

Epigenetic control of Pollen Ole e 1 allergen and extensin family gene expression in *Arabidopsis thaliana*

Bo Hu · Boyu Liu · Ling Liu · Chunlin Liu ·
Lin Xu · Ying Ruan

Received: 6 January 2014 / Revised: 17 March 2014 / Accepted: 26 May 2014 / Published online: 8 June 2014
© Franciszek Górski Institute of Plant Physiology, Polish Academy of Sciences, Kraków 2014

Abstract Pollen Ole e 1 domain-containing proteins were initially identified as major allergens. However, recent studies have suggested that these proteins also function as developmental regulators in many other plant tissues. To understand the possible roles of these proteins in plant development, we characterized 28 Pollen Ole e 1 allergen and extensin (AtPOE1) family proteins from *Arabidopsis thaliana*, and analyzed the transcript levels of the genes encoding these proteins in various tissues. Our analysis revealed that most of the AtPOE1 family genes showed spatial transcription patterns, suggesting that they may be involved in various events at diverse developmental stages. Using one member of the family, *AtPOE1;26*, as a model, we showed that it is mainly expressed in the root. In addition, the *AtPOE1;26* gene locus has a significant histone H3 lysine 27 trimethylation (H3K27me3) modification, which is the typical epigenetic regulation marker of the Polycomb group (PcG) pathway. *AtPOE1;26* transcription was upregulated in the PcG mutant *curly leaf*

(*clf*), and the H3K27me3 level at the *AtPOE1;26* locus was reduced in *clf*, suggesting that its expression is controlled by PcG. Besides *AtPOE1;26*, the gene loci of at least 13 AtPOE1 family members were modified by H3K27me3. Overall, these data indicate that AtPOE1 family members have spatial expression patterns in multiple plant tissues, and that their expression could be epigenetically controlled.

Keywords *Arabidopsis* · AtPOE1 · Polycomb · Epigenetics

Introduction

Pollen Ole e 1 domain-containing proteins are major allergens, and are present in many plants (Jimenez-Lopez et al. 2011). The first Ole e 1 protein was purified and characterized from *Olea europaea* (Batanero et al. 1994; Lauzurica et al. 1988; Villalba et al. 1990, 1993). These proteins contain the Ole e 1 domain, which is defined as a conserved amino acid sequence of E/Q/T-G-X-V-Y-C-D-T/N/P-C-R (where X represents any amino acid) (Jiang et al. 2005; Jimenez-Lopez et al. 2011; Sigrist et al. 2010; Villalba et al. 1993). The amount of Ole e 1 protein in pollen increases as the pollen tube forms, and this protein can be released into the culture medium when pollen germinates *in vitro* (de Dios et al. 2004). However, recent studies have suggested that Ole e 1 domain-containing proteins function not only in pollen, but also in other plant tissues. A detailed characterization of genes encoding Ole e 1 domain-containing proteins in rice showed that their transcripts were present in several different tissues (Jiang et al. 2005). Some Ole e 1 domain-containing genes were shown to be specifically expressed in root cells during root

Communicated by M. Stobiecki.

Electronic supplementary material The online version of this article (doi:10.1007/s11738-014-1597-6) contains supplementary material, which is available to authorized users.

B. Hu · B. Liu · L. Liu · C. Liu · Y. Ruan (✉)
Hunan Provincial Key Laboratory of Crop Germplasm
Innovation and Utilization, Hunan Agricultural University,
Changsha 410128, Hunan, China
e-mail: yingruan@hotmail.com

B. Hu · L. Xu (✉)
National Laboratory of Plant Molecular Genetics, Institute
of Plant Physiology and Ecology, Shanghai Institutes
for Biological Sciences, Chinese Academy of Sciences,
300 Fenglin Road, Shanghai 200032, China
e-mail: xulin01@sibs.ac.cn

hair formation (Bruex et al. 2012). In *Arabidopsis thaliana*, the gene encoding Arabinogalactan-protein 30 (AGP30), an Ole e 1 domain-containing protein, was specifically expressed in the root, and was required for root regeneration and seed germination (van Hengel and Roberts 2003). Therefore, genes encoding Ole e 1 domain-containing proteins can show specific expression patterns and may play multiple roles in controlling plant development and physiology.

Epigenetic regulations are usually required for spatial and temporal control of gene expression. Polycomb group (PcG) proteins function in an evolutionarily conserved epigenetic pathway that represses gene expression via histone modifications (Whitcomb et al. 2007). In *Arabidopsis*, PcGs are found in two multiprotein complexes; Polycomb Repressive Complex 1 (PRC1) and PRC2. PRC2 trimethylates histone H3 lysine 27 (H3K27me3) at target loci, and PRC1 recognizes the trimethylated histone and mono-ubiquitinates H2A (Calonje 2013; He et al. 2013; Molitor and Shen 2013; Schatlowski et al. 2008). CURLY LEAF (CLF) is a core component of PRC2, and serves as a methyltransferase that specifically targets H3K27 (Chanvivattana et al. 2004; Goodrich et al. 1997; Schubert et al. 2006). In *Arabidopsis* seedlings, there are more than 4,000

gene loci with the H3K27me3 modification (Zhang et al. 2007), suggesting that PRC2 controls the expression of many diverse genes.

In this study, we show that the *Arabidopsis* genome encodes at least 28 Ole e 1 domain-containing proteins in the *Arabidopsis thaliana* Pollen Ole e 1 allergen and extensin (AtPOE1) family. The genes in this family showed different expression patterns, suggesting that they may function in regulating different aspects and stages of plant development. Using *AtPOE1;26* as a model, we showed that PRC2-mediated H3K27me3 modification is required to regulate its expression.

Materials and methods

Plant materials and growth conditions

The *Arabidopsis thaliana* L. mutant *clf-50* is in the Wasilewskija (Ws) background (Chanvivattana et al. 2004). Seeds of this mutant were obtained from Dr. J. Goodrich. We used Ws as the wild-type plant in this study. Plants were grown as described previously (Xu et al. 2003).

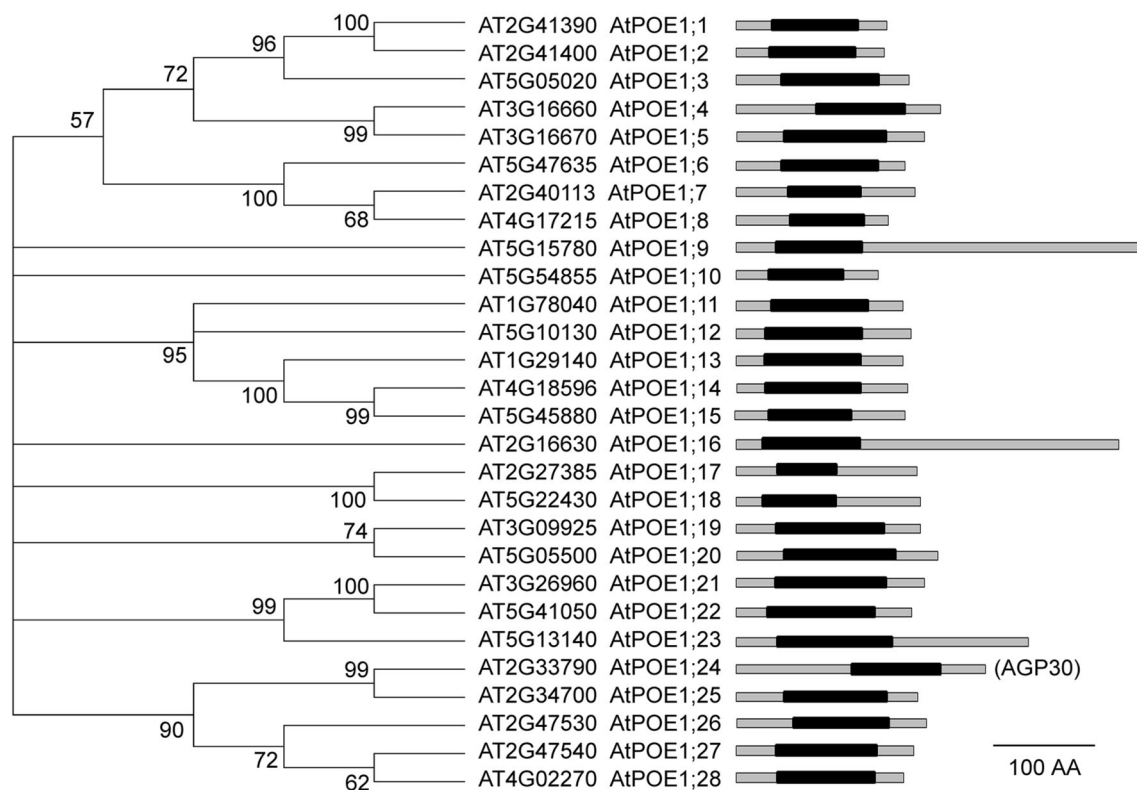


Fig. 1 Bioinformatic analysis of AtPOE1 family proteins. We identified 28 AtPOE1 family proteins in *Arabidopsis*. AtPOE1;24 corresponds to the previously characterized AGP30 (van Hengel and Roberts 2003). Full-length protein sequences were included in the

phylogenetic analysis using MEGA3.0 (Kumar et al. 2004). Grey bars indicate full-length protein; black boxes indicate Pollen Ole e 1 domain. Alignment of AtPOE1 protein sequences is shown in Fig. S1

Construction of transgenic plants and β -glucuronidase (GUS) staining

A 0.7-kb genomic fragment containing the *AtPOE1*;26 promoter was amplified by polymerase chain reaction (PCR) from wild-type plants. The amplified product was inserted before the *GUS* gene in the pBI101 vector, yielding *AtPOE1*;26_{pro}:*GUS*. This construct was introduced into wild-type plants by *Agrobacterium*-mediated transformation as described previously (Xu et al. 2003). The primers used for plasmid construction are listed in

Table S1. GUS staining was performed as described previously (He et al. 2012).

Reverse transcription-PCR (RT-PCR), quantitative RT-PCR (qRT-PCR), and chromatin immunoprecipitation (ChIP) analyses.

RNA extraction and reverse transcription were performed as described previously (He et al. 2012). Real-time PCR was conducted with gene-specific primers (listed in Table S1). All statistical analyses of real-time PCR results were based on three technical replicates. The ChIP assay was performed as described previously (He et al. 2012)

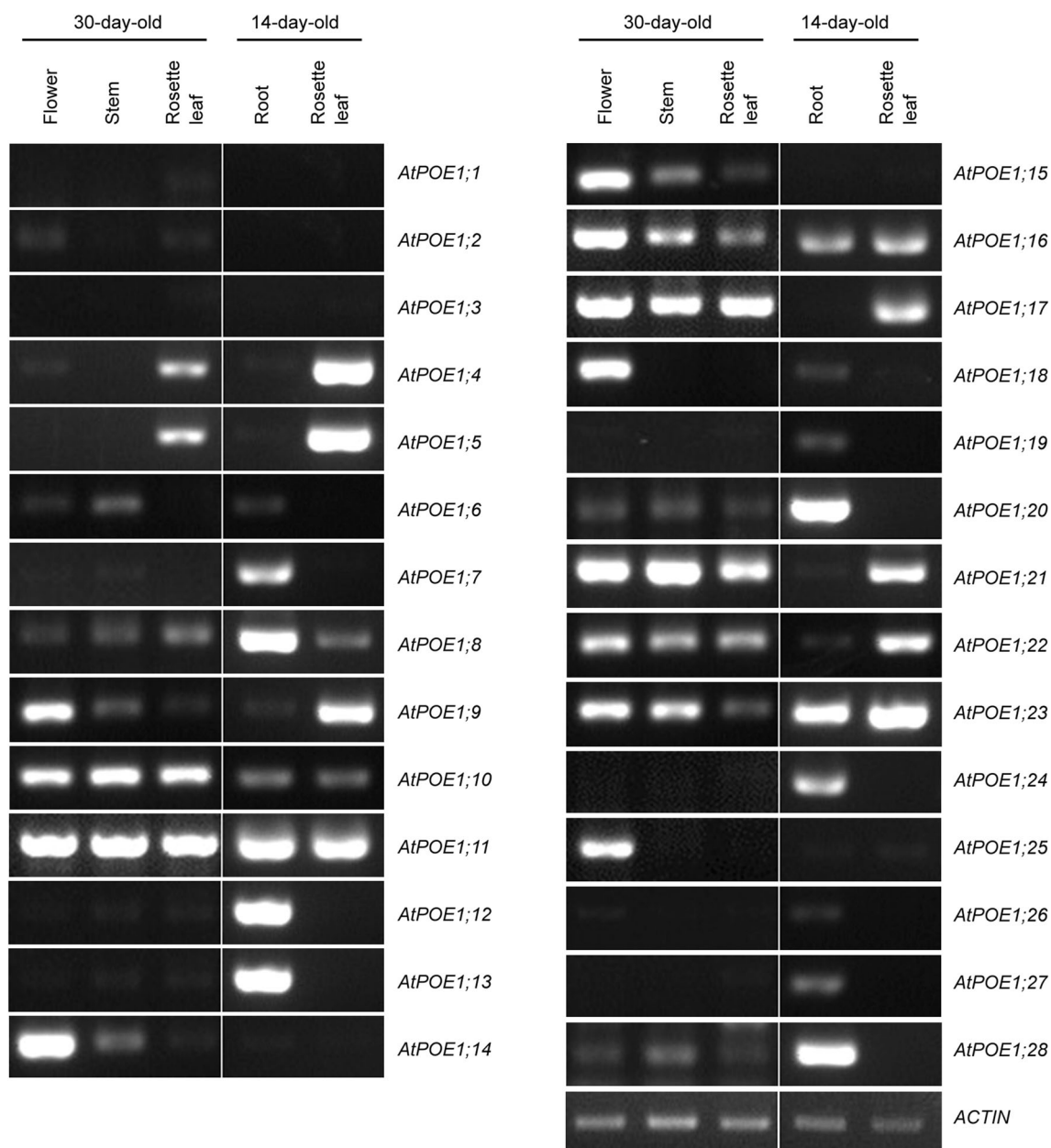
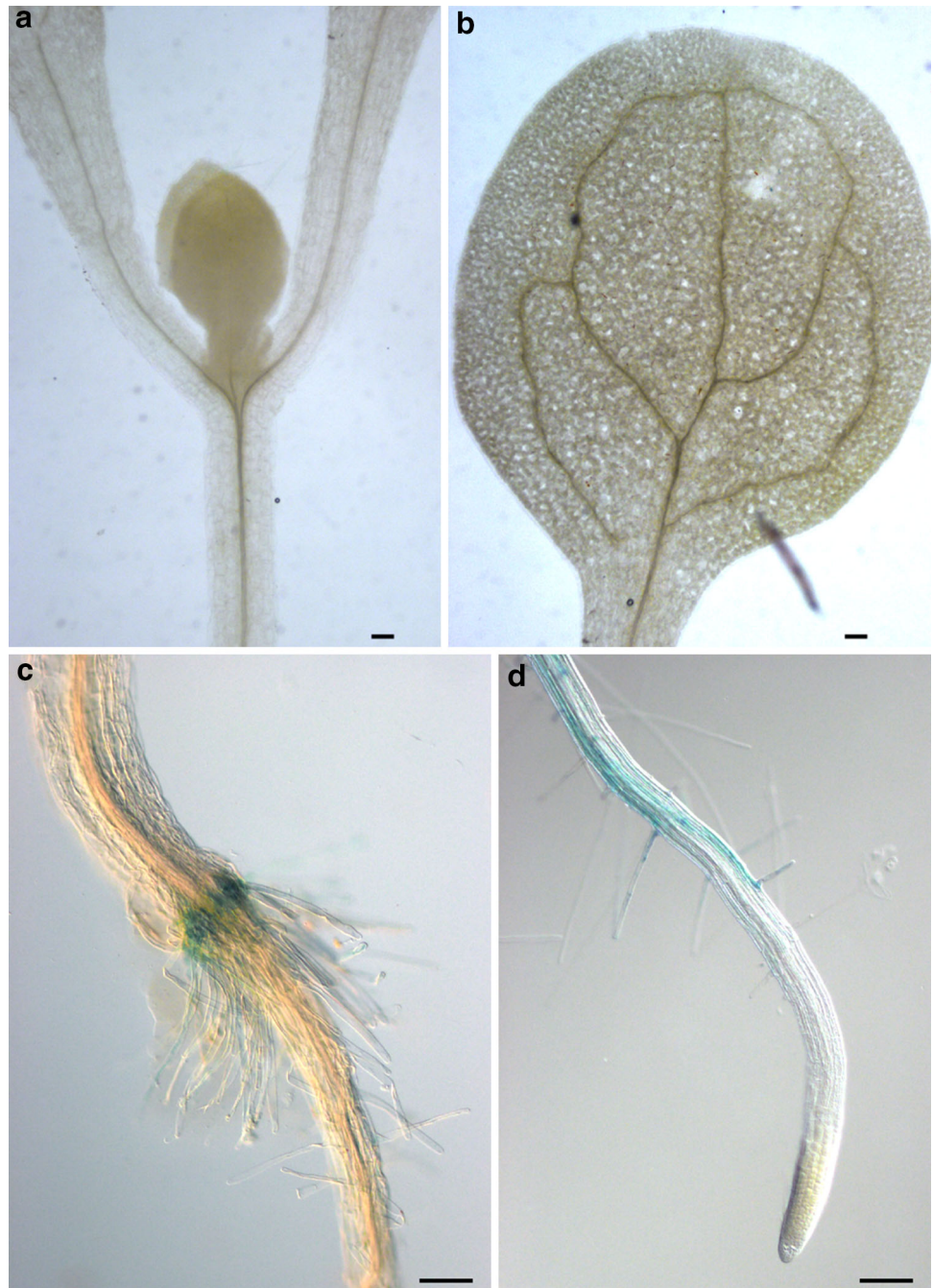


Fig. 2 RT-PCR analysis of transcript levels of *AtPOE1* family genes. RNA samples were extracted from flowers, stems, and rosette leaves of 30-day-old seedlings and from roots and rosette leaves of 14-day-

old seedlings. *ACTIN* served as the control. Note that *AtPOE1* family genes exhibited diverse expression patterns

Fig. 3 Spatial expression pattern of the *AtPOE1;26* gene. **a–c** GUS staining of rosette leaf (**a**), cotyledon (**b**), hypocotyl (**c**) and root (**d**) of a 9-day-old *AtPOE1;26_{pro}::GUS* transgenic plant (T2 generation). Note that no GUS signals were detected in above-ground tissues. GUS signals were detected at the junction region of the hypocotyl and the primary root (**c**), and in root hair regions (**d**). Bar 100 μ m



using an antibody against H3 trimethyl-Lys 27 (Millipore, Cat. 07-449).

Results and discussion

Pollen Ole e 1 domain-containing proteins are present in many plants (Jimenez-Lopez et al. 2011). To study the molecular features of these proteins in *Arabidopsis*, we first analyzed the whole *Arabidopsis* genome and identified at least 28 genes encoding proteins with a Pollen Ole e 1

domain (Fig. S1), based on TAIR annotations (<http://www.arabidopsis.org>). These genes were classified into the AtPOE1 family, and were named *AtPOE1;1* to 28 (Fig. 1). The previously characterized *AGP30* gene (van Hengel and Roberts 2003) corresponds to *AtPOE1;24* in the phylogenetic tree (Fig. 1).

To understand the transcription profiles of AtPOE1-family genes, we performed RT-PCR using multiple tissues from 14-day-old and 30-day-old seedlings of wild-type plants (Fig. 2). Most of the AtPOE1 family genes showed specific transcription patterns. For example, *AGP30*

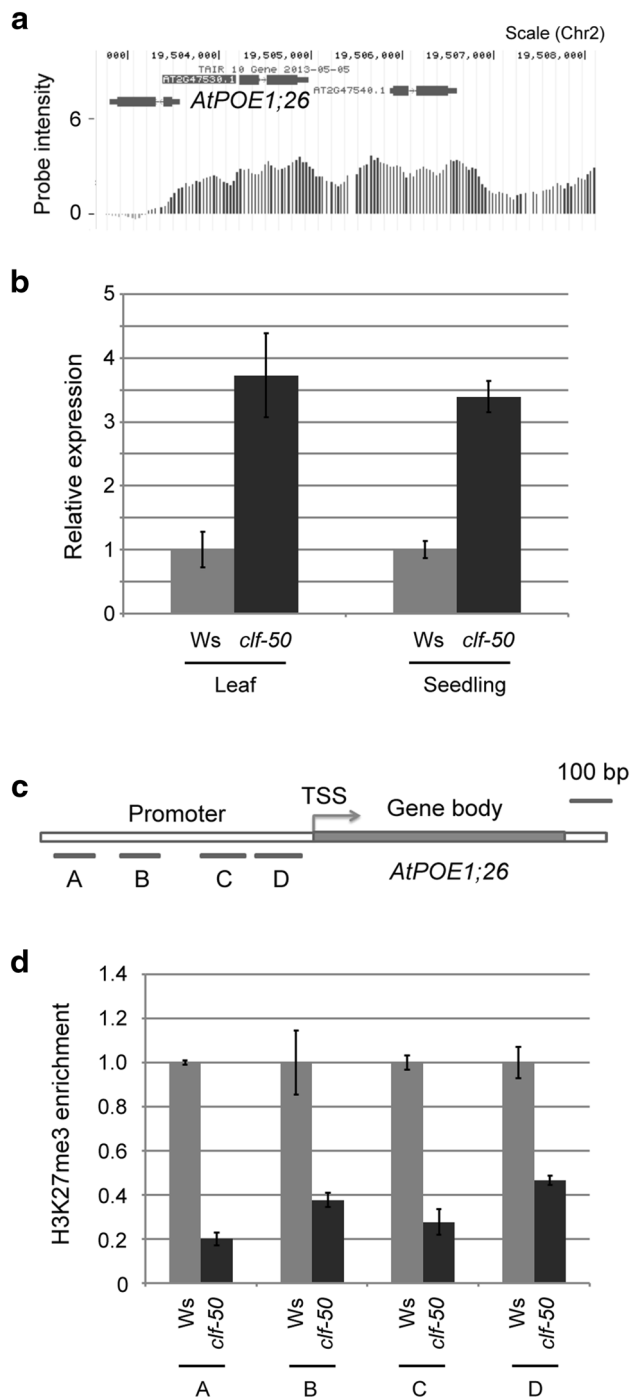


Fig. 4 Transcript level of *AtPOE1;26* is controlled by PRC2-mediated H3K27me3. **a** ChIP-chip analysis of the H3K27me3 modification at the *AtPOE1;26* locus in rosette leaves. **b** qRT-PCR analysis of *AtPOE1;26* transcript levels in rosette leaves of 14-day-old wild type (Ws) and *clf-50* plants. Transcript level of *AtPOE1;26* was higher in *clf-50* than in wild type (Ws). Results show relative transcript levels, which were normalized against that of *ACTIN*. Bars show s.e. from three technical repeats. **c** Diagram of the *AtPOE1;26* gene locus showing the positions of PCR-amplified fragments (A, B, C, and D) in ChIP analysis. TSS transcription start site. **d** ChIP analysis of H3K27me3 levels in 14-day-old seedlings of wild type (Ws) and the *clf-50* mutant. Note that methylation levels were lower in both promoter and gene body regions in *clf-50* than in Ws. Results show relative methylation levels, normalized against those produced by inputs. Bars show s.e. from three technical repeats

transcription patterns. These genes included *AtPOE1;7*, *8*, *12*, *13*, *20*, *24*, *26*, *27*, and *28* (Fig. 2). Consistent with this result, a recent study reported that *AtPOE1;8*, *20*, *27*, and *28* were specifically expressed in root cells during root hair formation (Brux et al. 2012). There were high transcript levels of many *AtPOE1* family genes, including *AtPOE1;9*, *10*, *11*, *14*, *15*, *16*, *17*, *18*, *21*, *22*, *23*, and *25*, in flowers (Fig. 2). This result suggested that the function of these *AtPOE1* proteins as allergens might be similar to that of the first-identified Ole e 1 protein from *Olea europaea* (Laururica et al. 1988). Together, these findings indicated that different *AtPOE1* family genes have diverse expression patterns, which may reflect specific developmental or physiological roles.

To further analyze the regulation of expression of *AtPOE1*-family genes, we focused on *AtPOE1;26* as a model gene. We selected this gene because our RT-PCR data showed that *AtPOE1;26* was only weakly expressed in the root under our experimental conditions (Fig. 2), suggesting that it might be a developmental regulator in root tissues. To analyze the tissue expression pattern of *AtPOE1;26* in detail, we constructed the *AtPOE1;26_{pro}:GUS* reporter line. GUS staining was only detectable in the hypocotyl and the root (Fig. 3a–d), consistent with the RT-PCR results.

It is known that PRC2-mediated H3K27me3 modification is usually required to repress the expression of developmental genes. Therefore, we checked the H3K27me3 modification status of *AtPOE1;26* in our previously published ChIP-chip database of *Arabidopsis* rosette leaves (He et al. 2012). There was a high level of the H3K27me3 modification on both the promoter region and the coding region of *AtPOE1;26* (Fig. 4a), suggesting that this gene is a potential target of PRC2.

To test whether the expression of *AtPOE1;26* was controlled by PRC2, we first performed qRT-PCR to compare the transcript levels of *AtPOE1;26* between the PRC2 mutant *clf-50* and the wild-type plants. In both 14-day-old rosette leaves and whole seedlings, the expression level of *AtPOE1;26* was significantly higher in *clf-50* than in wild-type plants (Fig. 4b). Next, we performed a ChIP assay to

AtPOE1;24 was transcribed only in the root of 14-day-old seedlings, and its transcripts were barely detected in rosette leaves of 14-day-old seedlings and flowers, stems and rosette leaves of 30-day-old seedlings (Fig. 2). This result is consistent with the previous report that *AGP30/AtPOE1;24* is a root-specific gene (van Hengel and Roberts 2003). Interestingly, many other *AtPOE1* family genes showed very similar transcription patterns to that of *AGP30/AtPOE1;24*, showing root-specific or primary

detect the H3K27me3 level at the *AtPOE1;26* locus in *clf-50* and wild type. The results showed that the H3K27me3 level at the *AtPOE1;26* locus was significantly lower in *clf-50* than in wild type (Fig. 4c, d). These data suggested that PRC2 represses the expression of *AtPOE1;26*.

We searched the H3K27me3 modification database of *Arabidopsis* seedlings (Zhang et al. 2007) and rosette leaves (He et al. 2012) and identified at least 13 AtPOE1 family genes that have the histone modification at their loci. These genes included *AtPOE1;1, 2, 3, 4, 5, 13, 15, 17, 19, 24, 25, 26, and 27* (Fig. S2). Therefore, PRC2-mediated H3K27me3 may regulate the expression of many members of the AtPOE1 gene family.

Conclusions

Our data showed that AtPOE1 family genes have various spatial expression patterns, suggesting that they may serve as developmental regulators in multiple tissues. Epigenetic pathways may play roles in controlling gene expression, resulting in the different expression patterns of the various family members. For example, our results showed that *AtPOE1;26* is repressed by the PcG pathway. This regulation might occur at not only the *AtPOE1;26* gene locus, but also loci of other AtPOE1 family genes, because at least 13 family members had H3K27me3 modifications at their gene loci. In future research, it would be interesting to characterize the diverse roles of AtPOE1 family genes in plant development and to reveal the function of *AtPOE1;26* in root development.

Author contribution CL, LX, and YR conceived and designed the experiments; BH, BL, and LL performed the experiments; LX and YR analyzed the data and wrote the paper. All the authors read and approved the final manuscript.

Acknowledgments We thank J. Goodrich for *Arabidopsis* mutant seeds. This work was supported by the grants from the National Basic Research Program of China (973 Program, 2012CB910500) and the National Natural Science Foundation of China (31071455).

Conflict of interest All authors declare that they have no conflict of interest.

References

- Batanero E, Villalba M, Rodriguez R (1994) Glycosylation site of the major allergen from olive tree pollen. Allergenic implications of the carbohydrate moiety. *Mol Immunol* 31:31–37
- Bruex A, Kainkaryam RM, Wieckowski Y, Kang YH, Bernhardt C, Xia Y, Zheng XH, Wang JY, Lee MM, Benfey P, Woolf PJ, Schiefelbein J (2012) A gene regulatory network for root epidermis cell differentiation in *Arabidopsis*. *PLoS Genet* 8:e1002446
- Calonje M (2013) PRC1 marks the difference in plant PcG repression. *Mol Plant*. doi:10.1093/mp/sst1150
- Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon YH, Sung ZR, Goodrich J (2004) Interaction of polycomb-group proteins controlling flowering in *Arabidopsis*. *Development* 131:5263–5276
- de Dios AJ, M'Rani-Alaoui M, Castro AJ, Rodriguez-Garcia MI (2004) Ole e 1, the major allergen from olive (*Olea europaea* L.) pollen, increases its expression and is released to the culture medium during in vitro germination. *Plant Cell Physiol* 45:1149–1157
- Goodrich J, Puangsomlee P, Martin M, Long D, Meyerowitz EM, Coupland G (1997) A polycomb-group gene regulates homeotic gene expression in *Arabidopsis*. *Nature* 386:44–51
- He C, Chen X, Huang H, Xu L (2012) Reprogramming of H3K27me3 is critical for acquisition of pluripotency from cultured *Arabidopsis* tissues. *PLoS Genet* 8:e1002911
- He C, Huang H, Xu L (2013) Mechanisms guiding Polycomb activities during gene silencing in *Arabidopsis thaliana*. *Front Plant Sci* 4:454
- Jiang SY, Jasmin PX, Ting YY, Ramachandran S (2005) Genome-wide identification and molecular characterization of Ole_e_I, Allerg_1 and Allerg_2 domain-containing pollen-allergen-like genes in *Oryza sativa*. *DNA Res* 12:167–179
- Jimenez-Lopez JC, Rodriguez-Garcia MI, Alché JD (2011) Systematic and phylogenetic analysis of the Ole e 1 pollen protein family members in plants. *Systems and computational biology—bioinformatics and computational modeling*. InTech, Rijeka, pp 245–260
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 5:150–163
- Lauzurica P, Gurbindo C, Maruri N, Galocha B, Diaz R, Gonzalez J, Garcia R, Lahoz C (1988) Olive (*Olea-Europea*) pollen allergens. 1. immunochemical characterization by immunoblotting, crie and immunodetection by a monoclonal-antibody. *Mol Immunol* 25:329–335
- Molitor A, Shen WH (2013) The polycomb complex PRC1: composition and function in plants. *J Genet Genomics* 40:231–238
- Schatlowski N, Creasey K, Goodrich J, Schubert D (2008) Keeping plants in shape: polycomb-group genes and histone methylation. *Semin Cell Dev Biol* 19:547–553
- Schubert D, Primavesi L, Bishopp A, Roberts G, Doonan J, Jenuwein T, Goodrich J (2006) Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *EMBO J* 25:4638–4649
- Sigrist CJ, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, Bairoch A, Hulo N (2010) PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res* 38:D161–D166
- van Hengel AJ, Roberts K (2003) AtAGP30, an arabinogalactan-protein in the cell walls of the primary root, plays a role in root regeneration and seed germination. *Plant J* 36:256–270
- Villalba M, Lopez-Otin C, Martin-Orozco E, Monsalve RI, Palomino P, Lahoz C, Rodriguez R (1990) Isolation of three allergenic fractions of the major allergen from *Olea europaea* pollen and N-terminal amino acid sequence. *Biochem Biophys Res Commun* 172:523–528
- Villalba M, Batanero E, Lopez-Otin C, Sanchez LM, Monsalve RI, Gonzalez de la Pena MA, Lahoz C, Rodriguez R (1993) The amino acid sequence of Ole e I, the major allergen from olive tree (*Olea europaea*) pollen. *Eur J Biochem* 216:863–869
- Whitcomb SJ, Basu A, Allis CD, Bernstein E (2007) Polycomb group proteins: an evolutionary perspective. *Trends Genet* 23:494–502

- Xu L, Xu Y, Dong A, Sun Y, Pi L, Xu Y, Huang H (2003) Novel *as1* and *as2* defects in leaf adaxial-abaxial polarity reveal the requirement for *ASYMMETRIC LEAVES1* and 2 and *ERECTA* functions in specifying leaf adaxial identity. *Development* 130:4097–4107
- Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE (2007) Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol* 5:e129