Biochimica et Biophysica Acta 1809 (2011) 395–406

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Polycomb-group mediated epigenetic mechanisms through plant evolution ☆

Yana Butenko [a], Nir Ohad [a,b,*]

[a] Department of Molecular Biology and Ecology of Plants, Tel-Aviv University, 69978, Israel
[b] Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Germany

ARTICLE INFO

Article history:
Received 3 March 2011
Received in revised form 23 May 2011
Accepted 26 May 2011
Available online 6 June 2011

Keywords:
PcG
Evolution
Gene expression
Epigenetic
Plant development
H3K27 methylation

ABSTRACT

Polycomb **G**roup (PcG) proteins form an epigenetic "memory system", conserved in both plants and animals, controlling global gene expression during development via histone modifications. The role of PcG proteins in plants was primarily explored in *Arabidopsis thaliana*, where PcG regulation of developmental processes was demonstrated throughout the plant life cycle. Our knowledge about the PcG machinery in terrestrial plants other than *Arabidopsis* began to accumulate only in recent years. In this review we summarize recent emerging data on the evolution and diversification of PcG mechanisms in various phyla, from early-diverging plants, including members of the Chlorophyte algae, through bryophytes and flowering plants. We describe the compositions of the PcG gene families, their so-far studied expression profiles, and finally summarize commonalities vs. differences among PcG functions across the various species. This article is part of a Special Issue entitled: Epigenetic control of cellular and developmental processes in plants.

1. Introduction

Polycomb group (PcG) proteins regulate gene expression epigenetically via chromatin remodeling. PcG proteins were initially identified in fly *Drosophila melanogaster* as factors necessary for the long-term repression of homeotic (*Hox*) genes [1–3]. However, in recent years it became apparent that transcriptional regulation by PcG proteins is a general mechanism, which operates in establishing and maintaining distinct patterns of gene expression in both animals ([4–6], reviewed by [7,8]) and plants ([9–12], reviewed by [13]). PcG proteins regulate key developmental pathways [7], presumably by suppressing particular genetic programs not required at a given time in a given cell. PcG proteins were shown to maintain the undifferentiated state of embryonic stem cells in both animals and plants [14–19]. Based on recent studies, up to 5% of the genes in vertebrates and *Drosophila* [20], and as much as 15% of the genes in the model plant *Arabidopsis thaliana* [21,22] are potentially regulated by PcG proteins. In plants, Polycomb proteins have been shown to take a role in numeral developmental processes, including flowering [16,23], seed development [12,24,25], regulation of shoot meristem development [26,27] and root patterning [28].

PcG proteins form multi-subunit **P**olycomb **R**epressive **C**omplexes (PRCs), which mediate transcription silencing. At least three distinct PcG-complexes were identified in metazoans: **P**olycomb **R**epressive **C**omplex **2** (PRC2), **Poly**c**o**mb-**l**ike PRC2 (Pcl-PRC2) and **P**olycomb **R**epressive **C**omplex 1 (PRC1) [29,30]. PRC2 and related Pcl-PRC2 complexes initiate gene silencing by catalyzing methylation of histone H3 at lysine 27 (H3K27me) [31–37]. Subsequently, PRC1 binds to the methylated histone [38] and proceeds to establish stable repression of PcG target genes, by catalyzing monoubiquitination of histone H2A at lysine 119 (H2AubK119) [39–41]. Histone modifications, such as H3K27me3 and H2Aub, play a key role in repressing gene expression, probably by preventing RNA-transcript elongation [42]. In addition to their enzymatic function, PRC1 and to a lesser extent PRC2, mediate compaction of the chromatin [30,43–45]. This chromatin-compaction may result in reduced accessibility of transcription-factors, such as SWI/SNF-class ATP-dependent chromatin remodelers [39,43], thereby the subsequent repression of target genes throughout consecutive cell divisions.

Interestingly, no proteins directly corresponding to PRC1 members were identified in plants [46]; hence, it is not fully clear how transcriptional repression is established and maintained. However, some degree of functional conservation among plants and animals was reported for PRC1. In *Drosophila*, the PRC1 subunit Polycomb (Pc) is responsible for the recognition of the H3K27me3 mark, placed by the PRC2 complex [38]. The *Arabidopsis* protein LIKE HETEROCHROMATIN PROTEIN 1 (LHP1) [23,47] was shown to bind H3K27me3 *in-*

Abbreviations: AG, AGAMOUS; CLF, CURLY LEAF; EMF1, EMBRYONIC FLOWER1; EMF2, EMBRYONIC FLOWER2; ESC, EXTRA SEX COMB; E(z), ENHANCER OF ZESTE; FIE, FERTILIZATION INDEPENDENT ENDOSPERM; FIS2, FERTILIZATION INDEPENDENT SEED2; KNOX, KNOTTED-like; MEA, MEDEA; MSI1, MULTICOPY SUPPRESSOR OF IRA1; PcG, Polycomb group; PRC1, Polycomb Repressive Complex 1; PRC2, Polycomb Repressive Complex 2; SWN, SWINGER; Su(z)12, SUPRESSOR OF ZESTE 12; VRN2, VERNALIZATION2

☆ This article is part of a Special Issue entitled: Epigenetic control of cellular and developmental processes in plants.

* Corresponding author at: Department of Molecular biology and Ecology of Plants, Tel-Aviv University, 69978, Israel. Tel.: +972 3 6407641; fax: +972 3 640 7642.
E-mail address: Niro@tauex.tau.ac.il (N. Ohad).

1874-9399/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbagrm.2011.05.013

vitro and *in-vivo* [21,22,48,49], and it is required for PcG-mediated repression of floral homeotic genes [47,50,51]. Furthermore, VERNALIZATION 1 [52] and EMBRYONIC FLOWER 1 [53] proteins were demonstrated to co-operate with the PRC2 complex in *Arabidopsis*, acting downstream to repress target genes. In addition, AtRING1a and AtRING1b [27] contain a RING-finger domain homologous to that found in *Drosophila* Posterior Sex Combs (PSC) protein, which is responsible for histone 2A mono-ubiquitination [54]. Both AtRING1a and AtRING1b interact with AtLHP1 and are necessary for repression of *KNOTTED-like (KNOX)* homeobox genes [27].

The PRC2 complex, on the other hand, is highly conserved among both plants [9–12,55] and animals [56–58]. The *Drosophila* PRC2 complex contains four core subunits: Enhancer of zeste (E(z)) [33], which serves as the catalytic subunit methylating H3K27 via its SET (**S**u(var) **E**(z) **T**hritorax) domain [33,34], **E**xtra **S**ex **C**omb (ESC) polypeptide containing seven WD40 domains [33], Suppressor of zeste 12 (Su(z)12) protein containing the C2H2 zinc-finger domain [59] and the Nucleosome remodeling factor 55-kDa subunit (Nurf55, known also as p55) [33,34,60].

In *Arabidopsis*, the best studied model plant, homologs of E(z) and Su(z)12 underwent gene duplication, giving rise to small gene families. Thus, the *Arabidopsis* genome encodes for three *E(z)* paralogs: *CURLY LEAF* (*CLF*) [9], *SWINGER* (*SWN*) [16,61] and *MEDEA* (*MEA*) [10,11]. The PRC2 member *Su(z)12* encoding for zinc-finger protein also diverged into a family of three, including *EMBRYONIC FLOWER 2* (*EMF2*) [62], *VERNALIZATION 2* (*VRN2*) [63] and *FERTILIZATION INDEPENDENT SEED 2* (*FIS2*) [11,64]. The remaining PRC2 complex members in *Arabidopsis* include two WD40 motif proteins, FERTILIZATION-INDEPENDENT ENDOSPERM (FIE) and MULTICOPY SUPRESSOR OF IRA 1 (MSI1), homologs of the *Drosophila* ESC and p55 proteins, respectively [12,24] (Table 1).

Based on genetic, molecular and biochemical evidence, at least three PRC2 complexes, harboring different paralogs of the E(z) and Su(z)12 proteins families, likely co-exist in *Arabidopsis*, each controlling a particular developmental program (Fig. 1) (reviewed in [13,46,65–69]). The reproductive FIS2-PRC2 complex is implicated in regulating the female gametophyte and seed development. Mutants impaired in members of this complex were shown to develop seed-like structures containing juvenile endosperm in the absence of fertilization and arrested embryos when fertilization took place, leading to seed abortion in both cases [10,70–73]. The EMF2-PRC2 complex suppresses premature transition from the vegetative to the reproductive stage and takes part in regulating floral organs development [16,74,75]. An additional vegetative PRC2 complex, VRN2-PRC2, regulates flowering time mediated by vernalization [63,76–78].

Consistent with PcG proteins function in metazoans, *Arabidopsis* PRC2-like complexes are required for the trimethylation of H3K27 at least at several loci [13,21,22,79–85]. In support of this hypothesis, intact SET domain was shown to be necessary for the functions of AtCLF and AtMEA proteins [83,84]. No *Hox* genes homologs were isolated in plants to date, however, the *Arabidopsis* PRC2 complexes repress homeotic transcription factors, e.g. members of the homeobox *KNOX*-family [26,27]. These findings point to a conserved role of the PcG complexes during ontogenesis in both kingdoms. Moreover, PcG proteins are also essential for establishing and then maintaining cell identity [13,19,83,86,87].

Despite being identified in *Arabidopsis* more than a decade ago [9], our knowledge about the composition and function of plant PcG complexes other than *Arabidopsis* is still limited. In recent years, homologs of PcG proteins were identified in both dicots and monocots, indicating conserved regulatory functions among flowering plants [13]; however, their evolution in Viridiplantae is still vague. Recently, we and others [19,88] have described the function of PcG proteins in the moss *Physcomitrella patens*, demonstrating that PRC2 epigenetic machinery evolved early in evolution of terrestrial plants and that PRC2 function was maintained along plant evolution. Here we review the current knowledge on the evolution, diversification and function of the PRC2 complex in the plant kingdom. As the nomenclature used to describe PcG genes across different species is diverse, which may complicate following the relations between them, we propose a uniform naming for members of the PRC2 complex. As shown in Table 1, the initials of the species are followed by the gene name based on its phylogenetic relations to the *Arabidopsis* PcG gene.

2. Evolution of PcG complexes from unicellular algae to early terrestrial plants

The origin of PcG genes is emerging in recent years, as more complete genomic sequences become available. As no homologs of E(z) or H3K27 methylation marks were identified in model unicellular fungi such as budding yeast [89,90], it was proposed that the emergence of PRC2 complex might have coincided with the development of multicellularity [13,91]. However, the conservation of PRC2 components and functions among extant eukaryotes, both in plants and metazoans, implies that PcG genes were already present in the last common unicellular ancestor before plants and animals diverged over ~1.6 billion years ago [92] (Fig. 2). Indeed, H3K27 methylation, presumed to be catalyzed by an E(z)-like protein, was found in early-diverging unicellular species, such as the green alga *Chlamydomonas reinhardtii* [93,94] (a chlorophyte, see Fig. 2) and in the ciliated protozoan *Tetrahymena thermophila* (chromalveolates), where it is associated with heterochromatin formation [95].

Recently, Shaver et al. [94] described a broad query in which they have identified genes encoding for PRC2 complex subunits in organisms from three of the six major eukaryotes groups [100]: Opisthokonta (fungi and animals), Chromalveolata (Alveolata and Stramenopiles) and Archaeplastida (red algae, green algae and land plants) (Fig. 2). The identification of PcG genes in unicellular species belonging to each of these eukaryotic super-groups supports the notion that the PRC2 complex had evolved in an ancestral unicellular eukaryote, where its function is yet unknown.

Reports employing sequence alignment demonstrated that overall structure of PcG proteins, and in particular their key domains, are remarkably conserved among evolutionarily distant organisms [19,55,94,97,98]. E(z) proteins show the highest degree of structural conservation among members of PRC2 complex and contain five domains positioned along its sequence: EZD1, EZD2, SANT, Cys-rich Pre-SET motif, and the SET-domain. The correlation between the genomic H3K27 methylation mark and the presence of the E(z) homolog in all eukaryotes tested so far, indicates a role for E(z)-like proteins in catalyzing the methylation reaction [94]. Potential homologs of *ESC* were also found in the genomes of numerous eukaryotes, and their phylogeny reflected their taxonomic relationships [19,94]. The comparison of PcG protein sequences between different organisms revealed high conservation of the plant ESC-like proteins, containing seven WD40 repeats and a putative nuclear localization signal (NLS) positioned near the N′ terminus [12,55]. On the other hand, the conservation of the Su(z)12 subunit is less extensive than that of E(z) and ESC, as Su(z)12 homologs were seemingly lost in a number of organisms (that do encode all other core PRC2 subunits). Sequence alignments demonstrated that most Su(z)12 polypeptides have a similar structural design, in which a C2H2 zinc-finger motif is followed by a VEFS-domain [97,98]. Finally, homologs of the p55 PRC2 subunit were found in all examined organisms, including those belonging to the Excavata (diplomonads and heterotrophic flagellates) and the Amoebozoa groups [94]. However, p55 homologs have take part in additional machineries other than PRC2, such as the chromatin assembly complex CAF-1 [60,101,102] and the nucleosome remodeling complex NURF [60], which may explain their widespread distribution across numerous Eukaryote phyla [94]. It appears that PcG genes were lost independently several times during the evolution of particular eukaryotes

Table 1
Summary of the PcG genes phylogenetic relations, expression and imprinting status: The above table summarizes the composition of PcG gene families across members of the Viridiplantae (all green plants), reviewed here. The PcG genes for each organism are organized according to their most conserved domain (first column on the left). Members of each PcG protein family are further subdivided according to their phylogenetic relations to the *Arabidopsis* genes. For each gene, available information regarding its expression pattern and its imprinting status are indicated. Exp'—Expression; Imp'—Imprinting; CC—central cell; N.I.—not imprinted; N.D.—not determined.

Protein domains	*Chlamydomonas reinhardtii*	*Physcomitrella patens*	*Arabidopsis thaliana*	*Petunia hybrida*	*Solanum lycopersicum*	*Hieracium species*	*Hordeum vulgare*	*Zea mays*	*Oryza species*
SET	**CrCLF (CrEZH)** [94]	**PpCLF** [19, 88] Exp': gametophyte meristematic cells; young leaves; gametes precursors; zygotes; sporophyte Imp': N.D.	**AtCLF** [9, 107] Exp': ubiquitous, except for CC and nuclear endosperm and suspensor Imp': N.I.	**PhCLF1** [127] (+ splice variants) Exp': leaves; floral organs Imp': N.D. **PhCLF2** [127] Exp': low levels in leaves and floral organs Imp': N.D.	**SlCLF1 (SlEZ2)** [128] Exp': vegetative tissues; flowers; young seeds **SlCLF2 (SlEZ3)** [128]			**Zm CLF (ZmMez1)** [55, 149] Exp': ubiquitous Imp': only the maternal allele is expressed in endosperm	**OsCLF** [97, 150, 151] Exp': ubiquitous Imp': N.I.
			AtSWN [104, 107] Exp': ubiquitous, except for nuclear endosperm and suspensor Imp': N.I.	**PhSWN (PhCLF3)** [127] Exp': N.D. Imp': N.D.	**SlSWN (SlEZ1)** [128] Exp': vegetative tissues; flowers; young seeds		**HvSWN (HvE(Z))** [141] Exp': ubiquitous Imp': N.D.	**ZmSWN1 (ZmMez2)** (3 splice variants) [55, 149] Exp': ubiquitous Imp': N.I. **ZmSWN2 (ZmMez3)** [55, 149] Exp': ubiquitous Imp': N.I.	**OsSWN (OsiEZ1)** [97, 150, 151] Exp': ubiquitous Imp': N.I.
			AtMEA [104, 107, 123] Exp': All cell of embryo sac, endosperm, embryo Imp': only maternal allele is expressed						
WD40	**CrMSI1** [94]	**PpMSI1**	**AtMSI1** [117, 153] Exp': ubiquitous Imp': N.I.		**SlMSI1** [154]			**ZmRBAP1, ZmRBAP2, ZmRBAP3** [154]	**OsRBAP1, OsRBAP2, OsRBAP3** [154]
Zinc-finger		**PpEMF2_1, PpEMF2_2, PpEMF2_3** [98]	**AtEMF2** [62] Exp': ubiquitous, including embryos and endosperm Imp': N.D.		**SlEMF2p** [98]		**HvEMF2a (HvSu(z)12a)** [141] Exp': young shoots Imp': N.D. **HvEMF2b (HvSu(z)12b)** [141] Exp': ubiquitous Imp': N.D. **HvEMF2c (HvSu(z)12c)** [141] Exp': young shoots, leaves; developing seeds Imp': N.D.	**ZmEMF2_1, ZmEMF2_2** [98]	**OsEMF2a** [97] Exp': ubiquitous Imp': N.I. **OsEMF2b** [97] Exp': ubiquitous Imp': N.I.
			AtVRN2 [155] Exp': ubiquitous Imp': N.D.						
			AtFIS2 [104] Exp': endosperm Imp': only maternal allele is expressed						
WD40	**CrFIE** [94]	**PpFIE** [19] Exp': gametophyte meristematic cells; young leaves; gametes precursors; zygotes; regenerating leaf cells Imp': N.D.	**AtFIE** [12] Exp': ubiquitous Imp': only the maternal allele is expressed in endosperm, bi-alellic during other stages			**HpFIE** [134] Exp': leaves; floral organs; ovaries Imp': N.D.	**HvFIE [141]** Exp': ubiquitous Imp': N.D.	**ZmFIE1** [55 , 144, 145] Exp': endosperm Imp': only the maternal allele is expressed **ZmFIE2** [55, 144, 145] Exp': ubiquitous; not expressed in sperm Imp': only the maternal allele is expressed in endosperm	**OsFIE1** [97, 150, 151] Exp': endosperm Imp': only the maternal allele is expressed **OsFIE2** [97, 150, 151] Exp': ubiquitous Imp': N.I.

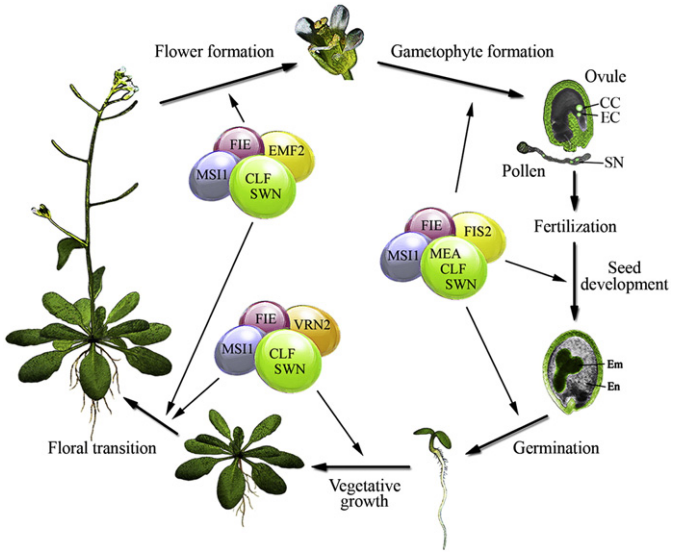

Fig. 1. Regulation of the *Arabidopsis* life cycle by the PRC2 complexes. The flower bears the male organs anthers, and female organ pistil, which contains ovules. Upon meiosis, female and male gametes are formed in the pollen and the ovules, respectively. Double fertilization leads to the formation of embryo and endosperm (nutritious tissue). Following germination of the seed, the seedling containing primary root, shoot and two embryonic cotyledons emerges. Further shoot development leads to the formation of rosette leaves, establishing the vegetative stage. Transition to the reproductive stage is marked by establishing of the inflorescence and flower formation. At least three types of PRC2 complexes control transitions through the vegetative and the reproductive phases of *Arabidopsis* development. *AtFIE* and *AtMSI1* are single copy genes; therefore, both proteins are expected to be found in all the PRC2 complexes. Different paralogs of the SET-domain proteins (AtMEA, AtCLF and AtSWN) and Zinc-finger proteins (AtVRN2, AtEMF2 and AtFIS2) may participate in distinct PRC2 complexes. "CC"—Central Cell nucleus; "EC"—Egg Cell nucleus; "SN"—Sperm Nucleui; "Em"—Embryo; "En"—Endosperm.

(with p55 set aside for reasons above). For instance, all PRC2 subunits but p55 are absent in yeast *(Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*) and in a number of the Amoebozoa members, whereas in the volvocine algae *C. reinhardtii* and *Volvox carteri*, *E(z)* and *ESC* homologs were identified but not *Su(z)12* homologs [94] (Fig. 2).

The high conservation of PcG proteins among various eukaryotic species indicates their crucial function. It is important to understand the significance of PcG genes diversification and the potential role of PRC2 composition for genomic targeting and/or mechanism of transcriptional repression.

The unicellular green alga *Chlamydomonas reinhardtii* (Fig. 2) is the earliest-diverged plant species in which Polycomb function was studied to date. A single homolog of E(z) protein, EZH (designated in this review CrCLF), was identified in *Chlamydomonas* [94], and it is probably the one responsible for the detected mono- and di-methylation of H3K27. Depletion of *CrCLF* transcript by RNAi-silencing resulted in global increase of H3K4 trimethylation and histone H4 acetylation marks, both modifications being generally associated with transcriptional activity and as marks antagonistic to the repressive E(z) activity. Moreover, the above mutant displayed release of transcriptional silencing of tandemly repeated transgenes and of retrotransposons. The above demonstrate PRC2 functionality early in the evolution of the unicellular Chlorophyte algae. Based on these findings, Shaver et al. [94] suggested that, as in all other examined contemporary eukaryotes, the algal CrCLF takes part in gene silencing through the modulation of repressive chromatin states. The authors [94] proposed that in the ancestral unicellular organisms, the PRC2-mediated silencing might have been a part of a mechanism protecting against DNA parasites, e.g. transposons. During the evolution of multicellular eukaryotes, where developmental programs are of much higher complexity, this global-silencing mechanism was further adapted for the regulation of intricate developmental programs.

In the moss *Physcomitrella patens*, member of the Bryophyta division of non-vascular plants (Fig. 2), a single copy orthologs of the *ESC* and *E(z)* genes were identified, designated *PpFIE* [19] and *PpCLF* [19,88], respectively. In addition, three *Su(z)12* homologous genes, *PpEMF2_1*, *PpEMF2_2* and *PpEMF2_3*, were identified in *P. patens* [98]. Bryophytes are the earliest-diverged group among terrestrial plants (Fig. 2). In contrast to the vascular plants, mosses have a gametophyte-dominant life cycle, whereas the diploid sporophyte is diminutive, short-lived and parasitizes on the gametophyte. During the gametophytic stage, juvenile filamentous protonemata grow and propagate as the apical meristematic cell continues to divide, while a subapical cell divides to produce an initial cell – stem cell that may further differentiate forming either a protonemal side branch or a bud, depending on internal cues, e.g. plant hormone cytokinin [103]. Buds are specialized meristematic apices that give rise to mature leafy gametophores which upon maturation will bear the sex organs. The moss PcG proteins, PpFIE and PpCLF were shown to interact *in-vivo* [19], suggesting that in mosses, as in *Arabidopsis*, these proteins may function within a putative PRC2 complex as well. In support of this hypothesis, both proteins were demonstrated to accumulate in all haploid meristematic gametophytic cells and in young leaves, as well as in gametes progenitor cells and unfertilized egg cells [19,88]. Both *PpFIE* and *PpCLF* ceased to express in the zygote [19,88]; however, expression of *PpCLF* is resumed temporarily in mid-stage of sporophyte development [88]. Interestingly, *PpFIE* accumulation was most prominent in cells undergoing cell-fate transition, such as buds formation or during the early stages of leaf regeneration [19].

Ppfie mutant moss developed abnormal buds, which gave rise to multiple secondary apices, resulting in formation of cone-like leafless masses. We suggest that this phenomenon might have resulted from failure of the primary bud cells to switch from meristematic to gametophore-development program, leading to repetitive initiation of new apices [19]. Furthermore, some of these masses continued to develop and formed sporophyte-like structures, which expressed sporophytic specific marker [19]. Thus, this might indicate that PcG complexes participate in restricting gametophore apical cells to pursue the sporophytic program. Absence of the PpCLF similarly resulted in abolition of leafy gametophores [88]. However, the initial cells in *Ppclf* mutants developed into sporophyte-like bodies that expressed sporophytic marker genes [88] instead of *Ppfie* multiapical masses. Resembling the *Ppfie* mutant phenotype, these bodies developed ectopic apices that grew out as branches [88], indicating that both PcG proteins are required for the restriction of meristems proliferation. When *PpCLF* expression was restored in these mutant bodies, they stopped proliferating and sporangium-like organs were formed [88]. Taken together the above, it is evident that early on in terrestrial plant evolution, moss PcG proteins PpFIE and PpCLF have acquired a regulatory role controlling proliferation and differentiation of the gametophyte stem cells. Moreover, PcG genes take part in restricting sporophyte initiation [19,88], as well as regulating its development [88]. Thus, it is expected that both proteins will be present in both gametophytic and sporophytic phases to control together common target genes, even so *PpFIE* was not detected so far in the sporophyte.

The degree of functional conservation of PRC2 members throughout land plant evolution is illustrated by the ability of the *Arabidopsis* and the moss FIE homologs to partially complement each other [19]. *Physcomitrella* plants expressing *AtFIE* instead of the endogenous *PpFIE* gene were able to develop leafy gametophores, though they failed to produce sex organs. This ability of AtFIE to substitute for the native protein, albeit partially, implies that it can functionally recognize components of the *P. patens* PcG complex [19]. However, additional components may be required to establish the reproductive phase in moss, which AtFIE did not recognize. In reciprocal experiments [19], PpFIE partially rescued the

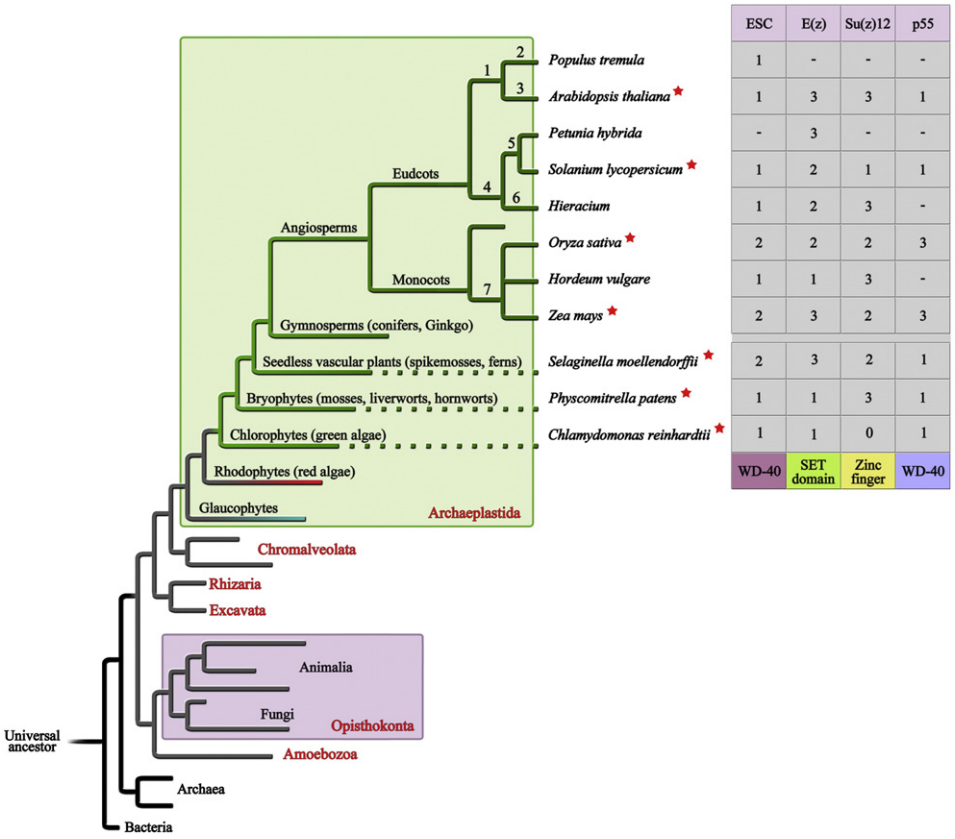

Fig. 2. Schematic illustration of the phylogenetic "tree of life", emphasizing the green plants. The six eukaryote kingdoms are marked in red. The green rectangular represents the Archaeplastida kingdom. Classes within the Viridiplantae phylum are marked by a green line. The phylogenetic classification of angiosperms is according to Angiosperm Phylogeny Group III report [96]. The tree depicts the phylogenetic relations between groups and species described in this review, and therefore not all clades and orders were included. 1 — clade rosids; 2 — clade fabids (eurosids I), Salicaceae family; 3 — clade malvids (eurosids II), Brassicaceae family; 4 — clade asterids; 5 — clade lamiids (euasterids I), Solanaceae family; 6 — clade campanulids (euasterids II), Asteraceae family; 7 — clade commelinids, Poaceae family. Asterisks indicate species for which genome sequences are available. Branch lengths are approximate. *Table*: the number of PRC2 proteins is shown for selected plant species. "–" marks unavailable data. The top row for each protein type lists the representative member from *D. melanogaster* PRC2 complex. The bottom row marks the key domain for each PcG protein (color-coded as in Fig. 1). Gene family members are presented according to the information provided in Table 1, as well as [97,98] for *Selaginella* and [99] for *Populus*.

seed-abortion phenotype [12] of *Arabidopsis fie/FIE* mutant plants. PpFIE was shown to interact *in-vivo* with *Arabidopsis* SET-domain proteins AtCLF and AtSWN, but not with AtMEA that takes part in regulating *Arabidopsis* central cell development [10,70,104]. The lack of interaction between PpFIE and AtMEA may explain the partial complementation displayed by PpFIE in *A. thaliana*, likely due to the inability of PpFIE to form a functional PRC2 complex within the *Arabidopsis* central cell. The above findings reveal that the fundamental function of FIE protein in regulating developmental programs along the plant life cycle, including phase transition from the gametophytic to the sporophytic phase was established early on in plant evolution.

In vascular plants, a major shift towards diploid-dominated life cycle resulted in reduction of the gametophyte to a merely several-cells organ. During this transition, functions of PcG proteins expanded to include regulation of sporophytic developmental programs as well. This was accompanied by multiple gene duplication events and diversification of the PRC2 genes, allowing multicomplexity of the PcG machinery in vascular plants. While the genomes of early-diverging species mostly contain single copies of the genes encoding PRC2 core subunits, in seed plants some PcG genes have diversified, forming small gene families with up to three members, such as the aforementioned three-membered SET-domain protein family in *Arabidopsis*. In the licophyte (club moss) *Selaginella moellendorfii*, a representative of the oldest extant division of vascular plants (Fig. 2), two *FIE* and three *CLF* homologous sequences were identified [97], as well as two putative *EMF2* homologs [98]. It remains to be discovered whether the duplicated PcG genes had acquired distinct individual functions in *Selaginella*.

3. Functions of PcG complexes in dicots

Accumulating evidence suggests that in the course of seed plants evolution ESC, Su(z)12 and E(z) paralogs have acquired specialized functions in certain lineages, such as a role in regulating the transition to the reproductive phase in response to cold temperature (termed vernalization), regulation of flowering time, flower organ formation and seed development [10–12,26,65,68,105–108].

3.1. Polycomb machinery in Arabidopsis

Arabidopsis thaliana, a member of Brassicaceae family, is the most studied species among the dicots. In Brassicaceae, large-scale genome duplication events [109–111] resulted in the establishment of several lineage-specific *E(z)* and *Su(z)12* paralogs (Table 1 and Fig. 2) [98,107,108]. The three *Arabidopsis* SET-domain proteins AtMEA, AtCLF, and AtSWN belong each to a separate clade of plant E(z) homologs [16]. In fact, while all three *Arabidopsis* SET-domain proteins are expressed in the ovule [9,10,16,104,112], *Atmea*-deficient mutants cannot be rescued by neither AtCLF nor AtSWN, implying a unique role for AtMEA during gametophyte and seed development [10,70,104]. In a similar fashion, despite the similarity between AtSWN and AtCLF, AtSWN is only partially redundant to AtCLF in regulating early vegetative development,

suggesting some functional divergence between both paralogs [16]. The current theory is that diversification of *SWN* from *CLF* occurred early in the evolution of land plants, whereas *MEA* has arisen from *SWN* by block duplication in the Brassicaceae, which occurred ~35–85 million years ago [107,108]. As MEA have acquired a novel function in regulating seed development, it became imprinted, probably following its neofunctionalization, as suggested by [107]. *MEA* evolved rapidly, likely by combination of positive, balancing and diversifying selections [106–108]. On the other hand, the Brassicaceae SWN probably retained its ancestral functions and evolved under purifying selection [107].

Similarly to the above, it was proposed that *VRN2* (homolog of the single zinc-finger protein Su(z)12) may have diversified from an ancestral *EMF2* gene [98] during a genome duplication event that took place in the Brassicaceae, followed by a second duplication of the *VRN2* chromosomal region giving rise to *FIS2* [97]. Consequently, no direct orthologues of *VRN2*, *FIS2* or *MEA* were identified in any species other than *Arabidopsis*. Sequences with VRN2-like domain arrangement have been identified in pepper, poplar, alfalfa, and soybean [98]; however, their function is yet unknown.

There is compelling evidence that *Arabidopsis* PcG proteins act together in PRC2-like complexes. Several studies demonstrated that the AtFIE protein can interact with the SET-domain proteins: AtMEA [24,113–116], AtCLF [19,26] and AtSWN ([113], Ohad lab, unpublished results), as well as with the AtMSI1 subunit [24]. Moreover, all *Arabidopsis* Su(z)12 homologs were shown to interact via a conserved VEFS domain with each of the SET-domain proteins [16]. Thus, the duplication of the PcG genes in *Arabidopsis* allows for variations in complex composition, depending on cell or tissue type. Since *AtFIE* is a single-copy gene, AtFIE is expected to take part in all the PRC2 complexes, regulating different aspects of the *Arabidopsis* life cycle [26], which will be further discussed below (Fig. 1).

3.1.1. Regulation of vernalization-dependent flowering

The pioneering work of F. De Lucia et al. led to the determination of the *Arabidopsis* vegetative VRN2-PRC2 complex composition [78]. This complex comprises of AtVRN2, AtFIE, AtMSI1 and AtCLF, and it takes part in vernalization-induced initiation of flowering, by maintaining the silenced state of the flowering repressor *FLOWERING LOCUS C* (*FLC*) after vernalization [16,63,66,76,81]. The VRN2-PRC2 complex interacts with additional proteins including AtVRN5, AtVEL1, and AtVIN3, which are recruited specifically to this complex during vernalization [9,63,77,78,82].

3.1.2. Regulation of flowering time and flower organs development

Several studies suggest that the *Arabidopsis* EMF2-PRC2 complex, containing AtEMF2, AtCLF/AtSWN, AtFIE, and AtMSI1, regulates the transition from vegetative to reproductive stage [9,14,16,26,62,80,117,118] by repressing expression of flowering promoting genes, such as *AGAMOUS-LIKE 19* (*AGL19*) [62,80] and *FLOWERING TIME* (*FT*) [118]. The EMF2-PRC2 complex is probably cooperated by LHP1 protein [23,119]. As one might expect, loss-of-function *Atclf* and *Atemf2* mutants and *Atfie*-silenced mutants undergo precocious flowering [9,14,26,62], as well as homeotic transformation of flower organs [14,26,62], indicating that the EMF2-PRC2 complex regulates the expression of floral organ-identity genes, as shown for the MADS-box *AGAMOUS* (*AG*) [9,53,120]. Furthermore, mutations in PcG genes lead to homeotic transformations of ovules into carpel-like structures in *Atfie* co-suppressed mutants [26].

Even more prominent illustration of the PcG proteins role in regulating organogenesis, not only during flower development, is the formation of disorganized callus-like plant body in both *Atclf*/*Atswn* double mutants and in *Atfie*-silenced plants [14,16].

3.1.3. Regulation of seed development

The FIS2-PRC2 complex, containing AtMEA, AtFIE, AtFIS2, and AtMSI1, is involved in regulating seed development, where it functions to restrain endosperm proliferation before and after fertilization [10–12,24,104] (Fig. 1). The complex subunits composition was determined genetically via the analysis of mutants displaying aspects of seed development in the absence of fertilization [10,71–73]. In addition, direct interactions between AtFIE and AtMSI1 [24], AtFIE and AtMEA [113–116], and AtFIE and AtFIS2 [104] were demonstrated. Subsequently, a 600 kDa complex containing AtFIE, AtMEA and AtMSI1 was isolated [24]. *AtMEA* and *AtFIS2* share several similarities, which allow proposing that these two genes had co-evolved. Both are unique to the Brassicaceae lineage, arising from genome block duplication [97,98,107]. Furthermore, the two genes are expressed similarly only in the central cell as well as during early endosperm development [104], where the maternal allele alone is expressed during early endosperm development [24,85,113,121–123]. In addition to *AtMEA* and *AtFIS2*, *AtFIE* gene is also expressed maternally during gametogenesis and early stage of embryo and endosperm development, while later on, its expression becomes bi-allelic, which persists during the entire plant life cycle [26,113,115,124].

Mutations of any of the four FIS2-PRC2 complex members result in the autonomous development of the endosperm without prerequisite for fertilization [10,70–72]. Furthermore, in fertilized seeds, dysfunction of the FIS2-PRC2 complex leads to abnormal development of the embryo and overproliferation of the endosperm [10,12,24,70,73,125]. In both cases, the seed-like structure or mutant seed eventually aborts [11,70,71]. In addition to its role in regulating endosperm proliferation, FIS2-PRC2 complex is involved in establishing the anterior–posterior polar axis in the endosperm [126].

Whole-genome profiling of H3K27met marks in *Arabidopsis* seedlings resulted in the identification of ±4400 genes containing the mark [21,22]. Yet it remains to be found, which additional genes are regulated by PRC2 complexes during gametogenesis and early seed development.

3.2. PRC2 homologs in dicots other than Arabidopsis

Studies of distantly related *Petunia hybrida* and *Solanum lycopersicum* (tomato) revealed duplication of the *CLF* gene in Solanaceae family. In *Petunia*, three *E(z)*-homologous genes, *PhCLF1*, *PhCLF2* and *PhCLF3* were isolated and partially characterized [127]. Sequence analysis demonstrated that *PhCLF1* and *PhCLF2* are closely related and are orthologs of *AtCLF*, whereas *PhCLF3* (*PhSWN*) is an ortholog of the *AtSWN* (Table 1). The tomato genome also contains two *CLF*-homologous sequences, termed *SlEZ2* (*SlCLF1*) and *SlEZ3* (*SlCLF2*), and one SWN homolog, termed *SlEZ1* (*SlSWN*) [128] (Table 1). Interestingly, How Kit et al. [128] reported that the similarity between *SlCLF1* and *PhCLF1*, or *SlSWN* and *PhSWN* is higher than between *CLF* and *SWN* homologs within each species. Based on the above, it was proposed that duplication of the *CLF* gene occurred in a common ancestor of Solanaceae lineage after its split from the *Arabidopsis* lineage, but before the speciation of tomato and *Petunia*, whereas the *CLF*/*SWN* divergence occurred before separation of rosids and asterids clades (Fig. 2) [127,128]. No orthologs of the *Arabidopsis MEA* gene were identified in these species [127,128], which is in agreement with the hypothesis that *MEA* evolved specifically in the Brassicaceae lineage.

There is evidence that in *Petunia*, the two *CLF* paralogs are under separate regulation. First, while both *PhCLF1* and *PhCLF2* genes are expressed in leaves and all floral organs, the pattern of transcripts accumulation levels among various plant organs is different for each gene [127]. Moreover, several *PhCLF1* splice variants were isolated from different tissues, whereas no alternative splicing was detected for *PhCLF2* [127]. The differences in regulation on *PhCLF1* and *PhCLF2* suggest that their corresponding proteins may have acquired at least some degree of separate functions, a hypothesis which could be tested in the future.

Unlike *Petunia* [127], tomato probably has only one functional *CLF* homolog, the *SlCLF1*, as *SlCLF2* is most likely a pseudogene [128]. The expression profile of the *SlCLF1* is notably different from that of *SlSWN*. While both genes are expressed in vegetative organs and enriched in

flower and fruit tissues [128], each of them is characterized by a distinct transcript accumulation patterns. Both *SlCLF1* and *SlSWN* were detected in young seeds and developing fruit tissues, but only *SlSWN* retained high expression levels during fruit ripening, suggesting distinct functions for the two proteins during this stage. Moreover, silencing of *SlSWN* by RNAi resulted in abnormal development of flower organs, indicating that SlCLF1 cannot compensate for the lack of SlSWN. Particularly interesting was the occasional styles fasciation and increase of locules number in fruit [128], the latter being determined by the number of carpels generated during early flower development [129]. In contrast to its role during flower development, no obvious phenotype was observed for *Slswn*-silenced plants in vegetative organs, which may indicate redundancy with SlCLF1 at this stage [128]. Taken together the above findings, How Kit et al. [128] suggested that tomato has at least two PRC2 complexes with partially overlapping functions, where SlSWN plays a distinct role in regulating floral organs growth and shape, and in establishing carpel number during early flower development.

There are several notable similarities between tomato and *Arabidopsis* PcG proteins functions. In both species the SET-domain proteins were found to localize to the nucleus, in agreement with their function in regulating transcription ([104,128]; Ohad lab, unpublished results). Likewise, the widespread expression and probable functional redundancy between tomato SET-proteins is highly reminiscent of that of *Arabidopsis*, in a similar manner to AtSWN and AtCLF [16]. Moreover, the multi-carpel gynoecia in tomato *Slswn* mutants [128] are reminiscent of the phenotype reported for *Atfie* plants [26], in which most ovaries comprised three fused carpels instead of two. Fasciation was also observed in *Arabidopsis fie* mutants; however, in inflorescence stems and not in stamens [26]. As SlSWN presumably functions within a PRC2 complex, it would be interesting to see whether mutation in tomato *FIE*-homolog will result in a similar flower phenotype. Unlike in *Arabidopsis* EMF2-PRC2 mutants, in *SlSWN* deficient tomato plants no homeotic transformations of the flower organs were reported [128], thus indicating that this gene is not involved in establishing floral organ identity. It is worth noticing that *Slswn* single mutant has a pronounced phenotype, as opposed the *Arabidopsis Atswn* mutant, which demonstrates a different balance in the roles of CLF and SWN between *Arabidopsis* and tomato, which are yet to be determined.

An additional putative regulatory role of PcG proteins was reported in the aspen tree *Populus tremula*, in which the *FIE* homolog is upregulated during the cambial dormancy in winter [99]. The transition of a cambial meristem from active to dormant state involves global alteration of gene expression that results in an overall decrease of transcriptome complexity. The correlation between *PtFIE* upregulation and global changes in gene transcription patterns led Schrader et al.[99] to propose that PtFIE may play a role in coordinating these changes during the induction of dormancy in poplar [99].

4. The role of Polycomb proteins in asexual reproduction

Given the functions of PcG complexes regulating seed and fruit development in *Arabidopsis* [10,26,71–73], the isolation of homologous PcG members in asexual species is intriguing. Certain members of the *Hieracium* genus (a Compositae; dicots), can form viable seeds asexually by a process called apomixis, which is characterized by the development of autonomous embryo and endosperm from either a somatic cell or a gamete within the ovule in a fertilization-independent manner [130–132]. Thus, apomixis produces clonal progenies, unlike sexual reproduction, which gives rise to genetically diverse offspring's. Understanding and controlling the production of genetically identical seeds have the potential for improving and manipulating seed production and plant breeding [133].

In *Hieracium*, apomixis is preceded by the initiation of normal ovule development up to the stage of megaspore mother cell (MMC) (the progenitor of female gamete). Furthermore, both sexual and asexual processes share a number of regulatory proteins [131]. For these reasons, it was suggested that apomixis may have evolved through deregulation of sexual reproductive programs, possibly including alterations in the epigenetic mechanisms [132].

Recently, an *AtFIE* homolog was characterized in both apomictic *Hieracium* (*H. piloselloides*) and non-apomictic *Hieracium* (*H. pilosella*) species [134]. *HpFIE* genes sequences and expression patterns are similar in the both species, resembling those of their *Arabidopsis* homolog [9,12,26,70]. The *HpFIE* transcripts were detected in leaves, floral organs, ovaries throughout the development of the embryo sac, and during early seed development in both embryo and endosperm [134] (Table 1). Furthermore, suppression of *HpFIE* resulted in a diminutive plant with curled leaves, resembling that of the *Arabidopsis Atclf* or *Atfie* mutants [9,26], possibly hinting for conserved functions of FIE proteins in *Hieracium* and *Arabidopsis*. This cross-family similarity is further supported by the interaction of HpFIE with AtCLF *in-vitro*, suggesting that HpFIE may form a PRC2-like complex together with a putative *Hieracium* CLF homolog [134]. In contrast to *Arabidopsis*, suppression of *HpFIE* expression in sexual plants did not lead to autonomous proliferation of endosperm. When mutant sexual *Hieracium* plants were fertilized, embryo and endosperm development was initiated, followed by seeds abortion at early globular stage. In apomictic *HpFIE*-supressed plants, formation of apomictic embryo sac and early seed occurred, whereas further development of embryo and endosperm was significantly inhibited. Interestingly, in both sexual and apomictic seeds, downregulation of *HpFIE* resulted in endosperm failing to cellularize [134], which is highly similar to that observed in *Arabidopsis Atfie* mutants [12,71]. Therefore, Rodrigues et al. [134] conclude that *HpFIE* is required for embryo and endosperm development in both sexual and apomictic *Hieracium* species, but has no role in repressing central cell proliferation.

In addition to *HpFIE*, homologs of *Arabidopsis AtMSI1* and *AtRBR1 (RETINOBLASTOMA-RELATED 1)* were identified in apomictic and sexual *Hieracium* species, which are both expressed in ovaries [134]. In *Arabidopsis*, AtRBR1 interact *in-vivo* with AtMSI1 and AtFIE, [135,136] while AtFIE and AtMSI1 were found to interact *in-vitro*[24]. Mutations in any of these genes in *Arabidopsis* lead to autonomous cell proliferation in the embryo sac [12,24,26,137]. Interestingly, AtFIE was shown to interact with HpRBR and HpMSI1 from both apomictic and sexual *Hieracium* plants [134], while HpFIEapo and HpFIEsex (FIE proteins from apomictic and sexual *Hieracium* species, respectively) failed to interact with neither the *Arabidopsis* nor the *Hieracium* RBR1 and MSI1 homologs. The authors [134] suggest that this disparity in protein interaction may result from structural differences, predicted from *Arabidopsis* and *Hieracium* FIE proteins structural models. Interestingly, when *HpFIEapo* was introduced into *Arabidopsis Atfie* mutants, it was able to complement the mutant phenotype, thus restoring the requirement for fertilization to allow endosperm development. Taken together the above, the authors [134] suggest that the dissimilarity of FIE functions in seed development between *Arabidopsis* and *Hieracium* relate to differences in PRC2-complex composition.

5. Functions of PcG complexes in monocots

The role of Polycomb-mediated epigenetic regulation in monocot species is emerging in recent years. In monocots, and especially in cereals, endosperm serves as a nutrients source, which supports the germinating seedling [138,139]. Thus, unlike most dicots, endosperm resources are not stored in the cotyledons or consumed by the embryo during seed development, but are mostly consumed post seed germination. Cereal crops, such as maize, rice and barley, accumulate proteins, starch, and fatty acids in the mature kernel, which provide up to half of the world's food supply [140] Thus, understanding the mechanisms governing endosperm development in monocots may allow us to manipulate and increase seed yield.

Homologs of PcG genes in barley (*Hordeum vulgare*) were identified through a query of the expressed sequences database [141]. The barley genome encodes single-copy homologs of *ESC* and *E(z)*, termed *HvFIE* and *HvE(Z)*, respectively. Phylogenetic analysis revealed that the *HvE(Z)* sequence could be grouped as part of the *SWN* clade, therefore in this review it will be termed *HvSWN*. In addition, three *Su(z)12-like* genes were identified, designated *HvSu(z)12a* (*HvEMF2a*), *HvSu(z)12b* (*HvEMF2b*) and *HvSu(z)12c* (*HvEMF2c*) (Table 1). Expression of all five barley PcG genes increased during seed development, yet highest levels were detected in young shoots [141]. *HvFIE*, *HvSWN* and *HvEMF2b*, were detected throughout plant development, in both vegetative and reproductive organs. These similar expression profiles suggest that the three protein subunits act together in a putative PRC2 complex. On the other hand, *HvEMF2c* expression was limited to young shoots, leaves and developing seeds, whereas *HvEMF2a* was undetectable but in shoots [141]. It is possible, therefore, that similarly to *Arabidopsis*, barley too has several PRC2 complexes, in which different zinc-finger subunits take part. Additional studies are required to corroborate this hypothesis.

Interestingly, expression of *HvFIE* and *HvSWN* was significantly elevated when the plant hormone abscisic acid (ABA) was added exogenously [141]. In *Arabidopsis*, ABA is involved in regulating various aspects of seed development [142]. Notably, during seed filling ABA plays a role in inducing the master regulator gene *FUS3* [143], while the PcG complexes repress its expression during both reproductive and vegetative phases [84]. Based on the above, the authors [141] hypothesized that the ABA-mediated processes in barely may also include activation of autoregulatory function, mediated by PcG complexes.

Finally, significant differences in *HvFIE* and *HvSWN* expression patterns during seed development were observed between two different barley cultivars with varying seeds size [141]. Expression of both genes starts to decline in the large-seeds Caresse cultivar approximately at the early stage of endosperm cellularization. However, in the small-seeds Ippolytos cultivar, the level of *HvSWN* expression remains high and that of *HvFIE* further increases. Therefore, the authors hypothesized that the variations in expression of *HvFIE* and *HvSWN* among the two cultivars may reflect the differences in seed filling and the resulting seed size [141].

Whereas a single *FIE* gene is present in dicots, two *FIE*-like genes were identified in the genomes of maize, rice (Fig. 2 and Table 1) and closely related *Sorghum bicolor* [55,97,144], suggesting that duplication of *FIE* occurred early in the Poaceae lineage evolution (Fig. 2, node 7). Contradictory, only one *FIE*-homologous sequence was identified in barley; however, as its genome is not fully mapped, additional genes may be discovered in the future.

The two *FIE* homologs of maize (*Zea mays*), *ZmFIE1* and *ZmFIE2*, show marked differences in expression and imprinting patterns. Similarly to the *Arabidopsis AtFIE* ([12], Ohad lab, unpublished results), *ZmFIE2* is broadly expressed in a bi-allelic manner, being detected in most organs throughout plant development except for sperm [55,144,145] (Table 1). During the reproductive stage, *ZmFIE2* is transcribed in the ovules prior to fertilization, where it is expressed in both egg and central cells [145]. Following fertilization, *ZmFIE2* is expressed throughout seed development, bi-allelic in the embryo and maternally in the endosperm [144]. Contrastingly to the *ZmFIE2*, *ZmFIE1* is expressed exclusively in the endosperm, starting ~24–32 h after fertilization, only from the maternal allele [144,145] probably due to DNA methylation of the parental loci [145–147]. Taking together this data, the authors [144] suggested that *ZmFIE2* may function similarly to *AtFIE*, regulating developmental processes in vegetative organs and repressing proliferation of the central cell within the gametophyte until fertilization takes place. Conversely, *ZmFIE1* may have acquired a separate role in regulating endosperm development, similar to that of *Arabidopsis* members comprising the FIS2-PRC2 complex. In both cases, lineage-specific *PcG* genes display endosperm specific expression and imprinting of the paternal allele [113].

Domesticated maize is a tetraploid, containing genes from two ancestral diploid species [148]. Interestingly, while the exons of both *ZmFIE1* and *ZmFIE2* are highly conserved, the noncoding genomic regions of the two genes show no similarity between them [55,144]. Moreover, the two genes are positioned on different chromosomes. Based on the fact that two *FIE* genes were identified in several Poacease family species, it is assumed that the duplication of *FIE* occurred prior to the polyploidization of maize. Therefore, it was suggested that two extant *ZmFIE* genes remained after a reciprocal deletion of the two additional ancestral paralogs occurred following polyploidization [97].

Three *E(z)* homologs were identified in maize: *ZmMez1*, *ZmMez2*, and *ZmMez3* [55,149]. Sequence analysis revealed that *ZmMez1* (*ZmCLF*) is homologous to *AtCLF*, whereas the closely related *ZmMez2* (*ZmSWN1*) and *ZmMez3* (*ZmSWN2*) are more similar to the *AtSWN*. As in the case of the maize *FIE* genes, *ZmCLF* and *ZmSWN1/2* are postulated to have distinct evolutionary origins, based on their low sequence similarity. Expression of all three *E(z)* homologs was detected throughout the maize life cycle. Interestingly, *ZmCLF* is imprinted in the endosperm with the maternal allele active, whereas both alleles are expressed in the developing embryo. Contrastingly, *ZmSWN1* and *ZmSWN2* are not imprinted. In addition to the three distinct *E(z)* maize homologs, two splice variants of *ZmSWN1* were detected, where each displaying distinct profile of expression [55]. Therefore, the multiplicity of PcG orthologs in maize may allow for the formation of diverse PRC2 complexes, each dedicated to regulate expression of distinct set of genes.

Rice (*Oryza sativa*) is an important cereal crop that underwent considerable cultivation. Rice PcG homologs of *Drosophila E(z)*, *Su(z)12* and *ESC* were identified and analyzed [97,98,150,151]. The rice genome contains two *E(z)*-like genes, designated *OsSWN* (*OsiEZ1* (or *OsSET1*)) and *OsCLF*, based on their sequence homology to the respective *Arabidopsis* genes [97,151]. Transient expression assays show that one of the rice E(z) homolog, OsSWN, localizes to the nucleus, supporting its predicted function in regulating transcription [150,151]. Rice genome also contains two homologs of *Su(z)12*: *OsEMF2a* and *OsEMF2b* [97,98]. No orthologs of the *Arabidopsis MEA*, *FIS2* or *VRN2* were identified in rice and in any of the cereals examined so far.

Similarly to maize [55,144,145], two rice ESC-like genes were identified, *OsFIE1* and *OsFIE2* [97]. *OsFIE1* displays high sequence similarity with the maize *ZmFIE1*, and similarly to maize the maternal copy is expressed restrictedly in the endosperm, while the paternal copy is not active [97]. Other rice PcG genes, including *OsFIE2*, are expressed in a wide range of vegetative and reproductive tissues and are not imprinted in the endosperm [150,151]. In contrast to maize, where the two *FIE* genes originated from separate ancestral genomes, in rice *OsFIE1* and *OsFIE2* mapped to a single chromosomal location, suggesting the two had evolved as a result of an intraspecies duplication event [97].

T-DNA insertion mutants of *OsFIE1*, *OsCLF* and *OsEMF2b* were characterized by Luo et al. [97]. Unlike in *Arabidopsis*, homozygous *Osclf* T-DNA mutant did not display any aberrant phenotype. This could be explained by partial redundancy between OsCLF and OsiEZ1, similarly to CLF and SWN in *Arabidopsis* [16]. Alternatively, it is possible that in rice the CLF homolog have no apparent developmental role. None of the above rice mutant T-DNA insertion lines, including a mutant in *Osfie1*, which is expressed specifically in the endosperm, displayed autonomous endosperm proliferation in the absence of fertilization. Based on these evidences the authors [97] suggested that these PcG genes (*OsFIE1*, *OsCLF* and *OsEMF2b*) are not involved in repression of central cell proliferation. Analysis of additional null-alleles, as well as double-mutant plants may help to elucidate the role of the PRC2 proteins during rice seed development. Homozygous *Osemf2b* T-DNA mutants demonstrated aberrant phenotype, characterized by early flowering in long-day conditions and abnormal development of floral organs [97], which is reminiscent of that of the *Arabidopsis emf2* mutant [62].

Multiple putative homologs of *AtEMF2* gene were identified in both rice and maize [98] (Table 1 and Fig. 2). The *EMF* homologs underwent

duplication in the monocots, probably in the grass lineage, as rice and maize have two *EMF2*-like genes [98], and barley encodes for three [141]. As other plants outside the Brassicaceae lineage, the cereals species do not encode *VRN2* and *FIS2* homologs.

6. Concluding remarks

During land plant evolution (Fig. 2), PcG genes "tended" to duplicate and subsequently diversify, allowing acquirement of versatile distinct functions. In unicellular algae species, single copy gene members of the PcG complex were identified. In early terrestrial plants, such as mosses and the ancestral vascular plants such as club mosses, certain PcG genes duplicated, to became a two to three-member gene family. In flowering plants, such as *Arabidopsis* and rice, nearly all the PRC2 core subunits underwent gene duplication. The expansion of PcG gene families along evolution was commonly coupled with functional diversification and speciation of orthologs, as deduced both from differences in the expression pattern of orthologs' as well as imprinting patterns (Table 1), and from mutants phenotypes. The current knowledge, summarized above, reveals that PcG proteins function in silencing gene expression through the modulation of chromatin state, developed early on in the evolution of eukaryotes. Moreover, the fundamental function of PcG proteins in regulating developmental programs along the plant life cycle, including phase transition from the gametophytic to the sporophytic phase, was demonstrated across diverse plant classes and phyla. This was exemplified in moss and *Arabidopsis*, which notably differ in their dominant life-cycle phase. Interestingly, it is notable that in several unrelated species of flowering plants, in which PcG genes duplicated to form multi-gene families, at least one the orthologs became endosperm-specific. This likely occurred independently, as members from different PcG genes families were designated to act in the endosperm, as in the case of *AtMEA* and *AtFIS2* (SET-domain protein and Zinc-finger protein, respectively) in *Arabidopsis* compared with Os*FIE1* and *ZmFIE1* (WD40 proteins) in monocot species. It is possible that these specialized genes allow formation of endosperm-specific PRC2 complexes, which may provide additional advantages in regulating endosperm development. Moreover, it is evident that the PRC2 genes that became endosperm-specific are also subjected to imprinting, allowing them to be expressed only from the maternal allele (Table 1) [55,104,107,123,144,145] [97,150,151]. The reoccurrence of PcG gene imprinting coincides with the kinship theory [152], reviewed by [124]), which allows control of the maternal nourishment of the embryo. It is feasible that in the near future identification of PcG target genes acting along the life cycle of different plant species will help to understand the functions and evolution and of the PRC2 complexes. Furthermore, elucidating the function of PcG proteins and their targets may serve in the future as tools to manipulate and improve important traits in crop plants.

Acknowledgements

We thank Daphna Michaeli for critical reading. Y.B. was partly supported by the Manna Foundation. N.O research was supported by grants from the German–Israeli Foundation for Scientific Research and Development (GIF, 154/2008); the German–Israeli Program (DIP, H 3. 1/030.6388) and the Israel Science Foundation (ISF, 767/09).

References

[1] G. Struhl, A gene product required for correct initiation of segmental determination in Drosophila, Nature 293 (1981) 36–41.
[2] G. Struhl, Role of the esc+ gene product in ensuring the selective expression of segment-specific homeotic genes in Drosophila, J. Embryol. Exp. Morphol. 76 (1983) 297–331.
[3] S.S. Sathe, P.J. Harte, The Drosophila extra sex combs protein contains WD motifs essential for its function as a repressor of homeotic genes, Mech. Dev. 52 (1995) 77–87.
[4] O.N. Denisenko, K. Bomsztyk, The product of the murine homolog of the Drosophila extra sex combs gene displays transcriptional repressor activity, Mol. Cell. Biol. 17 (1997) 4707–4717.
[5] R. Holdeman, S. Nehrt, S. Strome, MES-2, a maternal protein essential for viability of the germline in *Caenorhabditis elegans*, is homologous to a Drosophila Polycomb group protein, Development 125 (1998) 2457–2467.
[6] B. Schuettengruber, D. Chourrout, M. Vervoort, B. Leblanc, G. Cavalli, Genome regulation by polycomb and trithorax proteins, Cell 128 (2007) 735–745.
[7] Y.B. Schwartz, V. Pirrotta, Polycomb silencing mechanisms and the management of genomic programmes, Nat. Rev. Genet. 8 (2007) 9–22.
[8] Y.B. Schwartz, V. Pirrotta, Polycomb complexes and epigenetic states, Curr. Opin. Cell Biol. 20 (2008) 266–273.
[9] J. Goodrich, P. Puangsomlee, M. Martin, D. Long, E.M. Meyerowitz, G. Coupland, A polycomb-group gene regulates homeotic gene expression in Arabidopsis, Nature 386 (1997) 44–51.
[10] U. Grossniklaus, J.P. Vielle-Calzada, M.A. Hoeppner, W.B. Gagliano, Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis, Science 280 (1998) 446–450.
[11] M. Luo, P. Bilodeau, A. Koltunow, E.S. Dennis, W.J. Peacock, A.M. Chaudhury, Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 296–301.
[12] N. Ohad, R. Yadegari, L. Margossian, M. Hannon, D. Michaeli, J.J. Harada, R.B. Goldberg, R.L. Fischer, Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization, Plant Cell 11 (1999) 407–416.
[13] C. Kohler, C.B. Villar, Programming of gene expression by Polycomb group proteins, Trends Cell Biol. 18 (2008) 236–243.
[14] T. Kinoshita, J.J. Harada, R.B. Goldberg, R.L. Fischer, Polycomb repression of flowering during early plant development, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 14156–14161.
[15] D. O'Carroll, S. Erhardt, M. Pagani, S.C. Barton, M.A. Surani, T. Jenuwein, The polycomb-group gene Ezh2 is required for early mouse development, Mol. Cell. Biol. 21 (2001) 4330–4336.
[16] Y. Chanvivattana, A. Bishopp, D. Schubert, C. Stock, Y.H. Moon, Z.R. Sung, J. Goodrich, Interaction of Polycomb-group proteins controlling flowering in Arabidopsis, Development 131 (2004) 5263–5276.
[17] L.A. Boyer, K. Plath, J. Zeitlinger, T. Brambrink, L.A. Medeiros, T.I. Lee, S.S. Levine, M. Wernig, A. Tajonar, M.K. Ray, G.W. Bell, A.P. Otte, M. Vidal, D.K. Gifford, R.A. Young, R. Jaenisch, Polycomb complexes repress developmental regulators in murine embryonic stem cells, Nature 441 (2006) 349–353.
[18] T.I. Lee, R.G. Jenner, L.A. Boyer, M.G. Guenther, S.S. Levine, R.M. Kumar, B. Chevalier, S.E. Johnstone, M.F. Cole, K. Isono, H. Koseki, T. Fuchikami, K. Abe, H.L. Murray, J.P. Zucker, B. Yuan, G.W. Bell, E. Herbolsheimer, N.M. Hannett, K. Sun, D.T. Odom, A.P. Otte, T.L. Volkert, D.P. Bartel, D.A. Melton, D.K. Gifford, R. Jaenisch, R.A. Young, Control of developmental regulators by Polycomb in human embryonic stem cells, Cell 125 (2006) 301–313.
[19] A. Mosquna, A. Katz, E.L. Decker, S.A. Rensing, R. Reski, N. Ohad, Regulation of stem cell maintenance by the Polycomb protein FIE has been conserved during land plant evolution, Development 136 (2009) 2433–2444.
[20] L. Ringrose, Polycomb comes of age: genome-wide profiling of target sites, Curr. Opin. Cell Biol. 19 (2007) 290–297.
[21] F. Turck, F. Roudier, S. Farrona, M.L. Martin-Magniette, E. Guillaume, N. Buisine, S. Gagnot, R.A. Martienssen, G. Coupland, V. Colot, Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27, PLoS Genet. 3 (2007) e86.
[22] X. Zhang, O. Clarenz, S. Cokus, Y.V. Bernatavichute, M. Pellegrini, J. Goodrich, S.E. Jacobsen, Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis, PLoS Biol. 5 (2007) e129.
[23] V. Gaudin, M. Libault, S. Pouteau, T. Juul, G. Zhao, D. Lefebvre, O. Grandjean, Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in Arabidopsis, Development 128 (2001) 4847–4858.
[24] C. Kohler, L. Hennig, R. Bouveret, J. Gheyselinck, U. Grossniklaus, W. Gruissem, Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development, EMBO J. 22 (2003) 4804–4814.
[25] C. heBaroux, S. Pien, U. Grossniklaus, Chromatin modification and remodeling during early seed development, Curr. Opin. Genet. Dev. 17 (2007) 473–479.
[26] A. Katz, M. Oliva, A. Mosquna, O. Hakim, N. Ohad, FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development, Plant J. 37 (2004) 707–719.
[27] L. Xu, W.H. Shen, Polycomb silencing of KNOX genes confines shoot stem cell niches in Arabidopsis, Curr. Biol. 18 (2008) 1966–1971.
[28] H. Cui, P.N. Benfey, Interplay between SCARECROW, GA and LIKE HETEROCHROMATIN PROTEIN 1 in ground tissue patterning in the Arabidopsis root, Plant J. 58 (2009) 1016–1027.
[29] B. Papp, J. Muller, Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins, Genes Dev. 20 (2006) 2041–2054.
[30] J. Muller, P. Verrijzer, Biochemical mechanisms of gene regulation by polycomb group protein complexes, Curr. Opin. Genet. Dev. 19 (2009) 150–158.
[31] R. Cao, L. Wang, H. Wang, L. Xia, H. Erdjument-Bromage, P. Tempst, R.S. Jones, Y. Zhang, Role of histone H3 lysine 27 methylation in Polycomb-group silencing, Science 298 (2002) 1039–1043.
[32] A. Kuzmichev, K. Nishioka, H. Erdjument-Bromage, P. Tempst, D. Reinberg, Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein, Genes Dev. 16 (2002) 2893–2905.
[33] B. Czermin, R. Melfi, D. McCabe, V. Seitz, A. Imhof, V. Pirrotta, Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites, Cell 111 (2002) 185–196.

[34] J. Muller, C.M. Hart, N.J. Francis, M.L. Vargas, A. Sengupta, B. Wild, E.L. Miller, M.B. O'Connor, R.E. Kingston, J.A. Simon, Histone methyltransferase activity of a Drosophila Polycomb group repressor complex, Cell 111 (2002) 197–208.
[35] R. Cao, Y. Zhang, The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3, Curr. Opin. Genet. Dev. 14 (2004) 155–164.
[36] R. Cao, H. Wang, J. He, H. Erdjument-Bromage, P. Tempst, Y. Zhang, Role of hPHF1 in H3K27 methylation and Hox gene silencing, Mol. Cell. Biol. 28 (2008) 1862–1872.
[37] K. Sarma, R. Margueron, A. Ivanov, V. Pirrotta, D. Reinberg, Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo, Mol. Cell. Biol. 28 (2008) 2718–2731.
[38] W. Fischle, Y. Wang, S.A. Jacobs, Y. Kim, C.D. Allis, S. Khorasanizadeh, Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains, Genes Dev. 17 (2003) 1870–1881.
[39] Z. Shao, F. Raible, R. Mollaaghababa, J.R. Guyon, C.T. Wu, W. Bender, R.E. Kingston, Stabilization of chromatin structure by PRC1, a Polycomb complex, Cell 98 (1999) 37–46.
[40] M. de Napoles, J.E. Mermoud, R. Wakao, Y.A. Tang, M. Endoh, R. Appanah, T.B. Nesterova, J. Silva, A.P. Otte, M. Vidal, H. Koseki, N. Brockdorff, Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation, Dev. Cell 7 (2004) 663–676.
[41] H. Wang, L. Wang, H. Erdjument-Bromage, M. Vidal, P. Tempst, R.S. Jones, Y. Zhang, Role of histone H2A ubiquitination in Polycomb silencing, Nature 431 (2004) 873–878.
[42] J.K. Stock, S. Giadrossi, M. Casanova, E. Brookes, M. Vidal, H. Koseki, N. Brockdorff, A.G. Fisher, A. Pombo, Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells, Nat. Cell Biol. 9 (2007) 1428–1435.
[43] N.J. Francis, R.E. Kingston, C.L. Woodcock, Chromatin compaction by a polycomb group protein complex, Science 306 (2004) 1574–1577.
[44] R. Margueron, G. Li, K. Sarma, A. Blais, J. Zavadil, C.L. Woodcock, B.D. Dynlacht, D. Reinberg, Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms, Mol. Cell 32 (2008) 503–518.
[45] J.A. Simon, R.E. Kingston, Mechanisms of polycomb gene silencing: knowns and unknowns, Nat. Rev. Mol. Cell Biol. 10 (2009) 697–708.
[46] S. Pien, U. Grossniklaus, Polycomb group and trithorax group proteins in Arabidopsis, Biochim. Biophys. Acta 1769 (2007) 375–382.
[47] V. Exner, E. Aichinger, H. Shu, T. Wildhaber, P. Alfarano, A. Caflisch, W. Gruissem, C. Kohler, L. Hennig, The chromodomain of LIKE HETEROCHROMATIN PROTEIN 1 is essential for H3K27me3 binding and function during Arabidopsis development, PLoS One 4 (2009) e5335.
[48] X. Zhang, S. Germann, B.J. Blus, S. Khorasanizadeh, V. Gaudin, S.E. Jacobsen, The Arabidopsis LHP1 protein colocalizes with histone H3 Lys27 trimethylation, Nat. Struct. Mol. Biol. 14 (2007) 869–871.
[49] M. Libault, F. Tessadori, S. Germann, B. Snijder, P. Fransz, V. Gaudin, The Arabidopsis LHP1 protein is a component of euchromatin, Planta 222 (2005) 910–925.
[50] S. Sung, Y. He, T.W. Eshoo, Y. Tamada, L. Johnson, K. Nakahigashi, K. Goto, S.E. Jacobsen, R.M. Amasino, Epigenetic maintenance of the vernalized state in *Arabidopsis thaliana* requires LIKE HETEROCHROMATIN PROTEIN 1, Nat. Genet. 38 (2006) 706–710.
[51] J.S. Mylne, L. Barrett, F. Tessadori, S. Mesnage, L. Johnson, Y.V. Bernatavichute, S.E. Jacobsen, P. Fransz, C. Dean, LHP1, the Arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 5012–5017.
[52] Y.Y. Levy, S. Mesnage, J.S. Mylne, A.R. Gendall, C. Dean, Multiple roles of Arabidopsis VRN1 in vernalization and flowering time control, Science 297 (2002) 243–246.
[53] M. Calonje, R. Sanchez, L. Chen, Z.R. Sung, EMBRYONIC FLOWER1 participates in polycomb group-mediated AG gene silencing in Arabidopsis, Plant Cell 20 (2008) 277–291.
[54] L. Sanchez-Pulido, D. Devos, Z.R. Sung, M. Calonje, RAWUL: a new ubiquitin-like domain in PRC1 ring finger proteins that unveils putative plant and worm PRC1 orthologs, BMC Genomics 9 (2008) 308.
[55] N.M. Springer, O.N. Danilevskaya, P. Hermon, T.G. Helentjaris, R.L. Phillips, H.F. Kaeppler, S.M. Kaeppler, Sequence relationships, conserved domains, and expression patterns for maize homologs of the polycomb group genes E(z), esc, and E(Pc), Plant Physiol. 128 (2002) 1332–1345.
[56] I. Korf, Y. Fan, S. Strome, The Polycomb group in *Caenorhabditis elegans* and maternal control of germline development, Development 125 (1998) 2469–2478.
[57] F. Tie, T. Furuyama, J. Prasad-Sinha, E. Jane, P.J. Harte, The Drosophila Polycomb Group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3, Development 128 (2001) 275–286.
[58] S. Poux, R. Melfi, V. Pirrotta, Establishment of Polycomb silencing requires a transient interaction between PC and ESC, Genes Dev. 15 (2001) 2509–2514.
[59] A. Birve, A.K. Sengupta, D. Beuchle, J. Larsson, J.A. Kennison, A. Rasmuson-Lestander, J. Muller, Su(z)12, a novel Drosophila Polycomb group gene that is conserved in vertebrates and plants, Development 128 (2001) 3371–3379.
[60] M.A. Martinez-Balbas, T. Tsukiyama, D. Gdula, C. Wu, Drosophila NURF-55, a WD repeat protein involved in histone metabolism, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 132–137.
[61] D. Preuss, Chromatin silencing and Arabidopsis development: a role for polycomb proteins, Plant Cell 11 (1999) 765–768.
[62] N. Yoshida, Y. Yanai, L. Chen, Y. Kato, J. Hiratsuka, T. Miwa, Z.R. Sung, S. Takahashi, EMBRYONIC FLOWER2, a novel polycomb group protein homolog, mediates shoot development and flowering in Arabidopsis, Plant Cell 13 (2001) 2471–2481.
[63] A.R. Gendall, Y.Y. Levy, A. Wilson, C. Dean, The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in Arabidopsis, Cell 107 (2001) 525–535.
[64] A.M. Chaudhury, A. Koltunow, T. Payne, M. Luo, M.R. Tucker, E.S. Dennis, W.J. Peacock, Control of early seed development, Annu. Rev. Cell Dev. Biol. 17 (2001) 677–699.
[65] T.F. Hsieh, O. Hakim, N. Ohad, R.L. Fischer, From flour to flower: how Polycomb group proteins influence multiple aspects of plant development, Trends Plant Sci. 8 (2003) 439–445.
[66] S. Sung, R.M. Amasino, Vernalization and epigenetics: how plants remember winter, Curr. Opin. Plant Biol. 7 (2004) 4–10.
[67] N. Schatlowski, K. Creasey, J. Goodrich, D. Schubert, Keeping plants in shape: polycomb-group genes and histone methylation, Semin. Cell Dev. Biol. 19 (2008) 547–553.
[68] A.E. Guitton, F. Berger, Control of reproduction by Polycomb Group complexes in animals and plants, Int. J. Dev. Biol. 49 (2005) 707–716.
[69] D.H. Kim, M.R. Doyle, S. Sung, R.M. Amasino, Vernalization: winter and the timing of flowering in plants, Annu. Rev. Cell Dev. Biol. 25 (2009) 277–299.
[70] T. Kiyosue, N. Ohad, R. Yadegari, M. Hannon, J. Dinneny, D. Wells, A. Katz, L. Margossian, J.J. Harada, R.B. Goldberg, R.L. Fischer, Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 4186–4191.
[71] N. Ohad, L. Margossian, Y.C. Hsu, C. Williams, P. Repetti, R.L. Fischer, A mutation that allows endosperm development without fertilization, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 5319–5324.
[72] A.M. Chaudhury, L. Ming, C. Miller, S. Craig, E.S. Dennis, W.J. Peacock, Fertilization-independent seed development in Arabidopsis thaliana, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 4223–4228.
[73] A.E. Guitton, D.R. Page, P. Chambrier, C. Lionnet, J.E. Faure, U. Grossniklaus, F. Berger, Identification of new members of fertilisation independent seed polycomb group pathway involved in the control of seed development in Arabidopsis thaliana, Development 131 (2004) 2971–2981.
[74] C.H. Yang, L.J. Chen, Z.R. Sung, Genetic regulation of shoot development in Arabidopsis: role of the EMF genes, Dev. Biol. 169 (1995) 421–435.
[75] S.Y. Kim, T. Zhu, Z.R. Sung, Epigenetic regulation of gene programs by EMF1 and EMF2 in Arabidopsis, Plant Physiol. 152 (2010) 516–528.
[76] J. Chandler, A. Wilson, C. Dean, Arabidopsis mutants showing an altered response to vernalization, Plant J. 10 (1996) 637–644.
[77] C.C. Wood, M. Robertson, G. Tanner, W.J. Peacock, E.S. Dennis, C.A. Helliwell, The Arabidopsis thaliana vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 14631–14636.
[78] F. De Lucia, P. Crevillen, A.M. Jones, T. Greb, C. Dean, A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization, Proc. Natl. Acad. Sci. U.S.A. 105 (2008) 16831–16836.
[79] C. Kohler, L. Hennig, C. Spillane, S. Pien, W. Gruissem, U. Grossniklaus, The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1, Genes Dev. 17 (2003) 1540–1553.
[80] N. Schonrock, R. Bouveret, O. Leroy, L. Borghi, C. Kohler, W. Gruissem, L. Hennig, Polycomb-group proteins repress the floral activator AGL19 in the FLC-independent vernalization pathway, Genes Dev. 20 (2006) 1667–1678.
[81] R. Bastow, J.S. Mylne, C. Lister, Z. Lippman, R.A. Martienssen, C. Dean, Vernalization requires epigenetic silencing of FLC by histone methylation, Nature 427 (2004) 164–167.
[82] S. Sung, R.M. Amasino, Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3, Nature 427 (2004) 159–164.
[83] D. Schubert, L. Primavesi, A. Bishopp, G. Roberts, J. Doonan, T. Jenuwein, J. Goodrich, Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27, EMBO J. 25 (2006) 4638–4649.
[84] G. Makarevich, O. Leroy, U. Akinci, D. Schubert, O. Clarenz, J. Goodrich, U. Grossniklaus, C. Kohler, Different Polycomb group complexes regulate common target genes in Arabidopsis, EMBO Rep. 7 (2006) 947–952.
[85] P.E. Jullien, A. Katz, M. Oliva, N. Ohad, F. Berger, Polycomb group complexes self-regulate imprinting of the Polycomb group gene MEDEA in Arabidopsis, Curr. Biol. 16 (2006) 486–492.
[86] C. Kohler, L. Hennig, Regulation of cell identity by plant Polycomb and trithorax group proteins, Curr. Opin. Genet. Dev. 20 (2010) 541–547.
[87] L. Hennig, M. Derkacheva, Diversity of Polycomb group complexes in plants: same rules, different players? Trends Genet. 25 (2009) 414–423.
[88] Y. Okano, N. Aono, Y. Hiwatashi, T. Murata, T. Nishiyama, T. Ishikawa, M. Kubo, M. Hasebe, A polycomb repressive complex 2 gene regulates apogamy and gives evolutionary insights into early land plant evolution, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 16321–16326.
[89] M. Lachner, R. Sengupta, G. Schotta, T. Jenuwein, Trilogies of histone lysine methylation as epigenetic landmarks of the eukaryotic genome, Cold Spring Harb. Symp. Quant. Biol. 69 (2004) 209–218.
[90] C.S. Veerappan, Z. Avramova, E.N. Moriyama, Evolution of SET-domain protein families in the unicellular and multicellular Ascomycota fungi, BMC Evol. Biol. 8 (2008) 190.
[91] P. Trojer, D. Reinberg, Histone lysine demethylases and their impact on epigenetics, Cell 125 (2006) 213–217.
[92] J.L. Bowman, S.K. Floyd, K. Sakakibara, Green genes-comparative genomics of the green branch of life, Cell 129 (2007) 229–234.
[93] J.H. Waterborg, A.J. Robertson, D.L. Tatar, C.M. Borza, J.R. Davie, Histones of Chlamydomonas reinhardtii. Synthesis, acetylation, and methylation, Plant Physiol. 109 (1995) 393–407.
[94] S. Shaver, J.A. Casas-Mollano, R.L. Cerny, H. Cerutti, Origin of the polycomb repressive complex 2 and gene silencing by an E(z) homolog in the unicellular alga Chlamydomonas, Epigenetics 5 (2010) 301–312.

[95] Y. Liu, S.D. Taverna, T.L. Muratore, J. Shabanowitz, D.F. Hunt, C.D. Allis, RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in Tetrahymena, Genes Dev. 21 (2007) 1530–1545.
[96] G. THE ANGIOSPERM PHYLOGENY, An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III, Bot. J. Linn Soc. 161 (2009) 105–121.
[97] M. Luo, D. Platten, A. Chaudhury, W.J. Peacock, E.S. Dennis, Expression, imprinting, and evolution of rice homologs of the polycomb group genes, Mol. Plant 2 (2009) 711–723.
[98] L.J. Chen, Z.Y. Diao, C. Specht, Z.R. Sung, Molecular evolution of VEF-domain-containing PcG genes in plants, Mol. Plant 2 (2009) 738–754.
[99] J. Schrader, R. Moyle, R. Bhalerao, M. Hertzberg, J. Lundeberg, P. Nilsson, R.P. Bhalerao, Cambial meristem dormancy in trees involves extensive remodelling of the transcriptome, Plant J. 40 (2004) 173–187.
[100] S.M. Adl, A.G. Simpson, M.A. Farmer, R.A. Andersen, O.R. Anderson, J.R. Barta, S.S. Bowser, G. Brugerolle, R.A. Fensome, S. Fredericq, T.Y. James, S. Karpov, P. Kugrens, J. Krug, C.E. Lane, L.A. Lewis, J. Lodge, D.H. Lynn, D.G. Mann, R.M. McCourt, L. Mendoza, O. Moestrup, S.E. Mozley-Standridge, T.A. Nerad, C.A. Shearer, A.V. Smirnov, F.W. Spiegel, M.F. Taylor, The new higher level classification of eukaryotes with emphasis on the taxonomy of protists, J. Eukaryot. Microbiol. 52 (2005) 399–451.
[101] V. Exner, P. Taranto, N. Schonrock, W. Gruissem, L. Hennig, Chromatin assembly factor CAF-1 is required for cellular differentiation during plant development, Development 133 (2006) 4163–4172.
[102] J.K. Tyler, M. Bulger, R.T. Kamakaka, R. Kobayashi, J.T. Kadonaga, The p55 subunit of Drosophila chromatin assembly factor 1 is homologous to a histone deacetylase-associated protein, Mol. Cell. Biol. 16 (1996) 6149–6159.
[103] C.J. Harrison, A.H. Roeder, E.M. Meyerowitz, J.A. Langdale, Local cues and asymmetric cell divisions underpin body plan transitions in the moss Physcomitrella patens, Curr. Biol. 19 (2009) 461–471.
[104] D. Wang, M.D. Tyson, S.S. Jackson, R. Yadegari, Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in Arabidopsis, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 13244–13249.
[105] C. Kohler, G. Makarevich, Epigenetic mechanisms governing seed development in plants, EMBO Rep. 7 (2006) 1223–1227.
[106] A. Kawabe, R. Fujimoto, D. Charlesworth, High diversity due to balancing selection in the promoter region of the Medea gene in *Arabidopsis lyrata*, Curr. Biol. 17 (2007) 1885–1889.
[107] C. Spillane, K.J. Schmid, S. Laoueille-Duprat, S. Pien, J.M. Escobar-Restrepo, C. Baroux, V. Gagliardini, D.R. Page, K.H. Wolfe, U. Grossniklaus, Positive Darwinian selection at the imprinted MEDEA locus in plants, Nature 448 (2007) 349–352.
[108] T. Miyake, N. Takebayashi, D.E. Wolf, Possible diversifying selection in the imprinted gene, MEDEA, in Arabidopsis, Mol. Biol. Evol. 26 (2009) 843–857.
[109] D. Grant, P. Cregan, R.C. Shoemaker, Genome organization in dicots: genome duplication in Arabidopsis and synteny between soybean and Arabidopsis, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 4168–4173.
[110] C. Simillion, K. Vandepoele, M.C. Van Montagu, M. Zabeau, Y. Van de Peer, The hidden duplication past of *Arabidopsis thaliana*, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 13627–13632.
[111] G. Blanc, K. Hokamp, K.H. Wolfe, A recent polyploidy superimposed on older large-scale duplications in the Arabidopsis genome, Genome Res. 13 (2003) 137–144.
[112] W. Xiao, M. Gehring, Y. Choi, L. Margossian, H. Pu, J.J. Harada, R.B. Goldberg, R.I. Pennell, R.L. Fischer, Imprinting of the MEA Polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase, Dev. Cell 5 (2003) 891–901.
[113] M. Luo, P. Bilodeau, E.S. Dennis, W.J. Peacock, A. Chaudhury, Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 10637–10642.
[114] C. Spillane, C. MacDougall, C. Stock, C. Kohler, J.P. Vielle-Calzada, S.M. Nunes, U. Grossniklaus, J. Goodrich, Interaction of the Arabidopsis polycomb group proteins FIE and MEA mediates their common phenotypes, Curr. Biol. 10 (2000) 1535–1538.
[115] R. Yadegari, T. Kinoshita, O. Lotan, G. Cohen, A. Katz, Y. Choi, A. Katz, K. Nakashima, J.J. Harada, R.B. Goldberg, R.L. Fischer, N. Ohad, Mutations in the FIE and MEA genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms, Plant Cell 12 (2000) 2367–2382.
[116] K. Bracha-Drori, K. Shichrur, A. Katz, M. Oliva, R. Angelovici, S. Yalovsky, N. Ohad, Detection of protein-protein interactions in plants using bimolecular fluorescence complementation, Plant J. 40 (2004) 419–427.
[117] L. Hennig, P. Taranto, M. Walser, N. Schonrock, W. Gruissem, Arabidopsis MSI1 is required for epigenetic maintenance of reproductive development, Development 130 (2003) 2555–2565.
[118] D. Jiang, Y. Wang, Y. Wang, Y. He, Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the Arabidopsis Polycomb repressive complex 2 components, PLoS One 3 (2008) e3404.
[119] T. Kotake, S. Takada, K. Nakahigashi, M. Ohto, K. Goto, Arabidopsis TERMINAL FLOWER 2 gene encodes a heterochromatin protein 1 homolog and represses both FLOWERING LOCUS T to regulate flowering time and several floral homeotic genes, Plant Cell Physiol. 44 (2003) 555–564.
[120] Y.H. Moon, L. Chen, R.L. Pan, H.S. Chang, T. Zhu, D.M. Maffeo, Z.R. Sung, EMF genes maintain vegetative development by repressing the flower program in Arabidopsis, Plant Cell 15 (2003) 681–693.
[121] C. Baroux, V. Gagliardini, D.R. Page, U. Grossniklaus, Dynamic regulatory interactions of Polycomb group genes: MEDEA autoregulation is required for imprinted gene expression in Arabidopsis, Genes Dev. 20 (2006) 1081–1086.
[122] T. Kinoshita, R. Yadegari, J.J. Harada, R.B. Goldberg, R.L. Fischer, Imprinting of the MEDEA polycomb gene in the Arabidopsis endosperm, Plant Cell 11 (1999) 1945–1952.
[123] J.P. Vielle-Calzada, J. Thomas, C. Spillane, A. Coluccio, M.A. Hoeppner, U. Grossniklaus, Maintenance of genomic imprinting at the Arabidopsis medea locus requires zygotic DDM1 activity, Genes Dev. 13 (1999) 2971–2982.
[124] N. Ohad, Plant development: parental conflict overcome, Nature 447 (2007) 275–276.
[125] M. Ingouff, J. Haseloff, F. Berger, Polycomb group genes control developmental timing of endosperm, Plant J. 42 (2005) 663–674.
[126] M.B. Sorensen, A.M. Chaudhury, H. Robert, E. Bancharel, F. Berger, Polycomb group genes control pattern formation in plant seed, Curr. Biol. 11 (2001) 277–281.
[127] T. Mayama, E. Ohtsubo, S. Tsuchimoto, Isolation and expression analysis of petunia CURLY LEAF-like genes, Plant Cell Physiol. 44 (2003) 811–819.
[128] A. How Kit, L. Boureau, L. Stammitti-Bert, D. Rolin, E. Teyssier, P. Gallusci, Functional analysis of SlEZ1 a tomato enhancer of zeste (E(z)) gene demonstrates a role in flower development, Plant Mol. Biol. 74 (2010) 201–213.
[129] L.S. Barrero, B. Cong, F. Wu, S.D. Tanksley, Developmental characterization of the fasciated locus and mapping of Arabidopsis candidate genes involved in the control of floral meristem size and carpel number in tomato, Genome 49 (2006) 991–1006.
[130] A.M. Koltunow, Apomixis: embryo sacs and embryos formed without meiosis or fertilization in ovules, Plant Cell 5 (1993) 1425–1437.
[131] M.R. Tucker, A.C. Araujo, N.A. Paech, V. Hecht, E.D. Schmidt, J.B. Rossell, S.C. De Vries, A.M. Koltunow, Sexual and apomictic reproduction in Hieracium subgenus pilosella are closely interrelated developmental pathways, Plant Cell 15 (2003) 1524–1537.
[132] A.M. Koltunow, U. Grossniklaus, Apomixis: a developmental perspective, Annu. Rev. Plant Biol. 54 (2003) 547–574.
[133] A.M. Koltunow, R.A. Bicknell, A.M. Chaudhury, Apomixis: molecular strategies for the generation of genetically identical seeds without fertilization, Plant Physiol. 108 (1995) 1345–1352.
[134] J.C. Rodrigues, M.R. Tucker, S.D. Johnson, M. Hrmova, A.M. Koltunow, Sexual and apomictic seed formation in Hieracium requires the plant polycomb-group gene FERTILIZATION INDEPENDENT ENDOSPERM, Plant Cell 20 (2008) 2372–2386.
[135] P.E. Jullien, A. Mosquna, M. Ingouff, T. Sakata, N. Ohad, F. Berger, Retinoblastoma and its binding partner MSI1 control imprinting in Arabidopsis, PLoS Biol. 6 (2008) e194.
[136] A. Mosquna, A. Katz, S. Shochat, G. Grafi, N. Ohad, Interaction of FIE, a polycomb protein, with pRb: a possible mechanism regulating endosperm development, Mol. Genet. Genomics 271 (2004) 651–657.
[137] M. Ingouff, P.E. Jullien, F. Berger, The female gametophyte and the endosperm control cell proliferation and differentiation of the seed coat in Arabidopsis, Plant Cell 18 (2006) 3491–3501.
[138] O.A. Olsen, Nuclear endosperm development in cereals and *Arabidopsis thaliana*, Plant Cell 16 (2004) S214–S227 (Suppl).
[139] M. Gehring, Y. Choi, R.L. Fischer, Imprinting and seed development, Plant Cell 16 (2004) S203–S213 (Suppl).
[140] M.W. Rosegrant, N. Leach, R.V. Gerpacio, Alternative futures for world cereal and meat consumption, Proc. Nutr. Soc. 58 (1999) 219–234.
[141] A. Kapazoglou, A. Tondelli, D. Papaefthimiou, H. Ampatzidou, E. Francia, M.A. Stanca, K. Bladenopoulos, A.S. Tsaftaris, Epigenetic chromatin modifiers in barley: IV. The study of barley polycomb group (PcG) genes during seed development and in response to external ABA, BMC Plant Biol. 10 (2010) 73.
[142] E. Nambara, A. Marion-Poll, ABA action and interactions in seeds, Trends Plant Sci. 8 (2003) 213–217.
[143] Y. Kagaya, R. Okuda, A. Ban, R. Toyoshima, K. Tsutsumida, H. Usui, A. Yamamoto, T. Hattori, Indirect ABA-dependent regulation of seed storage protein genes by FUSCA3 transcription factor in Arabidopsis, Plant Cell Physiol. 46 (2005) 300–311.
[144] O.N. Danilevskaya, P. Hermon, S. Hantke, M.G. Muszynski, K. Kollipara, E.V. Ananiev, Duplicated fie genes in maize: expression pattern and imprinting suggest distinct functions, Plant Cell 15 (2003) 425–438.
[145] P. Hermon, K.O. Srilunchang, J. Zou, T. Dresselhaus, O.N. Danilevskaya, Activation of the imprinted Polycomb Group Fie1 gene in maize endosperm requires demethylation of the maternal allele, Plant Mol. Biol. 64 (2007) 387–395.
[146] J.F. Gutierrez-Marcos, L.M. Costa, M. Dal Pra, S. Scholten, E. Kranz, P. Perez, H.G. Dickinson, Epigenetic asymmetry of imprinted genes in plant gametes, Nat. Genet. 38 (2006) 876–878.
[147] J.F. Gutierrez-Marcos, P.D. Pennington, L.M. Costa, H.G. Dickinson, Imprinting in the endosperm: a possible role in preventing wide hybridization, Philos. Trans. R. Soc. Lond. B Biol. Sci. 358 (2003) 1105–1111.
[148] B.S. Gaut, Patterns of chromosomal duplication in maize and their implications for comparative maps of the grasses, Genome Res. 11 (2001) 55–66.
[149] W.J. Haun, S. Laoueille-Duprat, J. O'Connell Test, Springer, Genomic imprinting, methylation and molecular evolution of maize Enhancer of zeste (Mez) homologs, Plant J. 49 (2007) 325–337.
[150] J.K. Thakur, M.R. Malik, V. Bhatt, M.K. Reddy, S.K. Sopory, A.K. Tyagi, J.P. Khurana, A POLYCOMB group gene of rice (*Oryza sativa* L. subspecies indica), OsiEZ1, codes for a nuclear-localized protein expressed preferentially in young seedlings and during reproductive development, Gene 314 (2003) 1–13.
[151] Y.K. Liang, Y. Wang, Y. Zhang, S.G. Li, X.C. Lu, H. Li, C. Zou, Z.H. Xu, S.N. Bai, OsSET1, a novel SET-domain-containing gene from rice, J. Exp. Bot. 54 (2003) 1995–1996.
[152] D. Haig, M. Westoby, Parent-specific gene-expression and the triploid endosperm, Am. Nat. 134 (1989) 147–155.

Further reading

[153] O. Leroy, L. Hennig, H. Breuninger, T. Laux, C. Kohler, Polycomb group proteins function in the female gametophyte to determine seed development in plants, Development 134 (2007) 3639–3648.
[154] L. Hennig, R. Bouveret, W. Gruissem, MSI1-like proteins: an escort service for chromatin assembly and remodeling complexes, Trends Cell Biol. 15 (2005) 295–302.
[155] M. Schmid, T.S. Davison, S.R. Henz, U.J. Pape, M. Demar, M. Vingron, B. Scholkopf, D. Weigel, J.U. Lohmann, A gene expression map of *Arabidopsis thaliana* development, Nat. Genet. 37 (2005) 501–506.

Glossary

Fasciation: An abnormal form of growth in which a shoot or organ becomes enlarged giving rise to fusion of several shoots or organs together.
Flower organs/structure: A basic flower is made of four concentric rings (whorls) of modified leaves. The outer whorl is the sepals, followed by petals, stamens (the male reproductive structures containing pollen) and in the very centre the female reproductive organs, the ovary. The ovary contains one or more ovules. From the ovary extends the style, a tubular structure bearing at the top a receptive surface designated the stigma, on which pollen land and start germination.
Fruit locules: A compartment or chamber within an ovary. Depending on the number of locules in the ovary, carpels and fruits can be classified as uni-locular, bi-locular or multi-locular etc.
Homeotic mutation: A change in a segment or organ identity during development.
Gametophore: Is the bearer of the sex organs (gametangia), the female archegonia and the male antheridia in bryophytes.
Kernel: A softer part of a seed, nut or fruit stone contained within its shell. The kernel of maize has a pericarp of the fruit fused with the seed coat, typical of the grasses, and the entire kernel is often referred to as the seed.
Protonema: Is a thread-like chain of cells that forms the earliest stage (the haploid phase) of a bryophyte life cycle.
Seed development: Flowering plants are characterized by a double-fertilization event. The ovule generates the female gametophyte, which contains the gamets: **egg cell** (n) and **central cell**, the latter contains two daughter nuclei that fuse before fertilization ($2n$). Upon fertilization, the pollen tube delivers two sperm nuclei, one fertilizes the egg to form the zygote ($2n$), whereas another sperm nucleus fuses with the diploid central cell nucleus to form the triploid ($3n$) **endosperm** serving as nutritive tissue for the developing embryo or seedling.
Vernalization: Is the acquisition of a plant's ability to flower or germinate in the spring by exposure to the prolonged cold of winter.