Minireview

Evolutionary origins of the endosperm in flowering plants

Célia Baroux, Charles Spillane and Ueli Grossniklaus

Address: Institute of Plant Biology, Zollikerstrasse 107, CH-8008 Zürich, Switzerland.

Correspondence: Ueli Grossniklaus. E-mail: grossnik@botinst.unizh.ch

Published: 30 August 2002

Genome **Biology** 2002, **3(9)**:reviews1026.1–1026.5

Abstract

The evolutionary origin of double fertilization and the resultant endosperm tissue in flowering plants remains a puzzle, despite over a century of research. The recent resurgence of approaches to evolutionary developmental biology combining comparative biology with phylogenetics provides new understanding of endosperm origins.

In seeds of flowering plants, the embryo is surrounded by a nutritive tissue called endosperm. Embryo and endosperm are derived from individual fertilization events (double fertilization) and develop embedded in maternal tissues that form the seed coat. Despite the nutritional and economical importance of the endosperm, which makes up about 80% of a corn kernel or a wheat grain, the evolutionary origin of this crucial food storage tissue remains unclear. The triploid nature of the endosperm is typical for most flowering plants, including all important crops and the model system *Arabidopsis thaliana*. The notion that double fertilization and triploid endosperm are specific features of flowering plants tightly linked to their evolutionary origin has recently been challenged. A study of endosperm in primitive flowering plants, such as the waterlily family, suggests that their diploid endosperm may be the remnant of an ancestral state [1,2]. This new information, combined with novel and established ideas, allows a clearer understanding of the possible evolutionary and developmental origin of the endosperm. The application of functional genomics to the evolutionary developmental biology of the endosperm promises to shed further light onto this curious yet critical tissue.

Gametophytes and double fertilization

The plant life cycle alternates between a haploid and a diploid phase: the haploid gamete-producing gametophyte can be viewed as the functional equivalent of the animal germline, while the diploid spore-producing sporophyte includes the root and shoot systems. In primitive plants (such as mosses and ferns) the gametophytes are usually free-living whereas in seed-bearing plants the gametophytes are sexually dimorphic and develop within the sexual organs of the flower: the male gametophyte (pollen grain) develops within the anthers and the female gametophyte (embryo sac) develops within the ovules, which ultimately gives rise to seeds after double fertilization. The structure of the female gametophyte and the fertilization process are distinctive features of the two phylogenetic groups of seed plants, the non-flowering gymnosperms (including conifers and cycads) and the flowering angiosperms (Figure 1).

In gymnosperms, the female gametophyte is a large multicellular structure - found, for example, at the base of each scale in a pinecone - producing from one to five archegonia, each carrying one egg cell. The large female gametophyte nourishes the developing embryo after fertilization. In angiosperms, the female gametophytes are embedded within ovules that in turn are harbored in the ovary, which develops into the fruit after fertilization. Each embryo sac displays a highly polar structure and consists of a small number of distinct cell types: two fertilization-competent cells (the egg cell and the central cell) and a variable number of sterile cells (synergids and antipodals) depending on the species (Figure 1).

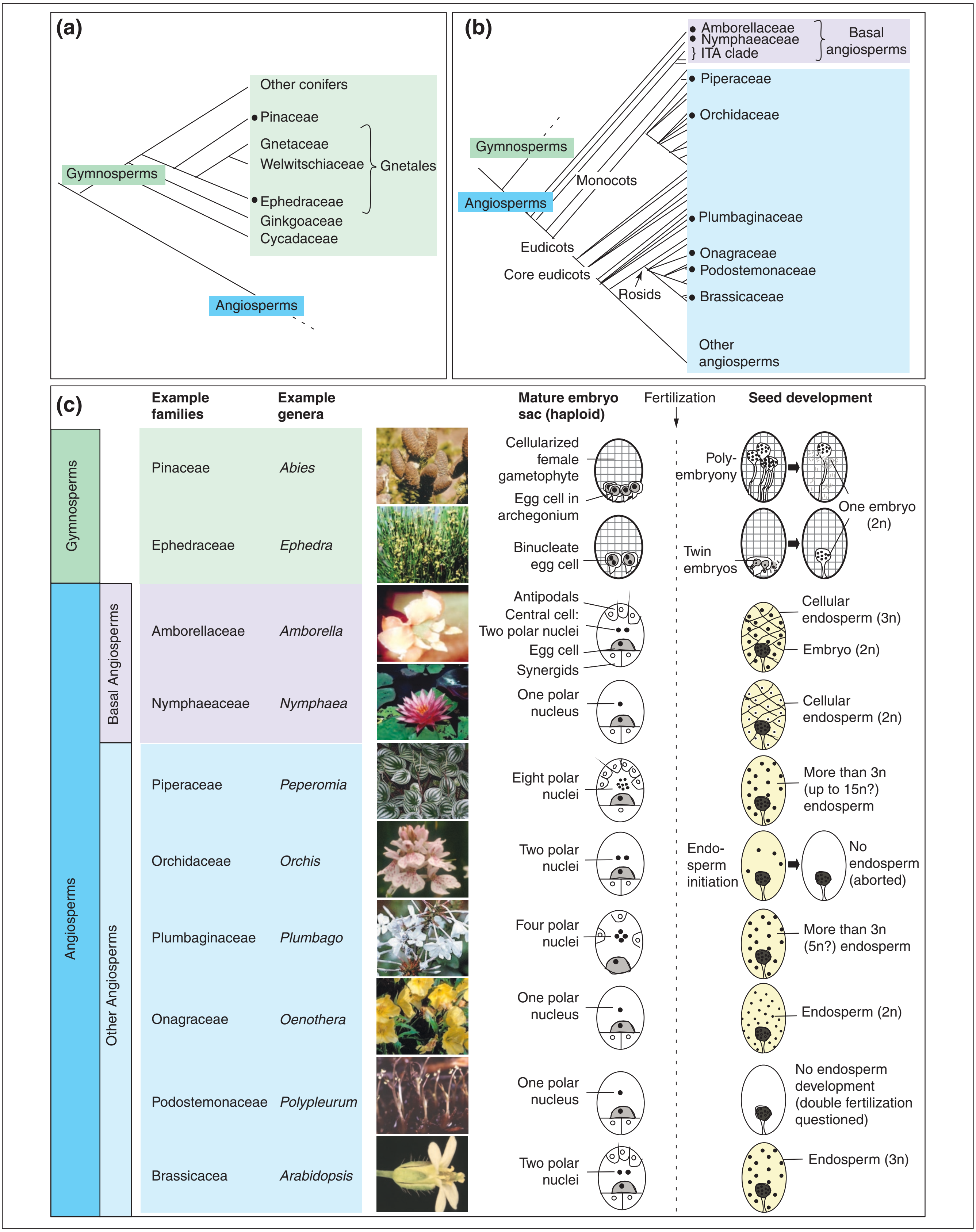

Figure 1 *(see the legend on the following page)*

Double fertilization is a distinctive feature of angiosperms. Each ovule receives a pollen tube that delivers two sperm cells to the embryo sac. One sperm fertilizes the egg cell, generating the diploid zygote, while the other sperm fertilizes the central cell giving rise to endosperm that is usually triploid. Like the gymnosperm female gametophyte, the endosperm is thought to have an embryo-nourishing function. Depending on the species, endosperm can be absent (as in Podostemonaceae, the river weed family), short-lived (as in the Orchidaceae, the orchid family), maintained in the mature seed (as in wheat, maize and palm), or absorbed by the developing embryo (as in pea, bean and *Arabidopsis*; see Figure 1 for examples).

Developmental origin of the endosperm

The structure of the embryo sac and the process of double fertilization provide the basis for endosperm development. The embryo sac is derived either from one (monosporic) or more (bisporic or tetrasporic) meiotic products (functional megaspores) [3]. Despite their clonal origin, the cells of a monosporic embryo sac differentiate along four separate developmental pathways. Nuclear migration and the relative position of the nuclei - which develop in a syncytium prior to cellularization - appear to play an important role in cell determination [4]. Differences in nuclear migration and division are likely to explain the variation in the number of nuclei present in the central cell (polar nuclei) and thus the variable ploidy of the endosperm between species. Most angiosperms have a Polygonum-type embryo sac with two polar nuclei and produce a triploid (3n) endosperm as in *Arabidopsis*. In a few species, more than two polar nuclei are present in the central cell leading to the formation of endosperm with a ploidy higher than 3n [3]. Interestingly, diploid endosperm originating from a central cell with only one polar nucleus was described in several families [1]. The higher angiosperm family Oenograceae (including the evening primrose) has diploid endosperm, as do basal angiosperms: Nymphaceae (including waterlilies); Cabombaceae (including fanworts); and members of the ITA clade (consisting of Illiciaceae such as star anise, Trimeniaceae, Austrobaileyaceae and in some phylogenies also Schisandraceae). In those species, the four-celled embryo sac reflects the minimal structure common to all angiosperms (Figure 1).

Hypotheses for the evolutionary origin of the endosperm

At the turn of the nineteenth century, fundamental discoveries about the female gametophyte and the double-fertilization process in flowering plants generated several hypotheses attempting to explain the evolutionary origin of the endosperm (reviewed in [2]). One hypothesis proposed a supernumerary embryo as the origin of endosperm. This model stipulates that the production of two genetically identical individuals ('organismal duplication') by double fertilization could favor functional divergence of the twin embryos. As a result, one 'altruist' embryo would acquire a novel nourishing function to the benefit of the sibling ('increased inclusive fitness through cooperative developmental behavior') [5]. The question then arises as to whether there are extant seed plants that have reproductive features reminiscent of this evolutionary transition. A recent revision of the phylogenetic position of the Gnetales (gymnosperms; see Figure 1a), which indicates that they are closely related to the conifers [6,7], has undermined the support this hypothesis would have gained from the discovery of a rudimentary double-fertilization process in *Ephedra* (Gnetales). The production of twin zygotes in *Ephedra* (Figure 1c, [8]) showed that fertilization of a binucleate egg can indeed produce two embryos in this species. Comparative embryology among basal angiosperms that are phylogenetically closer to a hypothesized ancestral species (Figure 1b) demonstrates that most basal angiosperms possess a four-celled embryo sac and produce a diploid endosperm. Could this be a remnant of the paleo-embryonic origin of the endosperm? It is difficult to accommodate this interpretation with the (predicted) occurrence of diploid endosperm among distant relatives

---

Figure 1 *(see the figure on the previous page)*
Reconstruction of the evolutionary ontogeny of the endosperm by phylogenetically anchored comparative embryology. A schematic representation of the phylogenetic trees of **(a)** gymnosperms and **(b)** angiosperms. Only species relevant for the text or the illustration in **(c)** are indicated. Trees were drawn according to [15] and [16], respectively. (c) A series of example species illustrating the diversity of embryo sac structures among extant plants and the subsequent diversity of endosperm ploidy. The embryo sac depicted for *Amborella* and *Arabidopsis* is of the Polygonum type. The variations in endosperm ploidy for Piperaceae and Plumbaginaceae, the transient presence of endosperm in Orchidaceae and its absence from Podostemonaceae are all only predicted, on the basis of cytological studies [3]. The ploidy number of the endosperm has been precisely measured in some basal angiosperms [1] and is known in *Arabidopsis*. For comparison to the angiosperm female gametophyte (embryo sac), the female gametophyte of two gymnosperm species is shown with copious cellularized haploid tissue and between one and five archegonia, each harboring one egg cell. In the case of *Ephedra*, egg cells are binucleate with a normal egg nucleus and a ventral canal nucleus [8]. In gymnosperms a situation of 'simple complex polyembryony' occurs [10] in which several egg cells can be fertilized (by mitotically distinct sperm nuclei), and each zygote then generates four embryo clones. The extent of this polyembryony varies between species, and a simplified form is depicted here. Ultimately, only one embryo will survive while the others degenerate (gray dashed lines in the example of *Abies*). In the case of *Ephedra*, both nuclei of the egg cell are fertilized by two sperm nuclei discharged by a single pollen tube (double fertilization) and polyembryony also applies to both fertilization products [2,8]. Embryo sac structures were drawn based on [3], and photographs are courtesy of the Missouri Botanical Garden [17]: *Abies balsamea, Ephedra viridis, Amborella trichopoda, Nymphaea odorata, Peperomia argyreia, Orchis mascula, Plumbago europaea and Oenothera macrocarpa.* The *Polypleurum dichotomum* and *Arabidopsis thaliana* photographs are courtesy of Rolf Rütishauser and John Bowman, respectively.

(Onagraceae) and the occurrence of triploid endosperm in the basal-most angiosperm *Amborella*.

A second hypothesis predicted that the endosperm is a homolog of the gymnosperm female gametophyte that became sexualized. In this scenario, the recruitment of the second sperm nucleus by a central cell, which is now fertilization-competent, provides some (unknown) fitness advantages to the embryo. This hypothesis is supported by the following observations and predictions: firstly, angiosperm endosperm development is as proliferative and invasive as that of the gymnosperm female gametophyte; secondly, the input of a paternal genome to the female precursor central cell may create hybrid vigor and allow biparental control over resource allocation to the embryo (see below); and finally, there could be an evolutionary advantage in forming nutritious storage tissue after fertilization as it ensures that no maternal resources are wasted on unfertilized gametophytes.

Although the female gametophyte has been at the center of hypotheses in the past, the male gametophyte merits further research from an evolutionary perspective, too. Given the energetic costs of gamete production [9], the presence of two sperm cells in the pollen of both gymnosperms and angiosperms is puzzling as it is generally accepted that double fertilization does not occur in the vast majority of gymnosperms [10] (except for *Ephedra* and *Gnetum*). Only one of the gymnosperm sperm nuclei fertilizes the egg cell, whereas the other is thought to degenerate. No role for the degenerating second sperm cell has yet been proposed, and the question remains as to why two sperm cells are maintained over long periods of evolutionary time in species that undergo only a single fertilization event. It seems a re-examination of the fertilization process in gymnosperms is overdue.

Triploidy of the endosperm

Apart from those exceptions producing endosperm with high ploidy, the widespread occurrence of triploid endosperm among angiosperms (with a 2:1 maternal:paternal genome ratio) raises the question of the evolutionary significance of this asymmetric parental genome balance. In 1989, Haig and Westoby [11] formulated the parental conflict theory, which states that the biparental contribution to the endosperm reflects the selfish interests of the mother and the father over resource allocation to the progeny. With an excess of maternal chromosomes over paternal ones, the mother can exert greater control over nutrient allocation to the progeny in relation to her own fitness. At the molecular level, parent-of-origin-specific mechanisms for the expression of selfish parental interests can act via cytoplasmically deposited determinants, dosage-sensitive gene products or genomic imprinting. Because of its role in embryo nourishment, the endosperm appears to be the battleground for such conflicts. In this evolutionary context, endosperm is viewed as the angiosperm homolog of the equivalent gymnosperm female gametophyte, but it also provides fitness advantages at the organismal level thanks to the acquisition of a paternal contribution (double fertilization). The current phylogenetic distribution of triploid and diploid endosperm does not indicate which was the basal state, but it does demonstrate that either diploidy or triploidy of the endosperm was gained or lost twice among flowering plants [1].

Integrating comparative embryology, phylogeny and functional genomics?

In the absence of any fossil record of an endosperm ancestor, elucidating the evolutionary ontogeny of this specialized embryo-nourishing fertilization product is difficult. Phylogenetically anchored studies of comparative embryology re-emerged a decade ago and are important for documenting the reproductive features underlying endosperm development in potential ‘missing links’ (basal angiosperms and exceptional cases). The application of functional genomics could allow a greater focus on key genes controlling endosperm developmental types and patterns. Genomic studies across extant seed plants are, however, limited to date by the lack of significant genome-sequence information or molecular genetic tools for the basal angiosperms and gymnosperms. Long generation times make ‘forward genetics’ for basal angiosperms and gymnosperms an unlikely proposition. Current proposals to develop genomics tools for the study of floral evolution in the angiosperms [12] might significantly contribute to the study of endosperm development and evolution.

Genomic research, for example, could be applied to detecting similarities between the various transcriptomes: the developing embryo and endosperm in angiosperms, or the endosperm and gymnosperm female gametophyte (perhaps using expression profiling based on cDNA-amplified full-length polymorphisms (cDNA-AFLP) or microarrays). One complementary approach would be to conduct systematic screens for gametophytic mutants that exhibit an altered development, for instance defects in nuclear migration and division that resemble features observed in basal angiosperms or gymnosperms. This can currently be done in model eudicot plant species such as *Arabidopsis* (see, for example, [13,14]). The unraveling of the genetic hierarchy controlling embryo sac and endosperm development will allow a focus on key developmental genes. These could then be analyzed in the basal angiosperms and gymnosperms with the prospect of reconstructing the evolutionary transitions relating to endosperm origin and development.

The puzzle of the evolutionary origin of the endosperm will continue to attract many scientists. Only the integration of comparative embryology, phylogeny and functional genomics within an evolutionary framework will allow the resolution of a series of enigmas that are already more than a century old.

Acknowledgements

We thank Peter Endress and Rolf Rütishauser for helpful discussions, and John Bowman and Rolf Rütishauser for providing photographs for Figure 1c. Our research is supported by the Kanton of Zürich, the Swiss National Science Foundation, Novartis, and a Searle Scholarship to U.G.

References

1. Williams JH, Friedman WE: **Identification of diploid endosperm in an early angiosperm lineage.** *Nature* 2002, **415:**522-526.
2. Friedman WE: **Developmental and evolutionary hypotheses for the origin of double fertilization and endosperm.** *CR Acad Sci III* 2001, **324:**559-567.
3. Maheshwari P: *An introduction to the embryology of angiosperms.* 15th edition. New Delhi: McGraw-Hill, 1995.
4. Grossniklaus U, Schneitz K: **The molecular and genetic basis of ovule and megagametophyte development.** *Semin Cell Dev Biol* 1998, **9:**227-38
5. Friedman WE: **Organismal duplication, inclusive fitness theory, and altruism: understanding the evolution of endosperm and the angiosperm reproductive syndrome.** *Proc Natl Acad Sci* 1995, **92:**3913-3917.
6. Chaw SM, Parkinson CL, Cheng Y, Vincent TM, Palmer JD: **Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of Gnetales from conifers.** *Proc Natl Acad Sci USA* 2000, **97:** 4086-4091.
7. Bowe, LM, Coat G, dePamphilis CW: **Phylogeny of seed plants based on all three genomic compartments: extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers.** *Proc Natl Acad Sci USA* 2000, **97:**4092-4097.
8. Friedman WE: **Evidence of a pre-angiosperm origin of endosperm: implications for the evolution of flowering plants.** *Science* 1992, **255:**336-339.
9. Snook RR, Markow TA: **Efficiency of gamete usage in nature: sperm storage, fertilization and polyspermy.** *Proc R Soc Lond B Biol Sci* 2002, **269:**467-473.
10. Hardev S: *Embryology of gymnosperms*. Stuttgart: Gebr. Borntraeger Science Publisher, 1978.
11. Haig D, Westoby M: **Parent-specific gene expression and the triploid endosperm.** *Am Nat* 1989, **134:**147-155.
12. Soltis DE, Soltis PS, Albert VA, Oppenheimer DG, dePamphilis CW, Ma H, Frohlich MW, Theissen G; Floral Genome Project Research group: **Missing links: the genetic architecture of flower and floral diversification.** *Trends Plant Sci* 2002, **7:**22-31.
13. Moore JM, Vielle-Calzada J-P, Gagliano W, Grossniklaus U: **Genetic characterization of *hadad*, a mutant disrupting megagametogenesis in *Arabidopsis thaliana*.** *Cold Spring Harb Symp Quant Biol* 1997, **62:**35-47.
14. Christensen CA, Subramanian S, Drews GN: **Identification of gametophytic mutations affecting female gametophyte development in *Arabidopsis*.** *Dev Biol* 1998, **202:**136-51.
15. Friedman WE, Floyd SK: **Perspective: the origin of flowering plants and their reproductive biology - a tale of two phylogenies.** *Evolution Int J Org Evolution* 2001, **55:**217-231.
16. **Angiosperm Phylogeny Website** [http://www.mobot.org/MOBOT/research/APweb/]
17. **Tropicos Image index** [http://mobot.mobot.org/W3T/Search/image/imagefr.html]