THE ROYAL SOCIETY

Received 15 July 2002
Accepted 16 September 2002
Published online 28 November 2002

The origin and maintenance of nuclear endosperms: viewing development through a phylogenetic lens

R. Geeta

Ecology & Evolution, State University of New York, Stony Brook, NY 11794-5245, USA
(*geeta@life.bio.sunysb.edu*)

The endosperm develops in fertilized ovules of angiosperms following fertilization of the central cell and nuclei in the female gametophyte. Endosperms differ in whether, and which, nuclear divisions are followed by cellular divisions; the variants are classified as cellular, nuclear or helobial. Functional correlates of this variation are little understood. Phylogenetic methods provide a powerful means of exploring taxonomic variation and phylogenetic patterns, to frame questions regarding biological processes. Data on endosperms across angiosperms were analysed in a phylogenetic context in order to determine homologies and detect biases in the direction of evolutionary transitions. Analyses confirm that neither all nuclear nor all helobial endosperms are homologous, raise the possibility that cellular development is a reversal in some derived angiosperms (e.g. asterids) and show that a statistically significant bias towards evolution of nuclear endosperms (and against reversals) prevails in angiosperms as a whole. This bias suggests strong selective advantages to having nuclear endosperm, developmental constraints to reversals or both. Homologies suggest that the microtubular cycle and cellularization pattern characteristic of reproductive cells across land plants may have been independently co-opted during multiple origins of nuclear endosperms, but information on cellular endosperms is essential to investigate further.

Keywords: endosperm development; evolution bias; constraint phylogeny

1. INTRODUCTION

The endosperm is a novel feature of developing angiosperm seeds. It results from the development of the central cell nuclei in the female gametophyte following fertilization, typically a triple fusion event involving two central cell nuclei and one sperm nucleus (but see Williams & Friedman 2002). Subsequent development of the endosperm is variable and falls into one of three classically defined modes: cellular, helobial or nuclear (e.g. Maheshwari 1950, but see Floyd & Friedman 2000). This variation occurs because cytokinesis may become uncoupled from the nuclear division cycle. In cellular endosperms, cell-wall formation follows the first division of the primary endosperm nucleus (PEN). In helobial endosperms, wall formation follows the first PEN division, producing two chambers, which vary in subsequent cellularization (Floyd & Friedman 2000). In nuclear endosperms, walls do not develop between the free nuclei (figure 1*a*). The origin and early evolution of the endosperm are widely discussed (Brink & Cooper 1940; Westoby & Rice 1982; Scheiner & Donoghue 1992; Friedman 1995; Williams & Friedman 2002), but questions regarding the functional significance and homologies of endosperm developmental modes remain (Wunderlich 1959).

Are there functional differences between seeds with cellular versus nuclear endosperms? This question has not been asked, and the data available to address it are limited. For example, most studies of embryo–endosperm relations in developing seeds have been conducted on nuclear endosperms, the type common among crop plants (Larkins & Vasil 1997; Berger 1999; Chaudhury *et al.* 2001); very few studies have considered cellular endosperms (Mogenesen 1985; Briggs 1996). Did nuclear (and cellular) endosperms arise just once, i.e. are they homologous? If so, then it is reasonable to use *Arabidopsis* (a rosid eudicot with nuclear mode) and tobacco (an asterid dicot with cellular mode) as representatives of the two modes in functional studies. Phylogenetic studies suggest that the ancestral angiosperm was cellular, and that nuclear endosperms arose multiple times from this state with few reversals (Bharathan 1999; Doyle & Endress 2000; Floyd & Friedman 2000; Albach *et al.* 2001). Such non-homology should be considered when sampling for functional studies. Does the phylogenetic pattern represent a prevalent bias in the direction of change? If so, it may indicate evolutionary advantages to nuclear endosperms, developmental constraints that prevent reversals to cellular endosperms or both. In any case, an evolutionary bias would suggest new approaches to poorly understood developmental variation.

I analyse data from a range of angiosperms to determine the phylogenetic patterns of transitions in endosperm development, to infer homologies and to detect biases in the direction of evolutionary change from one state to the other. The nuclear endosperms *Arabidopsis* and maize, for example, are shown not to be homologous, and homologies of cellular endosperms may be uncertain. A strong bias for evolutionary change from cellular to nuclear endosperms within angiosperms is demonstrated. This study illustrates the power of joining two disparate fields and the use of phylogenetic methods to frame questions regarding biological processes.

© 2002 The Royal Society

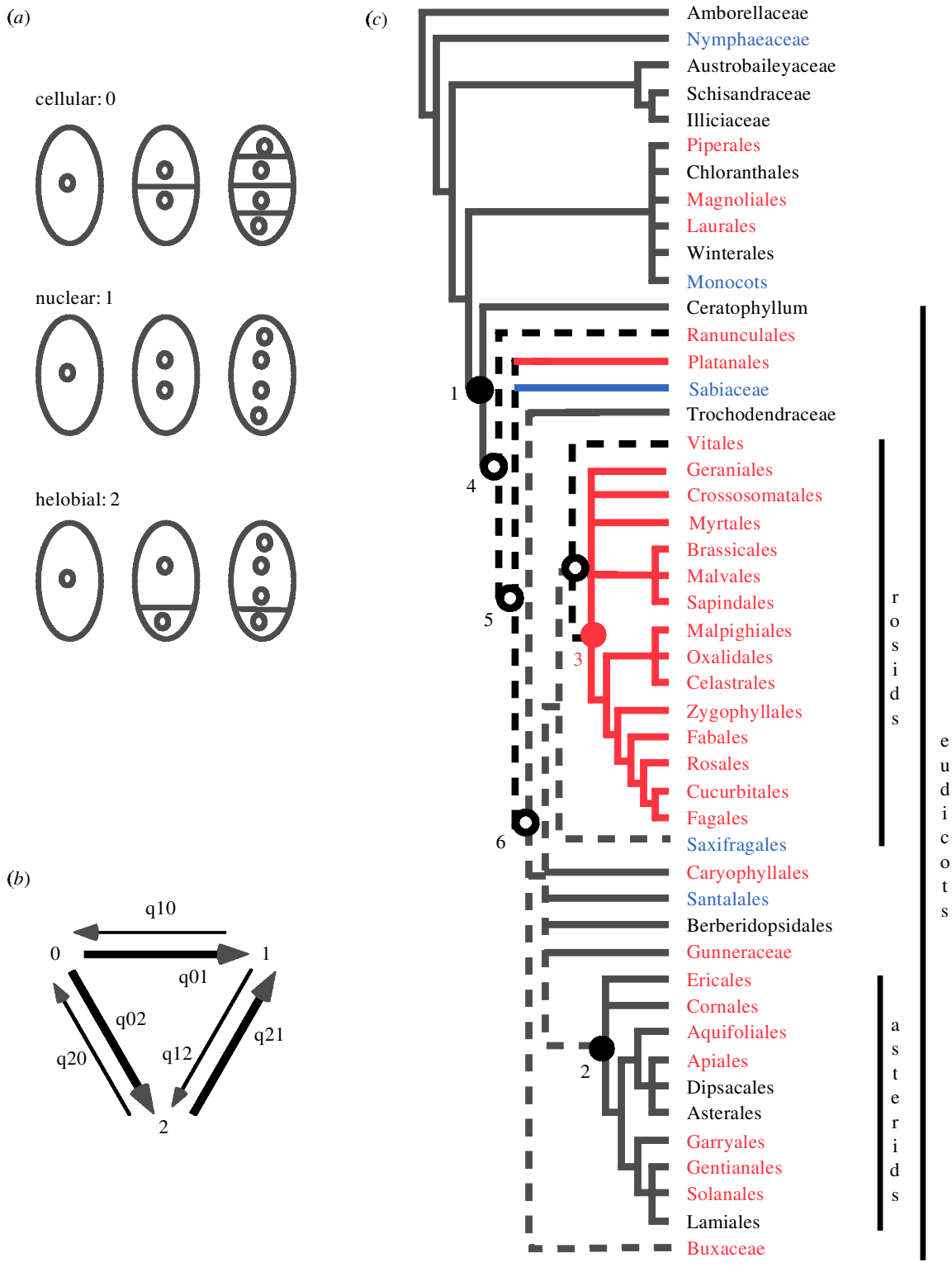

Figure 1. Development and evolution of endosperms. (*a*) Three modes of development, which vary in the pattern of cellularization. (*b*) Directions of change between the states indicated by arrows, along with rates of change (q01, q10 etc). (*c*) Summary of maximum-parsimony reconstructions across 100 resolutions of a strict consensus tree of 560 angiosperms (Soltis *et al.* 2000). The ancestral angiosperm had cellular endosperm, as did the eudicot ancestor (1) and asterid ancestor (2). Ancestors 4–6 within the eudicots were equivocally reconstructed under parsimony. If they are assumed to be cellular, then the cellular endosperm of asterids is homologous to that in ancestral angiosperms, but it is not if one of the ancestors is assumed to be not cellular. This result is supported by ML analyses of two-state data, but not three-state data. Taxa names in black, cellular; red, nuclear; blue, helobial (may include nuclear and/or cellular). Ancestral states: black lines and filled black circles, cellular; red lines and filled red circles, nuclear; hatched lines and open black circles, uncertain. Unless the colour of the subtending branch indicates otherwise, name colour should be taken to indicate that the corresponding state evolved within the named lineage.

2. MATERIAL AND METHODS

(a) *Data and coding*

Genera in recent angiosperm phylogenetic analyses (see § 2b) were used as sample taxa, assuming that they are monophyletic and that their inclusion in the phylogenetic studies was not based on their mode of reproductive development. Data on endosperm development are from Johri *et al.* (1992) and supplemented from other sources (see electronic appendix A available on The Royal Society's Publications Web site). Terminal taxa were coded according to the genus if information was available, otherwise, according to the family. Taxa in polyphyletic families (e.g. Saxifragaceae *sensu lato*) were coded only if information was available for the genus. Endosperm development was coded as cellular—0, nuclear—1 or helobial—2. Helobial endosperm was not scored in Boraginaceae and Solanaceae where, arguably, it occurs sporadically (Swamy & Parameswaran 1962). Statistical tests of character evolution required modification by removing taxa with missing information or polymorphisms and/or recoding as a binary trait. Recoding was carried out in four ways: (i) helobial and nuclear states combined; (ii) helobial and cellular states combined; (iii) helobial states combined with cellular states for non-monocots and with nuclear states for monocots, based on observed phylogenetic patterns (Bharathan 1999); and (iv) the helobial state was excluded.

(b) *Phylogenetic trees*

A strict consensus tree representing relationships among 560 angiosperm taxa was used (Soltis *et al.* 2000). The statistical tests required dichotomous trees, so polytomies in the strict consensus tree were resolved randomly to generate 100 or 1000 trees representing a range of phylogenetic hypotheses consistent with the strict consensus tree. Branch lengths were obtained from nucleotide data (18S rDNA, *rbc*L, *atp*B data matrix from multiple sources; Pam Soltis, personal communication) using maximum parsimony (MP) as implemented in PAUP (Swofford 2002). All analyses were done on these trees.

(c) *Patterns and trends in character evolution*

Phylogenetic patterns in character state transitions were assessed by MP and maximum likelihood (ML).

Parsimony reconstructions used unordered characters (Fitch parsimony: same 'cost' of transitions in all directions). Ancestral states at particular nodes were determined by MP reconstructions on 100–1000 trees using MacClade (Maddison & Maddison 2000) and ML estimation on five topologies using Discrete and Multi-state (Pagel 1994, 1997). Log-likelihoods (LogL) of alternative reconstructions were compared; a difference of greater than two log units represents significant support for the state with higher likelihood (Pagel 1999; Ree & Donoghue (1999) following a rule-of-thumb proposed by Edwards (1972)). Trends in evolution were detected by comparing 'opportunity for change' or frequencies of change as the proportion of the ancestral nodes with different states (Sanderson 1993). Proportions were calculated using the 'charting' option in MacClade (Maddison & Maddison 2000) (table 1).

The theory that there is a transition bias towards nuclear endosperms was tested by comparing ML results under models of character change that assume either equal or variously restricted transition rates given a phylogenetic hypothesis. Two ML methods were used, which differ in the type of model and the computation of transition rates: a discrete-time Markov model using ancestral-state reconstructions (Sanderson 1993), and a continuous-time Markov model that includes a parameter, κ, that scales change according to relative branch lengths, using all possible ancestral reconstructions (Pagel 1994, 1997). Sanderson's method permits inclusion of polymorphisms and missing data, and Pagel's method permits analysis of multi-state traits. Sanderson's method was applied using average numbers of transitions, obtained from MP reconstructions on 1000 topologies from two-state data in 560 taxa using MacClade. Pagel's method was applied to five topologies and reduced data (463 taxa, no polymorphisms or missing data) with two- or three-state coding using Discrete and Multi-state, respectively (Pagel 1994, 1997). Analyses of two-state data compared likelihoods under the unrestrained model with two rate parameters (gain: q01, loss: q10) and the restrained null model with one rate parameter (figure 1*b*). Analyses of three-state data compared likelihoods of the unrestrained model (six rate parameters) and variously restrained nested null models (*viz.* all equal rates, no reversals of nuclear or helobial, gains and losses of nuclear or helobial equal, and transitions between nuclear and helobial equal).

3. RESULTS

The distribution of character states was: cellular (139; 25%), nuclear (280; 50%), helobial (44; 8%), polymorphic (28; 5%) and missing (69: 12%). Nuclear endosperms dominated the dataset. However, this does not bias the results, as seen from consistent results in analyses of subsets of data that excluded monocots and rosids, which contain most of the nuclear endosperms in the dataset.

(a) *The ancestral endosperm was cellular in eudicots, asterids and monocots, and nuclear in rosids*

Parsimony reconstructions across 100 topologies with characters coded in multiple ways (taxa with missing data and polymorphisms included or not) indicate that the ancestral angiosperm, eudicot, asterid and monocot had cellular endosperms, while the ancestral rosid had nuclear endosperm (figure 1*c*). ML estimates supported these MP reconstructions with differences of more than eight log points between likelihoods of alternative states. Other reconstructions had lower levels of support. Nodes 4–6 were equivocally cellular or nuclear under MP (two- and three-state) and ML (two-state: LogL differences 1.04–4.11), but cellular under ML analysis of the three-state data (LogL differences 2.38–5.59). These results suggest that cellular endosperms in some derived groups (e.g. asterids) may not be homologous to those in 'basal' angiosperms. Within monocots, MP and ML reconstructions suggest multiple transitions from helobial to nuclear endosperms (not shown).

(b) *Once nuclear endosperm evolves it tends not to change*

The trend overall is that the state (cellular, nuclear or helobial) at an internal node is unchanged in descendant lineages (table 1). Nuclear endosperm is most stable, remaining unchanged in 97.6%, cellular is next (92.4%) and helobial is the least stable (87.7%). Gains of nuclear endosperms (q01) are significantly more frequent than losses (q10) regardless of topology, coding scheme, model

Table 1. Frequencies of transitions between endosperm modes in 560 angiosperms. Conservation (bold) is more frequent than change, nuclear endosperm is the most conserved and transitions to nuclear are the most frequent.
(Average number of changes in ancestral character state reconstructions on 1000 random resolutions of a strict consensus tree (Soltis *et al.* 2000). Numbers computed from ancestral-state reconstructions under MP using MacClade (Maddison & Maddison 2000). Numbers in parentheses are proportions of 'conservation' (diagonal) or 'change' (off-diagonal), calculated as a fraction of the opportunities for change (row totals) from cellular, nuclear or helobial.)

development to: from:	cellular (0)	nuclear (1)	helobial (2)
cellular (0)	**377.6 (0.924)**	26.7 (0.065)	4.4 (0.011)
nuclear (1)	5.7 (0.009)	**586.7 (0.976)**	8.8 (0.015)
helobial (2)	0.4 (0.004)	12.9 (0.119)	**94.8 (0.877)**

Table 2. 'Gains' and 'losses' of nuclear endosperm in angiosperms and subgroups. Data fit a two-rate model (gains, q01, more than losses, q10) better than a one-rate model (q01 = q10), as shown by a significant difference between the log-likelihoods (LogL) under the two models, in angiosperms except asterids analysed alone.
(Endosperm development coded as a two-state trait in modes (i)–(iv) (see § 2a), results of coding (iii) shown, other codings are generally consistent. (*a*) ML estimates of rates of gains and losses under a discrete-time Markov model are proportions, calculated as in table 1 for 100 trees (model I: Sanderson 1993). (*b*) ML estimates of instantaneous rates of gains and losses shown for one tree, calculated under a continuous-time Markov model using Discrete (Pagel 1994, 1997). Results from five trees were consistent. The statistic $G = -2 \times (\mathrm{LogL}_{\mathrm{model1}} - \mathrm{LogL}_{\mathrm{model2}})$ follows a χ^2 distribution, d.f. = 1, the difference in the number of parameters between the models compared.)

	one-rate model		two-rate model				
	q01 = q10	LogL	q01	q10	LogL	*G*	two-rate model better?
(*a*) discrete-time Markov model							
I. angiosperms (560 taxa)	0.032	−159.2	0.070	0.009	−143.78	32.29**	yes (q01 > q10)
II. rosids and monocots pruned (296 taxa)	0.058	−130.83	0.069	0.028	−128.75	4.17*	yes (q01 > q10)
III. asterids (148 taxa)	0.057	−65.07	0.064	0.031	−64.56	1.04	no (q01 = q10)
(*b*) continuous-time Markov model							
I. angiosperms (463 taxa)	0.043	−297.50	0.108	0.003	−252.31	90.39**	yes (q01 > q10)
II. rosids and monocots pruned (249 taxa)	0.06	−234.37	0.067	0.003	−228.57	11.60*	yes (q01 > q10)
III. asterids (137 taxa)	0.016	−126.74	0.018	0.005	−125.33	2.82	no (q01 = q10)

* $p < 0.025$, ** $p < 0.001$.

of character evolution or method of analysis. Both ML tests of two-state data revealed a significant bias towards the evolution of nuclear endosperms in all groups except asterids (table 2). The different coding schemes give generally consistent results; however, in analyses that excluded rosids and monocots differences in rates were not significant in some trees using coding (i) (helobial with nuclear). These results are not surprising because most helobial–nuclear transitions occur in monocots. Analysis of three-state data revealed that the different rates are not equal and that change from cellular and helobial to nuclear (q01, q21) is greater than the reverse (q10, q12). The likelihood of all rates being different is not differentiated from that of no reversals from nuclear or helobial (q10 = 0, q20 = 0) or from that of gains and losses of helobial being equal (q02 = q20) (table 3). Thus, all analyses point to a strong bias towards the evolutionary origin and maintenance of nuclear endosperms.

Comprehensive studies of this type are necessarily limited by the data and prevailing developmental and phylogenetic hypotheses. Taxa were excluded because either their phylogenetic position or their embryology is unknown. There are uncertainties in the positions of specific groups (Qiu *et al.* 1999) and questions about the typology (Floyd & Friedman 2000). However, the present results are expected to remain robust in future analyses because of the overwhelming nature of the bias, the fact that varying the positions of groups does not affect the results in preliminary tests using Sanderson's (1993) method (not shown), and because changes in the typology

Table 3. Transition rates of endosperm modes in 463 angiosperms. The unrestrained model gives rates of change as q01 = 0.0262, q02 = 0.0057, q10 = 0.0000 1, q12 = 0.0000 2, q20 = 0.0015 and q21 = 0.0548. The data best fit models under which the rate of transition to nuclear endosperms is higher than the reverse. Other restrictions of rates are indistinguishable from the model under which all rates are different.
(Endosperm development coded as three-state; 1–6 rate parameters and scaling parameter, κ, estimated under a continuous-time Markov model using MULTI-STATE (Pagel 1994, 1997). LogL under model (*a*), the alternative hypothesis (Ha), was compared, in turn, with LogL under different restrained models (*b–g*), the null hypotheses (Ho). ML estimates shown for one tree; results from five trees were consistent.)

model	parameters	LogL	d.f.	*G*	*p*	Ho rejected? ($\alpha = 0.05$)
Ha						
(*a*) all rates differ	6	−190.35				
Ho						
(*b*) all rates equal	1	−246.63	5	112.57	< 0.001	yes, rates not equal
(*c*) q10 = 0	5	−190.23	1	0.235	0.5–0.75	no
(*d*) q20 = 0	5	−189.92	1	0.856	0.25–0.50	no
(*e*) q01 = q10	5	−215.55	1	50.411	< 0.001	yes, gain > loss for nuclear
(*f*) q02 = q20	5	−192.98	1	3.259	0.05–0.10	no
(*g*) q12 = q21	5	−263.64	1	53.808	< 0.001	yes, gain > loss for nuclear

will affect classification of cellular and helobial but not nuclear endosperms.

4. DISCUSSION

Nuclear endosperms evolved multiple times and, therefore, are not homologous across angiosperms. The homologies of cellular endosperms are uncertain. The cellular endosperm of ancestral angiosperms was retained in 'basal' groups and the ancestral eudicot, ranunculid and monocot. However, some cellular endosperms (e.g. asterids) may represent retained ancestral states, reversals or newly evolved modes, although different analyses disagree on this.

There are several implications of these proposals of homology. Nuclear endosperms in rosids and monocots have similar microtubule cycles, decoupled from the cell cycle, with cellularization following radial microtubular patterning (unlike meristematic cells). These features are conserved in reproductive cells across angiosperms (Brown & Lemmon 2001; Olsen 2001), so nuclear endosperms may represent instances of parallel evolution (no information is available for helobial endosperms). Based on light microscopy, cellular endosperms could be expected to have coupled microtubule and cellular cycles of the meristematic type with no radial microtubular patterning; surprisingly, this has not been confirmed for any cellular endosperm (B. Lemmon, personal communication). Characterization of these features in cellular endosperms of 'basal' angiosperms and eudicots would enable the questions emerging here to be addressed. Are independent origins of nuclear from cellular endosperms the result of parallel evolution of the same processes? Do asterids and basal angiosperms have different cellular processes, supporting the theory of non-homology of cellular endosperms? Of broader significance is the possibility that if the 'basal' cellular endosperm has microtubule cycles with pre-prophase bands (PPBs) and cellularization typical of meristematic cells (Gunning 1982), then the hypothesis that the endosperm and embryo share a common evolutionary origin would be supported (Friedman 1995; Floyd & Friedman 2000). The cytoskeleton features of nuclear endosperms would represent mechanisms co-opted from reproductive cells.

These results may support previous theories that the asterid cellular endosperm is not homologous to other cellular endosperms (Dahlgren 1991). Non-homology may be reflected in a set of other ovular and seed features of asterids, such as micropylar and chalazal endosperm haustoria, tenuinucellate ovules and integumentary tapetum (Kapil & Tiwari 1978; Mikesell 1990; Albach *et al.* 2001). It is possible that these features result from correlated evolution, perhaps controlled by a common genetic pathway (Balasubramanian & Schneitz 2000).

Once a nuclear endosperm evolves, it is likely to persist, whereas cellular and helobial endosperms are likely to evolve into nuclear endosperms. This evolutionary pattern is consistent with two not mutually exclusive scenarios: (i) nuclear endosperms have a strong selective advantage that prevents reversals; and (ii) developmental constraints prevent reversals once nuclear endosperms evolve. There is little indication of the nature of any selective advantage, but two developmental aspects suggest mechanisms that could present directional constraints on evolution.

First, nuclear development is characterized by non-PPB microtubule cycles uncoupled from cellularization cycles. If these cycles were coupled in the ancestral cellular endosperm and evolution of nuclear endosperms entailed a loss of coupling, it is conceivable that an associated loss of critical (regulatory?) information poses a barrier to evolutionary reversal to cellular endosperm. However, such a shift is possible, as indicated by some nuclear endosperms switching to the PPB cycle late in development (Brown & Lemmon 2001). Second, it has been suggested that endosperms control embryo morphogenesis, particularly at the globular embryo stage, when changes critical for normal embryogenesis occur (Krishnamurthy 1988; Lester & Kang 1998). At this stage, cellular and nuclear endosperms secrete and accumulate polysaccharides and enzymes that may be important for further development;

most nuclear endosperms become cellularized, while cellular endosperms secrete polysaccharides into the space between endosperm and embryo (Mogenesen 1985; Briggs 1996; van Hengel *et al.* 1998; Otegui *et al.* 1999). It may be that genetic pathway(s) underlying these transitions differ across endosperm types or are altered to prevent the evolution of cellular and helobial endosperms from nuclear endosperms. Little is known about the genetic bases of these differences. This phylogenetic analysis suggests that comparative investigations of cellular, helobial and nuclear endosperms could yield new insights into the developmental processes that underlie this variation.

5. CONCLUSIONS

Phylogenetic patterns of variation in endosperm development reveal that: (i) nuclear and helobial endosperms are not homologous across angiosperms, and homologies of cellular endosperms are uncertain; and (ii) a strong transition bias favours evolution of nuclear development across angiosperms. These phylogenetic patterns suggest new lines of inquiry using genetic and comparative approaches that could advance our understanding of endosperm development, embryo–endosperm relations and evolution.

The author thanks Mark Pagel for being generous with his time and patient in his responses to queries regarding the use of DISCRETE and MULTI-STATE, Rick Ree for the file formatting program that made it all possible, Pam Soltis for the three-gene angiosperm dataset, and Roy Brown, Beatrice Grabowski, Yaffa Grossman, Betty Lemmon, Manuel Lerdau, Dianna Padilla, Michael Sanderson and two anonymous reviewers for helpful comments on previous versions of the manuscript.

REFERENCES

Albach, D. C., Soltis, P. S. & Soltis, D. E. 2001 Patterns of embryological and biochemical evolution in the asterids. *Syst. Bot.* **26**, 242–262.

Balasubramanian, S. & Schneitz, K. 2000 NOZZLE regulates proximal–distal pattern formation, cell proliferation and early sporogenesis during ovule development in *Arabidopsis thaliana*. *Development* **127**, 4227–4238.

Berger, F. 1999 Endosperm development. *Curr. Opin. Plant Biol.* **2**, 28–32.

Bharathan, G. 1999 Endosperm development in angiosperms: a phylogenetic analysis. In *Biodiversity, taxonomy, and ecology* (ed. R. K. Tandon & P. Singh), pp. 167–180. New Delhi: Scientific Publishers (India).

Briggs, C. 1996 An ultrastructural study of the embryo/endosperm interface in the developing seeds of Solanum nigrum L. zygote to mid torpedo stage. *Ann. Bot.* **78**, 295–304.

Brink, R. A. & Cooper, D. C. 1940 Double fertilization and development of the seed in angiosperms. *Bot. Gaz.* **102**, 1–25.

Brown, R. C. & Lemmon, B. E. 2001 The cytoskeleton and spatial control of cytokinesis in the plant life cycle. *Protoplasma* **215**, 35–49.

Chaudhury, A. M., Koltunow, A., Payne, T., Lu, M., Tucker, M. R., Dennis, E. S. & Peacock, W. J. 2001 Control of early seed development. *Ann. Rev. Cell Dev. Biol.* **17**, 677–699.

Dahlgren, G. 1991 Steps toward a natural system of the dicotyledons: embryological characters. *Aliso* **13**, 107–165.

Doyle, J. M. & Endress, P. K. 2000 Morphological phylogenetic analysis of basal angiosperms: comparison and combination with molecular data. *Int. J. Plant Sci.* **161**, S121–S153.

Edwards, A. W. F. 1972 *Likelihood*. Cambridge, UK: Cambridge University Press.

Floyd, S. K. & Friedman, W. E. 2000 Evolution of endosperm developmental patterns among basal flowering plants. *Int. J. Plant Sci.* **161**, S57–S81.

Friedman, W. E. 1995 Organismal duplication, inclusive fitness theory, and altruism: understanding the evolution of endosperm and the angiosperm reproductive syndrome. *Proc. Natl Acad. Sci. USA* **92**, 3913–3917.

Gunning, B. E. S. 1982 The cytokinetic apparatus: its development and spatial regulation. In *Cytoskeleton in plant growth and development* (ed. C. W. Lloyd), pp. 229–292. London: Academic Press.

Johri, B. M., Ambegaokar, K. B. & Srivastava, K. B. 1992 *Comparative embryology of angiosperms*. Berlin: Springer.

Kapil, R. N. & Tiwari, S. C. 1978 The integumentary tapetum. *Bot. Rev.* **44**, 457–490.

Krishnamurthy, K. V. 1988 Endosperm controls symmetry changes in the developing embryos of angiosperms. *Proc. Ind. Acad. Sci.* (*Plant Sci.*) **98**, 257–259.

Larkins, B. & Vasil, I. K. (eds) 1997 *Cellular and molecular biology of plant seed development*. Dordrecht, The Netherlands: Kluwer.

Lester, R. N. & Kang, J. H. 1998 Embryo and endosperm function and failure in *Solanum* species and hybrids. *Ann. Bot.* **82**, 445–453.

Maddison, D. R. & Maddison, W. P. 2000 *MACCLADE: interactive analysis of phylogeny and character evolution (computer program)*. Sunderland, MA: Sinauer.

Maheshwari, P. 1950 *An introduction to the embryology of angiosperms*. New York: McGraw-Hill.

Mikesell, J. 1990 Anatomy of terminal haustoria in the ovule of plantain (*Plantago major* L.) with taxonomic comparison to other angiosperm taxa. *Bot. Gaz.* **151**, 452–464.

Mogenesen, H. L. 1985 Ultracytochemical localization of plasma membrane-associated phosphatase activity in developing tobacco seeds. *Am. J. Bot.* **72**, 741–754.

Olsen, O. A. 2001 Endosperm development: cellularization and cell fate specification. *A. Rev. Plant Physiol. Plant Mol. Biol.* **52**, 233–267.

Otegui, M., Lima, C., Lima, C., Madonado, S. & de Lederkremer, R. M. 1999 Development of the endosperm of *Myrsine laetevirens* (Myrsinaceae). I. Cellularization and deposition of cell-wall storage carbohydrates. *Int. J. Plant Sci.* **160**, 491–500.

Pagel, M. 1994 Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. Lond.* B **255**, 37–45.

Pagel, M. 1997 Inferring evolutionary processes from phylogenies. *Zool. Scripta* **26**, 331–348.

Pagel, M. 1999 The maximum likelihood approach to reconstructing ancestral states of discrete characters on phylogenies. *Syst. Biol.* **48**, 612–622.

Qiu, Y.-L., Lee, J., Fabiana, B.-Q., Soltis, D. E., Soltis, P. S., Zanis, M., Zimmer, E. A., Chen, Z., Savolainen, V. & Chase, M. W. 1999 The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature* **401**, 404–407.

Ree, R. H. & Donoghue, M. J. 1999 Inferring rates of change in flower symmetry in asterid angiosperms. *Syst. Biol.* **48**, 633–641.

Sanderson, M. J. 1993 Reversibility in evolution: a maximum likelihood approach to character gain/loss bias in phylogenies. *Evolution* **47**, 236–252.

Scheiner, S. M. & Donoghue, M. J. 1992 The evolution of endosperm: a phylogenetic account. In *Ecology and evolution*

of plant reproduction (ed. R. Wyatt), pp. 356–389. New York: Chapman & Hall.

Soltis, D. E. (and 15 others) 2000 Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences. *J. Linn. Soc. Bot.* **133**, 381–461.

Swamy, B. G. L. & Parameswaran, N. 1962 The helobial endosperm. *Biol. Rev.* **38**, 1–50.

Swofford, D. L. 2002 *PAUP 4.0b10: phylogenetic analysis using parsimony (and other methods) (computer program)*. Sunderland, MA: Sinauer.

van Hengel, A. J., Guzzo, F., van Kemmen, A. & de Vries, S. C. 1998 Expression pattern of the carrot EP3 endochitinase genes in suspension cultures and in developing seeds. *Plant Physiol.* **117**, 45–53.

Westoby, M. & Rice, B. 1982 Evolution of seed plants and inclusive fitness of plant tissues. *Evolution* **36**, 713–724.

Williams, J. H. & Friedman, W. E. 2002 Identification of diploid endosperm in an early angiosperm lineage. *Nature* **415**, 522–526.

Wunderlich, R. 1959 Zur Frage der Phylogenie der Endospermtypen bei den Angiospermen. *Oester. Bot. Zeitsch.* **106**, 204–293.

As this paper exceeds the maximum length normally permitted, the author has agreed to contribute to production costs.

Visit http://www.pubs.royalsoc.ac.uk to see an electronic appendix to this paper.