

# SCIENTIFIC REPORTS



OPEN

## Testing the impact of effective population size on speciation rates – a negative correlation or lack thereof in lichenized fungi

Jen-Pan Huang<sup>1</sup>, Steven D. Leavitt<sup>2</sup> & H. Thorsten Lumbsch<sup>1</sup>

The effect of microevolutionary processes on macroevolutionary patterns, and *vice versa*, is an important but under-investigated question. Here we present an integrative population genetic and phylogenetic study using molecular sequence data from three lichen-forming fungal lineages to empirically test the potential correlation between effective population size – approximated by the parameter  $\theta$  – and estimated speciation rates using a phylogenetic tree ( $\lambda$ ). A negative association between  $\theta$  and  $\lambda$  was supported in one lineage of lichen-forming fungi, *Melanelixia* (Parmeliaceae), while no significant relationships was found for two other genera within the same family, *Melanohalea* and *Xanthoparmelia*. We discuss the significance of our results and the importance of considering microevolutionary processes when studying macroevolutionary patterns.

Whether macroevolutionary patterns can be explained by microevolutionary processes, and *vice versa*, is a critical, unsettled question in evolutionary biology<sup>1–4</sup>. Many studies have been devoted to bridge microevolutionary and macroevolutionary patterns<sup>2,3</sup>. For example, speciation has been hypothesized to be a driving force accelerating molecular evolution<sup>4</sup> (but see<sup>5</sup>), and a positive correlation between speciation rate and the rate of molecular evolution between lineages have been reported at both adaptive and neutral genomic regions<sup>2</sup>. In comparison to investigations focusing on the correlation between the rate of speciation and molecular evolution among species, the impact of effective population size<sup>1</sup> has seldom been empirically investigated within a macroevolutionary framework<sup>6,7</sup>. However, the potential effect of differences in effective populations sizes on variation in speciation rates across evolutionary lineages has been commonly discussed<sup>6,8</sup>.

Genetically structured populations can be in initial phases of becoming different species, and the rate to generate structured populations, which can be affected by size of the populations, may relate to the rate of speciation<sup>9,10</sup>. Predictions for positive, negative, and no correlation between effective population size and speciation rate have been proposed. For example, small effective population sizes may facilitate rapid genetic differentiation between subdividing populations<sup>1</sup>. However, based on the neutral theory, the only demographic parameter that may affect the rate of organismal diversification is generation time<sup>11</sup>, if speciation results from the accumulation of fixed mutations between diverging taxa. Furthermore, if the diversification process is selection-driven, a larger effective population size may confer to a higher speciation rate. Empirical studies that specifically test these predictions and discuss the effects of effective population size on speciation rate are still scarce<sup>7,9–11</sup>.

Multiple factors can affect intraspecific genetic diversity, e.g.,  $\theta$ , which is often used to approximate effective population size, and the main determinant may vary among evolutionary lineages<sup>6</sup>. Note that  $\theta$  is a function of effective population size ( $N$ ) and the per-generation mutation rate ( $\mu$ ) ( $\theta = 4N\mu$ ). While  $\mu$  tends to vary by less than an order of magnitude in real data,  $N$  varies by several orders of magnitude<sup>12</sup>. Therefore, estimated  $\theta$  values can be used as proxies for population size. However, directly comparing and contrasting results among studies is often unrealistic and potentially problematic due to differences in methods proposed to estimate speciation rate from a phylogenetic tree (e.g., Pagel's  $\lambda$  versus Magallon & Sanderson's  $\lambda$ )<sup>13,14</sup>.

Lichenized fungi represent a highly-diversified group, for which species diversity has generally been under-estimated<sup>15,16</sup>. A broad sampling of DNA sequence data from previous molecular species delimitation studies allow us to estimate intraspecific genetic diversity ( $\theta$ ) as an approximation for effective population size, in

<sup>1</sup>Integrative Research Center, The Field Museum, Chicago, USA. <sup>2</sup>Department of Biology & M. L. Bean Museum, Brigham Young University, Provo, USA. Correspondence and requests for materials should be addressed to J.-P.H. (email: [jhuang@fieldmuseum.org](mailto:jhuang@fieldmuseum.org))

addition to phylogenetic reconstruction and the estimation of speciation rate ( $\lambda$ ). We used sequence data from three genera in the hyperdiverse family Parmeliaceae to test the correlation between  $\theta$  and  $\lambda$  – *Xanthoparmelia*, *Melanelixia*, and *Melanohalea*. These genera share a most recent common ancestor at ca. 68 Myr, and each lineage has a Neogene-dominated diversification history<sup>17</sup>. *Melanelixia* and *Melanohalea* have moderate species numbers (<30), and most species are represented by DNA sequence data and dense, intraspecific specimen sampling. In contrast, *Xanthoparmelia* is highly diverse, with over 800 species<sup>18</sup>, and dense specimen sampling is limited to a small number of well-sampled clades. Differences in species-level circumscriptions among taxonomists may have significant effect on the estimations of population genetic parameters, and our sampling strategy attempted to minimize the effect of such inconsistency on assessing the correlation between  $\theta$  and  $\lambda$ . Specifically, these three genera were selected because their taxonomic status has been recently revised using similar empirical species delimitation methods based on multi-locus sequence data<sup>19,20</sup>. Our study question concerns local  $\theta$ , which are the basic units of evolution that contain randomly mating individuals, and for most sampled species, specimens were collected from multiple, geographically distinct populations<sup>19,20</sup>.

In this study, we evaluate the association between  $\theta$  and  $\lambda$  using recently revised species-level taxonomy for three genera in Parmeliaceae, in addition to applying a molecular species delimitation approach (GMYC) that has been shown to effectively delimit local populations as independent evolving lineages/species should genetic structure exists<sup>21</sup>.  $\theta$  values were estimated for different species delineation scenarios based on the number of segregating sites per site, and  $\lambda$  values were estimated using Magallon and Sanderson's method<sup>14</sup>. Below we report the results and discuss implications of our study.

## Results

The quality of the internal transcribed spacer region (ITS) multiple sequence alignment, a maximum likelihood (ML) reconstruction of the ITS phylogeny, and a chronogram based on the ITS ML phylogeny can be found in the supplementary data (Figs S1, S2, and S3). The total length of the ITS sequences after alignment and quality filtering was 565 base pairs. The ITS region does not encode protein sequence and is presumably evolving neutrally. Therefore, variation in the ITS region, and thus the estimated  $\theta$  value, can be assumed to be an appropriate proxy for effective population size because it is not expected to show large variation in mutation rate across lineages.

A total of 74 species were delineated based on the results of previous multi-locus species delineation studies<sup>19,20</sup> – 22 species in *Xanthoparmelia*, 18 in *Melanelixia*, and 26 in *Melanohalea*. In addition to previous, multi-locus species circumscriptions, for this study species delineations were also determined using a Bayesian Generalized Mixed Yule Coalescent (bGMYC)<sup>21</sup> approach based on sequence data from the ITS region. A total of 437 species were delimited using bGMYC of the ITS topology (Fig. S4), with 102 in *Xanthoparmelia*, 110 in *Melanelixia*, and 192 in *Melanohalea*. The estimated  $\lambda$  varied among lineages, with species from the *Xanthoparmelia* lineage having the highest values (Figs S5 and S6). Some species from the *Melanelixia* lineage also exhibited high  $\lambda$  values but only under the bGMYC species delimitation scenario (Fig. S6).

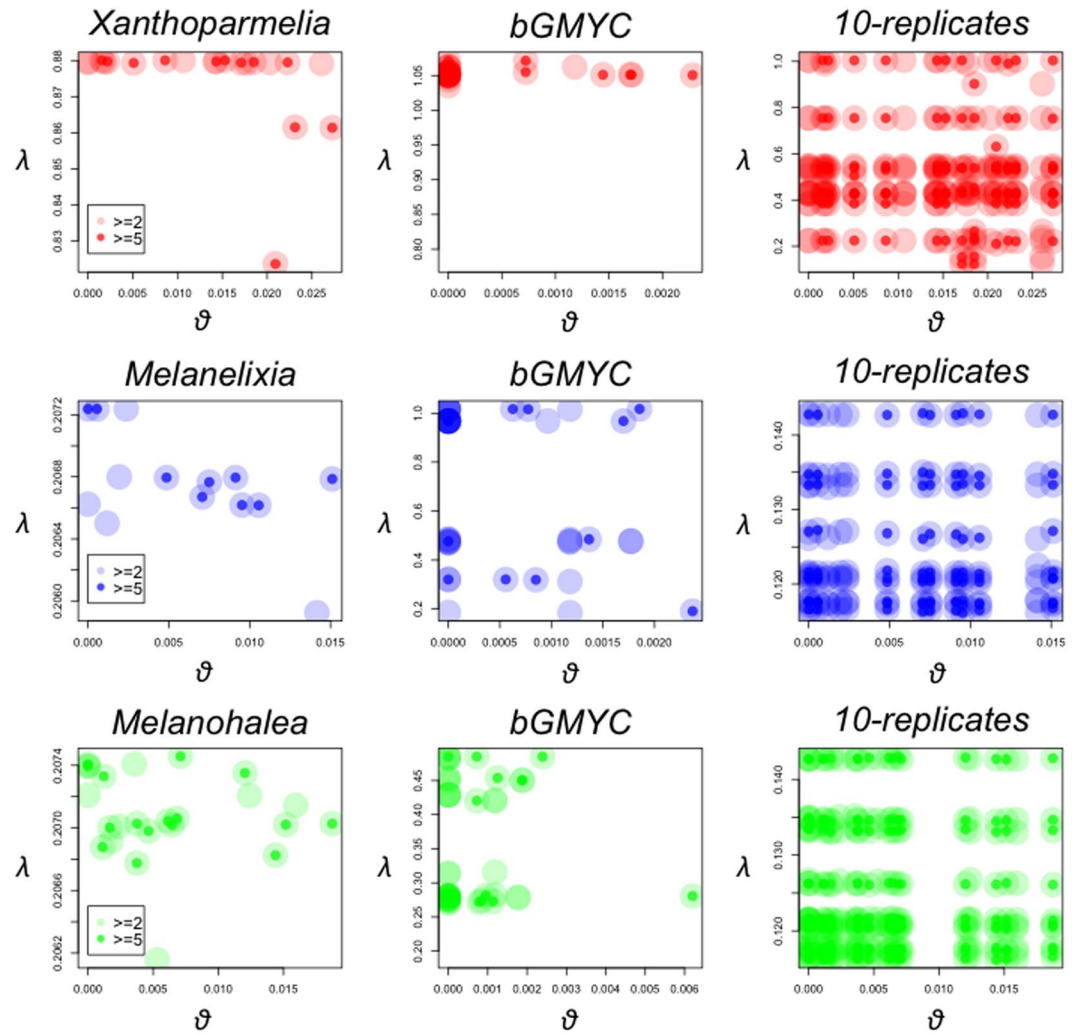
The distributions of estimated  $\theta$  from different species overlapped among the three lineages (Figs S7 and S8). The estimated  $\theta$ , inferred based on the number of segregating sites per site, was smaller when species were delineated via bGMYC than those based on results from previous multi-locus species delineation analyses (Figs S7 and S8). Statistically significant correlation between  $\theta$  and  $\lambda$  after controlling for phylogenetic relatedness (i.e., phylogenetic independent contrast; PIC) was found in the *Melanelixia* lineage, where species were delineated based on multi-locus sequence data<sup>19</sup> ( $P = 0.0002832$  and  $R^2 = 0.8885$  for comparison that only included species that have at least five sampled individuals [ $\geq 5$ ]; Fig. 1). All the other comparisons revealed insignificant results.

In the assessment of the impact of uncertainty in phylogenetic reconstruction and its effect on estimated divergence times and speciation rates, only a single significant correlation between  $\theta$  and  $\lambda$  was detected from ten randomly sampled post-burnin trees. The significant result was found in the *Melanelixia* lineage when species were delineated based on previous study<sup>19</sup> and for comparison that only included species that have at least five sampled individuals. All the other comparisons revealed insignificant results (Fig. 1). Note that an independent bGMYC analysis was not implemented for each of the ten replications.

## Discussion

Our study addressing the impact of effective population sizes on speciation rates in lichenized fungal lineages revealed a correlation between  $\theta$  and  $\lambda$  values in one of the three empirical cases investigated here. However, our results varied based on how species were delineated and by the number of individuals per species that were sampled. We emphasize that our use of the term 'species', as it relates to the different species delimitation scenarios, refers to putative evolutionarily independent species-level lineages. The two different species delimitation scenarios provide strikingly different perspectives into delineations of species-level lineage in these symbiotic fungi – 74 putative species-level lineages using an integrative approach vs. 437 putative species-level lineages based on a bGMYC analysis of the ITS multiple sequence alignment alone. No significant correlation between  $\theta$  and  $\lambda$  values was found when using species delineations based on the bGMYC analyses. In the single case supporting a significant correlation between  $\theta$  and  $\lambda$  – *Melanelixia* species delimited using an integrative approach (from a previous study<sup>19</sup>) and with each taxon represented by at least five samples to estimate the  $\theta$  value – a smaller  $\theta$  was associated with a higher  $\lambda$ . This result implies that mechanisms that generate microevolutionary diversity might also help shape macroevolutionary patterns in some cases. However, the limited scope of the number of lineages analyzed in this study and a lack of correlation between  $\theta$  and  $\lambda$  for the remaining clades limits our ability to make broad, generalizable inferences. Low  $\theta$  caused for example by founder events has been hypothesized to accelerate population subdivision and thus may lead to high speciation rate. A high speciation rate may in return cause low genetic diversity, e.g.,  $\theta$ , within species and high genetic divergence between species<sup>4,6</sup>.

Lineage-specific diversification patterns, where some evolutionary lineages accumulate more species than the others given the same time<sup>22</sup>, has been a main focus in macroevolutionary studies. Models have been



**Figure 1.** Scatter plots of estimated effective population size parameter ( $\theta$ , estimated based on the number of segregating sites per site;  $\theta = 4N\mu$ ) versus estimated speciation rate ( $\lambda$ , estimated based on reference<sup>14</sup> and the programs BAMM<sup>23</sup> and BAMMtools<sup>33</sup>) for tip taxon from three lineages of lichenized fungi. Left panels: results from species delineated based on previous species delimitation results. Middle panels: results from species delineated based on bGMYC species delimitation. Right panels: results from ten replicates using ten randomly chosen post burnin MrBayes trees and species delineated based on previous species delimitation.

developed not only to detect lineage-specific diversification rate<sup>22,23</sup>, but also to test if such pattern is associated with key innovations and geological/climatic events<sup>23,24</sup>. However, without statistically rejecting the effect of stochastic processes that may also lead to variation in diversification rate across evolutionary lineages, the test of trait-dependent speciation and discussion on adaptive radiation seem *ad hoc*<sup>24</sup>. Unlike deterministic mechanisms, such as key innovations and adaptive evolution, that have been the foci of macroevolutionary studies, the equally important stochastic mechanisms, e.g. demographic history, have rarely been incorporated in macroevolutionary models. Our results suggest that stochastic processes might result in differences in speciation rates among lineages in lichenized fungi. Therefore, we caution against testing and inferring trait-dependent diversification patterns without rejecting the effect from stochastic process.

In two of the three empirical cases investigated here, an association between  $\theta$  and  $\lambda$  could not be found (Fig. 1), implying that other processes governing the transition between micro- and macroevolutionary diversity patterns likely exist<sup>9</sup>. Mechanisms that can initiate population subdivision may not always result in speciation events<sup>9</sup>. For example, populations structured by geographic isolation may not acquire ecological differences or reproductive barriers, and a panmictic population may result if secondary contact occurs. Similarly, ecotypes formed by adaptation into different micro-habitats may not become distinct species if the selective pressure is weak<sup>10</sup>.

Alternatively, a lack of correlation between  $\theta$  and  $\lambda$  may, however, simply indicate that factors other than  $\theta$  may have a larger impact on  $\lambda$ . For example, it has been shown that generation time is highly correlated to the number of species in midges<sup>11</sup>. It is indeed unreasonable to assume that all the species in the three lichenized fungal lineages investigated here have the same generation time. These fungal lineages exhibit different life styles,

e.g., asexual versus sexual, and live in a wide range of habitat types, e.g., arid versus temperate regions. Crucial life history information is unfortunately incomplete in our study system, and the impact of generation time on  $\lambda$  cannot be properly addressed here. Additionally, symbiotic relationships with different algal lineages between fungal lineages may have a large effect on adaptive evolution<sup>18</sup>. The effect of population size, or stochastic process on speciation rate can be diminished when speciation is predominantly selective driven. Furthermore, how many lineages and samples per species are needed to confidently detect significant/non-significant correlation between  $\theta$  and  $\lambda$  is beyond the scope of our study. The possibility that our non-significant results are due to the lack of statistical power cannot be excluded.

In addition to the lack of complete life history data and power analysis mentioned previously, numerous assumptions are inherent in species delimitation using gene trees<sup>21</sup>, the use of  $\theta$  to represent effective population size<sup>6</sup> (especially the  $\theta$  value was estimated from only one genetic marker<sup>6</sup>), and the estimation of  $\lambda$  using the BAMM model<sup>14,23</sup>. How these assumptions may affect our results is not clear. The hypothesis addressed in our study concerns intraspecific, or local  $\theta$ , where individuals within each population/species mate randomly. Should within-species genetic structure exists (e.g., geographically or ecologically structured forms/populations), the estimated  $\theta$  can be inflated and lead to biased results. Nevertheless, when a species delimitation method that has been shown able to delimit populations, or fine scale genetic structure, effectively was applied to our dataset (i.e., GMYC), we obtained results consistent with those from integrative taxonomic circumscriptions.

The GMYC approach consistently delimited more species in our study than previous circumscriptions based on multi-locus sequence data and resulted in fewer individuals per species. The GMYC approach may not only lead to a smaller  $\theta$  value and a higher  $\lambda$ , but also further reduce the statistical power to detect correlation. An additional assumption made in this study is that the measurement of current genetic diversity can be related to the estimation of speciation rate from a phylogenetic tree. Although the BAMM analyses can be used to estimate speciation rate at each tip, the estimated rate is integrated over the past million years, depending on where in the phylogeny the rate shift was identified by the model. Although our novel, innovative approach provides important insight into the connection between microevolutionary and macroevolutionary processes, these results should be interpreted with caution, and future studies are needed to fully investigate whether the patterns reported in the lichen system are robust. Specifically, we did not reject the hypothesis that a small  $\theta$  may lead to a larger  $\lambda$ . Instead, our results indicate that with the currently available data, we can only detect a significant negative correlation between  $\theta$  and  $\lambda$  in one lichen lineage.

Our study demonstrates that diverse evolutionary lineages of lichenized fungi are great candidates to bridge micro- and macroevolutionary studies. Due to the prevalence of phenotypically cryptic species in many nominal lichenized fungal species, multiple samples per species are commonly sequenced, which enable both population genetic and phylogenetic approaches using the same system. This is a great advantage over phylogenetic studies that only sequence a single, or a limited number of samples per species, and population genetic studies, which often do not have a wide taxonomic sampling breadth. Future studies incorporating multi-locus and genome-scale data, as well as new models that account for population processes on speciation<sup>9</sup>, and life history data will help to understand not only how the genetic diversity was generated, but also maintained through evolutionary time in the natural system.

## Materials and Methods

A total of 1013 internal transcribed spacer region (ITS) sequences from *Xanthoparmelia* ( $n = 198$ ), *Melanohalea* (489), *Melanelixia* (264), and their close relatives, *Montanelia* (60) and *Emodomanelia* (2), were obtained from TreeBase (S11457 & S18577). A previous species delimitation and taxonomic study of *Melanelixia* and *Melanohalea* was based on sampling of worldwide populations, and approximately 20% of species in these two genera are represented by sequence data from specimens collected from multiple, geographically distinct populations<sup>19</sup>. Our sequence sampling of *Xanthoparmelia* species was largely based on populations in western North America (approximately 4% of the named species), with each species-level lineage represented by samples collected from multiple, geographically distinct populations<sup>21</sup>.

Sequences were aligned using MAFFT ver. 7.304<sup>25</sup> and the quality of the alignment was calculated using Guidance 2<sup>26</sup>. A phylogenetic reconstruction of ITS was performed using a maximum likelihood (ML) method via phyML<sup>27</sup>, specifying an automatic sequence evolutionary model selection. The resulted ML phylogeny was then imported into PATHd8<sup>28</sup> and converted to a chronogram with the root of the phylogeny specified as 68 MYA<sup>17</sup>.

Species delineations in this study were based on two different scenarios: (1) results from recent multi-locus species delineation studies<sup>19,20</sup> and (2) a Bayesian Generalized Mixed Yule Coalescent (bGMYC<sup>21</sup>) approach based on ITS sequence data alone. The taxonomy of the brown parmelioid genera, which includes *Melanohalea* and *Melanelixia*, was recently revised<sup>19</sup>, reflecting empirical species delimitation studies. However, while species boundaries in *Xanthoparmelia* has also been circumscribed using empirical species delimitation analyses<sup>20</sup>, a formal taxonomic revision has not yet been proposed. We were fully aware of potential over-splitting species when applying the GMYC processes for species delimitation<sup>29</sup>. However, the rationale here was to investigate the inconsistency of our results. We followed the suggested procedure by Talavera and colleagues<sup>29</sup> for our bGMYC analysis.

The MCMC searches for bGMYC were launched for  $5 \times 10^5$  generations while thinning every  $1 \times 10^3$  generations and a burnin of the first 10% of samples. Species were delineated using a threshold of 0.05 from the bGMYC result, which yielded a conservative estimate of species number<sup>21</sup>. Species with multiple samples (delineated under both scenarios) were pruned using the drop.tip function in R package ape<sup>30</sup>. The edited species trees, where each delineated species contains only one representative sample, were saved in newick format.

The speciation, extinction, and net diversification rates of the three lichenized fungal genera were estimated based on the edited species trees and a Bayesian Analysis of Macroevolutionary Mixtures (BAMM ver. 2.5)<sup>23</sup>. A recent study highlighted potential limitations of using BAMM to study macroevolutionary dynamics<sup>31</sup> (but see<sup>32</sup>).

However, the estimated rate has been shown to correlate with the true rate using simulation<sup>32</sup>. In our study, we focused on the relative rate difference between species, and BAMM represented a suitable model for our study quest. We used the function `setBAMMpriors` from the R package `BAMMtools`<sup>33</sup> to specify initial  $\lambda$ , extinction rate, and rate shift values. Sample fraction files that specified the proportion of missing species from the phylogenetic tree were provided for the BAMM analyses (available upon request). A total of  $1 \times 10^8$  MCMC searches were run for the BAMM analyses with samples stored every  $1 \times 10^4$  generations. After discarding the first 50% of samples as burnin, the remaining MCMC samples from the BAMM analyses were used to estimate  $\lambda$  values at tips of the trees via the `getTipRates` function.

To account for uncertainty in phylogenetic reconstruction and its effect on the estimated divergence times and speciation rates, we repeated the BAMM analyses ten times by using different trees as inputs. Specifically, a Bayesian phylogenetic reconstruction was performed using the program `MrBayes` (ver. 3. 2. 6)<sup>34</sup>. A GTR + I + G sequence evolution model (`nst = 6` and `rate = invgamma`) was specified for the analysis. The analysis was launched for  $1 \times 10^8$  MCMC generations with a sampling frequency of every 100 generations. The split frequency decreased below 0.01 after generation  $3 \times 10^7$ . The burnin period was thus set to  $3 \times 10^7$  generations. We randomly selected ten trees from the 70000 post burnin trees. The selected trees were converted into ultrametric trees with a root depth at 68 MYA using `PATHd8`. The converted ultrametric trees were then imported into BAMM for ten replicate analyses to estimate  $\lambda$  values at tips.

Sequences from multiple individuals of the same species (delineated under both scenarios) were used to calculate the number of segregating sites with a customized R script (indels '-' and missing data '?') were not counted as segregating sites). We subsequently used the function `theta.s` from the `pegas` package<sup>35</sup> to estimate  $\theta$  ( $\theta = 4N\mu$ ) for different species based on the number of segregating sites. The value of  $\theta$  represents the expected number of variable sites per site between two sequences randomly pulled from a sequence alignment. We estimated the  $\theta$  values for different species with two minimum numbers of individual as thresholds – at least 2 (size 2) and 5 (size 5) individuals have to present in a species for that species to be included in the analyses. A linear regression to test for significant correlation between  $\theta$  and  $\lambda$  for the three lichenized fungi lineages was statistically evaluated. We controlled for phylogenetic non-independence of  $\theta$  and  $\lambda$  values by using the phylogenetic independence contrast via the `pic` function implemented in `ape`<sup>30</sup>.

## References

1. Nei, M. & Takahata, N. Effective population size, genetic diversity, and coalescent time in subdivided populations. *J. Mol. Evol.* **37**, 240–244 (1993).
2. Zhang, G. *et al.* Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311–1320 (2014).
3. Eo, S. H. & DeWoody, J. A. Evolutionary rates of mitochondrial genomes correspond to diversification rates and to contemporary species richness in birds and reptiles. *Proc. Roy. Soc. B* **277**, 3587–3592 (2010).
4. Venditti, C. & Pagel, M. Speciation as an active force in promoting genetic evolution. *Trends Ecol. Evol.* **25**, 14–20 (2010).
5. Pennell, M. W. *et al.* Speciation is unlikely to drive divergence rates. *Trends Ecol. Evol.* **29**, 72–73 (2014).
6. Ellegren, H. & Galtier, N. Determinants of genetic diversity. *Nat. Rev. Genet.* **17**, 422–433 (2016).
7. Lnafear, R. *et al.* Population size and the rate of evolution. *Trends Ecol. Evol.* **29**, 33–41 (2014).
8. Lutzoni, F. & Pagel, M. Accelerated evolution as a consequence of transitions to mutualism. *Proc. Natl. Acad. Sci. USA* **94**, 11422–11427 (1997).
9. Dynesius, M. & Jansson, R. Persistence of within-species lineages: a neglected control of speciation rates. *Evol.* **68**, 923–934 (2014).
10. Cutter, A. D. & Gray, J. C. Ephemeral ecological speciation and the latitudinal biodiversity gradient. *Evol.* **70**, 2171–2185 (2016).
11. Oppold, A.-M. *et al.* Support for the evolutionary speed hypothesis from intraspecific population genetic data in the non-biting midge *Chironomus riparius*. *Proc. R. Soc. B* **283**, 20152413 (2016).
12. Lynch, M. Evolution of the mutation rate. *Trends Genet.* **26**, 345–352 (2010).
13. Pagel, M. Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884 (1999).
14. Magallon, S. & Sanderson, M. J. Absolute diversification rates in angiosperm clades. *Evolution* **55**, 1762–1780 (2001).
15. Crespo, A. & Lumbsch, H. T. Cryptic species in lichen-forming fungi. *IMA Fungus* **1**, 167–170 (2010).
16. Lumbsch, H. T. & Leavitt, S. D. Goodbye morphology? A paradigm shift in the delimitation of species in lichenized fungi. *Fungal Diversity* **50**, 59–72 (2011).
17. Divakar, P. K. *et al.* Evolution of complex symbiotic relationships in a morphologically derived family of lichen-forming fungi. *New Phytol.* **208**, 1217–1226 (2015).
18. Thell, A. *et al.* A review of the lichen family Parmeliaceae – history, phylogeny and current taxonomy. *Nord. J. Bot.* **30**, 641–664 (2012).
19. Leavitt, S. *et al.* Hidden diversity before our eyes: delimiting and describing cryptic lichen-forming fungal species in camouflage lichens (Parmeliaceae, Ascomycota). *Fungal Biol.* **120**, 1374–1391 (2016).
20. Leavitt, S. *et al.* Species delimitation in taxonomically difficult lichen-forming fungi: an example from morphologically and chemically diverse *Xanthoparmelia* (Parmeliaceae) in North America. *Mol. Phylogenet. Evol.* **60**, 317–332 (2011).
21. Reid, N. M. & Carstens, B. C. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evol. Biol.* **12**, 196 (2012).
22. Alfaro, M. E. *et al.* Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proc. Natl. Acad. Sci. USA* **106**, 13410–13414 (2009).
23. Rabosky, D. L. *et al.* Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat. Commun.* **4**, 1958 (2013).
24. Rabosky, D. L. & Huang, H. A robust semi-parametric test for detecting trait-dependent diversification. *Syst. Biol.* **65**, 181–193 (2016).
25. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
26. Sela, I. *et al.* GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucl. Acids Res.* **43**, W7–W14 (2015).
27. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
28. Britton, T. *et al.* Estimating divergence times in large phylogenetic trees. *Syst. Biol.* **56**, 741–752 (2007).
29. Talavera, G. *et al.* Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods in Ecology and Evolution* **4**, 1101–1110 (2013).

30. Paradis, E. *et al.* APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
31. Moore, B. R. *et al.* Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proc. Natl. Acad. Sci. USA* **113**, 9569–9574 (2016).
32. Rabosky, D.L. *et al.* Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Syst. Biol.* **syx037** (2017).
33. Rabosky, D. L. *et al.* BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods Ecol. Evol.* **5**, 701–707 (2014).
34. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755 (2001).
35. Paradis, E. Pegas: an R package for population genetics within an integrated-modular approach. *Bioinformatics* **26**, 419–420 (2010).

## Acknowledgements

This study is supported by a grant of the Negaunee Foundation to H. T. L.

## Author Contributions

J.-P. H. and S.D.L. conceived the project. S.D.L. collected the data and J.-P.H. analyzed the data. All authors wrote and read the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-24120-9>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018