



## REVIEW

# Genome-wide analyses of gene activity during seed development

John J. Harada\* and Julie Pelletier

Department of Plant Biology, University of California, One Shields Avenue, Davis, CA 95616, USA

(Received 24 July 2011; accepted after revision 26 July 2011)

## Abstract

The seed consists of three distinct regions, the zygotic embryo and endosperm and the maternal seed coat. Development of the seed requires co-ordinated growth and differentiation within each of these regions. Although the major events that occur during seed development have been studied extensively at the anatomical, physiological and molecular level, a comprehensive understanding of gene activity in each seed region has been lacking. DNA microarrays and extensive sequencing of cDNA libraries permit genome-wide analyses of mRNA populations of biological samples. We review recent studies of genome-wide gene activity that illustrate approaches used to provide new insights into the processes that occur during seed development.

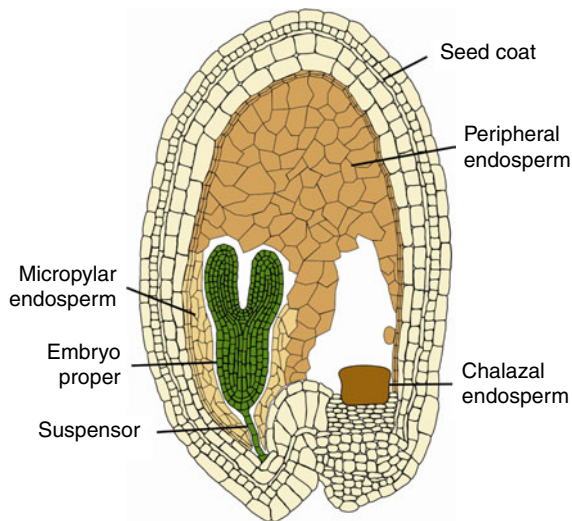
**Keywords:** *Arabidopsis*, DNA microarray, embryo, endosperm, RNA sequencing, seed coat

## Introduction

The seed is an elegant structure, consisting of three regions that each has a unique genotype: the diploid and zygotic embryo, the triploid and zygotic endosperm, and the diploid and maternal seed coat (Fig. 1) (Bewley and Black, 1994; Ohto *et al.*, 2008). Development of the zygotic compartments is initiated with the double fertilization event in which the egg cell and central cell of the female gametophyte each fuses

with one sperm cell of the pollen to give rise to the zygote and endosperm mother cell, respectively. Subsequent development of these cells can be divided conceptually into two phases, morphogenesis and maturation. During the morphogenesis phase early in seed development, the zygote of many plants, including the focus of this review *Arabidopsis thaliana* (L.) Heynh, undergoes a stereotypical pattern of cell divisions to generate the suspensor, an ephemeral structure that serves structural, nutritive and physiological roles in embryo development, and the embryo proper that will develop into the body of the vegetative plant (Jenik *et al.*, 2007; Capron *et al.*, 2009). Embryonic tissue and organ systems arise from the regulated division and differentiation of cells within the embryo proper. Endosperm development during the morphogenesis phase is syncytial in that nuclei divide without corresponding cell divisions (Brown *et al.*, 1999; Olsen, 2004). The resulting nuclear-cytoplasmic domains migrate within the endosperm cell and differentiate morphologically and molecularly to form the three endosperm domains: micropylar, peripheral and chalazal. Later in development, endosperm nuclear-cytoplasmic domains become cellularized in a wave-like fashion from the micropylar to the chalazal end of the seed. During the maturation phase late in seed development, morphogenetic processes, including cell division, are suppressed, and a major shift in embryo metabolism occurs (Harada, 1997; Baud *et al.*, 2008). Storage macromolecules, primarily proteins and lipids in *Arabidopsis*, are synthesized and accumulated to high levels within the embryo and endosperm. Other metabolic pathways are activated that allow the embryo to withstand the stresses of desiccation, which occurs at the end of seed development. By the end of seed development, the embryo and endosperm

\*Correspondence  
Email: jjharada@ucdavis.edu



**Figure 1.** Diagram of an *Arabidopsis* seed with a linear cotyledon-stage embryo, showing the different regions and compartments of the seed. (A colour version of this figure can be found online at <http://journals.cambridge.org/ssr>)

are quiescent metabolically and arrested developmentally, and they remain in this state until conditions favourable for germination are encountered.

Seed coat development from ovule integuments appears to be initiated concurrently with endosperm development. In *Arabidopsis*, the two cell layers of the outer integument and the three layers of the inner integument of the ovule undergo initial growth by cell division and expansion (Beeckman *et al.*, 2000; Haughn and Chaudhury, 2005). The innermost layer of the inner integument differentiates into the endothelium that accumulates proanthocyanidin flavonoids that form tannins thought to be involved in plant defence. The other two inner integument layers undergo programmed cell death. Epidermal cells of the outer integument form columella and secrete mucilage that is speculated to have roles in protection and/or dispersal of seedlings. In the mature seed, all seed coat cells are crushed and not viable.

A mechanistic understanding of seed development requires knowledge of the genes that are expressed at different stages of development and in different seed regions. Seed biologists have employed a variety of approaches to define genes required to make a seed. First, genes involved in prominent seed processes, such as storage protein and lipid biosynthesis and accumulation, have been identified (Pang *et al.*, 1988; Da Silva Conceição and Krebbers, 1994; Beisson *et al.*, 2003; Baud and Lepiniec, 2009). Second, genetic approaches have been used extensively to identify mutations that cause defects in embryo or seed development. In *Arabidopsis*, individual mutations in 481 genes are sufficient to cause defects in seed development (Meinke *et al.*, 2008; <http://www.seedgenes.org/>). Third, the expressed sequence tag

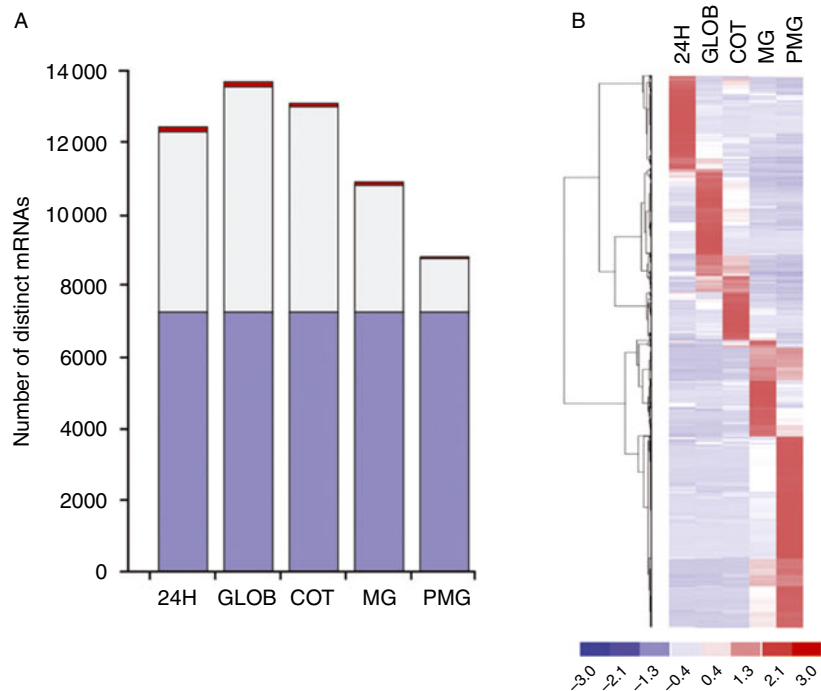
(EST) strategy of partially sequencing cloned cDNAs derived from seed RNAs has been used to identify genes expressed in seeds (White *et al.*, 2000). However, none of these approaches have identified all of the genes required to make a seed.

The advent of experimental methods to analyse gene expression on a whole genome or near-whole genome basis has provided new and powerful approaches to study biological processes, including seed development. DNA microarray technology allows the expression of genes to be detected and quantified on a whole-genome basis through the hybridization of RNA from a biological sample with DNA probes that have been synthesized or immobilized on a glass, plastic or silicon surface, which correspond to all, or most, genes in the genome (Schena *et al.*, 1995; Ruan *et al.*, 1998). The approach permits large numbers of biological samples to be analysed simultaneously, although a weakness of the method is that mRNAs present at very low levels are often not detected. Development of next-generation sequencing methods has allowed RNA populations to be characterized through extensive sequencing of cDNA libraries (Mardis, 2008; Mortazavi *et al.*, 2008). This approach enables mRNAs in a population to be detected with high sensitivity and mRNA levels to be quantified robustly. In recent years, whole genome analyses of mRNA populations have been applied to the study of seeds. In this short review, we discuss selected examples of how information provided by these studies has provided insight into the biological processes that underlie seed development in *Arabidopsis*.

### Gene expression during seed development

A comprehensive overview of gene expression throughout seed development was obtained from whole-genome DNA microarray analyses of mRNA populations in *Arabidopsis* seeds containing zygote, globular, cotyledon, mature green and post-mature green stage embryos (Le *et al.*, 2010). To understand gene expression in seeds within the context of other phases of development, the seed mRNA populations were compared with those of ovules before seed development, seedlings after seed development, and vegetative and reproductive organs of the plant.

The study provided new information about the extent and specificity of gene activity during seed development that could only be obtained from a global perspective. First, similar numbers of distinct mRNAs accumulate in seeds at the early stages of development, but mRNA numbers decline at the later stages, as summarized in Fig. 2A. These findings suggest that the functional complexity of seeds remains relatively constant during the morphogenesis phase and that the



**Figure 2.** Gene expression during Arabidopsis seed development. (A) Bars show the total number of distinct mRNAs detected in developing seeds containing embryos at the indicated stages of development: 24H, zygote stage 24 h after pollination; GLOB, globular stage; COT, cotyledon stage; MG, mature-green stage; PMG, postmature-green stage. Red, grey and blue bars depict the number of mRNAs that are detected only at that given stage, are detected in seeds at two or more stages of development, and are detected in seeds at all stages of development, respectively. (B) mRNAs detected at all stages of seed development accumulate at highest levels primarily at one developmental stage. Heat map showing unsupervised hierarchical clustering (dChip; Li and Wong, 2001) of the 1500 mRNAs that display the highest coefficient of variation in levels during development. Each line on the heat map represents one mRNA, with colours indicating Z scores representing the relative level of that mRNA. Z score is the number of standard deviations from the mean signal intensity for each mRNA across the indicated stages. (A colour version of this figure can be found online at <http://journals.cambridge.org/ssr>).

diversity of cellular processes diminishes during the maturation phase as seeds progress to developmental arrest and metabolic quiescence. Second, although the cellular processes that occur at each developmental stage differ significantly, only approximately 15–100 mRNAs accumulate specifically at any given stage, at least at the level of detection permitted by DNA microarray experiments (Fig. 2A). Instead, the majority of all mRNAs are present at each stage of seed development. These ‘common’ mRNAs probably do not all represent constitutively expressed, house-keeping genes, because many of these shared mRNAs accumulate primarily at a specific stage of development. Figure 2B shows that the ‘common’ mRNAs that display the greatest coefficient of variation in levels during seed development accumulate at highest levels primarily at one stage of development. Third, comparison of mRNA populations showed that ~15,500 mRNAs are detected, at the sensitivity level of the DNA microarray, throughout seed development, whereas ~18,500 mRNAs are detected throughout the plant life cycle, including vegetative and reproductive development. Moreover,

most mRNAs detected during seed development are also present throughout the plant life cycle. Thus, a majority of expressed genes in the Arabidopsis genome are active and, presumably, required at all stages of development. Fourth, a small number of genes, ~300, are present specifically in seeds at the level of DNA microarray detection. Mutations in some of the seed-specific genes encoding transcription factors cause defects in seed development, indicating the importance of these genes. These findings suggest that gene expression is highly temporally regulated during seed development. The cellular processes that underlie morphological and physiological functions that occur at each stage of seed development are governed by small sets of genes that are expressed either specifically or primarily at a given stage.

### Endosperm, parent-of-origin effects and imprinting

A key to understanding the co-ordination of biological processes that occur during seed development is to

define the activities of genes in distinct seed regions. The endosperm plays critical roles in seed development, even in plants such as *Arabidopsis* in which only a few layers of endosperm cells remain in the mature seed. The endosperm is a nutritive source for the developing embryo and/or seedling and is a major determinant of seed mass (Lopes and Larkins, 1993; Boissard-Lorig *et al.*, 2001; Garcia *et al.*, 2005). An overview of endosperm gene expression was obtained from DNA microarray analysis of early, syncytial-stage endosperm that was captured from histologically fixed and sectioned seeds using laser-capture microdissection (Day *et al.*, 2007). DNA microarray experiments showed the dominant expression of genes associated with cell cycle, DNA processing, chromatin assembly, protein synthesis, cytoskeleton- and microtubule-related processes, and cell/organelle biogenesis and organization (Day *et al.*, 2008).

The endosperm has received substantial attention in recent years as the major site of gene imprinting in angiosperms (Huh *et al.*, 2008; Berger and Chaudhury, 2009; Raissig *et al.*, 2011). Following double fertilization, imprinted genes are expressed in zygotic tissues predominantly from either maternally-derived or paternally-derived alleles, and these genes are thought to underlie the differential influence of maternal and paternal genomes on seed development, known as parent-of-origin effects. Parent-of-origin effects are best illustrated by crosses between *Arabidopsis* plants with different ploidy levels (Scott *et al.*, 1998). Progeny of interploidy crosses that have an excess of maternal genomes, i.e. from a cross of a tetraploid female with a diploid male ( $4x \times 2x$ ), have seeds that are smaller than self-fertilized progeny of either diploid or tetraploid plants. By contrast, progeny with an excess of paternal genomes, e.g. progeny from a  $2x \times 4x$  interploidy cross, produce seeds larger than self-fertilized seeds. The parental conflict theory has been proposed to explain seed phenotypes caused by maternal and paternal genome excesses (Haig and Westoby, 1991). It posits that the father in polygamous organisms will try to enhance resource allocation specifically to his progeny to promote their growth, whereas a mother will try to distribute resources equally among all of her offspring to equalize their growth. By this hypothesis, factors that enhance resource allocation to progeny will be expressed from the paternal genome, whereas those that restrict the distribution of resources to the embryo will be expressed from the maternal genome. The finding that the endosperm is the major site of gene imprinting is consistent with its role in nourishing the developing embryo.

Parent-of-origin effects on seed development have been studied using mutations in maternal components of the Polycomb-repressive complex (Huh *et al.*, 2008; Berger and Chaudhury, 2009; Raissig *et al.*, 2011). For example, mutations of the maternal allele of the

*MEDEA* (*MEA*) gene that encodes a SET-domain component of the Polycomb-repressive complex result in mutant phenotypes that resemble those caused by interploidy crosses resulting in paternal genome excess, i.e. enlargement of the endosperm cavity, increase in chalazal endosperm size, and delay of endosperm cellularization (Grossniklaus and Vielle-Calzada, 1998; Kiyosue *et al.*, 1999). To explore the role of the Polycomb-repressive complex in parent-of-origin effects, mRNA populations of siliques derived from a *mea* mutant and from crosses that yield progeny with balanced genomes, paternal genome excess, and maternal genome excess were profiled (Tiwari *et al.*, 2010). The mRNA populations of progeny with maternal *mea* mutant alleles resembled those of progeny from paternal genome excess crosses, with both showing enhanced expression of genes encoding MADS-box transcription factors, cell cycle proteins and hormone metabolic enzymes. Because *MEA* is an imprinted gene expressed from maternal alleles, the results confirm the relationship between the Polycomb-repressive complex, imprinting and parent-of-origin effects.

To identify genes that contribute to parent-of-origin effects, a mutation that permits endosperm development to occur in the absence of a paternal genome was exploited (Shirzadi *et al.*, 2011). Pollen carrying a mutation in the *CYCLIN DEPENDENT KINASE A;1* (*CDKA;1*) gene will result in a single fertilization event in which the egg cell becomes fertilized, but the central cell does not (Nowack *et al.*, 2006). The central cells of many of these seeds undergo endosperm proliferation, although the endosperm remains underdeveloped and the seed aborts. Comparison of the mRNA populations of early stage seeds derived from crosses of wild-type female plants with *cdka;1* male plants identified sets of genes that are either downregulated or upregulated in the progeny relative to crosses with a wild-type male. More than 60% of the genes encoding type-I My MADS-box transcription factors were among the downregulated genes, and detailed analysis of one of these genes, *AGL36*, showed that it is imprinted in wild-type progeny and expressed from the maternal allele. This result was interpreted to suggest that a paternally expressed gene is likely required for expression of the maternal *AGL36* allele.

Imprinted endosperm genes were identified directly by Hsieh *et al.* (2011). They analysed mRNA populations from endosperm that was either hand dissected or isolated using laser-capture microdissection of seeds derived from a cross between parents of different *Arabidopsis* ecotypes. Endosperm mRNA populations were characterized by extensive sequencing of cDNA libraries. The nucleotide sequence sensitivity of the method permitted mRNAs derived from maternal and paternal alleles to be distinguished primarily by single-nucleotide polymorphisms on a

genome-wide basis. Genes expressed predominantly from maternal or paternal alleles were confirmed to be imprinted if their allele-specific expression was compromised by mutations that affect molecular processes known to underlie imprinting, specifically DNA methylation, DNA demethylation and Polycomb complex-mediated gene repression. On the basis of this study, the list of known Arabidopsis imprinted genes was expanded approximately fourfold, and it was estimated that there are 30–50 paternally expressed and approximately 200 maternally expressed imprinted genes. Moreover, imprinted genes were shown to encode proteins involved in maintenance of DNA methylation, methylation and demethylation of histones, and ubiquitination. In relation to the parental conflict theory, these findings open the possibility that parent-of-origin effects may be mediated at several regulatory levels.

### **Embryo gene expression profiles provide information about developmental processes and the influence of maternal and paternal genomes**

The embryo is the focus of two recent studies that addressed different questions about seed development by characterizing mRNA populations. A developmental analysis of embryonic gene activity was conducted in which mRNA populations from hand-dissected embryos at seven stages, ranging from the zygote to mature stages, were characterized using DNA microarray analyses (Xiang *et al.*, 2011). Co-expressed sets of genes whose mRNAs accumulate primarily at specific stages of seed development were identified by clustering analysis. The annotated functions of genes that accumulate specifically at a given stage provide insight into the biological processes that characterize that stage. For example, sets of mRNAs that accumulate primarily at sequential developmental stages, the zygote and quadrant stages, globular and heart stages, torpedo and bent stages, and mature stage are associated with the biological processes of auxin stimulus and signalling, meristem and morphogenesis, carbohydrate, fatty acid and storage protein synthesis, and abscisic acid (ABA) response and dehydration, respectively. Thus, the study identified genes associated with stage-specific processes during embryo development. Another contribution of this study was to relate changes in gene expression with metabolic networks during embryo development. A network of metabolites and the reactions involved in their interconversion was established, and mRNAs encoding metabolic enzymes whose levels change during transitions between developmental stages were mapped on to this network. This analysis showed that genes encoding enzymes involved in related pathways are

co-ordinately regulated. Moreover, concerted changes in the levels of mRNAs encoding enzymes in the same metabolic pathway predicted the timing at which pathways become activated and deactivated during embryo development.

In another study, embryo mRNA populations were profiled to address a controversy about the relative contributions of maternal and paternal genomes to early embryo development. The activation of paternally derived alleles of embryonic genes marks the transition from a strictly maternal influence over embryo development to zygotic control. In animals, the paternal genome becomes active after a few or many cell divisions (reviewed in Baroux *et al.*, 2008). This parental allele-specific expression occurs only transiently and, thus, differs from gene imprinting. Some studies have shown that the maternal alleles of many plant genes are predominantly expressed in early stage embryos, whereas other studies have shown that both maternal and paternal alleles are expressed. To clarify these results, RNA populations of seeds derived from a cross of two different Arabidopsis ecotypes were analysed using deep sequencing of cDNA libraries (Autran *et al.*, 2011). Nucleotide sequence polymorphisms were used to distinguish mRNAs derived from maternal and paternal alleles. Very early in embryo development, at the two- to four-cell embryo-proper stage, mRNAs from maternal alleles dominate the population, with 30% and 2% of mRNAs, respectively, being derived exclusively from the maternal and paternal genome. Furthermore, genes expressed from both parental alleles display a strong maternal bias. Repression of paternal genes occurs, at least in part, epigenetically through the chromatin small interfering RNA pathway, although it is not clear if maternal transcripts represent mRNAs stored during female gametophyte development or if they are expressed in the zygote. By the globular stage, the representation of paternally derived mRNAs in the population increases significantly. The results were interpreted to indicate that processes that occur at the earliest stages of embryo development are determined in large part by the maternal genome, but the influence of the paternal genome gradually becomes apparent. This expression pattern potentially explains how maternal-effect genes can influence embryo development.

### **Using mutations to identify genes involved in seed coat development**

The seed coat derives from terminal differentiation of the ovule inner and outer integuments. To gain insight into the genes controlling seed coat development, mRNA populations in seed coats from wild-type seeds and two mutants with defects in seed coat structure

were profiled (Dean *et al.*, 2011). *apetala2* mutants were used because the wild-type gene is required for differentiation of outer integuments, particularly the epidermis (Jofuku *et al.*, 1994; Western *et al.*, 2001), and *transparent testa16* mutants are defective in differentiation of the endothelium in the inner integument (Nesi *et al.*, 2002; Debeaujon *et al.*, 2003). RNA profiling was done with hand-dissected seed coats at several stages of development, corresponding to major events in epidermal and endothelial differentiation. The comparisons identified sets of genes whose expression is significantly downregulated and upregulated in mutant seed coats relative to wild type. The relationship of these mRNAs with epidermal and endothelial differentiation was validated by showing that genes known to be affected by the *apetala2* or *transparent testa16* mutations are included in these gene sets. The authors concluded that the study enables the discovery of new genes involved in seed coat development.

## Summary

Genome-wide analyses of gene activity have enabled new information about seed development to be obtained. The global nature of the approach provides a comprehensive view of gene expression. This is particularly important for comparisons of seeds with different genotypes or at different developmental stages, because it permits all differentially expressed genes to be identified. The nucleotide sequence resolution of the approaches allows the contributions of maternal and paternal alleles of genes to be distinguished. Although mRNA profiling data alone provides a powerful tool that can be used to address biological questions, it is likely that its integration with other data, such as metabolite levels, protein levels, binding sites for transcription factors and histone marks through network analyses, will provide even greater insights into seed development (Yuan *et al.*, 2008; Burow *et al.*, 2010; Moreno-Risueno *et al.*, 2010; Lucas *et al.*, 2011).

## Acknowledgements

We thank Ryan Kirkbride for his comments about the manuscript. Work cited in this review from the authors' laboratory was supported by grants from the National Science Foundation.

## References

Autran, D., Baroux, C., Raissig, M.T., Lenormand, T., Wittig, M., Grob, S., Steimer, A., Barann, M., Klostermeier, U.C., Leblanc, O., Vielle-Calzada, J.P., Rosenstiel, P.,

- Grimanelli, D. and Grossniklaus, U. (2011) Maternal epigenetic pathways control parental contributions to Arabidopsis early embryogenesis. *Cell* **145**, 707–719.
- Baroux, C., Autran, D., Gillmor, C.S., Grimanelli, D. and Grossniklaus, U. (2008) The maternal to zygotic transition in animals and plants. *Cold Spring Harbor Symposia on Quantitative Biology* **73**, 89–100.
- Baud S., Dubreucq B., Miquel M., Rochat C. and Lepiniec L. (2008) Storage reserve accumulation in Arabidopsis: metabolic and developmental control of seed filling. *The Arabidopsis Book*, e0113. The American Society of Plant Biologists. <http://www.bioone.org/loi/arbo.j>. doi: 10.1199/tab.0113.
- Baud, S. and Lepiniec, L. (2009) Regulation of de novo fatty acid synthesis in maturing oilseeds of Arabidopsis. *Plant Physiology and Biochemistry* **47**, 448–455.
- Beeckman, T., De Rycke, R., Viane, R. and Inzé, D. (2000) Histological study of seed coat development in *Arabidopsis thaliana*. *Journal of Plant Research* **113**, 139–148.
- Beisson, F., Koo, A.J.K., Ruuska, S., Schwender, J., Pollard, M., Thelen, J.J., Paddock, T., Salas, J.J., Savage, L., Milcamps, A., Mhaske, V.B., Cho, Y. and Ohlrogge, J.B. (2003) Arabidopsis genes involved in acyl lipid metabolism. A 2003 census of the candidates, a study of the distribution of expressed sequence tags in organs, and a web-based database. *Plant Physiology* **132**, 681–697.
- Berger, F. and Chaudhury, A. (2009) Parental memories shape seeds. *Trends in Plant Science* **14**, 550–556.
- Bewley, J.D. and Black, M. (1994) *Seeds: physiology of development and germination*. New York, Plenum Publishing Corporation.
- Boisnard-Lorig, C., Colon-Carmona, A., Bauch, M., Hodge, S., Doerner, P., Bancharel, E., Dumas, C., Haseloff, J. and Berger, F. (2001) Dynamic analyses of the expression of the HISTONE:YFP fusion protein in Arabidopsis show that syncytial endosperm is divided in mitotic domains. *Plant Cell* **13**, 495–509.
- Brown, R.C., Lemmon, B.E., Nguyen, H. and Olsen, O.A. (1999) Development of endosperm in *Arabidopsis thaliana*. *Sexual Plant Reproduction* **12**, 32–42.
- Burow, M., Halkier, B.A. and Kliebenstein, D.J. (2010) Regulatory networks of glucosinolates shape *Arabidopsis thaliana* fitness. *Current Opinion in Plant Biology* **13**, 347–352.
- Capron, A., Chatfield, S., Provart, N. and Berleth, T. (2009) Embryogenesis: pattern formation from a single cell. *The Arabidopsis Book*. pp. 1–28, The American Society of Plant Biologists.
- Da Silva Conceição, A. and Krebbers, E. (1994) A cotyledon regulatory region is responsible for the different spatial expression patterns of Arabidopsis 2S albumin genes. *Plant Journal* **5**, 493–505.
- Day, R.C., McNoe, L.A. and Macknight, R.C. (2007) Transcript analysis of laser microdissected plant cells. *Physiologia Plantarum* **129**, 267–282.
- Day, R.C., Herridge, R.P., Ambrose, B.A. and Macknight, R.C. (2008) Transcriptome analysis of proliferating Arabidopsis endosperm reveals biological implications for the control of syncytial division, cytokinin signaling, and gene expression regulation. *Plant Physiology* **148**, 1964–1984.
- Dean, G., Cao, Y., Xiang, D., Provart, N.J., Ramsay, L., Ahad, A., White, R., Selvaraj, G., Datta, R. and Haughn, G.

- (2011) Analysis of gene expression patterns during seed coat development in Arabidopsis. *Mol. Plant* **4**, 1074–1091.
- Debeaujon, I., Nesi, N., Perez, P., Devic, M., Grandjean, O., Caboche, M. and Lepiniec, L.** (2003) Proanthocyanidin-accumulating cells in Arabidopsis testa: regulation of differentiation and role in seed development. *Plant Cell* **15**, 2514–2531.
- Garcia, D., Fitz Gerald, J.N. and Berger, F.** (2005) Maternal control of integument cell elongation and zygotic control of endosperm growth are coordinated to determine seed size in Arabidopsis. *Plant Cell* **17**, 52–60.
- Grossniklaus, U. and Vielle-Calzada, J.-P.** (1998) ... response: Parental conflict and infanticide during embryogenesis. *Trends in Plant Science* **3**, 328.
- Haig, D. and Westoby, M.** (1991) Genomic imprinting in the endosperm: its effect on seed development in crosses between species, and between different ploidies of the same species, and its implications for the evolution of apomixis. *Philosophical Transactions of the Royal Society London* **333**, 1–13.
- Harada, J.J.** (1997) Seed maturation and control of germination. pp. 545–592 in Larkins, B.A.; Vasil, I.K. (Eds) *Advances in cellular and molecular biology of plants*. Dordrecht, Kluwer Academic.
- Haughn, G. and Chaudhury, A.** (2005) Genetic analysis of seed coat development in Arabidopsis. *Trends in Plant Science* **10**, 472–477.
- Hsieh, T.F., Shin, J.Y., Uzawa, R., Silva, P., Cohen, S., Bauer, M.J., Hashimoto, M., Kirkbride, R.C., Harada, J.J., Zilberman, D. and Fischer, R.L.** (2011) Regulation of imprinted gene expression in Arabidopsis endosperm. *Proceedings of the National Academy of Sciences USA* **108**, 1755–1762.
- Huh, J.H., Bauer, M.J., Hsieh, T.F. and Fischer, R.L.** (2008) Cellular programming of plant gene imprinting. *Cell* **132**, 735–744.
- Jenik, P.D., Gillmor, C.S. and Lukowitz, W.** (2007) Embryonic patterning in *Arabidopsis thaliana*. *Annual Review of Cellular and Developmental Biology* **23**, 207–236.
- Jofuku, K.D., Boer, B., Montagu, M.V. and Okamoto, J.K.** (1994) Control of Arabidopsis flower and seed development by the homeotic gene *APETALA2*. *Plant Cell* **6**, 1211–1225.
- Kiyosue, T., Ohad, N., Yadegari, R., Hannon, M., Dinneny, J., Wells, D., Katz, A., Margossian, L., Harada, J.J., Goldberg, R.B. and Fischer, R.L.** (1999) Control of fertilization-independent endosperm development by the *MEDEA* polycomb gene in Arabidopsis. *Proceedings of the National Academy of Sciences USA* **96**, 4186–4191.
- Le, B.H., Cheng, C., Bui, A.Q., Wagmaister, J.A., Henry, K.F., Pelletier, J., Kwong, L., Belmonte, M., Kirkbride, R., Horvath, S., Drews, G.N., Fischer, R.L., Okamoto, J.K., Harada, J.J. and Goldberg, R.B.** (2010) Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. *Proceedings of the National Academy of Sciences USA* **107**, 8063–8070.
- Li, C. and Wong, W.H.** (2001) Model-based analysis of oligonucleotide arrays: Expression index computation and outlier detection. *Proceedings of the National Academy of Sciences USA* **98**, 31–36.
- Lopes, M.A. and Larkins, B.A.** (1993) Endosperm origin, development, and function. *Plant Cell* **5**, 1383–1399.
- Lucas, M., Laplaze, L. and Bennett, M.J.** (2011) Plant systems biology: network matters. *Plant Cell & Environment* **34**, 535–553.
- Mardis, E.R.** (2008) Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics* **9**, 387–402.
- Meinke, D., Muralla, R., Sweeney, C. and Dickerman, A.** (2008) Identifying essential genes in *Arabidopsis thaliana*. *Trends in Plant Science* **13**, 483–491.
- Moreno-Risueno, M.A., Busch, W. and Benfey, P.N.** (2010) Omics meet networks – using systems approaches to infer regulatory networks in plants. *Current Opinion in Plant Biology* **13**, 126–131.
- Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. and Wold, B.** (2008) Mapping and quantifying mammalian transcripts by RNA-Seq. *Nature Methods* **5**, 621–628.
- Nesi, N., Debeaujon, I., Jond, C., Stewart, A.J., Jenkins, G.I., Caboche, M. and Lepiniec, L.** (2002) The *TRANSPARENT TESTA16* locus encodes the *ARABIDOPSIS* *BSISTER* *MADS* domain protein and is required for proper development and pigmentation of the seed coat. *Plant Cell* **14**, 2463–2479.
- Nowack, M.K., Grini, P.E., Jakoby, M.J., Lafos, M., Koncz, C. and Schnittger, A.** (2006) A positive signal from the fertilization of the egg cell sets off endosperm proliferation in angiosperm embryogenesis. *Nature Genetics* **38**, 63–67.
- Ohto, M., Stone, S.L. and Harada, J.J.** (2008) Genetic control of seed development and seed mass. pp. 1–24 in Bradford, K.J.; Nonogaki, H. (Eds) *Seed development, dormancy, and germination*. Oxford, Blackwell Publishing.
- Olsen, O.-A.** (2004) Nuclear endosperm development in cereals and *Arabidopsis thaliana*. *Plant Cell* **16**, S214–S227.
- Pang, P.P., Pruitt, R.E. and Meyerowitz, E.M.** (1988) Molecular cloning, genomic organization, expression and evolution of 12S seed storage protein genes of *Arabidopsis thaliana*. *Plant Molecular Biology* **11**, 805–820.
- Raissig, M.T., Baroux, C. and Grossniklaus, U.** (2011) Regulation and flexibility of genomic imprinting during seed development. *Plant Cell* **23**, 16–26.
- Ruan, Y., Gilmore, J. and Conner, T.** (1998) Towards Arabidopsis genome analysis: monitoring expression profiles of 1400 genes using cDNA microarrays. *Plant Journal* **15**, 821–833.
- Schena, M., Shalon, D., Davis, R.W. and Brown, P.O.** (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470.
- Scott, R.J., Spielman, M., Bailey, J. and Dickinson, H.G.** (1998) Parent-of-origin effects on seed development in *Arabidopsis thaliana*. *Development* **125**, 3329–3341.
- Shirzadi, R., Andersen, E.D., Bjerkan, K.N., Gloeckle, B.M., Heese, M., Ungru, A., Winge, P., Koncz, C., Aalen, R.B., Schnittger, A. and Grini, P.E.** (2011) Genome-wide transcript profiling of endosperm without paternal contribution identifies parent-of-origin-dependent regulation of *AGAMOUS-LIKE36*. *PLOS Genetics* **7**, e1001303.
- Tiwari, S., Spielman, M., Schulz, R., Oakey, R.J., Kelsey, G., Salazar, A., Zhang, K., Pennell, R. and Scott, R.J.** (2010) Transcriptional profiles underlying parent-of-origin effects in seeds of *Arabidopsis thaliana*. *BMC Plant Biology* **10**:72 <http://www.biomedcentral.com/bmcplantbiol/>.
- Western, T.L., Burn, J., Tan, W.L., Skinner, D.J., Martin-McCaffrey, L., Moffatt, B.A. and Haughn, G.W.** (2001)

- Isolation and characterization of mutants defective in seed coat mucilage secretory cell development in Arabidopsis. *Plant Physiology* **127**, 998–1011.
- White, J.A., Todd, J., Newman, T., Focks, N., Girke, T., de Ilárduya, O.M.n., Jaworski, J.G., Ohlrogge, J.B. and Benning, C.** (2000) A new set of Arabidopsis expressed sequence tags from developing seeds: the metabolic pathway from carbohydrates to seed oil. *Plant Physiology* **124**, 1582–1594.
- Xiang, D.Q., Venglat, P., Tibiche, C., Yang, H., Risseuw, E., Cao, Y.G., Babic, V., Cloutier, M., Keller, W., Wang, E., Selvaraj, G. and Datla, R.** (2011) Genome-wide analysis reveals gene expression and metabolic network dynamics during embryo development in Arabidopsis. *Plant Physiology* **156**, 346–356.
- Yuan, J.S., Galbraith, D.W., Dai, S.Y., Griffin, P. and Stewart, C.N. Jr** (2008) Plant systems biology comes of age. *Trends in Plant Science* **13**, 165–171.