C	T	22	0.85	222	CDS	Syn	R	R
Aa.chr8	10308489	Aa_G258210	AALP_AA8G144700	AT3G46610	Pentatricopeptide repeat (PPR-like) superfamily protein.	C	T	17	0.94	222	CDS	Nonsyn	R	K
Aa.chr8	11090942	Aa_G308210	AALP_AA8G153300	AT1G51172	Hypothetical protein	C	T	23	0.92	164	CDS	Syn	L	L
Aa.chr8	11236078	Aa_G304340	AALP_AA8G153900	AT5G15070	Phosphoglycerate mutase-like family protein; FUNCTIONS IN: oxidoreductase activity, transition metal ion binding, acid phosphatase activity; INVOLVED IN: oxidation reduction.	C	T	11	0.85	207	intronic/noncoding			
Aa.chr8	12717465	Aa_G530170	AALP_AAs56912U000300	NA	NA	C	T	15	0.88	190	intronic/noncoding			
Aa.chr8	13741890	Aa_G56670	AALP_AA2G144800	AT5G17320	HDG9, HOMEODOMAIN GLABROUS 9. Encodes a homeobox-leucine zipper family protein belonging to the HD-ZIP IV family.	C	T	18	1	222	CDS	Syn	K	K
Aa.chr8	14263283	Aa_G220420	AALP_AA8G187200	AT5G17540	HXXXD-type acyl-transferase family protein. Has transferase activity, transferring acyl groups other than amino-acyl groups.	C	T	11	0.79	220	CDS	Syn	T	T
Aa.chr8	14605246	Aa_G106560	AALP_AA8G190000	AT5G17760	P-loop containing nucleoside triphosphate hydrolases superfamily protein; FUNCTIONS IN: nucleoside-triphosphatase activity, nucleotide binding, ATP binding.	C	T	10	0.77	199	CDS	Nonsyn	A	T
Aa.chr8	15107402	Aa_G85860	AALP_AA8G193400	AT3G23890	ATTOPII, TOPII, TOPOISOMERASE II. Encodes a topoisomerase II that is highly expressed in young seedlings. The protein is localized in the nucleus and gene expression levels are increased in proliferative tissues.	C	T	22	0.96	222	intronic/noncoding			
Aa.chr8	16655981	Aa_G446950	AALP_AA8G203300	AT5G18660	PALE-GREEN AND CHLOROPHYLL B REDUCED 2, PCB2. Encodes a protein with 3,8-divinyl protochlorophyllide a 8-vinyl reductase activity.	C	T	15	1	222	CDS	Nonsyn	R	K

Aa.chr8	30896577	Aa_G84120	AALP_AA8G287400	AT5G51340	Tetratricopeptide repeat (TPR)-like superfamily protein.	C	T	19	0.83	222	CDS	Nonsyn	P	L
Aa.chr8	34012166	Aa_G206860	AALP_AA8G314200	AT3G44480	COG1, RECOGNITION OF PERONOSPORA PARASITICA 1, RPP1. Encodes a TIR-NB-LRR R-protein RPP1 that confers resistance to Peronospora parasitica (downy mildew).	C	T	11	0.92	222	CDS	Syn	V	V
Aa.chr8	35656431	Aa_G54540	AALP_AA8G327700	AT5G56810	F-box/RNI-like/FBD-like domains-containing protein	C	T	18	1.00	222	CDS	Nonsyn	P	S
Aa.chr8	36228954	Aa_G149040	AALP_AA8G334600	AT5G54910	DEA(D/H)-box RNA helicase family protein; FUNCTIONS IN: helicase activity, nucleic acid binding, ATP binding, ATP-dependent helicase activity.	C	T	11	0.79	226	CDS	Nonsyn	V	I
Aa.chr8	39539501	Aa_G227720	AALP_AA8G365800	AT5G57685	ARABIDOPSIS THALIANA GLUTAMINE DUMPER 3, ATGDU3, GDU3, GLUTAMINE DUMPER 3, LESS SUSCEPTIBLE TO BSCTV 1, LSB1. Encodes a member of the GDU (glutamine dumper) family proteins involved in amino acid export.	C	T	25	0.83	222	CDS	Nonsyn	D	N
Aa.chr8	39948203	Aa_G93730	AALP_AA8G374300	AT5G58440	Sorting nexin 2A (SNX2a); FUNCTIONS IN: phosphoinositide binding; INVOLVED IN: signal transduction, intracellular signaling pathway.	C	T	18	0.82	225	CDS	Nonsyn	E	K
Aa.chr8	39948203	Aa_G70060	AALP_AA4G167100	NA	NA	C	T	18	0.82	225	intronic/noncoding			
Aa.chr8	41003710	Aa_G482910	AALP_AA8G390800	AT5G59680	Leucine-rich repeat protein kinase family protein. INVOLVED IN: protein phosphorylation	C	T	11	0.85	168	CDS	Nonsyn	Q	*
Aa.chr8	41489315	Aa_G456980	AALP_AA8G400100	AT3G45330	Concanavalin A-like lectin protein kinase family protein; FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation.	C	T	16	0.8	174	CDS	Syn	L	L

(1) Chr: Chromosome. **(2)** Pos: Mutated nucleotide position. **(3)** Gene ID: Gene identifier. **(4)** NCBI ID: Gene identifier as on NCBI. **(5)** Arabidopsis gene: Orthologues of the *A. alpina* genes in *A. thaliana*. . **(6)** Gene description description of the *A. thaliana* orthologue as given by The Arabidopsis Information Resource (TAIR, www.arabidopsis.org), NA indicates an orthologue is not available. **(7)** R: Nucleotide in reference genome. **(8)** M: Nucleotide in mutant genome. **(9)** N: Number of reads supporting the mutation. **(10)** AF: Allele frequency. **(11)** Mq: Mapping quality (MAPQ), higher = more unique (min. 30, implies a 0.001 probability of false read alignment). **(12)** Region: Region affected by the mutation. **(13)** Type: Mutation type (nonsynonymous or synonymous). **(14)** AR: Amino acid in reference genome. **(15)** AM: Amino acid in mutant genome.

Supplementary Table 6. Genes with EMS changes within *eop091* mapped intervals

Chr [1]	Pos [2]	Gene ID[3]	NCBI ID[4]	Arabidopsis gene[5]	Gene description[6]	R[7]	M[8]	N[9]	AF[10]	Mq[11]	Region[12]	Type[13]	AR[14]	AM[15]
Aa.chr8	6406645	Aa_G98160	AALP_AA8G097000	AT5G44050	MATE efflux family protein; FUNCTIONS IN: antiporter activity, drug transmembrane transporter activity, transporter activity; INVOLVED IN: drug transmembrane transport, transmembrane transport.	C	T	53	0.76	225	CDS	Nonsyn	G	E
Aa.chr8	7046265	Aa_G765980	AALP_AA8G105800	AT5G11060	A member of Class II KN1-like homeodomain transcription factors (together with KNAT3 and KNAT5), with greatest homology to the maize knox1 homeobox protein. Expression regulated by light.	C	T	71	0.82	225	intronic/noncoding			
Aa.chr8	7540193	Aa_G116810	AALP_AA8G111100	AT5G11470	Bromo-adjacent homology (BAH) domain-containing protein; FUNCTIONS IN: DNA binding, nucleic acid binding; INVOLVED IN: biological_process unknown.	C	T	75	0.78	225	CDS	Syn	G	G
Aa.chr8	7563151	Aa_G116860	AALP_AA8G111600	AT5G11520	ASP3, ASPARTATE AMINOTRANSFERASE 3, YELLOW-LEAF-SPECIFIC GENE 4, YLS4. Encodes the chloroplastic isozyme of aspartate aminotransferase. Involved in aspartate biosynthesis and nitrogen metabolism. mRNA is expressed in senescing leaves.	C	T	80	0.83	226	intronic/noncoding			
Aa.chr8	7691667	Aa_G267480	AALP_AA8G113200	AT5G11640	Thioredoxin superfamily protein.	C	T	46	0.82	225	intronic/noncoding			
Aa.chr8	8272741	Aa_G256480	AALP_AA8G121100	AT5G12120	Ubiquitin-associated/translation elongation factor EF1B protein; INVOLVED IN: biological_process unknown.	C	T	113	0.89	222	CDS	Nonsyn	C	Y
Aa.chr8	8491434	Aa_G138250	AALP_AA8G124100	AT2G15690	Tetratricopeptide repeat (TPR)-like superfamily protein.	C	T	101	0.86	222	CDS	Nonsyn	A	T
chr8	8565397	Aa_G138370	AALP_AA8G125100	AT5G12420	O-acyltransferase (WSD1-like) family protein. INVOLVED IN: glycerolipid biosynthetic process	C	T	67	0.83	222	intronic/noncoding			
chr8	8565397	Aa_G952120	AALP_AAs50802U000100	NA	NA	C	T	67	0.83	222	intronic/noncoding			
Aa.chr8	11974246	Aa_G356670	AALP_AA8G166000	AT5G15850	ATCOL1, B-BOX DOMAIN PROTEIN 2, BBX2, COL1, CONSTANS-LIKE 1. Homologous to the flowering-time gene CONSTANS.	C	T	150	0.88	222	CDS	Syn	E	E
Aa.chr8	14606418	Aa_G106560	AALP_AA8G190000	AT5G17760	P-loop containing nucleoside triphosphate hydrolases superfamily protein; FUNCTIONS IN: nucleoside-triphosphatase activity, nucleotide binding, ATP	C	T	106	0.97	222	CDS	Nonsyn	E	K

chr8	15068358	Aa_G70350	AALP_AA8G192700	NA	NA	C	T	40	0.93	222	intronic/noncoding			
Aa.chr8	15906089	Aa_G410990	AALP_AA8G198200	AT5G18240	ARABIDOPSIS MYB-RELATED PROTEIN 1, ATMYR1, MYB-RELATED PROTEIN 1, MYR1.	C	T	114	0.95	222	CDS	Nonsyn	R	K
Aa.chr8	16980761	Aa_G354520	AALP_AA8G206700	AT5G18860	NSH3, NUCLEOSIDE HYDROLASE 3. Encodes a purine nucleoside hydrolase active in the apoplast. It might play a role in salvaging extracellular ATP.	C	T	42	0.93	222	intronic/noncoding			
Aa.chr8	22757391	Aa_G425350	AALP_AAs50386U000100	AT5G28640	AN3, ANGUSTIFOLIA 3, ARABIDOPSIS GRF1-INTERACTING FACTOR 1, ATGIF1, GIF, GIF1, GRF1-INTERACTING FACTOR, GRF1-INTERACTING FACTOR 1. Encodes a protein with similarity to mammalian transcriptional coactivator that is involved in cell proliferation during leaf and flower development.	C	T	54	0.89	222	CDS	Nonsyn	G	R
chr8	22993986	Aa_G90320	AALP_AAs60715U000400	AT3G04470	Ankyrin repeat family protein.	C	T	115	0.95	222	intronic/noncoding			
Aa.chr8	26452056	Aa_G382520	AALP_AA6G090500	AT5G49110	unknown protein; INVOLVED IN: biological_process unknown	C	T	63	0.95	222	CDS	Syn	L	L
Aa.chr8	27632236	Aa_G286990	AALP_AA6G099000	AT5G47960	ATRABA4C, RAB GTPASE HOMOLOG A4C, RABA4C, SMALL MOLECULAR WEIGHT G-PROTEIN 1, SMG1.	C	T	71	0.95	222	CDS	Nonsyn	R	Q
Aa.chr8	29326092	Aa_G442830	AALP_AA8G280600	AT5G50375	CPI1, CYCLOPROPYL ISOMERASE. Converts pentacyclic cyclopropyl sterols to conventional tetracyclic sterols.	C	T	76	0.92	222	CDS	Syn	F	F
Aa.chr8	31274952	Aa_G101870	AALP_AA8G292300	AT5G51690	1-AMINO-CYCLOPROPANE-1-CARBOXYLATE SYNTHASE 12, ACS12. Encodes an aminotransferase with broad specificity for aspartate and aromatic amino aids such as tyrosine and phenylalanine.	C	T	86	0.97	222	splice_site_change			
Aa.chr8	31823620	Aa_G375200	AALP_AA8G295900	AT5G51990	C-REPEAT-BINDING FACTOR 4, CBF4, DEHYDRATION-RESPONSIVE ELEMENT-BINDING PROTEIN 1D, DREB1D. Encodes a member of the DREB subfamily A-1 of ERF/AP2 transcription factor family (CBF4).	C	T	121	0.94	222	CDS	Nonsyn	V	M
chr8	35032991	Aa_G496120	AALP_AA8G325600	NA	NA	C	T	67	0.82	125	intronic/noncoding			
Aa.chr8	35145947	Aa_G141180	AALP_AA8G322400	AT5G53920	Ribosomal protein L11 methyltransferase-related; FUNCTIONS IN: protein methyltransferase activity; INVOLVED IN: protein amino acid	C	T	70	0.96	222	CDS	Nonsyn	P	S

Aa.chr8	36210157	Aa_G149090	AALP_AA8G334100	AT5G54860	Major facilitator superfamily protein; FUNCTIONS IN: transporter activity; INVOLVED IN: transport.	C	T	95	0.93	222	CDS	Nonsyn	Q	*
chr8	37056748	Aa_G673820	AALP_AA8G343500	NA	NA	C	T	54	0.84	79	intronic/noncoding			
Aa.chr8	39132239	Aa_G307370	AALP_AA8G362100	AT5G57330	Galactose mutarotase-like superfamily protein; FUNCTIONS IN: isomerase activity, carbohydrate binding, aldose 1-epimerase activity, catalytic activity; INVOLVED IN: galactose metabolic process, carbohydrate metabolic process.	G	A	92	0.87	222	CDS	Nonsyn	G	R
chr8	39942706	Aa_G70060	AALP_AA4G167100	NA	NA	C	T	52	0.85	222	intronic/noncoding			
Aa.chr8	40154866	Aa_G148050	AALP_AA8G377400	AT5G58784	Undecaprenyl pyrophosphate synthetase family protein; FUNCTIONS IN: dehydrodolichyl diphosphate synthase activity; INVOLVED IN: dolichol biosynthetic process.	C	T	68	0.88	222	intronic/noncoding			
chr8	40976101	Aa_G482940	AALP_AA8G391100	AT5G59680	Leucine-rich repeat protein kinase family protein. INVOLVED IN: protein phosphorylation.	C	T	59	0.87	222	CDS	Nonsyn	L	F

(1) Chr: Chromosome. **(2)** Pos: Mutated nucleotide position. **(3)** Gene ID: Gene identifier. **(4)** NCBI ID: Gene identifier as on NCBI. **(5)** Arabidopsis gene: Orthologues of the *A. alpina* genes in *A. thaliana*. . **(6)** Gene description description of the *A. thaliana* orthologue as given by The Arabidopsis Information Resource (TAIR, www.arabidopsis.org), NA indicates an orthologue is not available. **(7)** R: Nucleotide in reference genome. **(8)** M: Nucleotide in mutant genome. **(9)** N: Number of reads supporting the mutation. **(10)** AF: Allele frequency. **(11)** Mq: Mapping quality (MAPQ), higher = more unique (min. 30, implies a 0.001 probability of false read alignment). **(12)** Region: Region affected by the mutation. **(13)** Type: Mutation type (nonsynonymous or synonymous). **(14)** AR: Amino acid in reference genome. **(15)** AM: Amino acid in mutant genome.

Supplementary Table 7. CAPS/dCAPS markers used for genotyping

Genotype	SNP position	Gene ID	Marker ID	Primer ID	Primer sequence	Enzymes	Restricted product size
eop002	Chr2: 7322892	Aa_G62080	Chr2-M1	A178 A179	F: TCACTATGGAAGCTCAATTTGC R: GGCCGTTGTTGAAACTATTG	ScrFI	*pep1-1*: 335 + 136 bp *eop002* : 471 bp
eop002	Chr2: 15821900	Aa_G201310	Chr2-M2	A182 A183	F: CAATGATTACGGCCTTGGAC R: TTGAGTATGCGGACCATCG	HinfI	*pep1-1:* 535 bp *eop002* : 270 + 265 bp
eop002	Chr8: 9313334	Aa_G526410	Chr8-M1	A047 A048	F: GAGGCAAATGATGGGAGATTCTCAGT R: TGCCTGTCCAATATTCAAATACATGC	BsmAI	*pep1-1*: 30+263 bp *eop002*: 293 bp
eop002	Chr8: 14606687	Aa_G106560 (*EOP*)	Chr8-M4	A035 A036	F: TTCACAGCTTATGCATCAATGCCTG R: CTGATTCTTAACCTGACTGCGTCTG	ScrFI	*pep1-1*: 25+223 bp *eop002* : 248 bp
eop002	Chr8: 18145960	AaG255500	Chr8-M5	A043 A044	F: CTAGAAGCAGACAAGTCTCGACTTCGT R: TACAACGACGCATACCCAACATCTC	PflFI	*pep1-1*: 224+23 bp *eop002* : 247 bp
eop002	Chr8: 21989344	Aa_G251110	Chr8-M2	A188 A189	F: GGGGAATGATTGCAGGCTAG R: TCCCGGCATCTGACATCTTG	Cac8I	*pep1-1:* 14 + 72 + 230 bp *eop002:* 14 + 302 bp
eop002	Chr8: 39796839	Aa_G31530	Chr8-M3	A214 A215	F: GCTTTCTTAGTCACTGTCCGG R: CTCAAGATCGTCAAGCTCGC	Dpn1	*pep1-1*: 230 + 118 bp *eop002*: 348 bp
eop085	Chr8: 14605246	Aa_G106560 (*EOP*)	–	A041 A042	F: GTATAACAAGACTGTGTCTCATTAG R: GTAAATGTGCATATCCATTCGATC	ScrFI	*pep1-1:* 205+41 bp *eop085:* 246 bp
eop088	Chr8: 14606679	Aa_G106560 (*EOP*)	–	A037 A038	F: TCCAGACGCAGTCAGGTTAAG R: CAAACTAGACTTCCCGGTTCC	HinfI	*pep1-1:* 21+314+185 bp *eop088:* 21+499 bp
eop091	Chr8: 8491434	Aa_G138250	A1	A142 A143	F: TGTTTGGAGAGTGTAGGAGTGT R: GAATGATTTGGGTGGTGGGG	NsiI	*pep1-1:* 132 + 393 bp *eop091 :* 525 bp
eop091	Chr8: 11974246	Aa_G356670	A2	A144 A145	F: GTGTACTGCCGTGCTGATTC R: TCCATACTCGAACCGTAATCG	BseRI	*pep1-1:* 130 + 333 bp *eop091:* 463 bp
eop091	Chr8: 14606418	Aa_G106560	A3	A039 A040	F: ATGTCAATCTCTATCTCAGCACTTGA R: CAAGATCCTCCTCTCGTTCTTAATC	BpuEI	*pep1-1:* 40+217 bp *eop091:* 257 bp
eop091	Chr8: 22993986	Aa_G90320	A4	A216 A217	F: ACCACAAAACTTCCCACAGG R: GAGAATTCCTCGGATGGTTG	Hpy188I	*pep1-1:* 132 + 58 bp *eop091:* 201 bp
eop091	Chr8: 39132239	Aa_G307370	A5	A210 A211	F: TGTGTTCAGTGAAGTGCGAG R: TCCTTTCCACTCTTCACCGG	FokI	*pep1-1:* 325 + 236 bp *eop091:* 541 bp
eop101	Chr8: 1969022	*AaTFL1*	T1	A001 A002	F: CATGAGCTTTTCCCTTCCAC R: TGGTGTTGTGACATGTGATTAGG	ScrFI	*pep1-1*: 390+203 bp *eop101*: 593 bp
eop101	Chr8: 14605372	Aa_G106560 (*EOP*)	T2	A138 A139	F: ACCAAACCATAATCCGTAGCAAG R: TCATCTCCTCCATCTCCAACC	Hpy188I	*pep1-1*: 516+58 bp *eop101*: 384+132+58 bp
pep1-1 and *eops*	Chr8: 6186226	*PEP1*	P	985 986	F: CATGAGCTACTAGAAACCGTTGAAAG R: AAGTGCACAGTTGTTTTCGAC	Bs1I	Pajares : 257bp *pep1-1 and eops:* 280 bp
pep1-1 and *eops*	Chr8: 31595025	*GA20ox2*	Chr8-M7, A6 and T3	E193 E194	F: TCAAGAACACGAGCCAAAAC R: TAACCGGGAATATCATGAGC	Cac8I	Pajares: 217bp + 441 bp *pep1-1* and *eops:* 658 bp

c. F: Forward primer c. bp: Base pair
d. R: Reverse primer

Supplementary Table 8. Primers for plasmids construction

Primer ID	Primer name	Primer sequence	Comment
A014	Aa_G106560_1R	CCGGTCAAGATCTGCTTTAAC	Colony PCR and sequencing primer for insert verification
A015	Aa_G106560_2F	TGACCGTGGTCACAAGAAAG	Colony PCR and sequencing primer for insert verification
A016	Aa_G106560_2R	CTGTCTTCTAAGGGAGAAGCTG	Colony PCR and sequencing primer for insert verification
A017	Aa_G106560_3F	TCGATTGTGCAGTGGAGTTG	Colony PCR and sequencing primer for insert verification
A018	Aa_G106560_3R	CAGCATCCTCGCTCTTCATC	Colony PCR and sequencing primer for insert verification
A019	Aa_G106560_4F	TTATGGTCAAGTTGCGGAGAC	Colony PCR and sequencing primer for insert verification
A020	Aa_G106560_4R	GGGTCTTCAATGGTTCACAAC	Colony PCR and sequencing primer for insert verification
A051	SeLA_pDNOR201_F	TCGCGTTAACGCTAGCATGGATCTC	Colony PCR and sequencing primer for insert verification
A052	SeLB_pDNOR201_R	TGTAACATCAGAGATTTTGAGACAC	Colony PCR and sequencing primer for insert verification
A055	M13_pGWB406_FP	GTAAAACGACGGCCAGT	Colony PCR and sequencing primer for insert verification
A056	M13_pGWB406_RP	AACAGCTATGACCATG	Colony PCR and sequencing primer for insert verification
A057	attBF_pDNOR201_AaG106560	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTTTTCTTTCTCTAGGGATCTTC	Forward primer with attB-site for amplification of *AaG106560*
A058	attBR_pDNOR201_AaG106560	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAGGTGCAAAACAGAACAAGTTTC	Reverse primer with attB-site for amplification of *AaG106560*
A098	AttB1_pDNOR201_AaG106560	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGGCTTAACTGCCCAATGTCTCCATC	Forward primer with attB-site for amplification of AaG106560 promoter region
A099	AttB2_pDNOR201_AaG106560	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAGCCGGAGCAGGATATTGAGGTG	Reverse primer with attB-site for amplification of *AaG106560*
A100	ColonyPCR_AaG106560dwnstF	GTGAACGAAAATTACAACTTTATGC	Colony PCR and sequencing primer for insert verification
A101	ColonyPCR_AaG106560proR	TGGGATTTGTTGGAAGTGTTTGAGTG	Colony PCR and sequencing primer for insert verification
A103	AaG106560_Progeno_overlpR	GATCCCTAGAGAAAGAAAACATTGTAATCTCTCTGTTTTTGCTCTG	Reverse primer for amplification of AaG106560 promoter with genomic DNA overlapping sequence for PIPE cloning
A104	AaG106560-HR-PCR_F	CAGAGCAAAAACAGAGAGATTACAATGTTTTCTTTCTCTAGGGATC	Reverse primer for HR-PCR to insert promoter region during PIPE cloning
A105	AaG106560-HR-PCR_R	GAAGCCTGCTTTTTTGTACAAACTTGGC	Forward primer for HR-PCR to insert promoter region during PIPE cloning
A112	ColPCR-AaG106560pGWB406F1	CACAATCCCACTATCCTTCG	Colony PCR and sequencing primer for insert verification
A113	ColPCR-AaG106560pGWB406R1	CTCCTCCTCCTCCTTTCTTG	Colony PCR and sequencing primer for insert verification
A114	ColPCR-AaG106560pGWB406F2	GTCTGTTGAAGGCAAGAACC	Colony PCR and sequencing primer for insert verification
A115	ColPCR-AaG106560pGWB406R2	AACAGCTATGACCATG	Colony PCR and sequencing primer for insert verification
A116	AaG106560pGWB406_SeqFP	GTAAAACGACGGCCAGT	Colony PCR and sequencing primer for insert verification
A117	AaG106560pGWB406_SeqRP	CTTGCTCACCATGGATCCTC	Colony PCR and sequencing primer for insert verification
A118	ColPCR-AaG106560pGWB406F3	GATGAAGAGCGAGGATGCTG	Colony PCR and sequencing primer for insert verification
A119	ColPCR-AaG106560pGWB406R3	AATCCTGTTGCCGGTCTTG	Colony PCR and sequencing primer for insert verification
A128	AaG106560pLeSeq-1F	AAGACCGGCAACAGGATTC	Colony PCR and sequencing primer for insert verification
A129	AaG106560pLeSeq-1R	TCCGATCACGATTCATTGTC	Colony PCR and sequencing primer for insert verification
	p35s-F	CGCAAGACCCTTCCTCTATATAAGGAAG	Colony PCR and sequencing primer for insert verification
	Pleela R	CTGGTGATTTTTGCGGACTCTAGAAC	Colony PCR and sequencing primer for insert verification

Acknowledgement

I would foremost like to thank my supervisor, Prof. Dr. Maria Albani, for giving me the opportunity to work in her group, for her wonderful supervision and productive discussions.

A special thank you to Prof. Dr. Ute Hoecker for being part of my PhD committee and my second examiner, for the fruitfull discussions and suggestions about my project.

Another special thank you to Dr. Korbinian Schneeberger for being part of my PhD committee and the 'Beisitzer' at my defence, for the helpful discussions and guidelines, especially on the bioinformatics aspect of my project.

I am very grateful to Prof. Dr. Wolfgang Werr for being the chairperson at my defense, thank you very much.

My thanks and appreciation to Ana Lazaro for the support and advice, fruitful inputs about my project, and for the help in editing my thesis.

Thank you to Hequan Sun (AG. Schneeberger, MPIPZ Cologne) for the guidance with bioinformatics analysis.

Furthermore, I would like to thank the entire AG Albani, Alice Vayssières, William Hughes, Natanael Viñegra de la Torre, Yanhao Zhou and Priyanka Mishra, for their good collaboration, stimulating discussions and for providing a great working atmosphere in all aspects. Thank you to Yan Zeng for the various technical help and thank you to Margaret Kox for all the help with administrative procedures. It was a great pleasure working with you all and you made every bit of work time an enjoyable one.

Many thanks to the entire MPIPZ staff for the friendly, internationally rich and scientifically stimulating working environment.

Finally, I would like to extend my deepest gratitude to my husband, Rafal Zgadzaj, the entire Obeng-Hinneh and Zgadzaj family for their unconditional love, advice, support and words of encouragement. Thank you all very much, I could not have made it without you being there for me.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeiteinschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Maria Albani betreut worden.

Köln, March 2018

Evelyn Obeng-Hinneh

Teilpublikationen

Lazaro, A., **E. Obeng-Hinneh**, and M.C. Albani, Extended vernalization regulates inflorescence fate in *Arabis alpina* by stably silencing *PERPETUAL FLOWERING 1.* Plant Physiology, 2018.

Lebenslauf

Persönliche Daten:

Name	Obeng-Hinneh
Vorname	Evelyn
Geburtsort	Berekum, Ghana
Geburtsdatum	14/01/1986
Nationalität	Ghanaisch

Ausbildung:

Seit 11/2013	Doktorarbeit am Universität zu Köln und Max-Planck-Institut für Pflanzenzüchtungsforschung, Köln-Germany; Betreuer: Prof. Dr. Maria Albani Thema: Mechanisms regulating inflorescence development and flowering traits in *Arabis alpina*, an alpine perennial
10/2010 – 09/2013	Masterarbeit am Justus-Liebig University, Giessen-Germany. Betreuer: Prof. Friedt und Dr. Obermeier Thema: Production and analysis of genomic reduced representation libraries from a rapeseed mapping population
10/2005 – 09/2009	Bachelorarbeit am University of Ghana, Legon-Ghana. Betreuer: Prof. Samuel Kwame Offei Thema: Genetic diversity in selected *Dioscorea praehensilis* germplasm from forest savannah transition ecozones of Ghana