doi:10.1111/j.1420-9101.2005.00998.x

Tempo, mode and phylogenetic associations of relative embryo size evolution in angiosperms

M. VERDÚ

Centro de Investigaciones sobre Desertificación (CSIC-UV-GV), Apartado Oficial, Albal, Valencia, Spain

Keywords:

anagenesis;
cladogenesis;
comparative method;
evolutionary trends;
gradualism.

Abstract

Relative embryo size (E : S, the ratio of embryo to seed) is a key trait related to germination ecology and seed plant evolution. A small, underdeveloped embryo is a primitive feature of angiosperms, which has led to the hypothesis that an evolutionary trend towards increasing E : S has occurred. Here, I examine first the tempo and mode of E : S evolution in angiosperms; then I test for phylogenetic associations of E : S with traits hypothetically related to anagenetic (germination time) and cladogenetic (number of species per family and differential speciation) change, and finally I test the existence of a directional increasing trend in E : S. The analysis of the evolutionary tempo suggests that E : S changed very fast early in evolutionary time and remained stable later, which is consistent with early radiations and fits well with the history of angiosperms consisting of rapid spread associated with great diversification rates soon after their origin. E : S evolution in angiosperms has not followed a punctuational mode of evolution but a scaled-gradualism evolution in which stasis has occurred in longer branches of the phylogeny. An evolutionary trend towards increasing E : S has not been actively driven by anagenesis nor cladogenesis, although large E : S is associated with high levels of diversification (i.e. number of species per family). This rapid ecological diversification occurring in the early radiation probably produced an increasing phenotypic variance in the E : S. Because the ancestral embryo was so small, an increase in variance might have produced a passive trend towards the only direction allowed for the ancestral embryo to evolve. Thus, a passive diffusion away from a lower bound may explain the average increase in E : S.

Introduction

Since Simpson's (1944) seminal work, two key features of evolutionary change have been studied: the 'tempo' (the rate) and the 'mode' (the manner or pattern). Evolutionary tempo refers to the long-term rate at which evolution proceeds (i.e. fast or slow) whereas the mode refers to the temporal distribution of that change (i.e. gradual or punctuated) (Ridley, 1993). The comparative method allows estimation of statistical parameters related to the tempo (δ), mode (κ) and the phylogenetic signal (λ) of the evolution of a trait (Pagel, 1994, 1999).

Correspondence: Miguel Verdú, Centro de Investigaciones sobre Desertificación (CSIC-UV-GV), Apartado Oficial, 46470 Albal, Valencia, Spain. Tel.:+34-961220540; fax: +34-961270967;
e-mail: Miguel.Verdu@uv.es

Evolutionary trends are produced when the average value of a trait changes directionally within a clade. Trends may be driven by anagenetic change accumulated slowly within a species, with speciation (cladogenesis) being a mechanism to iterate this process across species (Gould, 2002). Thus, an evolutionary trend driven by anagenesis may be caused by natural selection within a lineage and also on average for a group of lineages if all experience the same directional selective pressure (i.e. natural selection favours the individuals holding a trait that confers greater fitness; Ridley, 1993). Cladogenesis may also produce evolutionary trends in a trait if the species having such trait speciate more often than others (Ridley, 1993; Mooers *et al.*, 1999; Gould, 2002). In the presence of bounds, a passive, nondriven evolutionary trend can also be produced, because a trait will evolve following a randomly generated diffusion away from that

bound (McShea, 1994, 1998; Wang, 2001; Gould, 2002). All these mechanisms producing trends are not exclusive, and all of them can act in different proportions. Wang (2001) has recently developed the skewness analysis to quantitatively decompose trends into driven and passive portions.

The study of evolutionary trends had been an exclusive topic for palaeontologists, who analysed the skewness of the trait distribution in monophyletic subclades (McShea, 1994, 1998; Wang, 2001), but the development of statistical techniques within the framework of the comparative method has put this topic in the arena of neontologists, who can trace the evolutionary history of a clade without fossils, using only information from extant species.

The evolution of angiosperms has traditionally been understood as a history of evolutionary trends (Stebbins, 1974). One of the classical evolutionary trends proposed in angiosperms is that of the relative embryo size, measured as the embryo to seed ratio (E : S hereafter). Small, underdeveloped embryos have been characterized as a primitive feature of angiosperms by many authors who have also suggested an evolutionary trend towards seeds in which the E : S increases at the expense of the endosperm (Martin, 1946; Stebbins, 1974; Baskin & Baskin, 1998). More recently, Forbis *et al*. (2002) have reconstructed the E : S value of the angiosperm ancestor with standard phylogenetic methods, confirming its small size but failing to detect a general directional increasing trend. It must be noted that standard phylogenetic models cannot detect directional trends because they assume an unbiased random walk model of evolution (Pagel, 1999). Instead, recent phylogenetic models allow detection of directional trends by examining the correlation between the taxa's trait values and the total phylogenetic distance of path length from the root of the phylogenetic tree (Pagel, 1994, 1999).

The evolution of E : S in angiosperms is intrinsically related to physiological processes like seed germination: species having large embryos with storage cotyledons and little endosperm germinate faster than species with small embryos that depend on copious endosperm (Vivrette, 1995). Rapid germination is a heritable and usually advantageous trait because it allows seedlings to establish early, enhancing their competitive ability relative to conspecifics (Geber & Griffen, 2003). Many case studies have shown that earlier germinators grow, survive and reproduce better than late germinators (see Verdú & Traveset, 2005 for a review). Thus, fast germination could be a selective force driving an anagenetic trend towards increasing E : S in angiosperms.

From the cladogenetic perspective, E : S has also been related to angiosperm diversification as species richness is greater in families with larger E : S (Vivrette, 1995). However, a phylogenetically informed analysis remains to be done to confirm the relationships between E : S and both germination rate and diversification.

This paper examines the evolution of angiosperm E : S reconstructed by Forbis *et al*. (2002) by estimating: (i) its evolutionary tempo to test slowed (consistent with early radiation) vs. accelerated (consistent with late radiation) evolutionary rates over time, (ii) its evolutionary mode to test gradual vs. punctuated change and (iii), the existence of a directional trend towards larger E : S driven by traits hypothetically related to anagenetic (germination time) and cladogenetic (diversification) change.

Material and methods

Data, phylogeny and sensitivity analyses

Data on E : S of angiosperms were obtained from Forbis *et al*. (2002), who reconstructed the embryo to seed ratio (E : S) of 179 family means after averaging 1218 species from the illustrations of Martin (1946). I restricted my analysis to angiosperms because few data on gymnosperm E : S were available and also because gymnosperms have a completely different history and seed anatomy than angiosperms, which could have led to different evolutionary mechanisms of E : S than those originally hypothesized for angiosperms. Thus, I used 175 family means from 1208 angiosperm species after pruning the 11 gymnosperm tips from Forbis *et al*. (2002) phylogenetic tree. I used the tree with maximum likelihood branch lengths that the authors reconstructed from the B-series tree of Soltis *et al*. (2000) with rearrangements according to Qiu *et al*. (1999) (Fig. 1).

Data on mean time of germination were obtained from the Thompson and Morgan seed germination database (http://www.backyardgardener.com/tm.html). This database provides, at the genus level, the range of days taken by the seeds to germinate under optimal conditions. I calculated the mean germination time for each family by averaging all the mean germination times of the genera contained within each family. After corrections for taxonomic synonymies following Mabberley (1997) and Angiosperm Phylogeny Group (1998), I calculated mean germination time for 765 genera from 160 families, but only used those families with E : S data ($n = 100$).

I tested the reliability of the Thompson and Morgan database by comparing the obtained values against the official sources (International Seed Testing Association and Association Official Seed Analysts). To do this, I took the total germination test duration of 64 families from the Handbook of Seed Technology for genebanks (Ellis *et al*., 1985) as a rough estimate of germination time. Then, I correlated these values with those from the Thompson and Morgan database. Both datasets were positive and significantly correlated (Pearson $r = 0.3$; d.f. $= 64$; $P = 0.01$), confirming that the Thompson and Morgan database is sufficiently reliable.

Data on the number of species per family were obtained from Mabberley (1997). The family

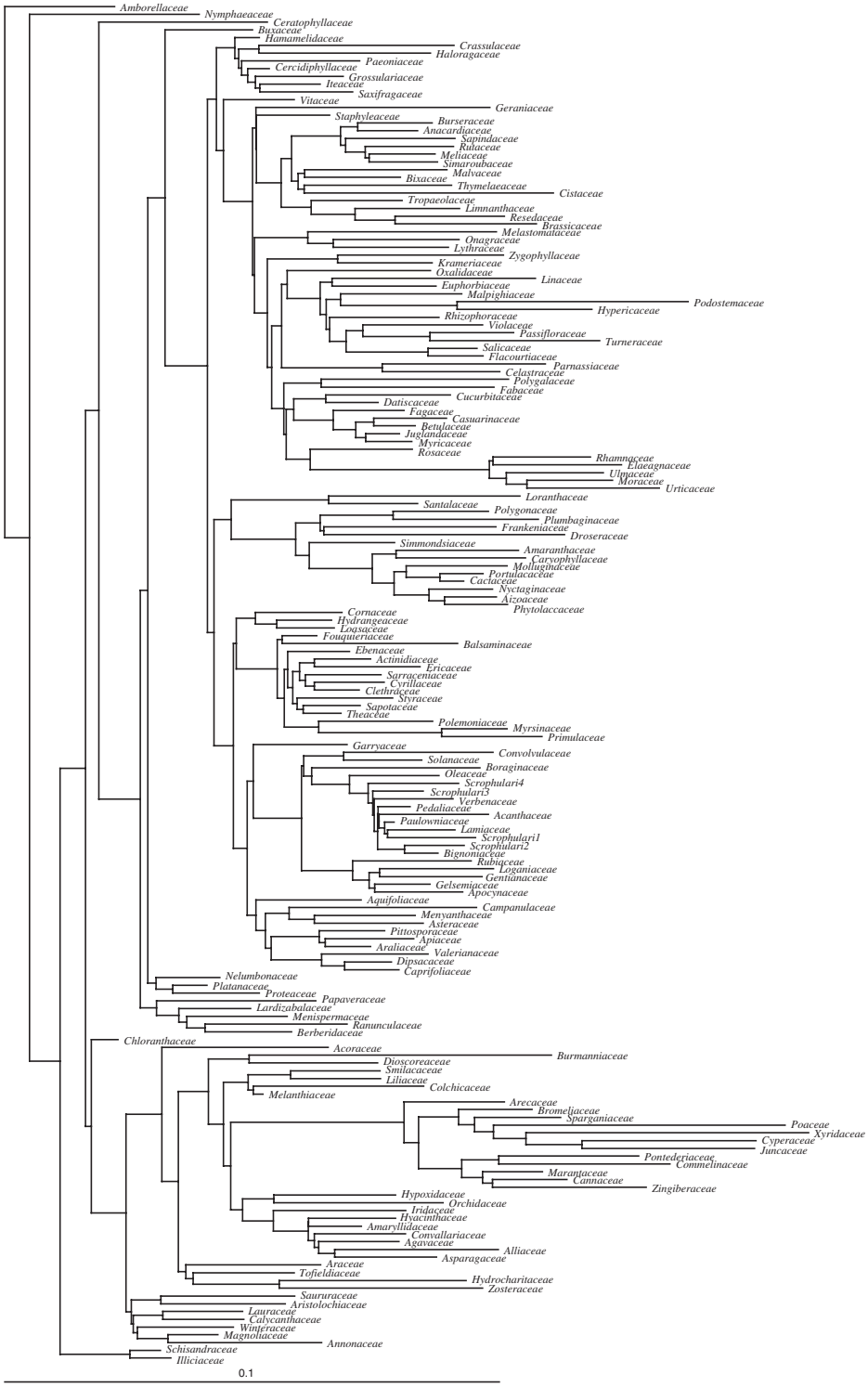

Fig. 1 Phylogenetic relationship of the families included in the dataset in which the E : S evolution has been studied. This tree is the same used in Forbis et al. (2002) except the gymnosperm clade that has been pruned.

Scrophulariaceae was excluded from this database because it is not a monophyletic family and it was not possible to assign the number of species belonging to each of the four families in which Scrophulariaceae is currently split. The entire database is available on request from the author.

Sensitivity analyses were run by repeating the tests after accommodating uncertainty on the topology of the phylogeny, the branch lengths and character scoring (Donoghue & Ackerly, 1996). Uncertainty on the topology of the phylogeny was accounted for by rerunning the tests in the alternative phylogenies provided by Forbis *et al*. (2002). The first alternative phylogeny rearranges the position of Ceratophyllaceae as the sister group of monocots whereas the second alternative phylogeny excludes Vitaceae from the tree. Uncertainty on branch lengths was accommodated by adding random noise to branch lengths of the tree (Díaz-Uriarte & Garland, 1998). The noise was normally distributed, with variance proportional to current branch length. A variance multiplier of 0.1 was entered to add to branch lengths. Thus, if branch length is 10, the noise added will have a variance of 1.0 (0.1×10). Uncertainty on character scoring (E : S and germination time) was accommodated by adding random noise to the trait mean values. Because E : S and germination time data were obtained at genus level and averaged to family level, I used the family means and variances of the original datasets to generate normally distributed random noise with these parameters. This method was not used to add noise to the number of species per family because this value has no associated variance in the original source (Mabberley, 1997). Isaac & Purvis (2004) have recently shown that the taxonomic uncertainty on species numbers does not compromise the validity of tests of correlates of species richness when the study traits are not associated to taxonomic splitting, as occurs with E : S and germination time. Thus, uncertainty on the species numbers was not considered. However, results must be interpreted with caution because it is difficult to ensure that these variables are not potentially correlated with the chance of a new species to be described throughout a third, unknown, biological factor or because of the appeal to taxonomists (Isaac & Purvis, 2004). For example, if seed mass is a taxonomic criterion and if seed mass is correlated to E : S ratio, this may bias the analysis.

Tempo and mode of evolution

I estimated the three scaling parameters (λ,κ and δ) in the phylogeny to test for the contribution of the phylogeny, the mode and the tempo of E : S evolution with the Continuous software (Pagel, 1999). The contribution of the phylogeny to the observed values of E : S was tested by comparing models with different λ values in a likelihood ratio (LR) test. The parameter λ reveals if the phylogeny correctly predicts the patterns of covariance among taxa on a given character (Freckleton *et al*., 2002). This parameter is similar to the measure of the phylogenetic signal of Blomberg *et al*. (2003), that is, it tests the tendency for related taxa to resemble each other, with no implication as to the cause of such a resemblance. If λ is not significantly different from zero, then related taxa are not more similar than expected by chance, and therefore the trait is evolving among the species as if they were independent (i.e. as in star-like phylogenies or non-phylogenetic informed cross-species analyses). If λ differs from zero, but it is small (i.e. $1 > \lambda > 0$), then the tree topology over-estimates the covariance among species. In this situation some phylogenetic correction is needed, although it is expected that the phylogenetic history will have a minimal effect. Finally, if $\lambda = 1$, then the default phylogeny can be used to inform the analyses because the trait is evolving as expected given the topology of the tree (i.e. following the constant variance random walk model assumed by many comparative methods).

The mode of E : S evolution was studied by comparing the likelihoods of models with different κ values. This parameter differentially stretches or compresses individual branch lengths in the phylogenetic tree. If κ does not differ from zero, then trait evolution has been independent on the branch lengths, which is consistent with punctuated evolution. If trait evolution has followed the branch lengths of the tree, then κ will not differ from the unity, which is consistent with default gradualism. If κ is significantly greater than one, then the long branches need to be stretched more than shorter ones, indicating that the rate of evolution accelerates within a long branch. Finally, if κ is significantly smaller than one (and greater than zero), then the long branches need to be compressed more than shorter ones, indicating stasis on the longer branches.

The tempo of E : S evolution was tested by comparing the likelihoods of models with different δ values. The parameter δ is the power at which the pathlengths (root to tip distances in the tree) are raised, and it detects whether the rate of character evolution has slowed ($\delta < 1$; shorter paths contribute much more to trait evolution) or accelerated ($\delta > 1$; longer paths contribute much more to trait evolution) over time from the roots to the tips of the tree. As shorter paths represent earlier evolution in the phylogeny, slowed evolution is consistent with early radiation. Likewise, as longer paths represent later evolution in the phylogeny, accelerated evolution is consistent with late radiation (Pagel, 1994, 1999; see Harmon *et al*., 2003 for an empirical example).

Evolutionary trend

Driven by anagenesis

The existence of an evolutionary trend towards increasing E : S driven by anagenesis was assessed in the Continuous software by means of a generalized least

squares approach (GLS) in which the log-likelihood of a directional random-walk model was statistically compared to that of the standard constant–variance random walk model (Pagel, 1997, 1999). The directional model has, in addition to the variance of the evolution parameter of the standard model, another parameter that measures the regression of trait values across taxa against total path length from the root to the tips of the phylogeny. Both models were statistically compared by means of a LR test statistic, which is asymptotically distributed as a χ^2 random variable with 1 d.f.

To explore possible adaptive forces associated to anagenesis, I tested the phylogenetic association between E : S and mean germination time by means of a GLS approach under the fittest evolutionary model (directional vs. standard constant-variance random walk models) in which the phylogeny was scaled with the three scaling parameters explained above. The GLS correlated evolution test compares the likelihood of an evolutionary model in which the covariances between traits have been set to zero (i.e. correlation = 0) against another model in which the covariances take the maximum likelihood values (i.e. correlation ≠ 0).

If the evolution of a trait is constrained by the presence of bounds (like E : S, that is a proportion of embryo to seed and therefore bounded between 0 and 100%), an evolutionary trend driven by anagenesis produces skewed distributions of that trait in the same direction (right-skewed for increasing trends and left-skewed for decreasing trends) for both the whole clade and its constituent subclades lying away from the bound (McShea, 1994). Thus, right-skewed distributions may be caused by selection if increases occur more often than decreases on account of the advantage to individuals of higher values of the dimension in question. Alternatively, passive (nondriven) trends may be produced when the trait originates close to that bound and therefore the most probably evolution follows a randomly generated temporal drift away from that wall. Under this situation, the clade's distribution is skewed by the presence of the bound but the subclades from the tail of that distribution, far from that bound, will have no tendency to be skewed (McShea, 1994).

I tested if the presence of bounds results in an anagenetic driven or a passive trend towards increasing E : S by calculating the symmetry of the E : S distribution (skewness) for the angiosperm clade and for each of its five largest monophyletic groups (subclades). These subclades were Asterids, Core Eudicots, Euasterids, Eurosids and Eumagnolids. Twenty-six families were excluded from the skewness analysis because they were not monophyletic groups containing a minimum sample size (sample size was <9 families for all of them). The shape of the density plot representing the E : S distribution in the angiosperm clade was not affected by the exclusion of these families. I tested whether the skewness estimates significantly departed from zero following the procedure outlined by Zar (1996). Density plots were drawn with the help of the ldahist (MASS) function in the R statistical software programme (Ihaka & Gentleman, 1996).

I also quantified the driven and passive portions of this evolutionary trend by calculating the skewness between and within subclades, and the skewness due to changes in variance among subclades, as suggested by Wang (2001). Driven trends are characterized by large within-group skewness because all the groups (subclades) are similarly skewed as a result of a common selective pressure. In contrast, passive trends may be characterized by large between-group skewness (because only subclades near the bounds are skewed) or by large heteroskedasticity skewness (because the variance of the subclades near the bounds is different to that of the clades far from that bound). Negative logs of the E : S were taken to fulfil the assumptions of the model regarding the non-negative skewness of each group and the increase of the variance with the mean. The analysis was run with a code written by Steve Wang for the R statistical software program (Ihaka & Gentleman, 1996).

Driven by cladogenesis

The number of species per family has previously been used as a proxy for angiosperm diversification rates after testing the significance of the correlation between this variable and net diversification rates, that takes into account the age of angiosperm taxa based on the fossil record (Magallón & Sanderson, 2001; Verdú, 2002). Large phylogenetic trees resolved at the species level are extremely rare, and we therefore have to work at present with trees resolved at higher taxonomic levels. Two different approaches to the study of cladogenesis have been used in studies dealing with diversification (speciation minus extinction) in these higher level phylogenies: the first approach uses the number of species per taxa (family, order, etc.) as a trait associated with each tip (e.g. Ricklefs & Renner, 1994; Tiffney & Mazer, 1995; Heilbuth, 2000; Verdú, 2002) whereas the second one focuses on the behaviour of each internal node (= speciation event) of the phylogeny (Savolainen *et al.*, 2002; Webster *et al.*, 2003). The species-counting method extends the analysis until the species level, but it does not take into account the phylogenetic relationships beyond the tips (i.e. it is like grafting a star-like phylogeny containing all the species per family onto each tip of the family level phylogeny). In contrast, the method of the nodes is a complete phylogenetic-informed analysis, but it does not extend the analysis to the species level. Here I use both methods to test if E : S is associated to cladogenesis. For the species-counting approach, I run a correlation between E : S and the number of species per family in the Continuous software as explained above for germination rate. Unlike other methods previously used to test for correlations of species richness (Agapow & Isaac, 2002), the GLS method used here does not rely on the reconstruction of ancestral states in internal nodes

and allows incorporation of the scaling parameters and the evolutionary model (directional or nondirectional) to the test (Pagel, 1994, 1999).

For the method of nodes, I followed Nee *et al.*'s (1992) procedure by testing if taxa with large E : S speciate more often than taxa with small E : S. If this pattern occurs, then a shorter daughter branch derived from each node will tend to end in the greater E : S clade. Node values were reconstructed by means of the phylogenetic independent contrasts implemented in the PDAP programme (Garland *et al.*, 1993). The hypothesis that shorter branches from each node lead to large E : S clades more often than expected by chance was tested with a χ^2 test.

Results

Tempo and mode of evolution

Phylogenetic correction is needed to study the evolution of E : S in angiosperms because the scaling parameter λ was significantly different from zero [$\lambda = 0.95$ (0.79–1); estimate and 95% confidence interval; LR = 30.67; $P << 0.01$]. The default phylogeny is appropriate to inform the subsequent analyses because λ was not significantly different from unity (LR = 0.53; $P = 0.30$). The value of λ is robust to different sources of uncertainty because strongly similar values and confidence intervals were obtained with alternative phylogenies, with noised branch lengths and with noised E : S values (results not shown).

The mode of E : S evolution was not consistent with either punctuational or gradual modes of change because the parameter κ (0.63; 0.32–0.97) was significantly greater than 0 (LR = 8.66; $P << 0.001$) and smaller than 1 (LR = −39.78; $P < 0.01$). A κ value between 0 and 1 is consistent with a mode of change in which stasis has occurred in longer branches of the phylogeny, as if the rate of evolution decelerates within a long branch. The estimated value of κ is robust to topological and branch length uncertainty but slightly sensible to character scoring uncertainty because the confidence interval included 1 (0.78; 0.45–1.13) when noise was added to E : S.

The tempo of E : S evolution estimated by the parameter δ (0.31; 0.02–0.84) indicates that evolution has slowed over time from the root to the tips of the phylogenetic tree, as the parameter was significantly smaller than 1 (LR = 3.16; $P \leq 0.01$). The estimation of δ is robust to character scoring and branch length uncertainty but slightly sensible to topological uncertainty because the confidence interval of δ in the phylogeny rearranging the position of Ceratophyllaceae included 1 ($\delta = 0.42$; 0.04–1.01).

Evolutionary trend

Driven by anagenesis

The existence of an evolutionary trend towards an increase in E : S driven by anagenesis was not supported by the GLS analysis, as the log-likelihood of the directional random-walk model ($L = -41.92$) did not differ significantly from that of the standard constant-variance random walk model ($L = -42.01$) (LR = 0.081; $P = 0.61$; Fig. 2). This result was robust to all the studied sources of uncertainty.

The presence of bounds did not support the existence of an anagenesis-driven trend towards increasing E : S, because the expected E : S right-skewed distribution of the angiosperms clade and its constituent subclades was not detected (Fig. 3). The E : S density plot of the angiosperm clade shows a tri-modal distribution, corresponding to the shape of its constituent subclades (Fig. 3, top). The first peak corresponds to the subclade with the smallest E : S average (Eumagnoliids), the second peak to the clades with medium E : S averages (Core Eudicots, Euasterids and Asterids) and the third to the clade with the higher E : S mean (Eurosids) (Fig. 3, bottom). The only significantly skewed subclades were those in the extremes of the E : S distribution, and the direction of the skewness was opposite: to the right for the subclade located in the smaller E : S extreme and to the left for the subclade located in the greater E : S extreme.

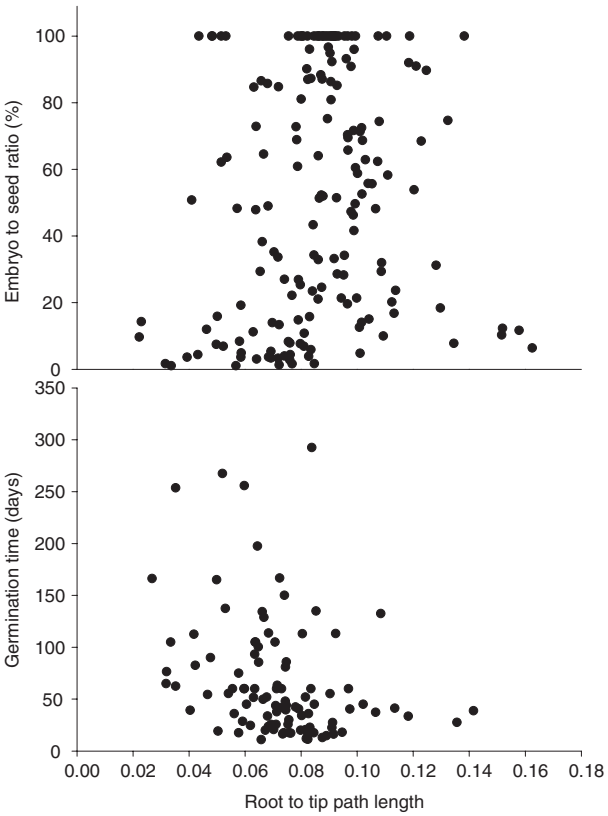

Fig. 2 Embryo size (top panel) has not followed an evolutionary trend whereas the time needed to germinate (bottom panel) has decreased since the origin of angiosperms, as shown when each trait value is plotted against the phylogenetic distance separating contemporary taxa (tips) from the ancestor (root).

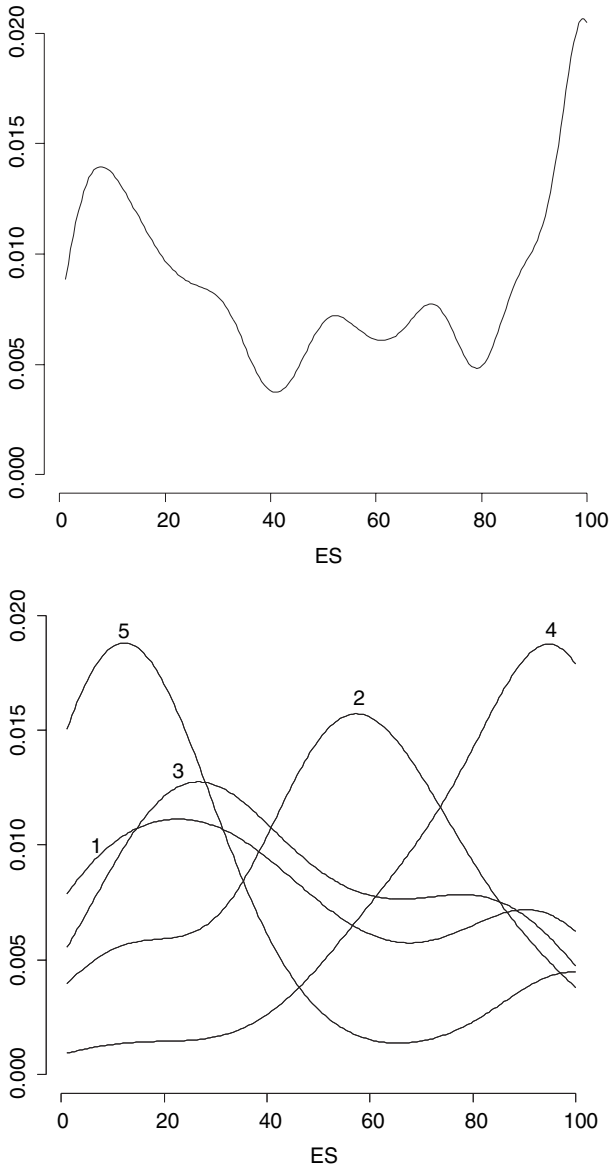

Fig. 3 Angiosperm density plot (top panel) shows that embryo size (E : S) is trimodal and nonskewed (skewness ± SE = −0.05 ± 0.20; $n = 149$, $P > 0.05$), but its constituent subclades close to the left and right walls (bottom panel) are skewed. 1 Asterids (0.39 ± 0.55; $n = 17$; $P > 0.05$); 2 Core Eudicots (−0.31 ± 0.58; $n = 15$; $P > 0.05$); 3 Euasterids (0.37 ± 0.43; $n = 29$; $P > 0.05$); 4 Eurosids (−1.67 ± 0.34; $n = 49$; $P < 0.01$); 5 Eumagnolids (1.47 ± 0.38; $n = 39$; $P < 0.01$).

The partitioning of skewness indicates that the heteroskedasticity skewness predominates (60%) over the within- and between-group skewness (39.4 and 0.6%, respectively), which is consistent with a passive trend produced by an increase in variance.

Mean germination time, a possible adaptive force associated to anagenesis, followed a decreasing evolutionary trend because the directional model ($L = -558.01$) explained the data better than the constant-variance model ($L = -563.72$) (LR test = 5.71; d.f. = 1; $P < 0.001$), suggesting that germination time has shortened since the origin of angiosperms (Fig. 2). However, germination time under such directional model was not related to E : S (correlation = −0.064; LR = 0.12; $P = 0.61$). The results were robust to the different sources of uncertainty analysed.

Driven by cladogenesis

The species counting and the node methods provided different results. E : S was positively related to the number of species per family (correlation = 0.22; LR test = 4.33; $P < 0.01$). In constrast, the number of nodes in which shorter branches ended in large E : S clades (32 of 55) was not more frequent than expected by chance ($\chi^2_1 = 1.473$; $P = 0.22$). These results did not change by using alternative phylogenies or after the addition of noise to branch lengths.

Discussion

The evolution of angiosperm E : S is consistent with a scaled-gradualism mode in which stasis has occurred in longer branches of the phylogeny and a decelerated tempo in which the rate of E : S evolution has slowed over time. An evolutionary trend towards increasing E : S has not been driven by anagenesis nor cladogenesis, although large E : S is associated with high levels of diversification (i.e. number of species per family).

The evolution of E : S in angiosperms cannot be understood without a phylogenetic framework because this trait shows a significant phylogenetic signal (*sensu* Blomberg *et al.*, 2003). The scaling parameter used to test the necessity of phylogenetic corrections (λ) also indicates that the angiosperm phylogeny used in the current analysis is appropriate because E : S evolution are evolving as expected given the tree topology, that is, following a Brownian constant-variance random walk model.

Even after ensuring that the phylogeny used here is appropriate to study E : S evolution, it is worthwhile to look at two main types of biases that could be affecting the results reported here. The first type is related to the uncertainty on different sources of information. The sensitivity analyses reported here show that results are very robust to topological, branch length and character scoring uncertainty, although many other untested scenarios (e.g. new phylogenies, new E : S data) could generate new uncertainty. The second type of bias is related to taxon sampling, which may affect the correlations shown here if the study traits (E : S, germination rate, and diversification rate) do not evolve independently of the sampling character (see Ackerly, 2000 for an explanation of the statistical consequences of sampling bias). The sampling character can be found in Martin (1946) because he provided the 'ancestral' source of the

E : S data used here, and he only recognises a slight geographical (USA) sampling bias. As long as this bias is not expected to be correlated with my study traits, results shown here may be considered robust to taxon sampling bias.

A change from E : S = 14% in the angiosperm ancestor to the average tip values of around 52% has led to the hypothesis that an increasing directional trend has been produced in E : S (Forbis *et al.*, 2002). However, such a directional trend was not detected in an anagenetic scenario after comparing constant-variance vs. directional random-walk models. Ancestor state reconstruction has been shown to be wrong under nonBrownian motion or in the absence of phylogenetic signal or in the presence of evolutionary trends (Oakley *et al.*, 2000; Laurin, 2004). None of these situations is followed by E : S evolution, and therefore the ancestral state of E : S is not expected to be erroneously reconstructed.

Then, if the reconstructed ancestral E : S value is correctly estimated in 14%, and the average tip values are around 52%, why is a directional trend not detected? Anagenetic and cladogenetic processes have been invoked to explain such a trend. The anagenetic process leading to increasing E : S was based on the relationship between E : S and another trait that is under directional selection (germination time). Vivrette (1995) found in a nonphylogenetically informed analysis that the larger the E : S, the faster the germination. Because rapid germination is a heritable trait that usually enhances plant fitness by means of the competitive advantages of the earlier seedlings, this trait is subjected to directional selection (Geber & Griffen, 2003; Verdú & Traveset, 2005). Supporting this hypothesis on germination time evolution, I have found here that germination time follows a directional trend towards a reduction in the number days required to germinate. However, a concomitant trend in the E : S evolution has not been produced because both variables are not related once phylogeny is accounted for. It should also be noted that other characters affecting germination velocity, such as other types of seed dormancy (i.e. physiological, physical and morphophysiological) and the type of habitat may mask the relationship between E : S and germination time (see Forbis *et al.*, 2002 for a discussion on the role of these characters in the evolution of E : S).

A cladogenetic process could also explain an evolutionary E : S trend if the families with large E : S diversify more than others. This process has not occurred as long as shorter branch lengths do not end in greater E : S clades more often than expected by chance. However, when the number of species per family was used as the proxy for diversification, then a highly significant correlation between E : S and diversification arises, similar to that found by Vivrette (1995) . It should be noted that using the number of species within families as a trait associated to each tip is like grafting star-like phylogenies onto the tree, and therefore the phylogenetic relationships are not properly accounted for (Nee *et al.*, 1992). Thus, both methods may be testing diversification rates in different parts of the tree, with the species counting method being affected greatly by cladogenesis nearer the tips. Alternatively, these apparently contradictory findings may be explained if large E : S is not the cause, but the consequence of successful cladogenesis in the presence of E : S bounds. This is because rapid ecological diversification occurring in evolutionary radiations usually leads to a greater variety of phenotypic states among terminal taxa, or in other words, to increasing phenotypic variance (Schluter, 2000; Harmon *et al.*, 2003; Ackerly & Nyfeller, 2004). And an increase in variance may produce passive trends in the presence of bounds (Wang, 2001).

Both premises (rapid diversification and the presence of bounds) are present in the history of angiosperm E : S evolution. Early rapid radiation characterized by a rapid increase in the diversification rate followed by a later decline has been confirmed for angiosperms (Midgley & Bond, 1991a, b; Crane *et al.*, 1995; Davies *et al.*, 2004). A similar pattern has been shown here for the evolutionary tempo of E : S because the phylogenetic scaling parameters indicate that this trait changed very fast early in evolutionary time and remained stable later, which is consistent with early radiations (Pagel, 1999). If radiations are accompanied by phenotypic diversification (Schluter, 2000; Harmon *et al.*, 2003; Ackerly & Nyfeller, 2004), the unique direction for the ancestral small embryo to diversify was towards increasing values. Thus, the increase in the average E : S size obeys a mere statistical, passive mechanism and not a driven trend. This mechanism, named 'diffusion within a structured design space' (Fisher, 1986), or the 'drunkard's walk' (Gould, 2002), shows that trends (more specifically, the skewed distributions) may arise in a random system because of the presence of a constraining boundary and despite equal probabilities for movement either to the right or to the left. The embryo of angiosperms originated and evolved very close to the left wall of the E : S. This left wall is 0% because seeds without an embryo are not compatible with life, and the initial angiosperm relative embryo size was 14%. According to the statistical model described above, the embryo in angiosperms has evolved away from the left wall toward an increase in the average size. Likewise, a right wall (E : S = 100%) is bounding the evolution of embryos in angiosperms because embryos larger than the seed are, again, not compatible with life. This right wall was far from the E : S of the angiosperms ancestor and therefore it should not limit the early passive trend produced towards an increase in the average E : S. However, once large embryos evolved, the right wall may have acted as a boundary producing passive trends in the opposite direction. Indeed, only the angiosperm subclades close to the left and right walls are skewed (Fig. 3), indicating that branching bias is not operating uniformly across the state space and therefore

that a driven trend does not exist (McShea, 1994). Additional evidence against a directional trend towards increasing E : S in the presence of bounds is the fact that the angiosperms E : S heteroskedasticity skewness predominates over the within and between groups skewness, which is a typical feature of passive trends produced by bounds (Wang, 2001).

In summary, the increase of E : S in angiosperms cannot be explained by the advantages of fast germination or successful cladogenesis of large E : S families. Instead, a diffusion away from morphological boundaries may explain this E : S increase produced in the early radiation that allowed the angiosperms to colonize new habitats.

Acknowledgments

I am especially indebted to T. Forbis for providing and helping me with the original database of E : S, and to M. Pagel and A. Diniz-Filho for discussions on the assumptions of distinct evolutionary processes to run the analyses. Special thanks are also given to S. Wang who gave me his R code for the analyses of evolutionary trends and advised me about variable transformations to fulfil the premises of the analysis. A. de Queiroz, A.Mooers, G. Gleiser, M. Geber, M. Pagel, P. García-Fayos, T. Forbis, S. Renner and S. Wang greatly improved the manuscript with their comments. I was granted a contract from the program Ramón y Cajal del Ministerio de Ciencia y Tecnología Español during this study.

References

Ackerly, D.D. 2000. Taxon sampling, correlated evolution, and independent contrasts. *Evolution* **54**: 1480–1492.

Ackerly, D.D. & Nyfeller, R. 2004. Evolutionary diversification of continuous traits: phylogenetic tests and application to seed size in the California flora. *Evol. Ecol.* **18**: 249–272.

Agapow, P.M. & Isaac, N.J.B. 2002. MacroCAIC: revealing correlates of species richness by comparative analysis. *Divers. Distrib.* **8**: 41–43.

Angiosperm Phylogeny Group. 1998. An ordinal classification for the families of flowering plants. *Ann. Mo. Bot. Gard.* **85**: 531–553.

Baskin, C.C. & Baskin, J.M. 1998. *Seeds: Ecology, Biogeography, and Evolution of Dormancy and Germination*. Academic Press, San Diego.

Blomberg, S.P., Garland, T. & Ives, A.R. 2003. Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution* **57**: 717–745.

Crane, P.R., Friis, J.M. & Pedersen, K.R. 1995. The origin and early diversification of angiosperms. *Nature* **374**: 27–33.

Davies, T.J., Barraclough, T.G., Chase, M.W., Soltis, P.S., Soltis, D.E. & Savolainen, V. 2004. Darwin's abominable mystery: insights from a supertree of the angiosperms. *Proc. Natl. Acad. Sci. USA* **101**: 1904–1909.

Díaz-Uriarte, R. & Garland, T. 1998. Effects of branch-length errors on the performance of phylogenetically independent contrasts. *Syst. Biol.* **47**: 654–672.

Donoghue, M.J. & Ackerly, D.D. 1996. Phylogenetic uncertainties and sensitivity analyses in comparative biology. *Philos. Trans. R. Soc.* **351**: 1241–1249.

Ellis, R.H., Hong, T.D. & Roberts, E.H. 1985. *Handbook of Seed Technology for Genebanks Vol II: Compendium of Specific Germination Information and Test Recommendations*. Handbooks for Genebanks N 3, International Board for Plant Genetic Resources, Rome.

Fisher, D.C. 1986. Progress in organismal design. In: *Patterns and Processes in the Histoy of Life* (D. M. Raup & D Jablonski, eds), pp. 99–117. Springer, Berlin.

Forbis, T.A., Floyd, S.K. & de Queiroz, A. 2002. The evolution of embryo size in angiosperms and other seed plants: implications for the evolution of seed dormancy. *Evolution* **56**: 2112–2125.

Freckleton, R., Harvey, P.H. & Pagel, M. 2002. Phylogenetic analysis and ecological data: a review of the evidence. *Am. Nat.* **160**: 712–726.

Garland, T., Dickerman, A.W., Janis, C.M. & Jones, J.A. 1993. Phylogenetic analysis of covariance by computer-simulation. *Syst. Biol.* **42**: 265–292.

Geber, M.A. & Griffen, L.R. 2003. Inheritance and natural selection on functional traits. *Int. J. Plant Sci.* **164**: S21–S42.

Gould, S.J. 2002. *The Structure of Evolutionary Theory*. The Belknap Press of Harvard University Press, Cambridge.

Harmon, L.J., Schulte, J.A., Larson, A. & Losos, J.B. 2003. Tempo and mode of evolutionary radiation in iguanian lizards. *Science* **301**: 961—964.

Heilbuth, J.C. 2000. Lower species richness in dioecious clades. *Am. Nat.* **156**: 221–241.

Ihaka, R. & Gentleman, R. 1996. R: a language for data analysis and graphics. *J. Comput. Graph. Stat.* **5**: 229–314.

Isaac, N.J.B. & Purvis, A. 2004. The 'species problem' and testing macroevolutionary hypotheses. *Divers. Distrib.* **10**: 275–281.

Laurin, M. 2004. The evolution of body size, Cope's Rule and the origin of amniotes. *Syst. Biol.* **53**: 594–622.

Mabberley, D.J. 1997. *The Plant-book. A Portable Dictionary of Vascular Plants*. Cambridge University Press, Cambridge.

Magallón, S. & Sanderson, M.J. 2001. Absolute diversification rates in angiosperm clades. *Evolution* **55**: 1762–1780.

Martin, A.C. 1946. The comparative internal morphology of seeds. *Am. Midl. Nat.* **36**: 513–660.

McShea, D.W. 1994. Mechanisms of large-scale evolutionary trends. *Evolution* **48**: 1747–1763.

McShea, D.W. 1998. Possible largest-scale trends in organismal evolution: Eight 'live hypotheses'. *Ann. Rev. Ecol. Syst.* **29**: 293–318.

Midgley, J.J. & Bond, W.J. 1991a. Ecological aspects of the rise of angiosperms – a challenge to the reproductive superiority hypotheses. *Biol. J. Linn. Soc.* **44**: 81–92.

Midgley, J.J. & Bond, W.J. 1991b. How important is biotic pollination and dispersal to the success of the angiosperms. *Philos. Trans. R. Soc. Lond. B* **333**: 209–215.

Mooers, A.O., Vamosi, S.M. & Schluter, D. 1999. Using phylogenies to test macroevolutionary hypotheses of trait evolution in Cranes (Gruinae). *Am. Nat.* **154**: 249–259.

Nee, S., Mooers, A.Ø. & Harvey, P.H. 1992. The tempo and mode of evolution revealed from molecular phylogenies. *Proc. Natl. Acad. Sci. USA* **89**: 8322–8326.

Oakley, T.H. & Cunningham, C.W. 2000. Independent contrasts succeed where ancestor reconstruction fails in a known bacteriophage phylogeny. *Evolution* **54**: 397–405.

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. Scr.* **26**: 331–348.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. *Nature* **401**: 877–884.

Qiu, Y.L., Lee, J., Bernasconi-Quadroni, F., 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* **402**: 404–407.

Ricklefs, R.E. & Renner, S.S. 1994. Species richness within families of flowering plants. *Evolution* **48**: 1619–1636.

Ridley, M. 1993. *Evolution*. Blackwell, Boston.

Savolainen, V., Heard, S.B., Powell, M. & Mooers, A.O. 2002. Is cladogenesis heritable?. *Syst. Biol.* **51**: 835–843.

Schluter, D. 2000. *The Ecology of Adaptive Radiations*. Oxford University Press, Oxford.

Simpson, G.G. 1944. *Tempo and Mode in Evolution*. Columbia University Press, New York.

Soltis, D.E., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., Kress, W.J., Nixon, K.C. & Farris, J.S. 2000. Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences. *Bot. J. Linn. Soc.* **133**: 381–461.

Stebbins, G.L. 1974. *Flowering Plants. Evolution Above the Species Level*. The Belknap Press of Harvard University Press, Cambridge.

Tiffney, B.H. & Mazer, S.J. 1995. Angiosperm growth habit, dispersal and diversification reconsidered. *Evol. Ecol.* **9**: 93–117.

Verdú, M. 2002. Age at maturity and diversification in woody angiosperms. *Evolution* **56**: 1352–1361.

Verdú, M. & Traveset, A. 2005. Early emergence enhances plat fitness: a phylogenetically controlled meta-analysis. *Ecology* **86**: 1385–1394.

Vivrette, N.J. 1995. Distribution and ecological significance of seed-embryo types in Mediterranean climates in California, Chile, and Australia. In: *Ecology and Biogeography of Mediterranean Ecosystems in Chile, California and Australia* (M. K. T. Arroyo, P. H. Zedler & M. D. Fox, eds), pp. 274–288. Springer Verlag, New York.

Wang, S.C. 2001. Quantifying passive and driven large-scale evolutionary trends. *Evolution* **55**: 849–858.

Webster, A.J., Payne, R.J.H. & Pagel, M. 2003. Molecular phylogenies link rates of evolution and speciation. *Science* **301**: 478.

Zar, J.H. 1996. *Biostatistical Analysis*. Prentice-Hall, Upper Saddle River.

Received 31 May 2005; revised 17 June 2005; accepted 21 June 2005