Environmental energy and evolutionary rates in flowering plants

T. Jonathan Davies¹, Vincent Savolainen², Mark W. Chase², Justin Moat³ and Timothy G. Barraclough¹,²*

¹Department of Biological Sciences and NERC Centre for Population Biology, Imperial College London, Silwood Park Campus, Ascot SL5 7PY, UK
²Molecular Systematics Section, Jodrell Laboratory, Royal Botanic Gardens, Kew TW9 3DS, UK
³GIS unit, Herbarium, Royal Botanic Gardens, Kew TW9 3AB, UK

The latitudinal gradient in species richness is a pervasive feature of the living world, but its underlying causes remain unclear. We evaluated the hypothesis that environmental energy drives evolutionary rates and thereby diversification in flowering plants. We estimated energy levels across angiosperm family distributions in terms of evapotranspiration, temperature and UV radiation taken from satellite and climate databases. Using the most comprehensive DNA-based phylogenetic tree for angiosperms to date, analysis of 86 sister-family comparisons shows that molecular evolutionary rates have indeed been faster in high-energy regions, but that this is not an intermediate step between energy and diversity. Energy has strong, but independent effects on both species richness and molecular evolutionary rates.

Keywords: environmental energy; evolutionary rates; species richness; flowering plants

1. INTRODUCTION

The latitudinal gradient in species richness is one of the most cited but least understood biological patterns. The trend for decreasing species richness with increasing latitude has been documented for a wide range of taxa in terrestrial and aquatic environments (Currie & Paquin 1987; Cardillo 1999; Macpherson 2002). Although a multitude of hypotheses have been proposed to explain this pattern, there is little consensus upon their relative importance or critical data on the underlying causes (Willig et al. 2003).

One possible explanation is that high levels of environmental energy promote higher species richness nearer the equator (Wright 1983; Rohde 1992; Allen et al. 2002). This idea is supported by observations that energy-rich regions tend to support more species than energy-poor regions (Turner et al. 1988; Currie 1991; Roy et al. 1998; Francis & Currie 2003). Putative mechanisms fall into two broad categories. First, the ‘biomass’ theory states that increased energy availability enables a greater biomass and thereby more individuals to be supported in a given area (Currie 1991; Willig et al. 2003). This assumes that energy-rich environments support more species rather than simply more individuals per species. Second, the ‘faster evolution’ theory states that energy speeds up the rate of evolution and thereby speciation rates, either by reducing generation times or via direct effects on mutation rates (Rohde 1992; Allen et al. 2002). This assumes that micro-evolutionary rates influence speciation rates, which need not be the case: for example, if speciation rates depended primarily on the rate of origin of geographical isolation (Barraclough & Savolainen 2001). Despite widespread interest in the latter theory, there have been no comprehensive tests of its importance (but see Barraclough & Savolainen 2001; Bromham & Cardillo 2003).

We evaluate the faster evolution theory in flowering plants (angiosperms). Angiosperms are a major radiation of recent geological times and are the dominant primary producers of terrestrial ecosystems, completely dependent on solar energy input. Previous work demonstrated a significant correlation between the general rates of molecular evolution at plastid and nuclear loci and species richness in angiosperms (Barraclough et al. 1996; Savolainen & Gouret 1998; Barraclough & Savolainen 2001; see also Webster et al. 2003), consistent with one step in the faster evolution theory but, to our knowledge, no studies have explored the evolutionary consequences of environmental energy in angiosperms. We use sister-group comparisons and measures of energy exposure from contemporary geographical information systems (GIS) data to investigate the links between energy, molecular evolutionary rates and species richness predicted by the faster evolution theory and the competing biomass theory.

Different aspects of environmental energy might influence evolutionary rates versus biomass more strongly in plant taxa. Therefore, we use three alternative measures of environmental energy in our analyses: UV radiation, actual evapotranspiration (AET) and temperature. UV radiation is a known mutagen and has been shown to cause genetic change in a variety of organisms, but it appears to have little impact on overall primary productivity or biomass (Caldwell et al. 1995). AET represents the amount of evapotranspiration permitted by the available soil moisture, which is an index of the potential biomass an area can support. Acting as a general measure of energy, mediated through interactions with other aspects of the environment, it is often strongly correlated with plant...
species richness among areas (Currie & Paquin 1987; Currie 1991; Wylie & Currie 1993). Temperature may have effects on both biomass and rates of molecular evolution (Rohde 1992; Allen et al. 2002), a recent paper proposed that high temperature speeds up biochemical kinetics and thereby rates of evolution and speciation (Allen et al. 2002). Therefore, if the faster evolution theory were correct, we might expect UV and/or temperature to display the strongest relationship with species richness, via an intermediate link with molecular rates.

Records of energy levels and family ranges over evolutionary time are not available, therefore we estimate mean exposure levels from current GIS data. This assumes that contemporary estimates of exposure to environmental energy reflect exposure levels experienced by plant lineages over evolutionary time. The time since the divergence between sister families is of the order of tens of millions of years whereas major climatic shifts have occurred in the order of every 10 000–100 000 years. However, range expansion and contraction as a response to environmental change is well documented (Huntley & Webb 1989), in which case environmental conditions experienced by each lineage might be relatively constant over time. If contemporary measures did not provide evolutionarily meaningful measures of energy load, this would most probably confound attempts to detect true relationships, particularly those with evolutionary variables such as rate of molecular evolution; thus our analyses are conservative. We discuss the implications of this assumption further in § 4.

2. MATERIAL AND METHODS

(a) Estimation of molecular rates

All pairs of terminal sister taxa were identified from the most comprehensive DNA-based phylogenetic tree of angiosperms to date (Soltis et al. 1999, 2000), yielding a total of 86 sister-family comparisons. Sister-family comparisons are evolutionarily independent and by relying on terminal clades we avoid difficulties of reconstructing ancestral values of study variables for much older nodes deep in the phylogenetic tree (Barracough et al. 1998). The DNA sequence matrix was trimmed to one representative taxon per family before calculating branch lengths to remove possible bias due to the node density effect (Fitch & Beintema 1990) and divided into the following partitions: nuclear ribosomal gene (18S rDNA), second position sites for the plastid protein-coding genes (rbcL and atpB, changes that predominantly leave the amino acid sequence unaffected), and third position sites for the plastid protein-coding genes (changes that predominantly leave the amino acid sequence unaffected). For each partition the maximum-likelihood branch length leading to each sister clade was then estimated in PAUP* 4.0b10 (Swofford 2001) using the HKY85 model of DNA evolution with a gamma distribution to account for heterogeneity among sites.

Because sister clades are the same age, the branch length contrast represents the relative rate of molecular evolution of the respective families. If environmental energy were driving rates of molecular evolution via effects on generation time or mutation rate, any effect should be most apparent for selectively neutral changes. Therefore, we present the results using substitution rate contrasts estimated from third position sites of the coding genes, sites at which the majority of substitutions do not affect amino acid sequence. Results for other partitions are presented in electronic Appendix A.

(b) Measurement of environmental energy

Global datasets for the environmental variables were obtained from a range of online databases selected for comprehensive cover over the longest time-period (see electronic Appendix A). A mean was calculated for each cell reference for the complete time-period represented in the separate datasets. Maps of the distribution for the sister families were obtained from Heywood (1993); distributions for families not included in this data source, or those in which major taxonomic revisions had occurred, were obtained from herbarium records at the Royal Botanic Gardens, Kew. Distribution maps were digitized by hand directly into ArcView (GIS 3.2, Environmental Systems Research Institute Inc.) as polygon themes and are available from the authors. For each separate global land coverage (temperature, UV, AET, elevation, latitude (the latter measured as distance from the equator and hence always positive)) the sum of the cell values, multiplied by area, contained within each distribution map was calculated. This was then divided by the sum of the area coverage, yielding a figure of mean exposure per unit area.

(c) Construction of linear models

Independent contrasts were calculated for all study variables as follows. For each sister pair, A and B, we calculated the species richness contrast as:

$$\log(\text{number of species in A} \, - \, \log(\text{number of species in B})).$$

This measure should reflect relative diversification rates and therefore be independent of the age of the split between the sister families (Isaac et al. 2003), but for our data there was a residual relationship between the magnitude of this contrast and the age of split. To standardize the variance among contrasts, we divided each contrast by the age of the split between the sister families estimated from the gene sequence data as described in Wikström et al. (2001). Other ways of including node age in the models led to similar results, but division performed better in model criticism (see electronic Appendix A). Contrasts in environmental variables and molecular branch lengths leading to each sister family were calculated as $X_A - X_B$, where $X$ is either the mean coverage of the environmental variable or molecular branch length. Because absolute molecular branch lengths are greater for older nodes, the branch length contrasts were standardized by dividing by the mean of the branch lengths leading to both sister families. The geographical area of each family was cube-root-transformed before calculation of contrasts because this transformation proved best to linearize the relationship between family values of species richness and area.

Contrasts of the different environmental variables were correlated with one another, as would be expected since all are linked to input solar radiation ($r^2$ for pairwise comparisons among the energy measures ranged from 0.40 to 0.89). However, there appeared to be sufficient unique variation for some potential to distinguish the importance of different measures of energy. To check that our conclusions were not affected by problems of collinearity among the energy variables, we repeated the analyses including just one of the energy variables in turn.

We used least-squares regression through the origin (Harvey & Pagel 1991) to explore the relationships between species richness, environmental energy, and rates of molecular evolution, implemented in the statistical package R (R Development Core Team 2004). In addition to the direct measures of environmental energy, we included the size of the area occupied by each family and two indirect measures of environmental energy (mean latitude and elevation). To account for the possible effects of
Table 1. Multiple regressions between species richness, molecular rates and various combinations of explanatory variables. (Area was included in all starting models: energy, both direct and indirect measures of environmental energy; SR, species richness; MR, molecular substitution rate estimated from third position sites of the protein coding genes; temp, temperature; lat, latitude. Significant explanatory variables retained following model simplification are shown. All models are significant with $p < 0.001$.)

<table>
<thead>
<tr>
<th>model</th>
<th>response variable</th>
<th>explanatory variables</th>
<th>$r^2$ coefficients</th>
<th>$r^2$</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SR</td>
<td>energy</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>temp</td>
<td>0.19</td>
<td>6.55</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>area</td>
<td>0.54</td>
<td>11.14</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>MR</td>
<td>energy</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>lat</td>
<td>0.15</td>
<td>−3.90</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
</tbody>
</table>

We constructed a set of models to explore the sensitivity of our results to the presence/absence of different variables, inclusion of interaction terms, different transformations of the data and to assess the relative importance of different environmental variables.

### 3. RESULTS

Species richness correlates with environmental energy, supporting the broad predictions of the species–energy theory (table 1, model 1). Temperature emerges as the most predictive of the energy measures, explaining an additional 19% of the variation in species richness compared to area alone (species richness against area: $r^2 = 0.44$, $p < 0.001$). The other measures of environmental energy correlate almost as strongly with species richness (species richness against AET and area, $r^2 = 0.59$; latitude and area, $r^2 = 0.59$; UV and area, $r^2 = 0.58$; elevation and area, $r^2 = 0.54$; all $p < 0.001$), but they dropped out during model simplification. The same model was obtained when pairwise interaction terms were included in the starting model.

The first step of the faster evolution theory is supported by our analyses. The rate of molecular evolution correlates significantly with environmental energy, with latitude emerging as the single most important main effect (table 1, model 2). The other measures of environmental energy also correlate with substitution rate, but although UV and temperature perform almost as well as latitude, the relationship with AET was not significant (substitution rate against AET, $r^2 = 0.09$, $p = 0.09$; temperature, $r^2 = 0.12$, $p < 0.001$; UV and area, $r^2 = 0.12$, $p < 0.001$). When we included interactions in the model, the explanatory power of the model greatly improved through retention of multiple interaction terms, but omitting each environmental variable in turn from the starting model indicated that UV was the most important underlying variable (electronic Appendix A, tables 2 and 3).

Confirming previous analyses by Barraclough & Savolainen (2001) and others, the second step of the faster evolution theory is also supported. Species richness correlates significantly with substitution rate ($r^2 = 0.08$, $p = 0.004$). However, we find no evidence for the faster evolution hypothesis: that the relationship between energy

---

and species richness is mediated by evolutionary rates. When substitution rate and energy variables were added simultaneously as explanatory variables for species richness, substitution rate dropped out of the simplified model (again yielding model 1; see table 1). This is the outcome predicted if the main effect of energy on species richness is direct, rather than via an intermediate effect on molecular rates (figure 1). The relationship between species richness and substitution rate appears to be an artefact of both variables being correlated with energy measures.

To check the generality and possible mechanism of these findings, we repeated all analyses for substitution rates derived from the entire DNA data, second positions (sites at which most substitutions lead to amino acid changes) and 18S rDNA in turn. Substitution rates of the different DNA partitions were found to correlate with one another and with species richness by Barraclough & Savolainen (2001). In all cases, substitution rates correlated strongly with measures of environmental energy (electronic Appendix A, table 4). In no cases did substitution rates displace the energy variables as explanatory variables of species richness. The only partition retained in the models was 18S rDNA, but this displayed a negative relationship with species richness (opposite to the predictions of the faster evolution hypothesis) and was not robust to the sensitivity analyses (see electronic Appendix A).

4. DISCUSSION

Our analyses investigated the relationship between species richness and energy in flowering plants and its possible evolutionary causes. Area explained most variation in species richness regardless of the other parameters within the models. Geographical range may put an upper limit on diversification of a clade (Owens et al. 1999; Ricklefs 2003), whereas a minimum range size may be required for speciation to occur (Rosenzweig 1992; Losos & Schluter 2000). After controlling for area, all measures of environmental energy were strong predictors of species richness, providing broad support for species–energy theory and confirming recent macroecological studies at large spatial scales (Francis & Currie 2003). We then performed a series of tests to distinguish the two main theories for explaining species–energy relationships: the faster evolution and biomass theories.

The faster evolution hypothesis proposes that environmental energy speeds up evolutionary rates and thereby speeds up speciation rates. Our results demonstrate a relationship between molecular rates and environmental energy, the first step in this process. Until now a link between energy and molecular rates has not been demonstrated in any group (Bromham & Cardillo 2003). The result reflects general rates of molecular evolution, in both nuclear and plastid regions and protein-coding and ribosomal genes. Because both synonymous and nonsynonymous substitutions are affected, the underlying mechanisms must affect both neutral and functional changes. Shorter generation times and/or faster mutation rates in high-energy environments would provide a simple, general explanation for the observed patterns (Rohde 1992; Barraclough & Savolainen 2001).

The direct measures of environmental energy that best explained mutation rates were UV and temperature. The mutagenic effects of UV radiation are well documented and have been implicated in influencing mutation rates among lineages of marine protists (Pawlowski et al. 1997). Temperature might affect mutation rates, either through decreasing development times or increasing metabolic rate and the production of DNA-damaging metabolites (Allen et al. 2002). However, latitude was identified as the single most important predictor and, therefore, the exact relationship between the different components of environmental energy and molecular rates remains unclear. Further analysis at finer spatial and taxonomic scales might provide greater resolution required to differentiate between alternate measures.

Our results also confirm previous studies that species richness correlates with molecular evolutionary rates in angiosperms, consistent with the second step of the faster evolution theory. However, despite finding evidence for both faster molecular evolution and higher species richness in higher energy environments, there is no evidence that faster evolution explains the relationship between energy and species richness. The main effects of environmental energy on both species richness and molecular rates are direct (figure 1). The relationship between species richness and molecular rates, needed for the faster evolution hypothesis, is lost or extremely weak when environmental variables are added to the model.

The faster evolution hypothesis relies on the effects of environmental energy on generation time, mutation, or other factors affecting general evolutionary rates; therefore, we would have detected any strong effect with our measures of molecular rates. That such an effect is lacking leads us to reject the faster evolution hypothesis for flowering.
plants. The direct link between species richness and environmental energy is more consistent with the biomass theory: either speciation rates are faster or extinction rates are lower in regions supporting greater biomass. Future availability of species-level trees for a representative set of plant clades might allow investigation of the roles of speciation and extinction in generating diversity patterns (Losos & Schluter 2000; Barraclough & Nee 2001), but until then we cannot distinguish their relative importance.

A critical assumption of our explanations is that contemporary energy levels across family ranges reflect conditions experienced over evolutionary time-scales, which from knowledge of past major climate changes and extensive range movements might seem unlikely. However, we can think of no artefact that would cause a strong relationship between environmental measures and molecular rates if current and past conditions were entirely incongruent. Tracking of the environment by plant taxa is well documented and could provide broad constancy of environment even in the face of environmental changes (Webb 1986; Huntley & Webb 1989). Many studies have assumed environmental constancy of plants as well as foraminifera and other organisms to reconstruct past climates but without means to test the assumption (see Huntley 2001). Evidence for correlated range dynamics between disjunct sister taxa indicate a high degree of ecological niche conservatism over a time-scale of up to tens of millions of years (Ricklefs & Latham 1992; Peterson et al. 1999; Qian & Ricklefs 2004). Unfortunately, data are not available to evaluate this across the taxonomical and geographical breadth encompassed by the present study. If the assumption of broad environmental constancy proved incorrect, so would our explanations, but the result that species richness and molecular rates correlate with contemporary environment experienced by plant families would remain.

5. CONCLUSIONS
Species richness and molecular rates have both been proposed to depend on many interacting factors. Generation time, ecological factors affecting population size, and other biological attributes might all affect species richness and substitution rates, as well as environmental factors unrelated to energy, such as disturbance, geological complexity and biogeographic history (Brown & Lovomino 1998; de Queiroz 2002; Bromham & Cardillo 2003; Sims & McConway 2003). Amid this large number of putative factors, we have shown that environmental energy is a key variable independently explaining variation in both molecular evolutionary rates and species richness in flowering plants.

The authors thank Mick Crawley, Charles Godfray, Andy Purvis, Marcel Cardillo and two anonymous referees for comments and help. This work was supported by the Natural Environment Research Council, UK, the Royal Botanic Gardens, Kew, and a Royal Society University Research Fellowship to T.G.B.

REFERENCES


Rosenzweig, M. L. 1992 Species diversity gradients: we know more and less than we thought. J. Mamm. 73, 715–730.


Visit www.journals.royalsoc.ac.uk and navigate through to this article in Proceedings: Biological Sciences to see the accompanying electronic appendix.