Fernandez-Mazuecos, Mario; Blanco-Pastor, José Luis; Juan, Ana; [et al.]. «Macroevolutionary dynamics of nectar spurs, a key evolutionary innovation». The new phytologist, Vol. 222, Issue 2 (April 2019), p. 1123-1138. DOI 10.1111/nph.15654

Macroevolutionary dynamics of nectar spurs, a key evolutionary innovation

Mario Fernández-Mazuecos[1,2], José Luis Blanco-Pastor[2,3], Ana Juan[4], Pau Carnicero[5], Alan Forrest[6], Marisa Alarcón[7], Pablo Vargas[2], Beverley J. Glover[1]

[1]Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK

[2]Departamento de Biodiversidad y Conservación, Real Jardín Botánico (RJB-CSIC), Plaza de Murillo 2, 28014 Madrid, Spain

[3]INRA, Centre Nouvelle-Aquitaine-Poitiers, UR4 (URP3F), 86600 Lusignan, France

[4]Departamento de Ciencias Ambientales y Recursos Naturales (dCARN) & Instituto de la Biodiversidad (CIBIO), Universidad de Alicante, P.O. Box 99, 03080 Alicante, Spain

[5]Departament de Biologia Animal, Biologia Vegetal i Ecologia, Facultat de Biociències, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

[6]Centre for Middle Eastern Plants, Royal Botanic Garden Edinburgh, 20a Inverleith Row, Edinburgh EH3 5LR, UK

[7]Institut Botànic de Barcelona (IBB-CSIC-ICUB), Passeig del Migdia s/n, Parc de Montjuïc, 08038 Barcelona, Spain

*Correspondence: Mario Fernández-Mazuecos. E-mail: mfmazuecos@rjb.csic.es. Tel.: +34 914203017

Twitter (Mario Fernández-Mazuecos): @mfmazuecos

Mario Fernández-Mazuecos: https://orcid.org/0000-0003-4027-6477

José Luis Blanco-Pastor: https://orcid.org/0000-0002-7708-1342

Ana Juan: http://orcid.org/0000-0002-2929-3818

Pau Carnicero: https://orcid.org/0000-0002-8345-3309

Pablo Vargas: https://orcid.org/0000-0003-4502-0382

Beverley J. Glover: https://orcid.org/0000-0002-6393-819X

Total word count: 6908

Introduction: 849

Materials and methods: 2132

Results: 1209

Discussion and conclusions: 2601

Acknowledgements: 117

Number of figures: 5 (all of them in colour)

Number of tables: 4

Supporting information: 10 figures, 5 tables

Summary

- Floral nectar spurs are widely considered a key innovation promoting diversification in angiosperms by means of pollinator shifts. We investigated the macroevolutionary dynamics of nectar spurs in the tribe Antirrhineae (Plantaginaceae), which contains 29 genera and 300-400 species (70-80% spurred). The effect of nectar spurs on diversification was tested, with special focus on *Linaria*, the genus with the highest number of species.
- We generated the most comprehensive phylogeny of Antirrhineae to date and reconstructed the evolution of nectar spurs. Diversification rate heterogeneity was investigated using trait-dependent and trait-independent methods, and accounting for taxonomic uncertainty. The association between changes in spur length and speciation was examined within *Linaria* using model testing and ancestral state reconstructions.
- We inferred four independent acquisitions of nectar spurs. Diversification analyses revealed that nectar spurs are loosely associated with increased diversification rates. Detected rate shifts were delayed with respect to the acquisition of the trait. Active evolution of spur length, fitting a speciational model, was inferred in *Linaria*, which is consistent with a scenario of pollinator shifts driving diversification.
- Nectar spurs played a role in diversification of the Antirrhineae, but diversification dynamics can only be fully explained by the complex interaction of multiple biotic and abiotic factors.

Key words: Antirrhineae, *Antirrhinum*, diversification, flower, key innovation, *Linaria*, nectar spur, speciation.

Introduction

Key evolutionary innovations have been widely considered as fundamental drivers of biodiversity (Erwin, 1992; Heard & Hauser, 1995; Hunter, 1998; Rabosky, 2014). According to the definition given by Heard and Hauser (1995), a key innovation is an evolutionary change in an individual trait that is causally linked to an increased diversification rate in the resulting clade. This effect may result from the invasion of new adaptive zones, increased clade fitness and/or increased propensity for reproductive isolation. In plants, traits usually considered key innovations include animal pollination, floral zygomorphy and nectar spurs (Hodges, 1997; Dodd *et al.*, 1999; Sargent, 2004; Kay *et al.*, 2006). In particular, floral nectar spurs have come to constitute a textbook example of a plant key innovation thanks to long-term research on the genus *Aquilegia* (Hodges & Arnold, 1995; Whittall & Hodges, 2007; Puzey *et al.*, 2012).

A nectar spur is a tubular outgrowth of a floral organ (petal or sepal) that usually contains nectar. By enhancing pollinator specificity, pollination efficiency and reproductive success, nectar spurs may facilitate the transition to a new adaptive space, at the same time promoting reproductive isolation and thus speciation (Fulton & Hodges, 1999; Shivanna, 2014; Minelli, 2015). Indeed, nectar spurs have evolved independently in numerous angiosperm families, and spurred clades usually exhibit significantly higher species diversity than their sister clades, suggesting a consistent positive effect on diversification rates (Hodges, 1997; Kay *et al.*, 2006). Nevertheless, the inability of sister group comparisons to precisely pinpoint the location of diversification rate shifts has led some authors to cast doubt on a straightforward relationship between nectar spurs and diversification (Donoghue & Sanderson, 2015). Some other authors have argued that a positive effect of specialised floral traits (such as nectar spurs) on speciation is only one possible explanation for the association between floral specialisation and clade species diversity (Armbruster & Muchhala, 2009). Two alternative explanations have been proposed: first, rather than increasing speciation, specialisation may reduce extinction by diminishing the negative effects of interspecific pollination, which promotes tighter species packing in communities; and second, high species diversity may conversely cause floral specialisation by promoting character displacement (see details in Armbruster & Muchhala, 2009). Empirical evidence is still needed to determine the relative importance of these mechanisms.

Methods to identify key innovations and investigate their macroevolutionary dynamics not only include the classical sister clade comparisons (Slowinski & Guyer, 1993), but also increasingly sophisticated model-based approaches (FitzJohn *et al.*, 2009; Rabosky, 2014; Beaulieu & O'Meara, 2016). However, methodological controversy surrounds many of the methods dealing with diversification rates (Rabosky & Goldberg, 2015; Moore *et al.*, 2016). For example, Rabosky and Goldberg (2015) reported model inadequacies producing a high rate of false positives in commonly used tests to detect trait-dependent diversification, and this led to the development of more complex models to analyse diversification dynamics (Beaulieu & O'Meara, 2016). It is clear that a critical combination of methodological approaches is needed to provide fundamental insights into the drivers of biodiversity (see Igea *et al.*, 2017).

The snapdragons and relatives (tribe Antirrhineae, Plantaginaceae), including the model species *Antirrhinum majus*, are an ideal study system to investigate the evolution of nectar spurs, their role as a key innovation and their macroevolutinary dynamics. The Antirrhineae include 300-400 species classified into 29 genera distributed in the Old and the New World, and characterised by their specialised floral traits (Sutton, 1988; Vargas *et al.*, 2014; Guzmán *et al.*, 2015; Guzmán *et al.*, 2017). Of these, six genera display nectar spurs and make up 70-80% of species diversity (Fig. 1a). Spurred genera appear in several phylogenetically unrelated lineages (Vargas *et al.*, 2014; Guzmán *et al.*, 2015), suggesting independent origins of the trait. Unlike previously studied systems like *Aquilegia*, characterised by a single origin of spurs (Fior *et al.*, 2013), the Antirrhineae provide a unique opportunity to investigate potentially replicated effects of spurs on diversification rates in a shared phylogenetic background (see Maddison & FitzJohn, 2014).

As pointed out by Donoghue and Sanderson (2015), it is not just the presence of a key innovation that matters, but also the phylogenetic distribution of the variable linked to speciation by specific mechanisms, such as nectar spur length. According to the “pollinator shift” scenario, differences in spur length would influence pollinator specificity and therefore lead to premating isolation and ultimately speciation (Whittall & Hodges, 2007). If this were true, evolutionary changes in spur length would tend to be associated with speciation events. In the Antirrhineae, the spurred genus *Linaria* is the most diverse, with 150-200 species, and displays remarkable variation in spur length (Sutton, 1988; Sáez & Bernal, 2009), providing a suitable study system to test the association between speciation and spur length evolution.

In this study, our objective was to investigate the macroevolutionary dynamics of nectar spurs in the tribe Antirrhineae with the aim of understanding their potential role as a key innovation. Two hypotheses were tested: (1) that independent acquisitions of nectar spurs during the evolution of the tribe are consistently linked to significant increases in diversification rates; and (2) that evolutionary changes in spur length in *Linaria* are significantly associated with speciation events.

Materials and methods

Taxonomic sampling and DNA sequencing

To make full use of available sequence data, we adopted a supermatrix approach (De Queiroz & Gatesy, 2007). We used a total of 650 DNA sequences from 304 named species of Antirrhineae (Supporting Information Table S1) belonging to the nuclear ribosomal internal transcribed spacers (ITS) and two plastid DNA (ptDNA) regions: *ndh*F and *rpl*32-*trn*L. These are the three DNA regions that have been most frequently used in phylogenetic analyses of Antirrhineae genera (Ghebrehiwet *et al.*, 2000; Oyama & Baum, 2004; Vargas *et al.*, 2004; Blanco-Pastor & Vargas, 2013; Fernández-Mazuecos *et al.*, 2013a; Fernández-Mazuecos *et al.*, 2013b; Rahmani *et al.*, 2014; Vargas *et al.*, 2014; Guzmán *et al.*, 2015; Yousefi *et al.*, 2016; Carnicero *et al.*, 2017). Five hundred and forty-five sequences of 262 Antirrhineae species from the referenced studies were retrieved from the GenBank database, and 113 sequences from 75 species were newly generated following the methods described in our previous publications (Fernández-Mazuecos *et al.*, 2013a; Fernández-Mazuecos *et al.*, 2013b; Vargas *et al.*, 2014) (see Supporting Information Table S1 for GenBank accession numbers and Supporting Information Table S2 for vouchers of newly sequenced species). Outgroup taxa were selected following the approach of Vargas *et al.* (2014), and included two species of the genus *Lafuentea* (sister to Antirrhineae; Albach *et al.*, 2005), 13 additional species of the family Plantaginaceae and 19 species representing 11 other families of the order Lamiales.

Sequences were assembled in Geneious version 5 (Kearse *et al.*, 2012) and aligned using MAFFT version 7 (Katoh & Toh, 2008). The final concatenated dataset comprised 338 taxa (including 304 species of Antirrhineae) and a total length of 3,916 bp. Within the Antirrhineae, taxon completeness was highest for ITS sequences (97% of species), and lower for *rpl*32-*trn*L (70%) and *ndh*F (50%). The outgroup comprised mostly *ndh*F sequences (100% of outgroup species).

Phylogenetic analyses and dating

The best-fitting substitution model was determined for each DNA region based on the Akaike Information Criterion (AIC) calculated in jModelTest 2.1.6 (Darriba *et al.*, 2012). To obtain a preliminary topology, a partitioned phylogenetic analysis was conducted in MrBayes 3.2.6 (Ronquist *et al.*, 2012) using two runs with four chains and 10 million generations each, and a sampling frequency of 1000. Then, a time-calibrated phylogenetic analysis was performed in BEAST 2.4.2 (Bouckaert *et al.*, 2014) with unlinked site models across partitions (as determined by jModelTest), unlinked clock models (uncorrelated relaxed clock in all cases), a birth-death process as tree prior, and uniform priors for substitution rates following Blanco-Pastor *et al.* (2012). *Plocosperma buxifolium* was set as the earliest-diverging species by constraining the remaining taxa as a monophyletic group (see Schäferhoff *et al.*, 2010). A secondary calibration for the time to most recent common ancestor (TMRCA) of all taxa except *Plocosperma* was implemented using a normal prior with mean 74 Ma and standard deviation 2.5 Ma (Bell *et al.*, 2010). After revising the limited fossil record of the Antirrhineae (Supporting Information Table S3), two fossil calibrations within the tribe were implemented: (1) fossil seeds identified as *Linaria vulgaris* (although indistinguishable from other species of *Linaria* sect. *Linaria*) from the Upper Pliocene of Russia (Dorofeev, 1963) were employed to calibrate the stem age of the *Linaria* sect. *Linaria* + sect. *Speciosae* clade (where species of both sections are intermixed) using a log-normal prior with offset = 2.6 Ma, mean = 1.0 and standard deviation = 1.25; and (2) fossil seeds designated as the extinct species *Asarina ruboidea* from the Serravallian (Middle Miocene) of Germany (Mai, 2001) were employed to calibrate the stem age of *Asarina* using a log-normal prior with offset 11.6 Ma, M = 1.0 and S = 1.25. Five additional fossil calibrations outside the Antirrhineae were included mostly following Vargas *et al.* (2014) (see Supporting Information Table S4 for details). The monophyly of *Linaria* sect. *Supinae* (except *L. latifolia*) was constrained following the results of Blanco-Pastor *et al.* (2012) (see also Fernández-Mazuecos *et al.*, 2013b). Results from seven MCMC chains with 200 million generations each were combined in LogCombiner after removing chain-specific burn-in fractions determined by examining trace plots in Tracer 1.6 (Rambaut *et al.*, 2014). Effective sample sizes >200 were obtained for all parameters. A maximum clade credibility (MCC) tree with common ancestor heights was calculated in TreeAnnotator. All non-Antirrhineae taxa, except for the two *Lafuentea* species, were pruned from the tree for downstream analyses.

Acquisitions of nectar spurs

We reconstructed the number of evolutionary transitions between absence and presence of nectar spurs in the Antirrhineae using maximum likelihood (ML) and stochastic character mapping (SCM), both implemented in the R package *phytools* (Revell, 2012). Presence/absence of nectar spurs was scored based on taxonomic descriptions (Sutton, 1988, among others) and our own knowledge of Antirrhineae genera. Two evolutionary models were tested: an equal rates (ER) model and a different rates (DR) model. The best model was selected based on AIC values. ML reconstructions were performed using the re-rooting method of Yang *et al.* (1995). SCM was conducted with 1000 simulations. Additional reconstructions were performed under trait-dependent diversification models (see below) (Goldberg & Igić, 2008).

Taxonomic treatments

The number of species that are recognised in a clade can strongly influence the outcome of diversification rate analyses (Faurby *et al.*, 2016). The last worldwide taxonomic treatment of the tribe Antirrhineae (Sutton, 1988) recognised 326 species in 27 genera. Since then, taxonomists have described many new species (particularly in the genera *Linaria* and *Chaenorhinum*; see The International Plant Names Index, http://www.ipni.org/) and even two new monotypic genera (*Pseudomisopates* and *Gadoria*; Güemes, 1997; Güemes & Mota, 2017). A number of additional taxonomic rearrangements have been suggested. Notably, a taxonomic revision of *Kickxia* sect. *Valvatae* proposed its separation as a different genus (*Nanorrhinum*) and a reduction in the number of species from 37 to 10 (Ghebrehiwet, 2000). Some revisions for the Iberian Peninsula, one of the centres of species diversity of Antirrhineae, also resulted in changes to species delimitation (Benedí & Güemes, 2009; Güemes, 2009; Sáez & Bernal, 2009). To account for uncertainty in species numbers, we defined three alternative taxonomic treatments of Antirrhineae with different species numbers based on available literature. In the *splitter* treatment, all species recognised and described in recent literature (since Sutton, 1988) were included, Sutton's (1988) treatment of *Nanorrhinum* (=*Kickxia* sect. *Valvatae*) was followed, and named subspecies were putatively considered as distinct species. In the *intermediate* treatment, all species recognised and described in recent literature were included and Sutton's (1988) treatment of *Nanorrhinum* was followed, but subspecies were not considered. In the *lumper* treatment, species described after Sutton (1988) were not considered (except for those of the two newly described monotypic genera), Ghebrehiwet's (2000) treatment of *Nanorrhinum* was followed, and subspecies were not considered (see Supporting Information Table S5 for details). Phylogenetic trees consistent with the three treatments were generated by pruning those species not recognised by each treatment from the original phylogeny.

Diversification rates: trait-dependent models

We applied a range of methods to test the hypothesis that nectar spurs positively influence diversification rates under the three alternative taxonomic treatments. Selected methods were of two types: trait-dependent and trait-independent.

Three methods to detect trait-dependent diversification rates were applied: BiSSE (FitzJohn *et al.*, 2009), FiSSE (Rabosky & Goldberg, 2017) and HiSSE (Beaulieu & O'Meara, 2016). BiSSE (Binary State Speciation and Extinction) is a model-based method to investigate the effect of a single binary trait on diversification rates. BiSSE analyses were conducted in the R package *diversitree* (FitzJohn, 2012) using the MCC tree for each taxonomic treatment. To account for incomplete sampling, clade-specific sampling fractions according to each taxonomic treatment were included. A model with state-dependent speciation and extinction and asymmetrical transition rates was compared against nested models with speciation rate (λ), extinction rate (μ) and transition rate (q) parameters constrained to be equal for both states. ML parameter values were calculated for each model, and model differences were assessed by AIC values. To obtain an estimate of parameter uncertainty, the full BiSSE model was additionally explored using Bayesian inference, with exponentially distributed priors based on ML values. Each MCMC comprised 10,000 steps, of which the first 1000 were discarded as burn-in. A marginal reconstruction of ancestral states was conducted based on ML parameter values under the full BiSSE model.

FiSSE (Fast, intuitive State-dependent Speciation and Extinction) is a simple nonparametric test with the same aim as BiSSE, but considered robust to some of the issues described for that method, such as the sensitivity to model inadequacy and phylogenetic pseudoreplication. FiSSE analyses were conducted using the R functions published by the original authors (https://github.com/macroevolution/fisse; Rabosky & Goldberg, 2017). We accounted for incomplete sampling by generating a distribution of 1000 completely sampled phylogenies by randomly adding unsampled species to the corresponding clades of the empirical phylogeny (using the add.species.to.genus function of *phytools*). FiSSE tests were conducted for the 1000 simulated phylogenies with standard specifications (reps=1000; tol=0.1; qratetype=mk). A histogram of two-tailed *P*-values was plotted for each taxonomic treatment.

HiSSE (Hidden State Speciation and Extinction) is a model-based method that extends the BiSSE framework to account for unmeasured factors (“hidden” states) that could impact diversification rates in addition to the trait of interest. HiSSE analyses were performed using the *hisse* package (Beaulieu & O'Meara, 2016). Incomplete sampling was accounted for by including state-specific sampling fractions according to each taxonomic treatment. Four models were tested: (1) a character-independent diversification model with two hidden states (CID-2); (2) a character-independent diversification model with four hidden states (CID-4); (3) a full binary-state speciation and extinction model (full BiSSE); and (4) a full hidden-state speciation and extinction model (full HiSSE). Marginal reconstructions of ancestral states and diversification rates under the four models were estimated. To incorporate uncertainty in model choice, reconstructions under the four models were averaged using AIC weights, and model-averaged rates for all tips and nodes of the phylogeny were obtained. Spur presence/absence at nodes was inferred based on marginal probabilities (a probability >0.5 was interpreted as spur presence). Then, differences in diversification rates between spurred and spurless tips and nodes were assessed using beanplots (Kampstra, 2008).

Diversification rates: trait-independent models

Two trait-independent methods to detect diversification rate shifts were employed: MEDUSA (Alfaro *et al.*, 2009) and BAMM (Rabosky, 2014). MEDUSA (Modeling Evolutionary Diversification Using Stepwise AIC) is a likelihood-based method employing a stepwise AIC procedure. It was implemented in the R package *MEDUSA* (https://github.com/josephwb/turboMEDUSA). The 1000 completely sampled simulated phylogenies generated for FiSSE analyses were analysed to account for incomplete sampling. Results were summarised on a single randomly chosen tree.

BAMM (Bayesian Analysis of Macroevolutionary Mixtures) is a Bayesian approach using reversible-jump MCMC. This method was implemented in BAMM version 2.5.0. Incomplete sampling was accounted for by specifying clade-specific sampling fractions. Appropriate prior values were generated using the setBAMMpriors function of the R package *BAMMtools* (Rabosky *et al.*, 2014). Four Metropolis-coupled MCMC chains were run for 10 million generations, with a sampling frequency of 10,000. Results were processed using *BAMMtools*, including the visualisation of mean phylorate plots and clade-specific rate-through-time (RTT) plots.

Spur length evolution in Linaria

We explored the timing of spur length evolution in *Linaria*, the genus with the highest number of species in the Antirrhineae, using the phylogeny of the genus obtained after pruning all other genera from our empirical Antirrhineae phylogeny. First, we tested the correlation between log-transformed spur length and corolla length using phylogenetic generalised least squares (PGLS; Grafen, 1989) in the R package *caper* (Orme, 2012), with log(spur length) as dependent variable, log(corolla length) as explanatory variable, and phylogenetic signal estimated by ML. Data were log-transformed to reduce heteroscedasticity and analyse relative rather than absolute variation. Trait values were taken from taxonomic literature (midpoints of given intervals; Sutton, 1988; Sáez & Bernal, 2009; among others). Given the positive correlation observed, we then analysed the evolution of both the log-transformed spur length and the spur length / corolla length ratio (to control for the effect of corolla length). Ancestral state reconstructions were performed by ML in *phytools* (contMap function; Revell, 2013). Rates of phenotypic evolution were analysed in BAMM using the approach described above for diversification rates. Finally, to examine whether phenotypic evolution occurs preferentially at speciation events, we tested four evolutionary models in the CoMET package (Lee *et al.*, 2007; implemented in Mesquite, Maddison & Maddison, 2011): (1) a gradual model, where the amount of phenotypic change depends on branch lengths (“distance, pure phylogenetic” in CoMET terminology); (2) a speciational model, where the amount of phenotypic change depends on the number of speciation events (“equal, pure phylogenetic”); (3) a punctuated model, where change also depends on speciation events, but only one of the daughter species changes at each split, while its sister retains the state of the parent (“equal, punctuated”); and (4) a non-phylogenetic model, where closely-related species are no more similar to each other than to distant relatives (“equal, non-phylogenetic”).

Results

Phylogenetic analyses and dating

Major clades of Antirrhineae were strongly supported (posterior probability, PP≈1) by both the MrBayes (Supporting Information Fig. S1) and BEAST (Fig 1b, c; Supporting Information Fig. S2) analyses. Species with a nectar spur were found in four separate clades: (1) the clade formed by *Anarrhinum*, *Kickxia* and *Nanorrhinum* (although not all species of *Anarrhinum* have a nectar spur); (2) the *Cymbalaria* clade; (3) the *Chaenorhinum* clade; and (4) the *Linaria* clade. The TMRCA of all Antirrhineae lineages estimated by the BEAST analysis (Fig. 1c; Supporting Information Fig. S2) was 36-52 Ma (95% highest posterior density interval, HPD). Therefore, a diversification of Antirrhineae since the Eocene was estimated.

Acquisitions of nectar spurs

The DR model had the lowest AIC value (AIC=46.5), closely followed by the ER model (AIC=46.7; ΔAIC=0.2). Under the DR model, both ancestral state reconstruction methods (ML and SCM) estimated that the absence of nectar spur is the ancestral condition in Antirrhineae, and clearly supported four convergent acquisitions of nectar spurs (Fig. 1b, c; Supporting Information Fig. S3a, b). A single loss was inferred within the genus *Anarrhinum*. Similar results were obtained under the ER model, but with more uncertainty at ancestral nodes (Supporting Information Fig. S3c, d).

Taxonomic treatments

A total of 501, 398 and 297 species of Antirrhineae (plus *Lafuentea*) were recognised respectively by the *splitter*, *intermediate* and *lumper* treatments (Table 1; Supporting Information Table S5). Of these, 306 (61%), 296 (74%) and 248 (84%) species were included in our phylogenetic analysis.

Diversification rates: trait-dependent models

Under the *splitter* and *intermediate* taxonomic treatments, BiSSE analyses in *diversitree* supported models where speciation rates are higher for spurred than for spurless lineages of Antirrhineae versus models with a single speciation rate (Tables 2, 3). Under both treatments, the strongest support (ΔAIC<2) was obtained for models with different λ for the two character states ($\lambda_0 \neq \lambda_1$), and models with equal λ received low support (ΔAIC>2; Table 2). Under the *lumper* taxonomic treatment, results were less clear. The set of supported models included models with $\lambda_0 \neq \lambda_1$, but also a model with $\lambda_0 = \lambda_1$ (but $\mu_0 \neq \mu_1$). The Bayesian analysis under the full model provided consistent results. Higher estimates of speciation rates for spurred lineages were obtained under all three taxonomic treatments, with no overlap of 95% HPD intervals for λ_0 and λ_1 under the *splitter* treatment, and progressively higher overlap under the *intermediate* and *lumper* treatments (Fig. 2a). Marginal reconstructions of ancestral states under the BiSSE model supported the absence of nectar spur as ancestral condition, four convergent acquisitions of nectar spurs in the Antirrhineae and a single loss in *Anarrhinum* (Supporting Information Fig. S3e-g).

FiSSE analyses only achieved statistical significance (P<0.05) for one of the 1000 simulated phylogenies under the *splitter* treatment. All remaining FiSSE tests under the three taxonomic treatments were non-significant (Fig. 3).

In *hisse* analyses, the full HiSSE model was supported under the *splitter* treatment, with all other models being significantly worse (ΔAIC>2) (Table 2). Under the *intermediate* and *lumper* treatments, higher uncertainty about the optimal model was obtained. Model-averaged marginal reconstructions revealed higher diversification rate heterogeneity under the *splitter* treatment than under the *intermediate* and *lumper* treatments (Fig. 4; Supporting Information Fig. S4). An ancestral absence of nectar spur, four convergent acquisitions of the trait and a single loss (in *Anarrhinum*) were inferred in all cases. The beanplot for the *splitter* treatment clearly showed a higher mean diversification rate in spurred lineages than in spurless lineages, but with a large overlap of values and a wide dispersion in spurred lineages (Fig. 2b; see also Table 3). The difference in mean diversification rate resulted from a combination of higher speciation rates and lower extinction rates estimated for spurred lineages than for spurless lineages (Supporting Information Fig. S5). Under the *intermediate* and *lumper* treatments, smaller differences in mean diversification rates and a larger overlap in values of spurred and spurless lineages were obtained.

Diversification rates: trait-independent models

MEDUSA and BAMM analyses revealed similar patterns of diversification rate heterogeneity across the Antirrhineae (Fig. 5a, b; Supporting Information Figs. S6, S7). Multiple increases in diversification rates were detected by both analyses under the three taxonomic treatments, with higher rate heterogeneity detected under the *splitter* and *intermediate* treatments than under the *lumper* treatment, as shown by mean phylorate plots and macroevolutionary cohort matrices (Supporting Information Figs. S6, S7). Increases in diversification rate did not generally coincide with the acquisition of nectar spurs, although the majority of increases occurred within clades displaying nectar spurs (numbers 1-4 in Fig. 5; see also Supporting Information Figs. S6, S7):

1. In the *Anarrhinum-Kickxia-Nanorrhinum* clade (number 1 in Fig. 5), a rate increase was detected under the *splitter* and *intermediate* treatments, either at the base of *Nanorrhinum* or at the base of *Nanorrhinum+Kickxia*.
2. In *Cymbalaria* (number 2 in Fig. 5), no shift was detected by MEDUSA, but a possible subtle rate increase was detected by BAMM at the base of the clade under the *splitter* treatment.
3. In *Chaenorhinum* (number 3 in Fig. 5), a shift was detected at the base of a predominantly western Mediterranean clade under the *splitter* and *intermediate* treatments.
4. In the highly diversified *Linaria* (number 4 in Fig. 5), two likely rate increases were found under the three taxonomic treatments, one of them at the base of *Linaria* subsect. *Versicolores*, and the other at the base of a large clade formed by species of the following sections: *Linaria* sect. *Linaria*, *Linaria* sect. *Speciosae*, *Linaria* sect. *Diffusae* and *Linaria* sect. *Supinae*.

The only shift affecting a spurless lineage was a rate increase at the base of *Antirrhinum* (number 8 in Fig. 5; see also Supporting Information Figs. S6, S7).

Rate-through-time plots estimated by BAMM (see Fig. 5c for results under the *splitter* treatment) depicted a similar pattern for three of the four spurred clades, with an initial phase of 5-15 million years with low diversification rate (similar to that of most spurless lineages) followed by a burst of diversification that extends to the present. The exception to this pattern among spurred clades is the recently originated *Cymbalaria*, for which no burst was inferred. Most spurless lineages maintained a constantly low diversification rate, the exception being *Antirrhinum*, an Old World clade with a much higher diversification rate than the closely related *Sairocarpus* clade from the New World (Fig. 5c). On average, higher diversification rates were inferred for spurred clades than for spurless clades in BAMM analyses, and the difference increased over time according to RTT plots, more markedly under the *splitter* than under the *intermediate* and *lumper* treatments (Fig. 2c).

Spur length evolution in Linaria

A significant positive correlation between log-transformed spur length and corolla length in *Linaria* was inferred by PGLS ($F_{1,152} = 191.9$, $R^2 = 0.558$, $P < 2.2\times10^{-16}$;

Supporting Information Fig. S8). Ancestral state reconstructions (Supporting Information Fig. S9) showed recurrent changes in both log-transformed spur length and spur/corolla ratio, particularly conspicuous in some of the most diversified clades. Rates of phenotypic evolution estimated by BAMM (Supporting Information Fig. S10) were largely homogeneous across *Linaria*, with substantial rate increases detected only at a limited number of small terminal clades. According to the CoMET analysis, the best-fitting models of character evolution were the speciational model for spur length and the non-phylogenetic model for spur/corolla ratio (Table 4).

Discussion

A comprehensive phylogenetic framework for the Antirrhineae

Through the combination of previously published and newly generated sequence data, we have generated the most comprehensive phylogenetic hypothesis for the Antirrhineae published to date, comprising 84% of the 297 species recognised under the lumper treatment (Table 1; Fig. 1; Supporting Information Figs. S1, S2). Compared to the phylogeny of Guzmán *et al.* (2015), our species sampling represents an increase of 125% in the number of taxa, as well as a more balanced representation of clades and geographical regions. In addition, our analysis (like that of Guzmán *et al.*, 2015) is based on a carefully curated set of DNA sequences, avoiding taxonomic misidentifications that led to phylogenetic misplacements in some earlier studies (misplacement of *Galvezia fruticosa* in Vargas *et al.*, 2004; misplacement of *Gambelia speciosa* and *Schweinfurthia pterosperma*, and misnaming of *Gambelia juncea* as *Galvezia juncea* in Ogutcen & Vamosi, 2016; Ogutcen *et al.*, 2017) (see also Guzmán *et al.*, 2015).

Despite the relatively fragmentary nature of the DNA sequence matrix (resulting from the combination of sequences of different DNA regions used in previous partial studies), we recovered the 17 major generic lineages of Antirrhineae (Fig. 1b), and phylogenetic relationships among them were highly resolved and mostly consistent with those inferred by Guzmán *et al.* (2015). Phylogenetic dating estimated that crown diversification of the Antirrhineae started in the Eocene, although most of the extant species diversity seems to have been generated since the late Miocene (Fig. 1c; see also Vargas *et al.*, 2014). As expected, many recent divergences among closely related species were poorly supported, probably as a result of rapid radiation. Although genome-wide data may be necessary to further resolve recent radiations (see Fernández-

Mazuecos *et al.*, 2018), the extensive time-calibrated phylogeny presented here provides a robust framework for ongoing research into the evolution and development of snapdragons and relatives (e.g. Hileman *et al.*, 2003; Feng *et al.*, 2009; Box *et al.*, 2011; Bradley *et al.*, 2017).

Nectar spurs originated multiple times during Antirrhineae evolution

Our phylogenetic hypothesis highlights the heterogeneous diversification of the Antirrhineae, with closely related generic lineages accounting for contrasting numbers of extant species (Fig. 1b). This observation leads to the search for biotic and abiotic factors potentially driving diversification rate variation (Donoghue & Sanderson, 2015). Floral nectar spurs have long been suggested as a key innovation promoting diversification in angiosperms (Hodges & Arnold, 1995; Hodges, 1997), and their presence in the (by far) most diverse genus of Antirrhineae (*Linaria*) would suggest a crucial role in diversification of the tribe.

Just as nectar spurs evolved independently in numerous angiosperm families (Hodges, 1997; Fernández-Mazuecos & Glover, 2017), ancestral state reconstructions support the idea that spurred lineages originated four times from spurless ancestors during diversification of the Antirrhineae (Fig. 1c; Supporting Information Figs. S3, S4). This result is robust to the use of alternative models and approaches. While developmental mechanisms generating nectar spurs seem to be different in distantly related families (Box *et al.*, 2011; Puzey *et al.*, 2012; Yant *et al.*, 2015; Cullen *et al.*, 2018), nothing is known about the degree to which the same genetic and developmental changes may have underlain the multiple origins of spurs in Antirrhineae (parallelism; Scotland, 2011). Future evo-devo studies may shed light on this question.

Nectar spurs are loosely associated with increased diversification in Antirrhineae

The multiple acquisitions of nectar spurs confirmed by ancestral state reconstructions make the Antirrhineae an ideal system to investigate the macroevolutionary dynamics of the trait, and particularly to test the hypothesis of a recurrent positive effect of spurs on diversification rates. This recurrent effect would support the role of nectar spurs as a key innovation (Hodges, 1997; Kay *et al.*, 2006). The use of alternative taxonomic treatments (Table 1; Supporting Information Table S5) had obvious effects on the results of our diversification rate analyses in Antirrhineae, with higher levels of rate heterogeneity recovered by treatments recognising higher numbers of species (Figs. 2, 4, 5; Supporting Information Figs. S4-S7) (see Faurby *et al.*, 2016). Nevertheless, the *splitter* and *intermediate* treatments are probably the most realistic based on our knowledge of Antirrhineae diversity, and they produced qualitatively similar results, leading to the same conclusions regarding the effect of spurs on diversification.

Models of trait-dependent speciation were clearly supported against simple constant-rate models under the BiSSE framework implemented in *diversitree*, with higher speciation rates inferred for spurred than for spurless lineages, in agreement with the key innovation hypothesis (Tables 2, 3; Fig. 2a). However, it is well known that BiSSE analyses are prone to false positives due to phylogenetic pseudoreplication and the use of trivial null models (Maddison & FitzJohn, 2014; Rabosky & Goldberg, 2015). Indeed, when some of these issues are accounted for under the HiSSE framework (Beaulieu & O'Meara, 2016), the clear effect found in BiSSE analyses becomes blurred. For the *splitter* and *intermediate* treatments, a BiSSE model (in which diversification rates exclusively depend on the presence or absence of nectar spurs) is rejected against models in which rate heterogeneity depends on a combination of nectar spurs and other unmeasured factors or is character-independent (Tables 2, 3). On average, diversification rates are still higher for spurred than for spurless lineages, but there is a large overlap in values estimated for the two character states across the phylogeny (Fig. 2b; Table 3). The nonparametric FiSSE tests, also robust to some of the issues described for BiSSE, failed to support an effect of spurs on diversification (Fig. 3), although the statistical power of this method is known to be low (Rabosky & Goldberg, 2017).

When applying trait-independent methods to investigate diversification, rate heterogeneity across the Antirrhineae phylogeny was clearly detected (Fig. 5; Supporting Information Figs. S6, S7). Rate-through-time plots obtained in BAMM for spurred and spurless lineages revealed an early period of overlap in estimated diversification rates, followed by a period of increasingly higher rates for spurred than for spurless lineages (Fig. 2c). Plots for three of the four spurred clades also depicted a pattern of delayed radiation after spur acquisition (Fig. 5c). Indeed, detected rate shifts did not generally coincide with the acquisition of nectar spurs (except for a possible subtle increase at the base of *Cymbalaria*), but several rate increases were nested within spurred clades (Fig. 5a, b; Supporting Information Figs. S6, S7). A similar pattern is depicted by the model-averaged *hisse* reconstruction under the *splitter* treatment, with low diversification rates estimated for early-diverging lineages within spurred clades, and higher rates obtained for several recently-diversified lineages (Fig. 4). A lag between the evolution of a putative key innovation and radiation has been frequently observed, not only for nectar spurs (*Halenia*: von Hagen & Kadereit, 2003; *Impatiens*: Janssens *et al.*, 2009), but also for other traits including the angiosperm flower itself (Tank *et al.*, 2015).

As explanation for this pattern of delayed radiation, several authors have proposed that additional intrinsic and extrinsic factors may be needed in conjunction with the trait of interest to trigger diversification (Bouchenak-Khelladi *et al.*, 2015). These factors include developmental robustness, additional phenotypic traits, and ecological opportunities (Donoghue & Sanderson, 2015; Melzer & Theißen, 2016). First, it is likely that a robust developmental determination of spur length is required before an effect on diversification rates can be observed (Melzer & Theißen, 2016). Second, additional traits possibly interacting with nectar spurs in promoting pollinator specialisation and diversification include the personate corolla, with different levels of occlusion and tube length in Antirrhineae (Sutton, 1988; Guzmán *et al.*, 2015; Guzmán *et al.*, 2017); breeding systems also seem to influence diversification, at least in *Linaria* (Blanco-Pastor & Vargas, 2013). And third, ecological opportunities triggering diversification in the Antirrhineae may include those provided by historical climate changes in the Mediterranean basin, as well as migration to previously unoccupied regions in the New World and Asia (Vargas *et al.*, 2009; Fernández-Mazuecos & Vargas, 2011; Blanco-Pastor & Vargas, 2013; Fernández-Mazuecos *et al.*, 2013a; Fernández-Mazuecos *et al.*, 2013b; Vargas *et al.*, 2014; Carnicero *et al.*, 2017; Vargas *et al.*, 2018). Since floral divergence is rarely sufficient to drive speciation in sympatry (Kay & Sargent, 2009), these historical events have probably been critical in promoting differentiation in allopatry, as indicated by the non-overlapping distributions of closely related narrow endemics of many clades of Antirrhineae (Sutton, 1988). For example, the Quaternary climatic cycles are thought to have promoted geographical isolation accompanied by divergent selection on floral traits driven by geographical differences in pollinator fauna, as suggested for *Linaria* (Blanco-Pastor *et al.*, 2015).

Ultimately, it is clear that a combination of factors (i.e. "confluence" sensu Donoghue & Sanderson, 2015) needs to be invoked to explain the increased diversification rates in certain clades of Antirrhineae (see also Sauquet & Magallón, 2018; Vamosi *et al.*, 2018). For example, the high diversification rates of *Linaria* subsect. *Versicolores* and *Linaria* subsect. *Supinae* (Fig. 5) may have been favoured by their specialised, predominantly self-incompatible flowers with both an occluded personate corolla and a nectar spur, together with Mediterranean conditions and climate changes since the late Miocene (Fernández-Mazuecos & Vargas, 2011; Blanco-Pastor *et al.*, 2012; Blanco-Pastor & Vargas, 2013; Fernández-Mazuecos *et al.*, 2013a; Blanco-Pastor *et al.*, 2015; Fernández-Mazuecos *et al.*, 2018). High diversification rates are also possible in the absence of nectar spurs, as shown by *Antirrhinum*, where geographic speciation under Mediterranean conditions and pollinator specialisation by evolution of corolla length are proposed as main drivers (Vargas *et al.*, 2009; Vargas *et al.*, 2010; Wilson & Hudson, 2011; Vargas *et al.*, 2017). In a similar way, nectar spurs are not the only driver of diversification in the Ranunculaceae genus *Aquilegia*. While spur length changes and pollinator shifts were crucial in diversification of the American clade studied by Whittall and Hodges (2007), that is not the case of the similarly diverse Eurasian lineages, where geographical isolation and habitat shifts played a more important role (Bastida *et al.*, 2010).

Changes in spur length are associated with speciation events in Linaria

The phylogenetic distribution of spur length, the variable putatively related to speciation, provides an additional test for the key innovation hypothesis (Bouchenak-Khelladi *et al.*, 2015; Donoghue & Sanderson, 2015). In *Linaria*, the most diverse genus in the Antirrhineae, a speciational model best explains evolution of spur length (Table 4), implying that changes in this trait preferentially occurred at speciation events. This result, consistent with a “pollinator shift” scenario, is similar to that reported for American *Aquilegia* by Whittall and Hodges (2007), in which a punctuated model was supported. Recurrent changes in spur length are depicted by the ancestral state reconstruction (Supporting Information Fig. S9), and rates of spur length change seem to have remained relatively homogeneous throughout the diversification of the Antirrhineae (Supporting Information Fig. S10). Unlike in American *Aquilegia*, where changes in spur length were mainly driven by shifts between bee, hummingbird and hawkmoth pollination (Whittall & Hodges, 2007), major changes in pollination syndrome do not seem to have been relevant in *Linaria*. Most studied species are bee-pollinated, with some pollinated by lepidopterans and a few generalists (Arnold, 1982; Sánchez-Lafuente, 2007; Fernández-Mazuecos *et al.*, 2013a; Blanco-Pastor *et al.*, 2015; Guzmán *et al.*, 2017). A role of spurs in pollinator specialisation and species differentiation has been shown in *Linaria* subsect. *Supinae*, where species with the most slender spurs have evolved repeatedly and are pollinated by bees with a longer proboscis (Blanco-Pastor *et al.*, 2015). Similarly, in *Linaria* subsect. *Versicolores*, spur length plays a role, in conjunction with tube width, in determining pollinator strategies in closely related species (Fernández-Mazuecos *et al.*, 2013a; Fernández-Mazuecos *et al.*, 2018).

Spur length not only evolves in response to pollinators. It also seems to be developmentally constrained to some degree by corolla size, and therefore evolves in correlation with this trait (Supporting Information Fig. S8). For example, some of the shortest spurs in *Linaria* are found in species with tiny corollas that appear to have evolved as a result of self-fertilisation (Segarra-Moragues & Mateu-Andrés, 2007; Blanco-Pastor & Vargas, 2013). After accounting for corolla size, evolution of the spur/corolla ratio also displays a pattern of recurrent changes with relatively homogeneous rates (Supporting Information Figs. S9, S10), and model testing supports a non-phylogenetic model (Table 4), indicating a high evolutionary lability. Although spur length (following a speciational model) is probably more relevant to pollinator specialisation than the spur/corolla ratio, the interaction between spur length, spur width, corolla size and corolla shape deserves further developmental and evolutionary research.

Is floral specialisation the cause or the consequence of species diversity?

Several possible mechanisms have been proposed to account for the observed correlation between specialised floral traits (such as nectar spurs) and clade species diversity (Armbruster & Muchhala, 2009; Armbruster, 2014): (1) specialisation may promote the establishment and reinforcement of pre-pollination reproductive barriers through floral isolation, leading to increased speciation rates and thus to high species diversity; (2) specialisation can also increase reproductive success and enable the occupation of narrower pollination niches, which will diminish the negative effects of interspecific pollination and enable the packing of more species into communities, leading to decreased extinction rates and thus to high species diversity; and conversely, (3) high clade species diversity may cause selection for partitioning of pollinator fauna and character displacement between sympatric relatives, therefore leading to floral specialisation. While the first explanation is the basis for the key innovation hypothesis as applied to nectar spurs, the remaining two have rarely been considered. Our results provide some insights to determine the relative importance of these three mechanisms in the Antirrhineae. Specialisation being the consequence, and not the cause, of high diversification (mechanism 3) can be ruled out on the basis of the delayed radiation of spurred clades. Effects of specialisation on speciation (mechanism 1) and extinction (mechanism 2) are hard to distinguish given the difficulties in estimating extinction rates from molecular phylogenies of extant taxa (Rabosky, 2010; Beaulieu & O'Meara, 2016). On the one hand, BiSSE analyses detected significant differences in speciation rates, but not in extinction rates, between spurred and spurless lineages (Table 2; Fig. 2a). On the other hand, HiSSE analyses under the *splitter* taxonomic treatment suggest that the higher mean diversification rate of spurred lineages may be the result of a combination of higher speciation rates and lower extinction rates (Supporting Information Fig. S5). The speciational evolution of spur length in *Linaria* additionally supports a role of this trait in speciation (Table 3). Taken together, our evidence is consistent with an effect of floral specialisation (in combination with other factors) on speciation, although an additional effect on extinction cannot be ruled out. These results confirm that macroevolutionary studies can provide key insights into the relationship between floral specialisation and evolutionary success (Armbruster, 2014; O'Meara *et al.*, 2016).

Conclusions

Here we provided a comprehensive and robust phylogenetic framework for evolutionary studies in snapdragons and relatives. Multiple acquisitions of nectar spurs during Antirrhineae diversification were supported by evolutionary reconstructions. Although nectar spurs are widely considered a key innovation promoting diversification in flowering plants, they are only loosely associated with increased diversification rates in Antirrhineae. Still, the fact that spur length evolves following a speciational model in *Linaria*, the most diverse genus, is consistent with a “pollinator shift” scenario, supporting a relevant role of spurs in diversification. Diversification rate heterogeneity in Antirrhineae is likely determined by a complex interaction of biotic and abiotic factors, including nectar spurs and other specialised floral traits, breeding systems, developmental robustness, historical climate changes and biogeographic events causing geographical isolation. The concept of “key innovation” is useful as a starting point of diversification analyses, but a more nuanced approach incorporating a variety of biotic and abiotic factors, as proposed by Donoghue and Sanderson (2015), is required to fully understand diversification dynamics in flowering plants.

Acknowledgements

We thank Ana Otero, Javier Igea and Rafael Rubio de Casas for useful discussion on diversification analyses; Mathew Dorling for technical support; Cecilia Martínez for providing photographs; S.M.M. Hamdi for providing information on Iranian species; Wayne Elisens, Jaime Güemes, Jorge Alfredo Reyes Betancort, Isabel Liberal and the MA, E, W, K, BCN, BC, B and RSA herbaria for providing plant materials. This work was supported by the Marie Curie Intra-European Fellowship *LINARIA-SPECIATION* (FP7-PEOPLE-2013-IEF, reference 624396), an Isaac Newton Trust Research Grant (Trinity College, Cambridge), a Juan de la Cierva fellowship to M.F.-M. (Spanish Ministry of Economy and Competitivity, reference IJCI-2015-23459) and a Generalitat Valenciana postdoctoral grant to A.J. (Ministry of Education, Culture and Sport, reference BEST/2014/264).

Author Contributions

M.F.-M. and B.J.G. designed the research; M.F.-M., J.L.B.-P., A.J., P.C., A.F., M.A. and P.V. collected data; M.F.-M. analysed data; M.F.-M., P.V. and B.J.G. interpreted results; M.F.-M. wrote the paper with feedback from all authors.

References

Albach DC, Meudt HM, Oxelman B. 2005. Piecing together the "new" Plantaginaceae. *American Journal of Botany* **92**: 297-315.

Alfaro ME, Santini F, Brock C, Alamillo H, Dornburg A, Rabosky DL, Carnevale G, Harmon LJ. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 13410-13414.

Armbruster WS. 2014. Floral specialization and angiosperm diversity: phenotypic divergence, fitness trade-offs and realized pollination accuracy. *AoB Plants* **6**: plu003.

Armbruster WS, Muchhala N. 2009. Associations between floral specialization and species diversity: cause, effect, or correlation? *Evolutionary Ecology* **23**: 159-179.

Arnold RM. 1982. Pollination, predation and seed set in *Linaria vulgaris* (Scrophulariaceae). *American Midland Naturalist* **107**: 360-369.

Bastida JM, Alcántara JM, Rey PJ, Vargas P, Herrera CM. 2010. Extended phylogeny of *Aquilegia*: the biogeographical and ecological patterns of two simultaneous but contrasting radiations. *Plant Systematics and Evolution* **284**: 171-185.

Beaulieu JM, O'Meara BC. 2016. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology* **65**: 583-601.

Bell CD, Soltis DE, Soltis PS. 2010. The age and diversification of the angiosperms re-revisited. *American Journal of Botany* **97**: 1296-1303.

Benedí C, Güemes J 2009. *Chaenorhinum* (DC.) Rchb. In: Castroviejo S, Herrero A, Benedí C, Rico E, Güemes J, eds. *Flora iberica XIII (Plantaginaceae – Scrophulariaceae)*. Madrid: CSIC, 167-198.

Blanco-Pastor JL, Ornosa C, Romero D, Liberal IM, Gómez JM, Vargas P. 2015. Bees explain floral variation in a recent radiation of *Linaria*. *Journal of Evolutionary Biology* **28**: 851-863.

Blanco-Pastor JL, Vargas P. 2013. Autecological traits determined two evolutionary strategies in Mediterranean plants during the Quaternary: low differentiation and range expansion versus geographical speciation in *Linaria*. *Molecular Ecology* **22**: 5651-5668.

Blanco-Pastor JL, Vargas P, Pfeil BE. 2012. Coalescent simulations reveal hybridization and incomplete lineage sorting in Mediterranean *Linaria*. *PLOS ONE* **7**: e39089.

Bouchenak-Khelladi Y, Onstein RE, Xing Y, Schwery O, Linder HP. 2015. On the complexity of triggering evolutionary radiations. *New Phytologist* **207**: 313-326.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10**: e1003537.

Box MS, Dodsworth S, Rudall PJ, Bateman RM, Glover BJ. 2011. Characterization of *Linaria* KNOX genes suggests a role in petal-spur development. *The Plant Journal* **68**: 703-714.

Bradley D, Xu P, Mohorianu II, Whibley A, Field D, Tavares H, Couchman M, Copsey L, Carpenter R, Li M et al. 2017. Evolution of flower color pattern through selection on regulatory small RNAs. *Science* **358**: 925-928.

Carnicero P, Sáez L, Garcia-Jacas N, Galbany-Casals M. 2017. Different speciation types meet in a Mediterranean genus: the biogeographic history of *Cymbalaria* (Plantaginaceae). *Taxon* **66**: 393-407.

Cullen E, Fernández-Mazuecos M, Glover BJ. 2018. Evolution of nectar spur length in a clade of *Linaria* reflects changes in cell division rather than in cell expansion. *Annals of Botany* **122**: 801-809.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772-772.

De Queiroz A, Gatesy J. 2007. The supermatrix approach to systematics. *Trends in Ecology & Evolution* **22**: 34-41.

Dodd ME, Silvertown J, Chase MW. 1999. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. *Evolution* **53**: 732-744.

Donoghue MJ, Sanderson MJ. 2015. Confluence, synnovation, and depauperons in plant diversification. *New Phytologist* **207**: 260-274.

Dorofeev PI. 1963. *Tretichnye flory Zapadnoĭ Sibiri*. Moscow: Izdatel'stvo Akademii Nauk SSSR.

Erwin DH. 1992. A preliminary classification of evolutionary radiations. *Historical Biology* **6**: 133-147.

Faurby S, Eiserhardt WL, Svenning JC. 2016. Strong effects of variation in taxonomic opinion on diversification analyses. *Methods in Ecology and Evolution* **7**: 4-13.

Feng X, Wilson Y, Bowers J, Kennaway R, Bangham A, Hannah A, Coen E, Hudson A. 2009. Evolution of allometry in *Antirrhinum*. *The Plant Cell* **21**: 2999-3007.

Fernández-Mazuecos M, Blanco-Pastor JL, Gómez JM, Vargas P. 2013a. Corolla morphology influences diversification rates in bifid toadflaxes (*Linaria* sect. *Versicolores*). *Annals of Botany* **112**: 1705-1722.

Fernández-Mazuecos M, Blanco-Pastor JL, Vargas P. 2013b. A phylogeny of toadflaxes (*Linaria* Mill.) based on nuclear internal transcribed spacer sequences: systematic and evolutionary consequences. *International Journal of Plant Sciences* **174**: 234-249.

Fernández-Mazuecos M, Glover BJ. 2017. The evo-devo of plant speciation. *Nature Ecology & Evolution* **1**: 0110.

Fernández-Mazuecos M, Mellers G, Vigalondo B, Sáez L, Vargas P, Glover BJ. 2018. Resolving recent plant radiations: power and robustness of genotyping-by-sequencing. *Systematic Biology* **67**: 250-268.

Fernández-Mazuecos M, Vargas P. 2011. Historical isolation *versus* recent long-distance connections between Europe and Africa in bifid toadflaxes (*Linaria* sect. *Versicolores*). *PLOS ONE* **6**: e22234.

Fior S, Li M, Oxelman B, Viola R, Hodges SA, Ometto L, Varotto C. 2013. Spatiotemporal reconstruction of the *Aquilegia* rapid radiation through next-generation sequencing of rapidly evolving cpDNA regions. *New Phytologist* **198**: 579-592.

FitzJohn RG. 2012. Diversitree: comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution* **3**: 1084-1092.

FitzJohn RG, Maddison WP, Otto SP. 2009. Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic Biology* **58**: 595-611.

Fulton M, Hodges SA. 1999. Floral isolation between *Aquilegia formosa* and *Aquilegia pubescens*. *Proceedings of the Royal Society B: Biological Sciences* **266**: 2247-2252.

Ghebrehiwet M. 2000. Taxonomy, phylogeny and biogeography of *Kickxia* and *Nanorrhinum* (Scrophulariaceae). *Nordic Journal of Botany* **20**: 655-690.

Ghebrehiwet M, Bremer B, Thulin M. 2000. Phylogeny of the tribe Antirrhineae (Scrophulariaceae) based on morphological and *ndh*F sequence data. *Plant Systematics and Evolution* **220**: 223-239.

Goldberg EE, Igić B. 2008. On phylogenetic tests of irreversible evolution. *Evolution* **62**: 2727-2741.

Grafen A. 1989. The phylogenetic regression. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **326**: 119-157.

Güemes J. 1997. *Pseudomisopates* (Scrophulariaceae), un nuevo género ibérico. *Anales del Jardín Botánico de Madrid* **55**: 492-493.

Güemes J 2009. *Antirrhinum* L. In: Castroviejo S, Herrero A, Benedí C, Rico E, Güemes J, eds. *Flora iberica XIII (Plantaginaceae – Scrophulariaceae).* Madrid: CSIC, 134-166.

Güemes J, Mota JF. 2017. *Gadoria* (Antirrhineae, Plantaginaceae): a new genus, endemic from Sierra de Gádor, Almería, Spain. *Phytotaxa* **298**: 201-221.

Guzmán B, Gómez JM, Vargas P. 2015. Bees and evolution of occluded corollas in snapdragons and relatives (Antirrhineae). *Perspectives in Plant Ecology, Evolution and Systematics* **17**: 467-475.

Guzmán B, Gómez JM, Vargas P. 2017. Is floral morphology a good predictor of floral visitors to Antirrhineae (snapdragons and relatives)? *Plant Biology* **19**: 515-524.

Heard SB, Hauser DL. 1995. Key evolutionary innovations and their ecological mechanisms. *Historical Biology* **10**: 151-173.

Hileman LC, Kramer EM, Baum DA. 2003. Differential regulation of symmetry genes and the evolution of floral morphologies. *Proceedings of the National Academy of Sciences* **100**: 12814-12819.

Hodges SA. 1997. Floral nectar spurs and diversification. *International Journal of Plant Sciences* **158**: 81-88.

Hodges SA, Arnold ML. 1995. Spurring plant diversification: are floral nectar spurs a key innovation? *Proceedings of the Royal Society B: Biological Sciences* **262**: 343-348.

Hunter JP. 1998. Key innovations and the ecology of macroevolution. *Trends in Ecology & Evolution* **13**: 31-36.

Igea J, Miller EF, Papadopulos AST, Tanentzap AJ. 2017. Seed size and its rate of evolution correlate with species diversification across angiosperms. *PLOS Biology* **15**: e2002792.

Janssens SB, Knox EB, Huysmans S, Smets EF, Merckx VSFT. 2009. Rapid radiation of *Impatiens* (Balsaminaceae) during Pliocene and Pleistocene: result of a global climate change. *Molecular Phylogenetics and Evolution* **52**: 806-824.

Kampstra P. 2008. Beanplot: A boxplot alternative for visual comparison of distributions. *Journal of Statistical Software, Code Snippets* **28**: 1-9.

Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* **9**: 286-298.

Kay KM, Sargent RD. 2009. The role of animal pollination in plant speciation: integrating ecology, geography, and genetics. *Annual Review of Ecology, Evolution, and Systematics* **40**: 637-656.

Kay KM, Voelckel C, Yang JY, Hufford KM, Kaska DD, Hodges SA 2006. Floral characters and species diversification. In: Harder LD, Barrett SCH, eds. *Ecology and evolution of flowers*. Oxford: Oxford University Press, 311-325.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647-1649.

Lee C, Blay S, Mooers AØ, Singh A, Oakley TH. 2007. CoMET: A Mesquite package for comparing models of continuous character evolution on phylogenies. *Evolutionary Bioinformatics* **2006**: 183-186.

Maddison WP, FitzJohn RG. 2014. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic Biology* **64**: 127-136.

Maddison WP, Maddison DR 2011. Mesquite: a modular system for evolutionary analysis. [WWW document] URL http://mesquiteproject.org [accessed 26 June 2018].

Mai HD. 2001. Die mittelmiozänen und obermiozänen Floren aus der Meuroer und Raunoer Folge in der Lausitz. Teil III: Fundstellen und Paläobiologie. *Palaeontographica Abteilung B* **258**: 1-85.

Melzer R, Theißen G. 2016. The significance of developmental robustness for species diversity. *Annals of Botany* **117**: 725-732.

Minelli A. 2015. Species diversity vs. morphological disparity in the light of evolutionary developmental biology. *Annals of Botany* **117**: 781-794.

Moore BR, Höhna S, May MR, Rannala B, Huelsenbeck JP. 2016. Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the National Academy of Sciences* **113**: 9569-9574.

O'Meara BC, Smith SD, Armbruster WS, Harder LD, Hardy CR, Hileman LC, Hufford L, Litt A, Magallón S, Smith SA et al. 2016. Non-equilibrium dynamics and floral trait interactions shape extant angiosperm diversity. *Proceedings of the Royal Society B: Biological Sciences* **283**: 20152304.

Ogutcen E, Theriault J, King DB, Vamosi JC. 2017. Diversification rates in Antirrhineae (Plantaginaceae): the contribution of range shifts and pollination modes. *Perspectives in Plant Ecology, Evolution and Systematics* **26**: 39-52.

Ogutcen E, Vamosi JC. 2016. A phylogenetic study of the tribe Antirrhineae: genome duplications and long-distance dispersals from the Old World to the New World. *American Journal of Botany* **103**: 1071-1081.

Orme D 2012. The *caper* package: comparative analysis of phylogenetics and evolution in R. [WWW document] URL http://caper.r-forge.r-project.org/ [accessed 26 June 2018].

Oyama RK, Baum DA. 2004. Phylogenetic relationships of North American *Antirrhinum* (Veronicaceae). *American Journal of Botany* **91**: 918-925.

Puzey JR, Gerbode SJ, Hodges SA, Kramer EM, Mahadevan L. 2012. Evolution of spur-length diversity in *Aquilegia* petals is achieved solely through cell-shape anisotropy. *Proceedings of the Royal Society B: Biological Sciences* **279**: 1640-1645.

Rabosky DL. 2010. Extinction rates should not be estimated from molecular phylogenies. *Evolution* **64**: 1816-1824.

Rabosky DL. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLOS ONE* **9**: e89543.

Rabosky DL, Goldberg EE. 2015. Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic Biology* **64**: 340-355.

Rabosky DL, Goldberg EE. 2017. FiSSE: A simple nonparametric test for the effects of a binary character on lineage diversification rates. *Evolution* **71**: 1432-1442.

Rabosky DL, Grundler M, Anderson CL, Shi JJ, Brown JW, Huang H, Larson JG. 2014. BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution* **5**: 701-707.

Rahmani A, Nejadsatari T, Hamdi SMM, Mehregan I, Assadi M. 2014. A phylogenetic analysis of *Linaria* (Plantaginaceae) species from Iran based on ITS sequence data. *European Journal of Experimental Biology* **4**: 127-134.

Rambaut A, Suchard MA, Xie D, Drummond AJ 2014. Tracer v1.6. [WWW document] URL http://tree.bio.ed.ac.uk/software/tracer/ [accessed 26 June 2018].

Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**: 217-223.

Revell LJ. 2013. Two new graphical methods for mapping trait evolution on phylogenies. *Methods in Ecology and Evolution* **4**: 754-759.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**: 539-542.

Sáez L, Bernal M 2009. *Linaria* Mill. In: Castroviejo S, Herrero A, Benedí C, Rico E, Güemes J, eds. *Flora iberica XIII (Plantaginaceae – Scrophulariaceae)*. Madrid: CSIC, 232-324.

Sánchez-Lafuente AM. 2007. Corolla herbivory, pollination success and fruit predation in complex flowers: an experimental study with *Linaria lilacina* (Scrophulariaceae). *Annals of Botany* **99**: 355-364.

Sargent RD. 2004. Floral symmetry affects speciation rates in angiosperms. *Proceedings of the Royal Society B: Biological Sciences* **271**: 603-608.

Sauquet H, Magallón S. 2018. Key questions and challenges in angiosperm macroevolution. *New Phytologist* **219**: 1170-1187.

Schäferhoff B, Fleischmann A, Fischer E, Albach DC, Borsch T, Heubl G, Müller KF. 2010. Towards resolving Lamiales relationships: insights from rapidly evolving chloroplast sequences. *BMC Evolutionary Biology* **10**: 352.

Scotland RW. 2011. What is parallelism? *Evolution & Development* **13**: 214-227.

Segarra-Moragues JG, Mateu-Andrés I. 2007. Levels of allozyme diversity in closely related toadflaxes (*Linaria*, Plantaginaceae) and their correspondence with the breeding systems of the species. *Conservation Genetics* **8**: 373-383.

Shivanna KR 2014. Biotic pollination: how plants achieve conflicting demands of attraction and restriction of potential pollinators. In: Ramawat KG, Mérillon JM, Shivanna KR, eds. *Reproductive Biology of Plants*. Delhi: CRC Press, 218-267.

Slowinski JB, Guyer C. 1993. Testing whether certain traits have caused amplified diversification: an improved method based on a model of random speciation and extinction. *The American Naturalist* **142**: 1019-1024.

Sutton DA. 1988. *A revision of the tribe Antirrhineae*. Oxford: Oxford University Press.

Tank DC, Eastman JM, Pennell MW, Soltis PS, Soltis DE, Hinchliff CE, Brown JW, Sessa EB, Harmon LJ. 2015. Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist* **207**: 454-467.

Vamosi JC, Magallón S, Mayrose I, Otto SP, Sauquet H. 2018. Macroevolutionary patterns of flowering plant speciation and extinction. *Annual Review of Plant Biology* **69**: 685-706.

Vargas P, Carrió E, Guzmán B, Amat E, Güemes J. 2009. A geographical pattern of *Antirrhinum* (Scrophulariaceae) speciation since the Pliocene based on plastid and nuclear DNA polymorphisms. *Journal of Biogeography* **36**: 1297-1312.

Vargas P, Fernández-Mazuecos M, Heleno R. 2018. Phylogenetic evidence for a Miocene origin of Mediterranean lineages: species diversity, reproductive traits and geographical isolation. *Plant Biology* **20**: 157-165.

Vargas P, Liberal I, Ornosa C, Gómez JM. 2017. Flower specialisation: the occluded corolla of snapdragons (*Antirrhinum*) exhibits two pollinator niches of large long-tongued bees. *Plant Biology* **19**: 787-797.

Vargas P, Ornosa C, Ortiz-Sanchez FJ, Arroyo J. 2010. Is the occluded corolla of *Antirrhinum* bee-specialized? *Journal of Natural History* **44**: 1427-1443.

Vargas P, Rosselló JA, Oyama R, Güemes J. 2004. Molecular evidence for naturalness of genera in the tribe Antirrhineae (Scrophulariaceae) and three independent evolutionary lineages from the New World and the Old. *Plant Systematics and Evolution* **249**: 151-172.

Vargas P, Valente LM, Blanco-Pastor JL, Liberal I, Guzmán B, Cano E, Forrest A, Fernández-Mazuecos M. 2014. Testing the biogeographical congruence of palaeofloras using molecular phylogenetics: snapdragons and the Madrean-Tethyan flora. *Journal of Biogeography* **41**: 932-943.

von Hagen KB, Kadereit JW. 2003. The diversification of *Halenia* (Gentianaceae): ecological opportunity versus key innovation. *Evolution* **57**: 2507-2518.

Whittall JB, Hodges SA. 2007. Pollinator shifts drive increasingly long nectar spurs in columbine flowers. *Nature* **447**: 706-709.

Wilson Y, Hudson A. 2011. The evolutionary history of *Antirrhinum* suggests that ancestral phenotype combinations survived repeated hybridizations. *The Plant Journal* **66**: 1032-1043.

Yang Z, Kumar S, Nei M. 1995. A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics* **141**: 1641-1650.

Yant L, Collani S, Puzey JR, Levy C, Kramer EM. 2015. Molecular basis for three-dimensional elaboration of the *Aquilegia* petal spur. *Proceedings of the Royal Society B: Biological Sciences* **282**: 20142778.

Yousefi N, Zarre S, Heubl G. 2016. Molecular phylogeny of the mainly Mediterranean genera *Chaenorhinum*, *Kickxia* and *Nanorrhinum* (Plantaginaceae, tribe

Antirrhineae), with focus on taxa in the Flora Iranica region. *Nordic Journal of Botany* **34**: 455-463.

Figures

Fig. 1 Phylogeny and evolution of nectar spurs in Antirrhineae. (a) Species diversity in spurred vs. spurless genera of Antirrhineae. The pie chart on the left represents the proportion of spurred and spurless species of Antirrhineae according to the *splitter* taxonomic treatment (see Table 1). Species number of the six spurred genera is represented on the right. (b) Phylogeny of Antirrhineae at generic level obtained in BEAST. All nodes had a posterior probability (PP) ≥ 0.95. For each generic lineage, a range of estimated species diversity according to different taxonomic treatments is indicated. Pie charts at tips indicate proportions of spurred and spurless species. The four spurred clades are numbered (1-4). (c) Time-calibrated phylogeny of Antirrhineae at species level obtained in BEAST. The maximum clade credibility tree is shown. Pie charts at nodes and branch colours represent probabilities of ancestral states for spur presence/absence according to the stochastic character mapping analysis under the different rates (DR) model. Spur lengths are shown at tips, and flowers of representatives of major clades are shown on the right, with nectar spurs indicated with red arrows. Photos of *Kickxia*, *Mabrya* and *Chaenorhinum* by Cecilia Martínez; *Cymbalaria*, *Galvezia*, *Antirrhinum* and *Linaria* by Mario Fernández-Mazuecos.

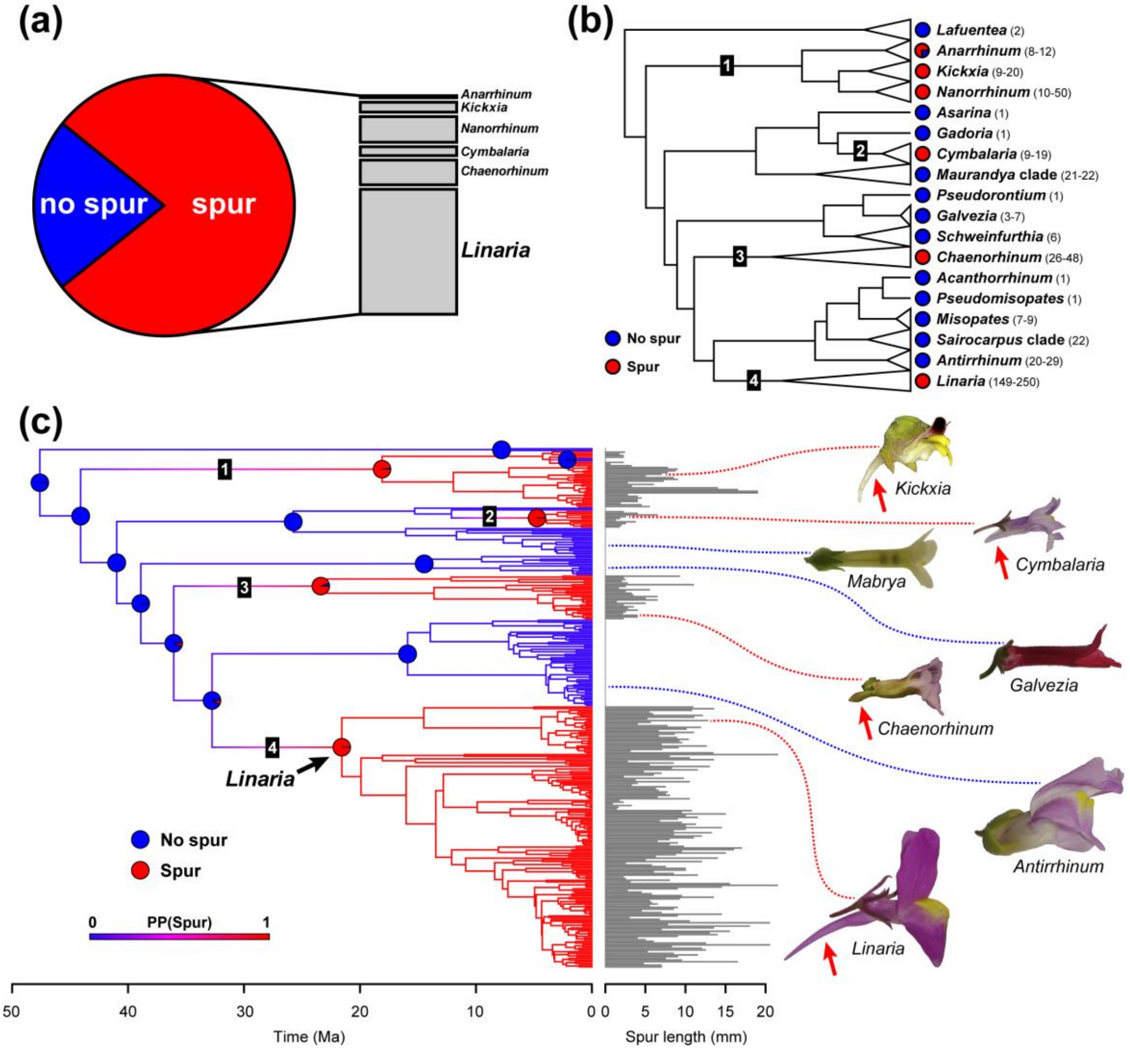
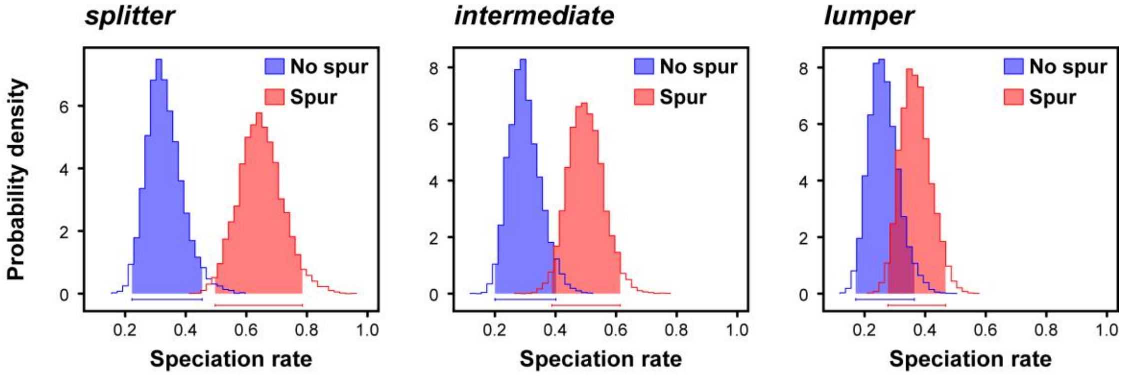
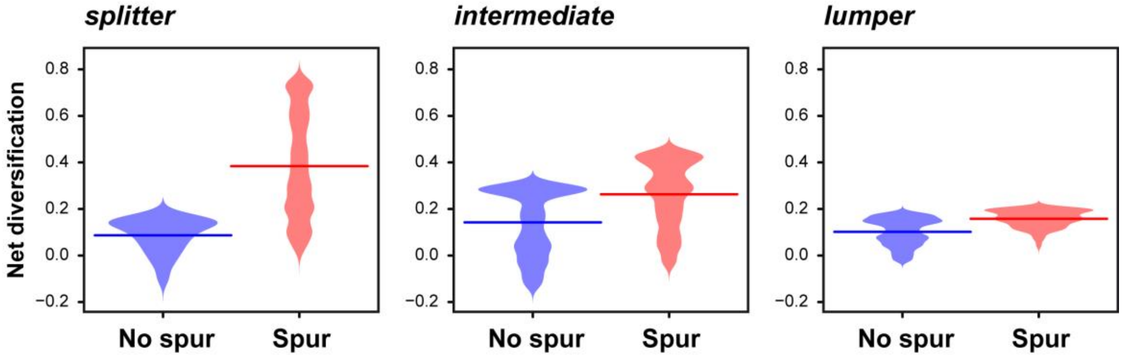

(a)
no spur
spur
Anarrhinum
Kickxia
Nanorrhinum
Cymbalaria
Chaenorhinum
Linaria
(b)
1
2
3
4
Lafuentea (2)
Anarrhinum (8-12)
Kickxia (9-20)
Nanorrhinum (10-50)
Asarina (1)
Gadoria (1)
Cymbalaria (9-19)
Maurandya clade (21-22)
Pseudorontium (1)
Galvezia (3-7)
Schweinfurthia (6)
Chaenorhinum (26-48)
Acanthorrhinum (1)
Pseudomisopates (1)
Misopates (7-9)
Sairocarpus clade (22)
Antirrhinum (20-29)
Linaria (149-250)
● No spur
● Spur
(c)
1
2
3
4
Linaria
● No spur
● Spur
0 PP(Spur) 1
50
40
30
20
10
0
Time (Ma)
0
5
10
15
20
Spur length (mm)
Kickxia
Cymbalaria
Mabrya
Galvezia
Chaenorhinum
Antirrhinum
Linaria

Fig. 2 Differences in diversification rates estimated for spurred and spurless lineages of Antirrhineae under three methods and three alternative taxonomic treatments (*splitter*, *intermediate*, *lumper*). (a) Results of BiSSE analyses implemented in *diversitree* considering nectar spur absence and presence as character states; the Bayesian posterior distributions of speciation rates under the full BiSSE model are shown; horizontal bars indicate 95% credibility intervals. (b) Results of HiSSE analyses implemented in *hisse*; beanplots represent variation in net diversification estimated across tips and nodes after averaging four models (CID-2, CID-4, full BiSSE, full HiSSE); horizontal bars indicate mean values. (c) Diversification rate-through-time plots estimated by BAMM analyses; shading represents confidence intervals.

(a) *diversitree* (full BiSSE)



(b) *hisse* (model average)



(c) BAMM

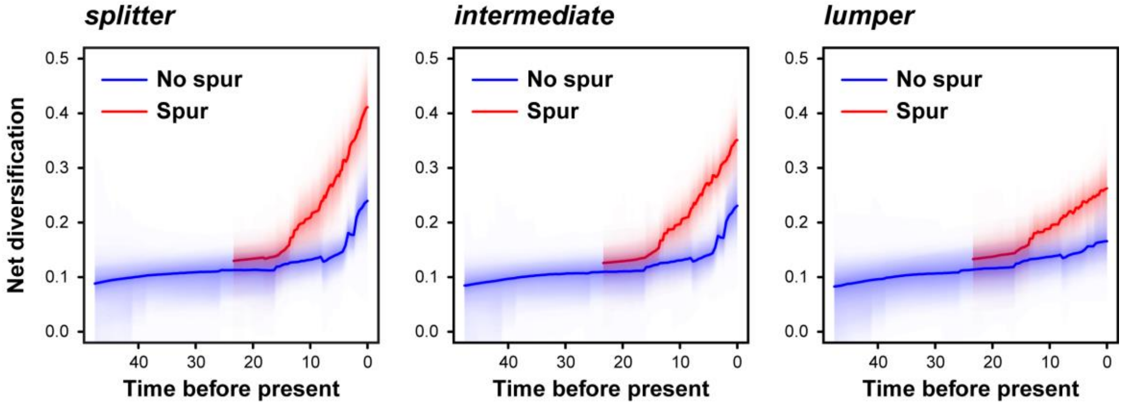


Fig. 3 Summary of FiSSE tests in Antirrhineae considering nectar spur absence and presence as character states. The histogram represents the distribution of two-tailed *P*-values obtained for 1000 completely sampled phylogenies simulated under each of three alternative taxonomic treatments. A significance level α=0.05 is indicated. *P*-values above this level indicate a lack of significant differences in diversification rates between spurred and spurless lineages.

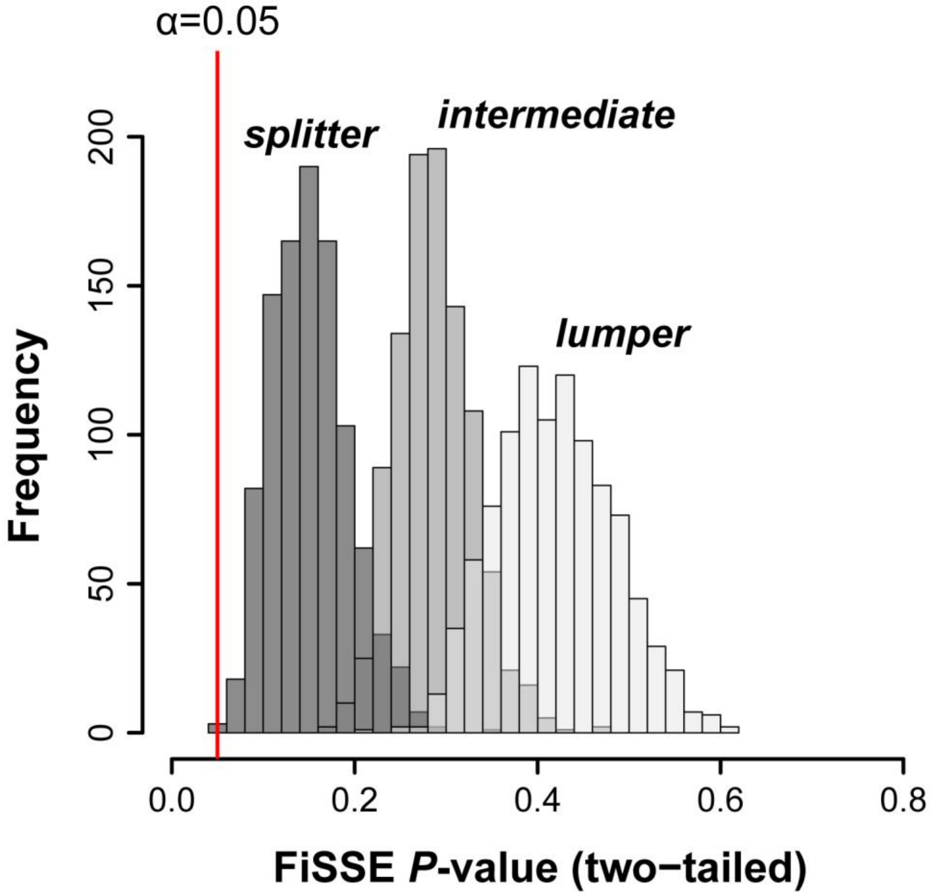


Fig. 4 Model-averaged marginal reconstruction of diversification rates and nectar spur evolution in Antirrhineae obtained using the *hisse* package under the *splitter* taxonomic treatment. Four models are averaged: CID-2, CID-4, full BiSSE and full HiSSE (see text for details). Diversification rates are represented as colour shading along branch edges (blue to red). Spur presence/absence is represented as black/white shading inside branches. The inset represents the distribution of diversification rates and character states across the tree.

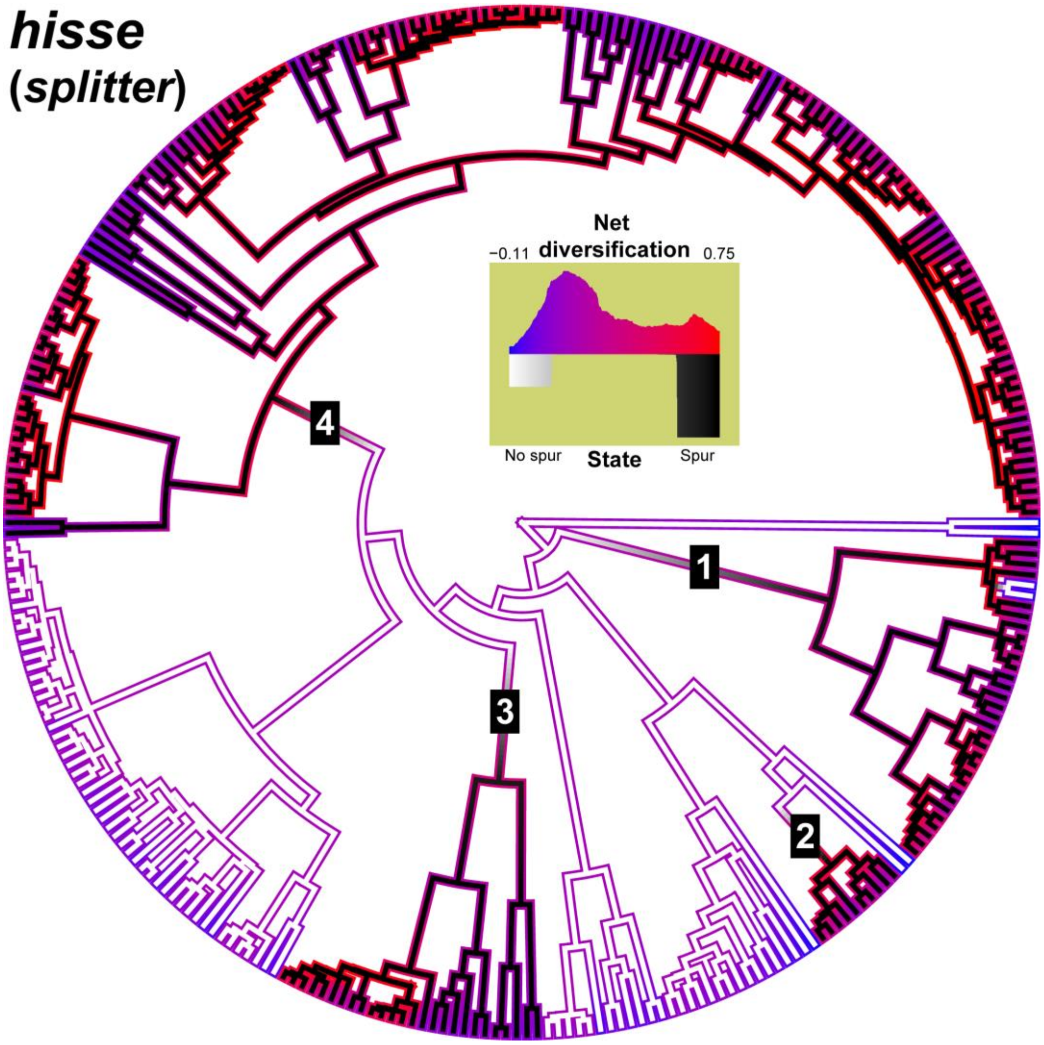
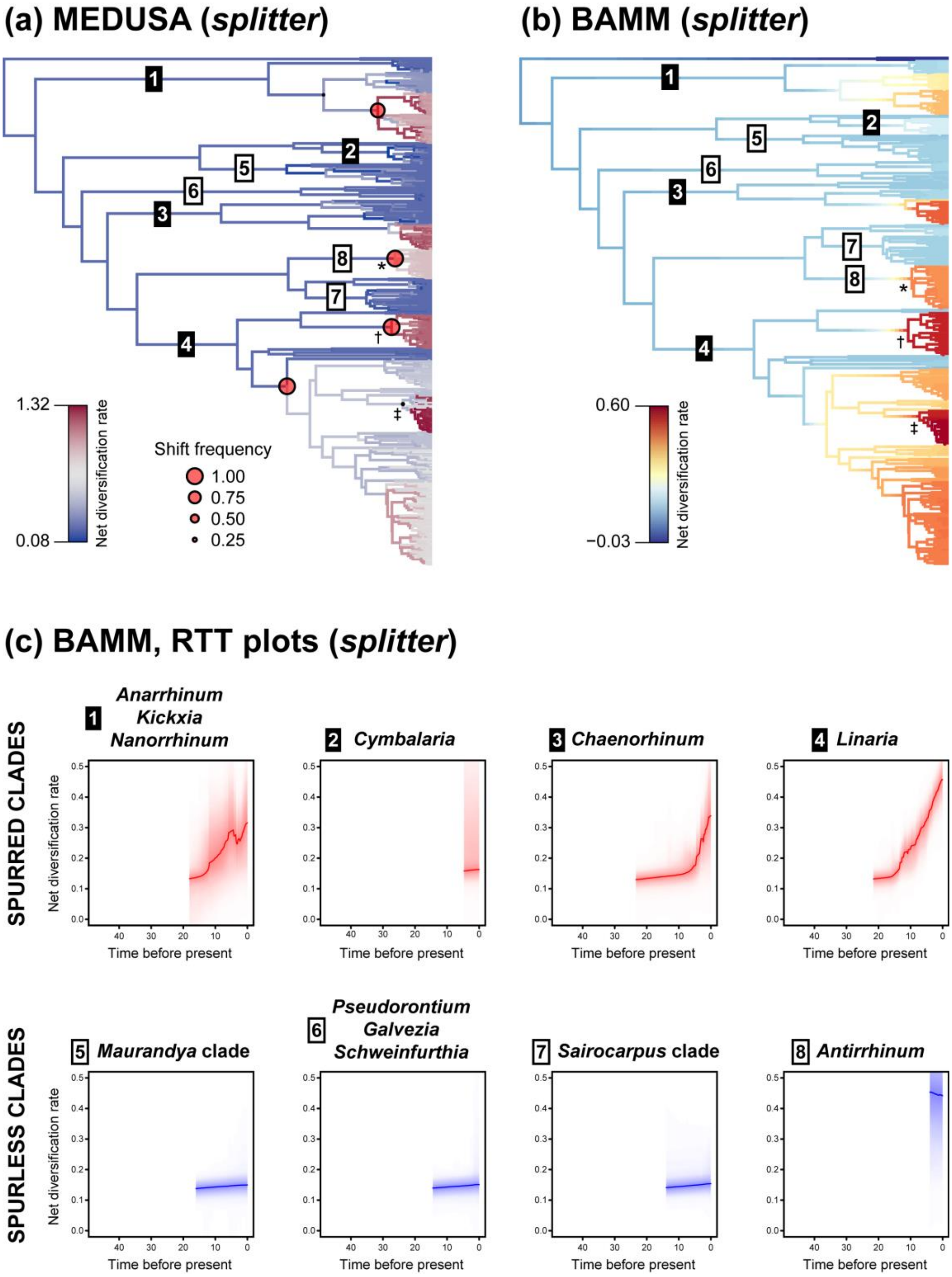


Fig. 5 Results of trait-independent analyses of diversification rates in Antirrhineae under the *splitter* taxonomic treatment. (a) MEDUSA analysis of 1000 completely sampled simulated phylogenies, summarised on a single randomly chosen tree. Branch colours represent estimates of diversification rates. Circles indicate rate shifts, with sizes representing their frequency. (b) Mean phylorate plot from the BAMM analysis. Colours represent mean, model-averaged diversification rates. In a and b, asterisks (*), daggers (†) and double daggers (‡) indicate the position of *Antirrhinum*, *Linaria* subsect. *Versicolores* and *Linaria* subsect. *Supinae* respectively (see Discussion). (c) Diversification rate-through-time plots estimated by BAMM for the four spurred clades and four selected spurless clades. Spurred (1-4) and spurless (5-8) clades are numbered in the three panels. See Supporting Information Figs. S6 and S7 for results under the *intermediate* and *lumper* treatments.


(a) MEDUSA (*splitter*)
(b) BAMM (*splitter*)
1
2
3
4
5
6
7
8
*
†
‡
1.32
0.08
Net diversification rate
Shift frequency
1.00
0.75
0.50
0.25
0.60
−0.03
(c) BAMM, RTT plots (*splitter*)
SPURRED CLADES
SPURLESS CLADES
1 *Anarrhinum Kickxia Nanorrhinum*
2 *Cymbalaria*
3 *Chaenorhinum*
4 *Linaria*
5 *Maurandya* clade
6 *Pseudorontium Galvezia Schweinfurthia*
7 *Sairocarpus* clade
8 *Antirrhinum*
0.0
0.1
0.2
0.3
0.4
0.5
40
30
20
10
0
Time before present

Tables

Table 1 Summary of three alternative taxonomic treatments of Antirrhineae (and *Lafuentea*) considered for diversification rate analyses.

Generic lineage	*splitter*		*intermediate*		*lumper*	
	Total No. of species	No. of species in phylogeny	Total No. of species	No. of species in phylogeny	Total No. of species	No. of species in phylogeny
Acanthorrhinum	1	1	1	1	1	1
Anarrhinum	12	8	8	8	8	8
Antirrhinum	29	26	27	26	20	19
Asarina	1	1	1	1	1	1
Chaenorhinum	48	26	35	25	26	20
Cymbalaria	19	10	12	10	9	8
Gadoria	1	1	1	1	1	1
Galvezia	7	5	4	4	3	3
Kickxia	20	10	10	8	9	8
Lafuentea	2	2	2	2	2	2
Linaria	250	154	194	150	149	124
Maurandya clade	22	16	21	16	21	16
Misopates	9	5	8	5	7	5
Nanorrhinum	50	15	44	13	10	6
Pseudomisopates	1	1	1	1	1	1
Pseudorontium	1	1	1	1	1	1
Sairocarpus clade	22	18	22	18	22	18
Schweinfurthia	6	6	6	6	6	6
TOTAL	501	306 (61%)	398	296 (74%)	297	248 (84%)

For each generic lineage and taxonomic treatment, total number of species and number of species sampled in our phylogenetic analysis are shown. Sampling percentage is indicated in brackets for each taxonomic treatment.

Table 2 Log-likelihood and AIC values of diversification models under three alternative taxonomic treatments of Antirrhineae evaluated using the *diversitree* and *hisse* packages.

Model	*splitter*			*intermediate*			*lumper*		
	lnL	AIC	ΔAIC	lnL	AIC	ΔAIC	lnL	AIC	ΔAIC
diversitree									
$\lambda_0=\lambda_1$, $\mu_0=\mu_1$, $q_{01}=q_{10}$ (constant rates)	-770.5	1546.9	20.7	-746.3	1498.6	15.3	-660.4	1326.7	9.1
$\lambda_0=\lambda_1$, $\mu_0\neq\mu_1$, $q_{01}\neq q_{10}$	-762.1	1534.1	7.9	-739.0	1488.0	4.7	-654.8	1319.6	2.0
$\lambda_0=\lambda_1$, $\mu_0=\mu_1$, $q_{01}\neq q_{10}$	-768.8	1545.5	19.3	-744.7	1497.4	14.1	-659.0	1326.0	8.4
$\lambda_0=\lambda_1$, $\mu_0\neq\mu_1$, $q_{01}=q_{10}$	-762.4	1532.9	6.6	-739.3	1486.7	3.4	**-655.1**	**1318.2**	**0.6**
$\lambda_0\neq\lambda_1$, $\mu_0=\mu_1$, $q_{01}\neq q_{10}$	**-758.9**	**1527.7**	**1.5**	**-736.8**	**1483.6**	**0.4**	**-653.9**	**1317.9**	**0.3**
$\lambda_0\neq\lambda_1$, $\mu_0=\mu_1$, $q_{01}=q_{10}$	**-759.6**	**1527.2**	**1.0**	**-737.6**	**1483.3**	**0.0**	**-654.8**	**1317.6**	**0.0**
$\lambda_0\neq\lambda_1$, $\mu_0\neq\mu_1$, $q_{01}=q_{10}$	**-758.9**	**1527.7**	**1.5**	**-737.6**	**1485.2**	**1.9**	**-654.6**	**1319.3**	**1.7**
$\lambda_0\neq\lambda_1$, $\mu_0\neq\mu_1$, $q_{01}\neq q_{10}$ (full BiSSE)	**-757.1**	**1526.3**	**0.0**	**-736.4**	**1484.9**	**1.6**	-653.9	1319.9	2.3
hisse									
CID-2	-736.7	1497.3	2.9	-729.6	1483.1	23.3	**-647.1**	**1318.2**	**0.6**
CID-4	-738.8	1499.6	5.1	**-718.9**	**1459.8**	**0.0**	**-647.8**	**1317.6**	**0.0**
Full BiSSE	-749.6	1511.2	16.8	-733.7	1479.4	19.6	**-653.2**	**1318.4**	**0.8**
Full HiSSE	**-731.2**	**1494.4**	**0.0**	**-714.7**	**1461.4**	**1.6**	-649.8	1331.6	14.0

For each *diversitree* model, parameters (λ, speciation rate; μ, extinction rate; q, character transition rate) were set to be equal or different between character states (0, no spur; 1, spur). Four models were tested in *hisse*: a character-independent diversification model with two hidden states (CID-2); a character-independent diversification model with four hidden states (CID-4); a full binary-state speciation and extinction model (full BiSSE); and a full hidden-state speciation and extinction model (full HiSSE). Models within 2 AIC units of the best model are shown in bold.

Table 3 Diversification parameter estimates (λ, speciation rate; μ, extinction rate; r, net diversification rate; 0, no spur; 1, spur) obtained using the *diversitree* and *hisse* packages under three alternative taxonomic treatments of Antirrhineae.

	splitter	*intermediate*	*lumper*
diversitree (best model)			
λ_0	0.357	0.356	0.283
λ_1	0.703	0.493	0.387
μ_0	0.300	0.311	0.228
μ_1	0.538	0.311	0.228
r_0	0.057	0.046	0.056
r_1	0.165	0.183	0.159
hisse (model average)			
λ_0	0.433 (0.100)	0.519 (0.143)	0.397 (0.078)
λ_1	0.631 (0.144)	0.626 (0.168)	0.524 (0.053)
μ_0	0.346 (0.028)	0.377 (0.005)	0.295 (0.020)
μ_1	0.247 (0.081)	0.363 (0.030)	0.366 (0.015)
r_0	0.087 (0.073)	0.142 (0.141)	0.102 (0.058)
r_1	0.384 (0.225)	0.263 (0.143)	0.158 (0.038)

For *diversitree*, maximum likelihood parameter estimates under the best-fitting BiSSE model (see Table 2) are shown. For *hisse*, reported values are means and standard deviations (in brackets) across tips and nodes obtained after model averaging.

Table 4 Comparison of evolutionary models for spur length and spur/corolla ratio in the genus *Linaria* obtained in CoMET.

Character	Model	lnL	AIC	ΔAIC	Scalar
ln (Spur length)	Gradual	1.8	-1.6	22.3	0.1
	Speciational	**13.0**	**-24.0**	**0.0**	**0.1**
	Punctuated	-5.2	12.6	36.6	0.4
	Non-phylogenetic	10.8	-19.6	4.4	0.3
Spur length / corolla length ratio	Gradual	63.0	-124.0	63.3	0.1
	Speciational	85.2	-168.4	19.0	0.0
	Punctuated	60.8	-122.1	65.2	0.2
	Non-phylogenetic	**94.7**	**-187.3**	**0.0**	**0.1**

The best-fitting models (lowest AIC values) are shown in bold.

Supporting Information

Fig. S1 Bayesian phylogenetic tree of Antirrhineae based on analysis of ITS, *ndh*F and *rpl*32-*trn*L sequences in MrBayes.

Fig. S2 Time-calibrated phylogeny of Antirrhineae obtained in BEAST.

Fig. S3 Ancestral state reconstructions of nectar spur presence/absence in Antirrhineae under different rates, equal rates and BiSSE models.

Fig. S4 Model-averaged marginal reconstructions of diversification rates and nectar spur evolution in Antirrhineae obtained using the *hisse* package under the *splitter*, *intermediate* and *lumper* taxonomic treatments.

Fig. S5 Speciation and extinction rates estimated by *hisse* analyses under the *splitter*, *intermediate* and *lumper* taxonomic treatments.

Fig. S6 MEDUSA analyses of diversification rates in Antirrhineae under the *splitter*, *intermediate* and *lumper* taxonomic treatments.

Fig. S7 BAMM analyses of diversification rates in Antirrhineae under the *splitter*, *intermediate* and *lumper* taxonomic treatments.

Fig. S8 Scatterplot from the phylogenetic generalised least squares (PGLS) analysis testing the correlation between log-transformed spur length and corolla length in *Linaria*.

Fig. S9 Ancestral state reconstructions of spur length and spur/corolla ratio in *Linaria*.

Fig. S10 BAMM analyses of phenotypic evolutionary rates for spur length and spur/corolla ratio in *Linaria*.

Table S1 GenBank accession numbers for both previously published and newly generated DNA sequences of Antirrhineae and the outgroup used in the present study.

Table S2 Voucher specimens for newly-sequenced species of Antirrhineae and the outgroup.

Table S3 Fossil record of Antirrhineae.

Table S4 Fossil calibrations used in the dating analysis of Antirrhineae.

Table S5 Species recognised under the *splitter*, *intermediate* and *lumper* taxonomic treatments of Antirrhineae used in diversification analyses.