SUPPLEMENTAL MATERIALS

Character data, character scoring, and sampling

Breeding system and cone type were primarily scored using compilation literature sources (Enright and Hill, 1995; Kelch, 1997; Farjon, 2005; Eckenwalder, 2009; Farjon, 2010) as well as additional information for species in some variable taxa (Barker & Kirkpatrick, 1994; Flores-Rentería et al., 2013). However, there is significant uncertainty in sources regarding the breeding system of some genera, particularly *Acmopyle*, *Araucaria*, and *Nageia*. In all, these taxa constitute a very small fraction of the dataset, but we tested the potential effect of these differences by fitting our best-supported diversification model (see below) to three sets of character scorings. Character Scoring 1 (whose results are reported in the main text and predominantly reported here) represents the coding of character states as reported in the current literature. Character Scoring 2 includes several ambiguous taxa (*Acmopyle* and *Nageia* species) scored as having the fleshy-monoecy character combination (Tim Brodribb, pers.comm.). Character Scoring 3 includes New Caledonian *Araucaria* as having the dry-dioecy combination (Patrick Knopf, pers. comm.). Parameter estimates using the different character scorings were generally similar in absolute and relative values (Figure S1*a*), at least within error, but using Character Scoring 2 (that is, including more species with the fleshy-monoecious state) did lower the inferred transition rate between fleshy-monoecy and fleshy-dioecy. This rate was still significantly higher than those of the other states (transition number four in Figure S1*b*), which is consistent with the other character scorings (see main text, Figure 1*c*).

Although not all extant conifer species are sampled in our data set, the relative frequencies of the various character state combinations is similar to the overall distribution of traits. For example, using estimates of total species numbers from the literature sources mentioned above, dry-monoecy occurs in 56% of all conifer species and fleshy-dioecy in 40%. In our dataset, these character combinations occur in 59% and 36% of the species respectively. We therefore expect little bias in our results due to sampling.

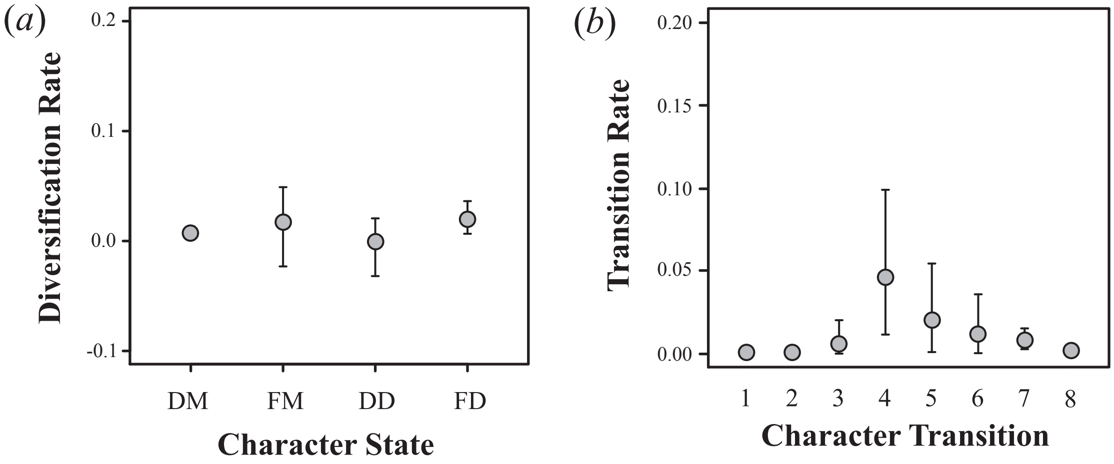

Figure S1. Diversification and transition rate estimates based on the best-fit MuSSE model using the second character scoring set. (*a*) Diversification rates for the character combinations dry-monoecy (DM), fleshy-monoecy (FM), dry-dioecy (DD), and fleshy-dioecy (FD). (*b*) Transition rates among character states, with numbers corresponding to those shown in figure 1*c* in the main text. Error bars were based on the posterior distribution of parameter estimates (see below).

Maximum likelihood analyses

In order to reconstruct cone type and breeding system evolution, we fit a variety of different maximum likelihood models to our data. These can be broadly divided into two classes of models, one in which all transition rates between states were governed by a single rate (equal rates, ER) and one in which transition rates were free to vary (all rates different, ARD). Within this context, we first tested models in which individual cone type and breeding system characters evolve independently (character independent models). In this case, the best-supported model for both traits was an ER model (Table S1).

Table S1. Model fits for breeding system and cone type considered individually. In both cases, the best model based on Akaike Information Criterion (AIC) was an equal rates model (in bold).

Model	**Parameters**	**-ln*L***	**AIC**	**ΔAIC**
Breeding system				
Equal rates	**1**	**-92.0**	**186.1**	**0**
All rates different	2	-91.7	187.4	1.3
Cone type				
Equal rates	**1**	**-26.2**	**54.4**	**0**
All rates different	2	-26.2	56.0	1.6

We next analyzed combined character states. Here, we tested five models of character evolution that differed in which transitions were allowed (figure S2). In the most unconstrained model ("Anything Goes"), transitions between all possible combinations may occur, including simultaneous changes in two traits (e.g., the transition from dry-monoecy to fleshy-dioecy). In the "Free" model, these simultaneous transitions were not allowed. We also tested three different sequences of character evolution in which the fleshy-dioecious state specifically evolves through a fleshy-monoecious intermediate state, as has been postulated in previous work (Givnish, 1980; Bawa, 1982). We refer to these models as "Path 1", "Path 2", and "Path 3" respectively, and they vary in the degree of irreversibility once a lineage transitions away from the dry-monecious character state (see Figure S2).

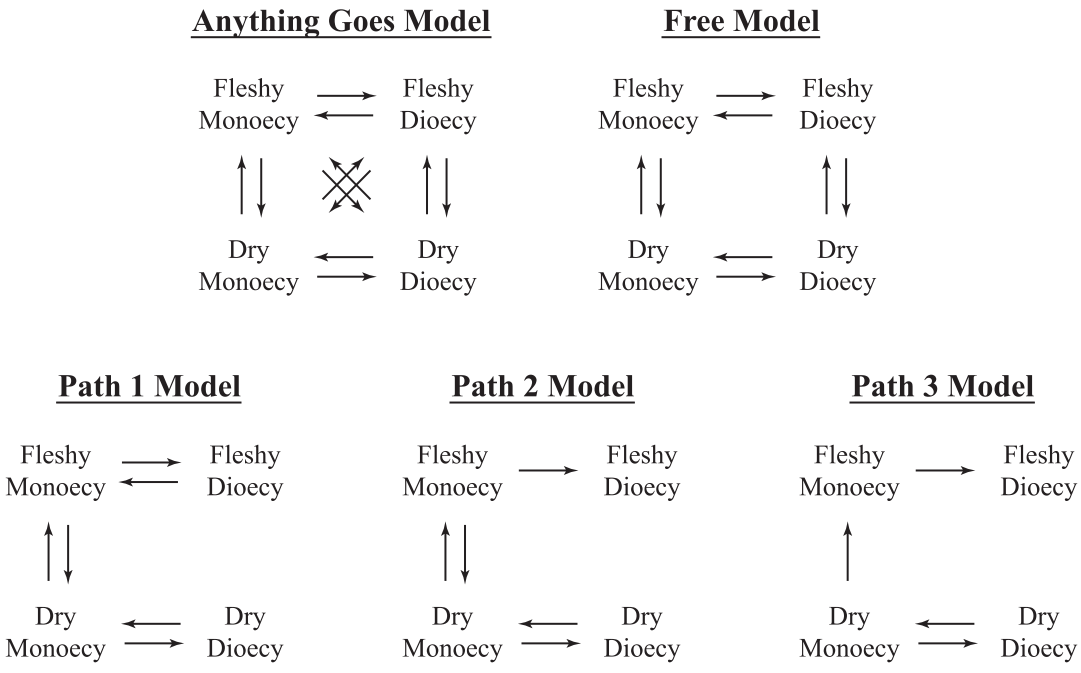

Figure S2. Schematic illustrations of the combined character state models used in this study. Arrows between states indicate transitions that are allowed under each model.

We first tested whether these combined character state models provided a better fit to the data than simply analyzing cone type or breeding system individually. In this case, the likelihood, *L*, of a given combined character state model is proportional to the probability of the data given a single model of evolution, **Q**, that defines the transitions among all possible character combinations. By contrast, the likelihood of an individual character model is defined as the product of the separate probabilities of observing the data given distinct **Q** matrices applied to each character trait separately (see Pagel, 1994).

The best-supported model, based on the Akaike Information Criterion (AIC), was a “Free” model with different transition rates between the various combined character states (Table S2). This model does not allow simultaneous trait transitions; in fact, the “Anything Goes” model that allowed them returned nonsensical likelihood values. Models specifying a particular path to the fleshy-dioecious state (“Path 1-3” models) were not generally well supported, although “Path 1” was the second best-supported model overall (Table S2). The results reported here were based on data from Character Scoring 1, but as noted above, parameter estimates were consistent among the three different character scorings. All analyses were carried out using the R package *corHMM* (Beaulieu et al., 2013).

Table S2. The fit of alternative models of breeding system and cone type evolution. The best fit model based on the Akaike Information Criterion (AIC) is shown in bold. ER=equal rates, ARD = all rates different. In the "Path 1-3" models, the numbers correspond to character states: 1= dry-monoecy, 2=fleshy-monoecy, 3=dry-dioecy, 4=fleshy-dioecy.

Model	parameters	-ln*L*	AIC	ΔAIC
Individual Character State				
Breeding, Cone = ER	2	-118.3	240.5	36.5
Breeding = ER; Cone = ARD	3	-118.0	242.0	38.0
Breeding = ARD; Cone = ER	3	-117.9	241.8	37.7
Breeding, Cone = ARD	4	-117.7	243.3	39.3
Combined Character State				
ER	1	-122.3	246.7	42.6
ARD Anything Goes	12	N/A	N/A	N/A
ARD Free	**8**	**-94.0**	**204.0**	**0**
ARD Path 1: $q_{3\rightarrow4}=q_{4\rightarrow3}=0$	6	-98.9	210.0	5
ARD Path 2: $q_{3\rightarrow4}=q_{4\rightarrow3}=q_{4\rightarrow2}=0$	5	-114.0	237.9	33.9
ARD Path 3: $q_{3\rightarrow4}=q_{4\rightarrow3}=q_{4\rightarrow2}=q_{2\rightarrow1}=0$	4	-114.0	235.9	31.9

Table S3. Parameter estimates ($\hat{q}$) for character state transitions obtained from the best-fit model ("ARD Free") with their approximate standard errors (SE). Mono = monoecy.

Transition	$\hat{q}$	SE (±)
Mono+Dry → Mono+Fleshy	0.00021	0.00022
Mono+Dry → Dioecy+Dry	0.00121	0.00048
Mono+Fleshy → Mono+Dry	0.00000	<0.0001
Mono+Fleshy →Dioecy+Fleshy	0.10762	0.03870
Dioecy+Dry → Mono+Dry	0.00216	0.00323
Dioecy+Dry → Dioecy+Fleshy	0.01069	0.00505
Dioecy+Fleshy → Mono+Fleshy	0.00497	0.00322
Dioecy+Fleshy → Dioecy+Dry	0.00024	0.00028

Diversification Analyses

The distribution of observed character states may not be due solely to differences in transition rate, but could also reflect differences in character dependent diversification. We therefore used BiSSE (Binary State Speciation and Extinction; Maddison et al., 2007) to examine diversification in relation to character states within in breeding system and cone type. For each character, we tested a model where net diversification rates for each state (monoecy/dioecy and dry/fleshy respectively) and transition rates between the states were free to vary, one in which transition rates were constrained to be the same, and one in which net diversification rates were constrained to be the same.

We used a Bayesian implementation of BiSSE (see Materials and Methods in main text) and computed the Deviance Information Criterion (DIC; Spiegelhalter et al., 2002) for each model to assess fit. The DIC is a Bayesian analog of Akaike's Information Criterion (AIC, Akaike, 1974) in that it assesses how well the model fits the data while penalizing for the number of parameters. More specifically, it is calculated as

$$DIC = \hat{d} - 2(\hat{d} - \bar{d})$$

where $\hat{d}$ is the deviance, or -2ln(L), calculated from likelihood of the means of the posterior sample of parameter estimates, and where $\bar{d}$ is the mean deviance calculated from likelihoods of the posterior sample. The second term in the calculation represents the penalty for the number of parameters in the focal model. As with AIC, the model with the smallest DIC value is estimated to be the model that would best predict a replicate dataset with the same structure as that which is currently observed.

For breeding system and cone type, the best-fit model was one in which diversification rates were the same for different character states (Table S4). This suggests that neither dioecy nor fleshy cones or seeds are associated with significantly higher net diversification rates in conifers.

Table S4. Comparisons of the BiSSE models. All models were compared with the Deviance Information Criterion (DIC), a Bayesian analog of AIC that assesses model fit while penalizing for the number of parameters (see text above). For breeding system and cone type, the best-fit model, shown in bold, was one in which net diversification rates (*r*; calculated as speciation rate [λ] – extinction rate [μ]) did not differ between the alternative character states (monecy/dioecy and dry/fleshy, respectively). For breeding system, state 0 = monoecy, state 1 = dioecy; for cone type, state 0 = dry cones, state 1 = fleshy seeds or cones. *r* = net diversification rate and q = transition rate between character states.

Model	**Parameters**	**DIC**	r_0	r_1	$q_{0\rightarrow 1}$	$q_{1\rightarrow 0}$
Breeding System						
Free	6	3780.3	0.0074	0.0136	0.0015	0.0044
$q_{0\rightarrow 1}=q_{1\rightarrow 0}$	5	3782.0	0.0087	0.0097	0.0022	0.0022
$\lambda_0=\lambda_1$; $\mu_0=\mu_1$	**4**	**3778.9**	**0.0081**	**0.0081**	**0.0022**	**0.0031**
Cone type						
Free	6	3651.1	0.0066	0.0156	0.0005	0.0014
$q_{0\rightarrow 1}=q_{1\rightarrow 0}$	5	3650.6	0.0069	0.0146	0.0006	0.0006
$\lambda_0=\lambda_1$; $\mu_0=\mu_1$	**4**	**3646.0**	**0.0080**	**0.0080**	**0.0006**	**0.0013**

It is also possible, however, that conifer clades show higher net diversification rates when in the dry-monoecy and fleshy-dioecy character combinations. We therefore analyzed our data using several different Multistate Speciation and Extinction models (MuSSE, Fitzjohn, 2012), which estimate character state-dependent speciation, extinction, and transition rates among the four possible breeding system and cone type character combinations. For the MuSSE analyses, we fit our previous models of combined character state evolution (see Figure S2). With MuSSE, however, each combined character state also has associated speciation (λ) and extinction (μ) rates. The most complex model, the MuSSE version of "Anything Goes", therefore contained four speciation rates, four extinction rates, and 12 transition rates between all possible character states. For each model we tested, we also fit versions in which net diversification rates (λ – μ) were constrained to be the same for some or all of the character states.

As for the BiSSE analyses, we used a Bayesian implementation of MuSSE to obtain a posterior distribution for the parameters estimated under each model. In this case, we chose a Bayesian approach over maximum likelihood because of the difficult in inferring precise parameter values for those character state combinations that contain few species (i.e., fleshy-monoecy and dry-dioecy). This approach also allowed us to examine the uncertainty associated with each parameter in any given model. For our analyses, we used a prior that followed an exponential distribution with a mean of 2*r*, where *r* is the net diversification rate that would produce our tree with 489 taxa (see FitzJohn et al., 2009).

The best model, based on DIC, was the MuSSE version of the "Free" model with different transition rates (Figure S2; Table S4). In this model, the posterior means of the diversification rates suggested some differences in net diversification among the character state combinations. There was, however, substantial overlap in the credibility intervals, implying that these differences were not significant. Interestingly, the next best model, with a DIC difference of ~2, was a "Free" model with equal diversification rates, that is, it assumed a single rate governing the diversification of all four character combinations (Table S5). These two models are consistent, however, in that they both estimate generally low rates for transition away from the inferred likeliest ancestral state of dry-monoecy (Tables S6, S7). Overall, MuSSE models indicate that transitions out of the dry-monoecious state have likely occurred very few times, but when lineages do transition away they are more likely to shift to the fleshy-dioecious state or back to dry-monoecious state. For instance, the two highest transition rates were inferred to be from fleshy-monoecy to fleshy-dioecy, followed by dry-dioecy to dry-monoecy (Table S6). MuSSE results are also generally consistent with transition rates estimated from the best-supported likelihood model (Table S3), which is not surprising given that net diversification rate does not appear to play a major role in influencing the evolution of these character states.

Finally, we used parameter estimates from the best-supported MuSSE model (the “Free” model) to calculate persistence times for each character state. This is the MuSSE analog to the persistence times presented in the main text (see Figure 1*e*), which were based on stochastic maps using transition rates from the best-supported likelihood model (also the “Free” model). We calculated the posterior distribution of persistence times for each of the combined character states as the inverse of the sum of the posterior sample of transition rates away from that state. As with the stochastic character mapping, MuSSE suggests that the preponderance of the dry-monoecious and fleshy-dioecious states results from their persistence once they evolve (Figure S3).

Table S5. Comparisons of the MuSSE (Multistate Speciation and Extinction) models. All models correspond to those presented in Figure S2, but also include speciation (λ) and extinction (μ) rates for each combined character state. For each model, we also tested versions where net diversification rates ($\lambda - \mu$) for some or all characters were constrained to be the same. All models were compared with the Deviance Information Criterion (DIC), and the best-fit model is shown in bold. Numbers correspond to each character state combination: 1=Dry-Monoecy; 2=Fleshy-Monoecy; 3=Dry-Dioecy; 4=Fleshy-Dioecy.

Model	**Parameters**	**DIC**	**ΔDIC**
Path 1: $q_{3\rightarrow 4}=q_{4\rightarrow 3}=0$			
No constraints	14	3799.2	10.2
$\lambda_2=\lambda_3$; $\mu_2=\mu_3$	12	3811.1	22.1
$\lambda_1=\lambda_4$, $\lambda_2=\lambda_3$; $\mu_1=\mu_4$, $\mu_2=\mu_3$	10	3806.3	17.3
$\lambda_1=\lambda_2=\lambda_3=\lambda_4$; $\mu_1=\mu_2=\mu_3=\mu_4$	8	3801.1	12.1
Path 2: $q_{3\rightarrow 4}=q_{4\rightarrow 3}=q_{4\rightarrow 2}=0$			
No constraints	13	3811.9	22.9
$\lambda_2=\lambda_3$; $\mu_2=\mu_3$	11	3826.7	37.7
$\lambda_1=\lambda_4$, $\lambda_2=\lambda_3$; $\mu_1=\mu_4$, $\mu_2=\mu_3$	9	3820.7	31.7
$\lambda_1=\lambda_2=\lambda_3=\lambda_4$; $\mu_1=\mu_2=\mu_3=\mu_4$	7	3829.1	40.1
Path 3: $q_{3\rightarrow 4}=q_{4\rightarrow 3}=q_{4\rightarrow 2}=q_{2\rightarrow 1}=0$			
No constraints	12	3812.8	23.8
$\lambda_2=\lambda_3$; $\mu_2=\mu_3$	10	3829.1	40.1
$\lambda_1=\lambda_4$, $\lambda_2=\lambda_3$; $\mu_1=\mu_4$, $\mu_2=\mu_3$	8	3824.8	35.8
$\lambda_1=\lambda_2=\lambda_3=\lambda_4$; $\mu_1=\mu_2=\mu_3=\mu_4$	6	3827.5	38.5
Free			
No constraints	**16**	**3789.0**	**0.0**
$\lambda_2=\lambda_3$; $\mu_2=\mu_3$	14	3802.0	13.0
$\lambda_1=\lambda_4$, $\lambda_2=\lambda_3$; $\mu_1=\mu_4$, $\mu_2=\mu_3$	12	3799.1	10.1
$\lambda_1=\lambda_2=\lambda_3=\lambda_4$; $\mu_1=\mu_2=\mu_3=\mu_4$	10	3791.1	2.1
Anything Goes			
No constraints	20	3793.8	4.8
$\lambda_2=\lambda_3$; $\mu_2=\mu_3$	18	3804.2	15.2
$\lambda_1=\lambda_4$, $\lambda_2=\lambda_3$; $\mu_1=\mu_4$, $\mu_2=\mu_3$	16	3801.3	12.3
$\lambda_1=\lambda_2=\lambda_3=\lambda_4$; $\mu_1=\mu_2=\mu_3=\mu_4$	14	3794.7	5.7

Table S6. Posterior means and associated 95% highest probability densities (HPD) for the parameters in the best-fit MuSSE model (the “Free” model with different transition rates). *r* represents the net diversification rate (speciation – extinction) for each character state. Mono = monoecy.

Parameter	Posterior Mean	95% HPD
r: Mono+Dry	0.0075	(0.0015, 0.0139)
r: Mono+Fleshy	0.0628	(-0.0010, 0.1162)
r: Dioecy+Dry	-0.0034	(-0.0350, 0.0187)
r: Dioecy+Fleshy	0.0145	(0.0030, 0.0266)
Mono+Dry → Mono+Fleshy	0.0003	(<0.0001, 0.0007)
Mono+Dry → Dioecy+Dry	0.0010	(0.0003, 0.0024)
Mono+Fleshy → Mono+Dry	0.0056	(0.0002, 0.0218)
Mono+Fleshy → Dioecy+Fleshy	0.0873	(0.0355, 0.1410)
Dioecy+Dry → Mono+Dry	0.0183	(0.0015, 0.0456)
Dioecy+Dry → Dioecy+Fleshy	0.0055	(0.0001, 0.0202)
Dioecy+Fleshy → Mono+Fleshy	0.0028	(0.0006, 0.0074)
Dioecy+Fleshy → Dioecy+Dry	0.0010	(0.0002, 0.0026)

Table S7. Posterior means and associated 95% highest probability densities (HPD) for the parameters in the second best-fit MuSSE model (the “Free” model with equal diversification rates). *r* represents the net diversification rate (speciation – extinction) for each character state. Mono = monoecy.

Parameter	Posterior mean	95% HPD
r : all	0.0080	(0.0030, 0.0131)
Mono+Dry → Mono+Fleshy	0.0005	(<0.0001, 0.0014)
Mono+Dry → Dioecy+Dry	0.0016	(0.0006, 0.0033)
Mono+Fleshy → Mono+Dry	0.0051	(0.0193, 0.0001)
Mono+Fleshy → Dioecy+Fleshy	0.0632	(0.0255, 0.1112)
Dioecy+Dry → Mono+Dry	0.0119	(0.0006, 0.0342)
Dioecy+Dry → Dioecy+Fleshy	0.0071	(0.0002, 0.0225)
Dioecy+Fleshy → Mono+Fleshy	0.0042	(0.0012, 0.0094)
Dioecy+Fleshy → Dioecy+Dry	0.0011	(0.0001, 0.0028)

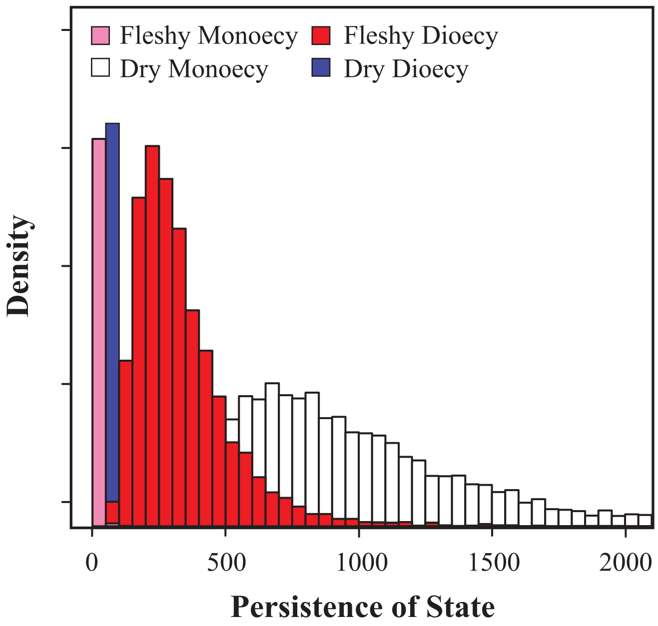

Figure S3. Persistence time for the four character state combinations. Persistence times are posterior distributions based on the inverse of transitions rates calculated from the best-fit MuSSE model (the "Free" model with different diversification rates).

References

Akaike, H. 1974 A new look at the statistical model identification. *IEEE T. Automat. Contr.* **19**, 716-723.

Barker, P.C.J. & Kirkpatrick, J.B. 1994 *Phyllocladus aspleniifolius*: variability in the population structure, the regeneration niche and dispersion patterns in Tasmanian forests. *Aust. J. Bot.* **42**, 163-190.

Bawa, K.S. 1982 Outcrossing and the incidence of dioecism in island floras. *Am. Nat.* **119**, 866-871.

Beaulieu, J.M., O'Meara, B.C. & Donoghue, M.J. 2013 Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Syst. Biol.* **62,** 725-737.

Eckenwalder, J.E. 2009 *Conifers of the World.* London, UK: Timber Press.

Enright, N.J. & Hill, R.S. 1995 *Ecology of the Southern Hemisphere Conifers.* Washington, DC, USA: Smithsonian Institution Press

Farjon, A. 2005 *A monograph of Cupressaceae and Sciadopitys*. Richmond, UK: Royal Botanic Gardens, Kew.

Farjon, A. 2010 *A Handbook of the World's Conifers*. Leiden, Netherlands: Brill Academic Publishers

Fitzjohn, R.G. 2012 Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol. and Evol.* **6**, 1084-1092.

FitzJohn, R.G., W.P. Maddison, & S.P. Otto. 2009 Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst. Biol.* **58**, 595-611.

Flores-Rentería, L., Molina-Freaner, F., Whipple, A.V., Gehring, C.A., Domínguez, C.A. 2013 Sexual stability in the nearly dioecious *Pinus johannis* (Pinaceae). *Am. J. Bot.* **100**, 602-612.

Kelch, D.G. 1997 The phylogeny of the Podocarpaceae based on morphological evidence. *Syst. Botany* **22**, 113-131.

Maddison, W.P., Midford, P.E. & Otto, S.P. 2007 Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* **56**, 701-710.

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.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. &van der Linde, A. 2002 Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B* **64**, 583-639.