A phylogeny of Antirrhinum reveals parallel evolution of alpine morphology

Mario Durán-Castillo
Andrew Hudson
Yvette Wilson
David L. Field
Edith Cowan University
Alex D. Twyford

Follow this and additional works at: https://ro.ecu.edu.au/ecuworkspost2013

Part of the Terrestrial and Aquatic Ecology Commons

Durán-Castillo, M., Hudson, A., Wilson, Y., Field, D. L., & Twyford, A. D. (2021). A phylogeny of Antirrhinum reveals parallel evolution of alpine morphology. *New Phytologist*. Advance online publication. https://doi.org/10.1111/nph.17581

This Journal Article is posted at Research Online. https://ro.ecu.edu.au/ecuworkspost2013/11389
A phylogeny of *Antirrhinum* reveals parallel evolution of alpine morphology

Mario Durán-Castillo1, Andrew Hudson2, Yvette Wilson2, David L. Field3 and Alex D. Twyford1,4

1Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, EH9 3FL, UK; 2Institute of Molecular Plant Sciences, University of Edinburgh, Edinburgh, EH9 3BF, UK; 3School of Science, Edith Cowan University, 270 Joondalup Drive, Joondalup 6027, Australia; 4Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh, EH3 5LR, UK

**Summary**

- Parallel evolution of similar morphologies in closely related lineages provides insight into the repeatability and predictability of evolution. In the genus *Antirrhinum* (snapdragons), as in other plants, a suite of morphological characters are associated with adaptation to alpine environments.
- We tested for parallel trait evolution in *Antirrhinum* by investigating phylogenetic relationships using restriction-site associated DNA (RAD) sequencing. We then associated phenotypic information to our phylogeny to reconstruct the patterns of morphological evolution and related this to evidence for hybridisation between emergent lineages.
- Phylogenetic analyses showed that the alpine character syndrome is present in multiple groups, suggesting that *Antirrhinum* has repeatedly colonised alpine habitats. Dispersal to novel environments happened in the presence of intraspecific and interspecific gene flow.
- We found support for a model of parallel evolution in *Antirrhinum*. Hybridisation in natural populations, and a complex genetic architecture underlying the alpine morphology syndrome, support an important role of natural selection in maintaining species divergence in the face of gene flow.

**Introduction**

Parallel phenotypic evolution – the repeated evolution of similar morphologies in closely related lineages – has long fascinated biologists, as it provides insight into the repeatability and predictability of evolution (Elmer & Meyer, 2011). In some animal groups, such as cichlid fish, there are numerous striking examples in which many aspects of body morphology, coloration and behaviour show similar phenotypic responses to particular environmental conditions (Salzburger, 2009; Elmer et al., 2014; Oke et al., 2017). In plants, morphological parallelism has been characterised in dwarf and tall *Eucalyptus* (Foster et al., 2007), sand dune and rocky headland *Senecio* (Roda et al., 2013), alpine and montane *Heliosperma* (Trucchi et al., 2017), and beach, estuary and spring *Cochlearia* (Brandrud et al., 2017). These studies revealed that certain suites of traits are often under selection in challenging and stressful environmental conditions such as high salinity, high elevation and extreme exposure, and more generally highlight the central role of natural selection in shaping phenotypic trait evolution.

The first stage in understanding the genetic basis of parallel evolution is to know the relationships between populations and species. A well resolved phylogeny is essential for confirming parallel evolution, while the integration of phenotypic data with molecular phylogeny can reveal the spatial and temporal context of phenotypic evolution (Rosenblum et al., 2014). However, many examples of parallel evolution come from closely related taxa in which phylogenetic reconstruction is hampered by low genetic divergence of species, incomplete lineage sorting and hybridisation (Twyford & Ennos, 2012; Fernández-Mazuecos et al., 2018). Next-generation sequencing approaches can generate a wealth of sequence data to address phylogenetic questions, but numerous phylogenetic issues remain problematic. For example, many genomic phylogenies overestimate branch support, while those branches with low support fail to distinguish low information content from the presence of multiple highly supported incongruent phylogenetic histories (Pease et al., 2018), both of which are expected in recent radiations. Characterising the sources of phylogenetic uncertainty can also help to identify the evolutionary mechanism of allele sharing. If phylogenetic uncertainty reflects hybridisation, then introgression may cause adaptive alleles to be shared among divergent species. However if introgression is rare, or can be ruled out, then parallel evolution is more likely to be due to independent mutations or adaptation from standing genetic variation within species (Wilding et al., 2014). To address these challenges we require tractable empirical systems with taxa that span a range of divergence times, together with broad geographic distributions that provide the opportunity for parallel evolution in allopatry.

Here we investigated parallel phenotypic evolution in snapdragons (*Antirrhinum*, Plantaginaceae). *Antirrhinum*, in particular *A. majus*, has long been a model system for the study of
pigment biosynthesis and photosynthetic pathways and their regulation, plant growth and development, transposons and self-incompatibility (Coen et al., 1986; Coen & Meyerowitz, 1991; Luo et al., 1996; Xue et al., 1996; Whibley et al., 2006; Hudson et al., 2008a). *Antirrhinum* is also becoming an important model system in evolutionary biology for the study of barriers to gene flow (Ringbauer et al., 2018). This includes work on genomic islands of divergence (Tavares et al., 2018) and the molecular genetic basis of traits responsible for reproductive isolation (Bradley et al., 2017). The genus *Antirrhinum* includes c. 20 species across the Mediterranean basin, and exhibits rich variation in flower, leaf-shape and branching traits. The ability of all species to form fertile hybrids (Hudson et al., 2008b), the presence of natural hybrid zones (Tavares et al., 2018), and the related morphologies of taxa, suggest a recent radiation of *Antirrhinum* species. Dated phylogenetic analysis of plastid sequence data suggested an origin for the radiation of *Antirrhinum* within the last 5 Myr (Vargas et al., 2009), with many species diverging recently and with relationships that have proved difficult to resolve (Zweitler et al., 2002; Jiménez et al., 2005; Mateu-Andrés & De Paco, 2005; Vargas et al., 2009; Wilson & Hudson, 2011; Liberal et al., 2014).

The genus *Antirrhinum* is separated into three morphological subsections (Rothmaler, 1956), with species from the two largest subsections, *Kickxiella* and *Antirrhinum*, showing suites of traits adaptive to contrasting environments. Species from *Antirrhinum* subsection *Kickxiella* are small prostrate alpines or xerophytes that have highly branched stems, small ovate hairy leaves, and small pale flowers (Fig. 1). Most of these species are endemics of a few mountains across Iberia where they grow as true alpines on rocky cliffs (including mountains over 2500 m elevation, Liberal et al., 2014). By contrast, species from *Antirrhinum* subsection *Antirrhinum* are large upright unbranched plants with large elongated leaves and magenta-pink or yellow flowers. These species are generally found in more competitive habitats with dense vegetation, including low elevation grasslands. The third, smaller subsection, *Antirrhinum* subsection *Streptosepalum*, are morphologically intermediate between the other two subsections, with a generally tall and upright growth habit, long thin leaves and large yellow flowers. Comparative morphological analyses show the morphological divergence of subsections *Kickxiella* and *Antirrhinum* represent the primary axis of morphological divergence in the genus (Wilson & Hudson, 2011).

Our study focuses on the case of putative parallel evolution of the *Kickxiella* morphology, which may have allowed species to colonise cliffs and rocky surfaces repeatedly. Previous population genetic analyses with amplified fragment length polymorphisms have suggested that species from subsection *Kickxiella* do not belong to a single genetic cluster, and instead fall within at least two divergent species groups (Wilson & Hudson, 2011). However, the previous study was unable to resolve relationships within this recent species radiation and could not identify the number of *Kickxiella* groups, the direction of morphological evolution, and whether introgression could underlie adaptation. The primary aims of this study were to resolve the relationships between *Antirrhinum* taxa and to test for the parallel phenotypic evolution of traits that characterise species of subsection *Kickxiella*. We reconstruct the phylogeny of *Antirrhinum* using dense restriction-site associated DNA (RAD) sequence data generated for species from across the genus. We then scored phenotypic characters and associated them with our phylogeny to reconstruct patterns of morphological evolution. Our evidence suggests that the genus *Antirrhinum* has repeatedly evolved suites of morphological traits allowing them to explore new ecological opportunities, with multiple independent species groups possessing alpine *Kickxiella* morphology allowing them to grow in challenging alpine

![Fig. 1](image1.png) Contrasting morphologies of the three *Antirrhinum* subsections. Pairs of images showing the growth habit and morphology of plants in the field and in the laboratory. From left to right: *A. pseudomajus* population L053 showing typical subsection *Antirrhinum* morphology; *A. molle* population E051 showing typical subsection *Kickxiella* morphology; *A. meananthum* population L118 showing typical subsection *Streptosepalum* morphology. Bars, 2 cm. [Correction added after online publication 1 August 2021: the middle-upper panel in the figure has been updated.]
conditions of dry rocky cliff faces. Signatures of hybridisation suggest that introgression may have been involved in parallel evolution of the *Kickxiella* morphology.

**Materials and Methods**

**Study species**

*Antirrhinum* species are short-lived perennial herbs or small shrubs that are diploids \((2n = 2x = 16)\), and most of which are self-incompatible (Schwarz-Sommer *et al.*, 2003). *Antirrhinum* are renowned for their bilaterally symmetrical flowers with large colourful petals, with the occluded corolla facilitating exclusive pollination by bees (Vargas *et al.*, 2010). This specialised pollination system is also characterised by conical epidermal cells, floral scent and flowers that often possess pollination guides. *Antirrhinum* species demonstrate extensive variation in organ shape and size (Feng *et al.*, 2009), with many of these traits under selection across the diverse habitats that *Antirrhinum* inhabits. For example, leaf surface area is associated with water limitation, with species inhabiting near desert environments producing smaller leaves than species found in wet or seasonally wet grasslands. The genus is exclusively found in Western Europe, with most diversity found in the Iberian Peninsula. Most taxa are ecologically specialised and geographically restricted with isolated populations (Forrest *et al.*, 2017), although some taxa are found more widely.

**Sampling and sequencing**

We sampled 120 individuals from 28 *Antirrhinum* taxa representing the geographic and taxonomic range of the genus. Most samples were collected from the wild, with these supplemented with additional samples from *Antirrhinum* collection holders (Supporting information Table S1). Our sampling represented seven main geographic regions: Portugal, Northern and Central Spain, the Sierra Nevada and South of Spain, Morocco, the Pyrenees, the Alps and Italy (Fig. 2). Wild-collected seeds were germinated and grown under the glasshouse conditions specified in Hudson *et al.* (2008b). Genomic DNA was extracted from 100 mg silica dried or fresh tissue frozen at \(-80^\circ\text{C}\), following a modified CTAB method (Doyle & Doyle, 1987). RAD libraries were prepared from *Pst*I-digested DNA following the method of Miller *et al.* (2007), with custom combinatorial indexing of P1 and P2 adaptors. Libraries were pooled and sequenced with 100-bp paired-end reads using the Illumina HiSeq-4000 platform at Edinburgh Genomics.

**Alignment of RAD data**

Raw reads were demultiplexed using the *process_radtags* script from the Stacks software (Catchen *et al.*, 2013). TRIMMOMATIC 0.36 (Bolger *et al.*, 2014) was used to remove adaptor sequences, clip sequences with a phred score of \(\leq 20\) and remove any reads shorter than 30 bp. Filtered reads were mapped to the *A. majus* genome (Li *et al.*, 2019) using *Bowtie2* (Langmead & Salzberg, 2012) and duplicate sequences were removed using Picard tools (Broad Institute, 2018). SNPs were called using *SAMtools* 1.6 and the multiallelic caller implemented in *Bcftools* 1.4 (Li, 2011), retaining invariant sites. This dataset was then filtered by mapping quality \((\geq 40)\), depth \((\geq 3x)\) and missing data, both per taxon (removing individuals with \(>70\%\) missing data) and per site (removing sites present in less than 50% of individuals). The final data included 16 061 293 sites from 86 samples corresponding to 24 taxa.

To root the phylogenetic trees we used available whole genome sequence data from *Misopates orontiunm* (A. Whibley & E. Coen, unpublished). *Misopates* has been shown to belong to the *Antirrhinum* clade, and diverged from the genus *Antirrhinum* in the last 10–15 Myr (Ogutcn and Vamosi, 2016). Variant calling was carried out as above retaining only the loci present in the alignment of *Antirrhinum* samples.

**Phylogenetic analyses**

To resolve species relationships in the recent radiation of *Antirrhinum* we used a combination of maximum likelihood and coalescent phylogenetic approaches. Maximum likelihood analysis of unpartitioned concatenated sequence data is a popular approach that scales well to large genomic datasets. However, such methods may lead to incorrect inferences in the presence of incomplete lineage sorting (Vachaspati & Warnow, 2018). Multispecies coalescent methods account for incomplete lineage sorting by directly inferring the species tree from the site patterns, and are often used to complement maximum likelihood analyses and understand sources of incongruence. We also used multiple methods to estimate phylogenetic support and potential causes of phylogenetic conflict.

A maximum likelihood analysis was conducted on the concatenated dataset of variant and invariant sites using RAxML (Stamatakis, 2014) with a GTR-GAMMA substitution model, as recommended by Stamatakis (2014). Initial branch support was assessed using the rapid bootstrap option with 100 replicates. As concatenated phylogenetic approaches using many nuclear loci often overestimate phylogenetic support (Chou *et al.*, 2015), we further explored levels of conflict in our data by comparing the topologies of a consensus tree obtained with RAxML from an independent run of 1000 bootstrap replicates, with 200 replicates of the quartet sampling method as described in Pease *et al.* (2018).

Quartet sampling (QS) evaluates for each node the observed unrooted topology of four taxa versus the two discordant topologies in terms of three quartet scores: concordance (QC), with a value of 1 when all quartets are concordant), differential (QD, with a value close to zero when one alternative topology is favoured over the other) and informativeness (QI, assessing whether any lack of support reflects low information content).

We also reconstructed phylogenetic relationships using SVDquartets (Chifman & Kubatko, 2015). The SVDquartets algorithm requires unlinked multilocus data and therefore we filtered the previous alignment using the function *thin* in *vcftools* (Danecek *et al.*, 2011) to keep sites separated by at least 100 bases. The phylogenetic analysis was run in *PAUP* 4.0, including all possible quartets of samples and with 500 bootstrap replicates. The resulting trees were visualised with *Figtree* v1.4.3.
ABBA-BABA tests of introgression

To test for an excess of shared derived polymorphisms indicative of hybridisation we calculated $D$-statistics based on four-taxon ABBA-BABA tests using the software DSUITE (Malinsky et al., 2020). This analysis used the same concatenated matrix as the RAxML analysis. For this analysis we aggregated individuals into species, with the exception of distinguishing populations of *A. barrelieri* from Morocco and the Sierra Nevada as they were placed in different groups in all our phylogenies (see Results). We tested for hybridisation between all trios of related species controlling for phylogenetic relatedness (the 'correct tree arrangement', *sensu* Malinsky et al., 2020), as well as alternative measures that produce a conservative estimate of hybridisation ('$D_{\text{min}}$ arrangement') or directly infer relatedness from the site patterns.

(Rambaut, 2016) and the topology compared with the RAxML tree using the R package PHYTOOLS (Revell, 2012).

**Fig. 2** Geographic map showing the locations of *Antirrhinum* populations used in this study. Colours and letters represent the seven geographic regions delimited for the analysis. These are Po, Portugal; NCS, North and Centre of Spain; SN, Sierra Nevada and South of Spain; Mo, Morocco; Py, Pyrenees; Al, Alps and It, Italy.
without using a phylogeny (‘BBA arrangement’). We fixed the species *M. orontium* as the outgroup for inferring the ancestral allele in the analysis of each ingroup trio. To better visualise introgression patterns inferred from the D-statistic test we plotted the results in a heatmap using a custom script available from www.github.com/mmatschiner.

**Ancestral state reconstruction of vegetative and reproductive characters**

We investigated morphological traits that are heritable and contribute the most to the differences in leaf and petal shape (allometry) between subsections *Kickxiella* and *Antirrhinum* (Wilson & Hudson, 2011). These traits were plant height (from cotyledons to inﬂorescence tip at anthesis of the ﬁrst ﬂower), leaf morphology, dorsal petal morphology and ﬂower colour. These characteristics were scored on plants from the same accessions used for phylogenetic analysis, growing under glasshouse conditions following the protocol described in Hudson et al. (2008b). Traits were scored on an average of two or three individuals per species. For measures of allometry, fully developed metamer four leaves were ﬂattened and imaged or the dorsal corolla excised, ﬂattened and scanned and points placed around their silhouettes using the software AAMTOOLBOX (http://cmpdartsvr1.cmp.uea.ac.uk/wiki/BanghamLab/index.php/Software), following the methods described in Langlade et al. (2005). A principal component analysis was then used to partition the variance between samples into main PCs. For each type of organ (leaf or ﬂower) PC1 was used for ancestral reconstruction. LePC1 accounts for 87% of the variation in shape and size of leaves and FsPC1 accounts for 82% of the variation in size, shape and angle of dorsal petals (Fig. 3). In addition to the allometric models, the traits plant height and flower colour were also used for reconstruction.

Colour patterns of the corolla were scored based on well characterised phenotypes and genotypes previously recorded in mutants and in natural populations of the species *A. majus* (Whibley et al., 2006). Flower colour was classiﬁed as white, yellow, magenta or restricted magenta (Fig. 3).

We performed maximum likelihood ancestral state reconstruction for all characteristics using fastAnc and contMap implemented in PHYTOOLS (Revell, 2012), assuming a Brownian model of evolution. To account for polymorphic states in ﬂower colour, we used the rayDisc function in the package corHMM (Beaulieu, et al., 2013). RayDisc deals with polymorphic data by assigning equal likelihood values to each state in a polymorphic sample. One individual was used per taxon and tested with both a symmetrical model, in which all transitions between character states are possible and forward and reverse transitions have equal rates, and an asymmetrical model in which all transitions are possible but forward and reverse transitions have different rates. We also ran constrained models for flower colour to account for the rarity of orange ﬂowers in the transition between yellow and magenta, which suggests a white intermediate (Ellis & Field, 2016). We evaluated the degree of support for each model using a likelihood ratio test and the Akaike Information Criterion (AIC) score.

**Results**

**Phylogenetic relationships**

To reconstruct phylogenetic relationships in *Antirrhinum* we generated RAD sequence data from species representative of the morphological and geographically diversity present in the genus. Illumina sequencing produced 104,031,701 paired-end sequencing reads, with a mean of 541,061 reads per sample. Thirty-four samples were removed in sample quality ﬁltering, including the only specimen of the taxon *A. subbaeticum*, leaving 86 individuals from 24 species. Alignment of sequence reads to the *A. majus* reference genome varied between 77% and 99% per sample, with an average of 93%, with a ﬁnal per-sample mean coverage of 17.7-fold.

The phylogenetic analyses resolve the relationships between major clades in *Antirrhinum* (Fig. 4), with maximum likelihood (ML) and coalescent trees showing similar overall topologies, although with some notable differences in parts of the phylogeny that were poorly resolved (discussed below). As expected from a concatenated genomic sequence alignment, most (90%) of all nodes across the ML phylogeny have bootstrap support values of 90% or higher, including the clades corresponding to the main *Antirrhinum* subsections. However, only 19% of nodes in the coalescent tree have support values of 90% or higher, with the node corresponding to what is traditionally considered as subsection *Antirrhinum* having a support value of 83%. Rooting the phylogeny with *Misopaetes* produced a long branch to the outgroup, with short branches separating early diverging *Antirrhinum* lineages in the ML tree.

Our phylogenies revealed notable discordance with the traditional taxonomic classiﬁcation of the genus based on morphology. For example, individuals traditionally classiﬁed as *Kickxiella* were distributed into at least four different groups across the phylogeny (Fig. 4). *Kickxiella* Group 1 was placed as early diverging in both phylogenies. Group 2 consisted of two endemic species from subsection *Kickxiella* with *A. meonanthum* (subsection *Streptosepalum*). Group 3 was composed exclusively of the species *A. molle* and was placed as a sister clade to the whole subsection *Antirrhinum*. Group 4 in the ML tree was nested within subsection *Antirrhinum* and was composed of the species *A. hispanicum*, *A. rupestris*, *A. mollissimum* and *A. charidemi* from the southeast of Spain. These species were split into two separate groups in the coalescent tree. Finally, the two *Streptosepalum* species were placed with *Kickxiella* *A. braun-blanquetii* within *Kickxiella* Group 1 in the ML phylogeny or as its sister in the coalescent tree, and *A. meonanthum* within *Kickxiella* Group 2.

In both phylogenetic trees, the topology within some clades showed a stronger relationship to geographic distributions than to morphology. This was the case for Group 4 of *Kickxiella*, which was nested within subsection *Antirrhinum* and grouped with other species distributed in the south of Spain and Morocco. Also, the accessions of *A. barrelieri* from the Sierra Nevada were more closely related to other species from southeast Spain than to the conspecific accessions from Morocco (Figs 3, 4).

Despite the well supported tree topology in the ML analysis, further characterisation of phylogenetic relationships revealed
substantial conflict across the tree (Fig. 5). The clade showing the highest level of conflict was the group of species distributed in the southeast of Spain that includes the taxa *A. australe* and *A. tortuosum*. The ML phylogeny failed to support the relationship of this group of species but the coalescent tree showed a division that matches the geographical distribution of each accession (Fig. 5). All accessions of *A. australe*, and the *A. tortuosum* accessions that clustered with the Moroccan *A. barrelieri* (L81, L93, L100, L102) were from the Sierra de Grazalema, west of Granada towards the Strait of Gibraltar. By contrast, *A. barrelieri* to the east of Granada (L148, L150, L205) clustered with the *Kickxiella* species found in the same region, consistent with local hybridisation. Overall, the high levels of conflict shown by quartet sampling and the bootstrap tree, suggest a biological process (e.g. hybridisation or incomplete lineage sorting (ILS)), as the main cause of conflict rather than a lack of informative characters.

Ancestral state reconstruction of morphological traits

Morphological subsections are defined by suites of morphological characters, particularly plant size, organ shape and size and flower colour (Wilson & Hudson, 2011; Feng *et al.*, 2009; Langlade *et al.*, 2005). However, morphological subsections are not supported as monophyletic in our phylogenetic analyses, suggesting that similar suites of characters have evolved in parallel. To test this, and to estimate ancestral morphologies, we reconstructed ancestral traits along the phylogeny (Fig. 6). These results suggested that the combination of size and shape traits associated with subsection *Kickxiella* evolved several times: in the early-diverged *Kickxiella* lineage (Group 1) and independently within subsection *Antirrhinum*, from an intermediate ancestral phenotype very similar to the species *A. siculum*. Additionally, the results for leaf and flower size showed variation between the taxa traditionally considered *Kickxiella*. The individuals placed within the *Kickxiella* Group 1 tended to have smaller and rounder leaves and flowers in comparison with the species within the other *Kickxiella* groups.

For ancestral state reconstruction of flower colour, the symmetrical model was chosen as it had the lowest AIC value, with no significant differences between constrained and unconstrained models (Table 1). Here we show the results obtained with corHMM (Beaulieu *et al.*, 2013) as it deals better with polymorphism in the data. In total, six taxa were found to be polymorphic. These were *A. boissieri*, *A. barrelieri* from Morocco, *A. mollissimum*, *A. pseudomajus*, *A. striatum* and *A. tortuosum*. The ML estimates of the transition rates under a symmetrical model

---

Fig. 3 Leaf and flower variation in *Antirrhinum*. (a) Representations of low, mean and high values for PC1 for leaf shape (LePC1). (b) Low, mean and high values for flower shape (FsPC1). (c) Front and side views of flowers showing colour patterns used in the ancestral state reconstruction: white, yellow, magenta and restricted magenta. PCAs of flowers and leaves follow Wilson (2010).
support the constrained evolution of flower colour in Antirrhinum (Table 2). Forward and reverse transitions between yellow and magenta (restricted or not) have a likelihood of zero (Table 2). Additionally, forward and reverse transitions from white to yellow are more likely to occur than transitions from white to magenta (restricted or not).

ABBA-BABA tests of introgression

Quartet analysis had suggested incongruence was the result of ILS or hybridisation. We tested the involvement of hybridisation further by calculating the excess of derived polymorphisms shared between lineages using the D-statistic. We found an average value
of $D = 0.07$ when accounting for phylogeny, with a maximum value of 0.37, suggesting hybridisation across several groups (Fig. 7). The species *A. braun-blanquetii* showed consistently high values of $D > 0.25$ in combination with species from subsection *Antirrhinum*, including species with overlapping geographic ranges and species that do not occur in sympatry. Within
subsection *Antirrhinum*, the species *A. tortuosum*, *A. australe*, and the Moroccan accessions of *A. barrelleri* have high values of \( D \) (c. 0.3) when compared with values for the rest of the genus, supporting widespread hybridisation. In some of these cases, hybridisation and introgression occur in sympatry, such as for *A. boissieri* and *A. tortuosum* in southern Spain, while a high value of \( D \) cannot be explained by current range overlap for other species such as *A. braun-blancuetii* and *A. pulverulentum*. Alternative measures of hybridisation were generally consistent with these results, with the conservative estimate of hybridisation \( D_{\text{min}} \) supporting widespread hybridisation of *A. braun-blancuetii*, but more limited hybridisation between other species combinations (Fig. S1), while hybridisation inferred directly from site patterns (the ‘BBAA arrangement’; Fig. S2) was broadly similar to that using the phylogenetic tree, above.

**Discussion**

Our analysis of genome-wide variation across *Antirrhinum* species has allowed us to reconstruct a phylogeny for the genus, and to test for parallel phenotypic evolution in the colonisation of alpine environments. We show that *Kickxiella* alpine morphology is present in multiple groups of *Antirrhinum* species, suggesting repeated colonisation of alpine environments through similar morphological changes. We also found evidence that the lowland morphology of subsection *Streptosepalum* species is likely to have...
Antirrhinum

We also identified a large group corresponding to subsection morphology that are endemic to different mountain ranges in the formation of narrow endemic with individual species clustering by geography, indicative of broad-scale clustering by taxonomic affinity and morphology, in age (Group 1), formed by species with a Kickxiella et al. (2005; Vargas et al., 2002; Jiménez et al., 2013; Butlin et al., 2014). First, there could be recurrent independent phenotypic evolution in separate populations after an initial colonisation event. This scenario can be further subdivided depending on the genetic basis of the phenotype, be it from independent origins of genetic variation in different populations or from standing genetic variation in the ancestral population(s). Second, there may be a single adaptive divergence event (i.e. a common genetic origin for a given phenotype), followed by widespread colonisation of intermingled or adjacent environments.

A hypothesis of independent transitions would require three transitions, one in each of the derived Kickxiella groups. While plausible, this is not the most parsimonious explanation given that Kickxiella Groups 1–3 are closely related. Instead, if the trait combination in Kickxiella Groups 1–3 had a common genetic origin, with Streptosepalum species evolving within this early evolved independently from alpine Kickxiella lineages. The evidence for hybridisation between taxa across the Antirrhinum phylogeny suggests that introgression may have played a role in adaptation, potentially facilitating the colonisation of new environments. Here, we discuss these results in the context of other studies of parallel evolution in plants, and consider the biogeographic scenarios and genetic architectures giving rise to morphological diversity in Antirrhinum.

### Table 1 Comparison between symmetrical, asymmetrical and constrained models of flower colour transition in Antirrhinum.

| Model                  | AIC      | Log, lik |
|------------------------|----------|----------|
| Ape                    |          |          |
| Symmetrical            | 181.09   | −84.54   |
| Asymmetrical           | 188.25   | −82.12   |
| Constrained            | 193.9    | −88.95   |
| corHMM                 |          |          |
| Symmetrical            | 70.68    | −29.34   |
| Symmetrical constrained| 72.5     | −28.25   |
| Asymmetrical           | 80.49    | −28.24   |
| Asymmetrical constrained| 80.41   | −28.2    |

Results are presented for the analyses in Ape and corHMM. Bold values are the lowest AIC and log-likelihoods (Log, lik).

### Table 2 Maximum likelihood point estimates of transition rates under the symmetrical model obtained in corHMM.

|       | White | Yellow | Magenta | Restricted magenta |
|-------|-------|--------|---------|--------------------|
| White | −     | 100.00 | 0.12    | 1.06               |
| Yellow| 100.00| −      | 0.00    | 0.00               |
| Magenta| 0.12  | 0.00   | −       | 14.89              |
| Restricted Magenta| 1.01  | 0.00   | 14.89   | −                  |

Despite the strong support found in the ML phylogeny, phylogenetic analyses using concatenated data can overestimate bootstrap support values and hide multiple equally supported conflicting topologies, especially in the presence of ILS (Gadagkar et al., 2005; Warnow, 2015). Our analyses of conflict showed high support values across the phylogeny; however, notably high levels of conflict were observed in the clade composed of the taxa *A. tortuosum*, *A. australis*, and the Moroccan accessions of *A. barrelieri*. This clade also showed low support in the ML and coalescence trees. Results from QS suggest that the topology observed is not caused by a lack of information in the data and point to another biological process as the most likely source of conflict. Together, a lack of support and low phylogenetic resolution, values of the hybridisation statistic $D$ close to 0.3, and a very similar morphology, suggest that this group of species is a species complex without clear reproductive, morphological or genome-wide genetic differences. The inclusion of *A. australis* within *A. tortuosum* has been proposed by Mateu-Andrés & De Paco (2005), who obtained a broadly consistent tree topology in an allozyme study of several species of *Antirrhinum*. Our phylogeny confirms these results and shows the presence of a geographical pattern of genetic structure in this species complex in the south of Spain and Morocco. Here, members of different species or even different morphologically subsections, can be genetically more similar to their neighbours rather than geographically more distant members of the same species or subsection, which is likely due to the homogenising effect of hybridisation in areas of sympatry (discussed below). Future taxonomic work should supplement this phylogenetic framework with additional samples and use this to identify monophyletic taxa that warrant continued recognition as distinct species.

### Parallel phenotypic evolution

The appearance of Kickxiella morphology in multiple groups across the Antirrhinum phylogeny supports parallel phenotypic evolution of this trait combination. Phenotypic parallelism can be the result of many different combinations of historical events and involves different underpinning genetic architectures, with a number of nonexclusive scenarios having been identified (Johanneson et al., 2010; Roda et al., 2013; Butlin et al., 2014). First, there could be recurrent independent phenotypic evolution in separate populations after an initial colonisation event. This scenario can be further subdivided depending on the genetic basis of the phenotype, be it from independent origins of genetic variation in different populations or from standing genetic variation in the ancestral population(s). Second, there may be a single adaptive divergence event (i.e. a common genetic origin for a given phenotype), followed by widespread colonisation of intermingled or adjacent environments.

A hypothesis of independent transitions would require three transitions, one in each of the derived Kickxiella groups. While plausible, this is not the most parsimonious explanation given that Kickxiella Groups 1–3 are closely related. Instead, if the trait combination in Kickxiella Groups 1–3 had a common genetic origin, with Streptosepalum species evolving within this early
diverging Kickxiella clade, there may be just a single transition, in Group 4. Moreover, given extensive hybridisation in Antirrhinum, it may be that a single introgression event from Groups 1–3 to 4 would be sufficient to explain this difference. A study of sequence variation in Hairy (Tan et al., 2020), a gene that represses trichrome fate and underlies trichrome differences between densely hairy Kickxiella species and largely hairless Antirrhinum species, showed that all Kickxiella species (except A. grosii) form a single clade. This suggests a common genetic basis for at least this component of the alpine Kickxiella phenotype.

Morphological and genetic data suggest that evolutionary transitions in Antirrhinum are driven by contrasting selection pressures, and different genetic mechanisms underlie each morphological trait. Our reconstructions of flower colour support the model of Ellis & Field (2016) where colour transitions, for example from yellow to magenta, predominantly occur through a white intermediate. Furthermore, the lack of orange coloured phenotypes, except in narrow hybrid zones (Tavares et al., 2018), is consistent with selection against double pigmented phenotypes. As such, evolutionary shifts in flower colour are likely to be due to mutations at few major effect colour loci.

In contrast with flower colour, traits such as height, leaf area, flower size and number have been shown to have a complex genetic architecture in Antirrhinum, with several loci responsible for each trait, with these spread throughout the genome (Feng et al., 2009). Ongoing quantitative trait locus (QTL) mapping between A. rupestris and A. barrelieri in the Sierra Nevada similarly suggests that multiple QTL spread across many chromosomes underlie trait divergence in this group (Duran-Castillo, 2019). As these loci are dispersed across chromosomes, adaptive divergence appears not to be solely maintained by a few regions of reduced recombination such as chromosomal inversions, which is a widely evoked mechanism to explain how traits can be maintained in the face of gene flow (Twyford & Friedman, 2015). Instead, selection on many regions of the genome are likely to maintain ecotypic divergence of alpine and grassland Antirrhinum species. Selection pressures on these morphological traits are likely to be complex and include both biotic and abiotic pressures, with smaller leaves and shorter stems adaptive to drier habitats on rocky surfaces, while longer stems and bigger leaves could be advantageous in the presence of a more competitive environments (Parkhurst & Loucks, 1972; Nicotra et al., 2011).

A particular challenge for studies of trait evolution in Antirrhinum is posed by the rapid burst of speciation experienced early in the origin of the genus. This has resulted in considerable divergences between Antirrhinum and its nearest relatives, such as New World Sairocarpus or other taxa in the wider Antirrhinum clade (such as Misopaetes, used here as an outgroup). This ‘evolutionary gap’ creates uncertainty in the ancestral state for the group, especially given the morphological diversity present in lineages related to Antirrhinum. Moreover, this burst of speciation makes it hard to know the exact relationships of early diverging lineages. While we...
are confident that the Kickxiella phenotype is present in multiple distinct lineages, the relationship between early diverging Kickxiella lineages and Streptosepalum warrants further study, particularly using long read sequencing and gene tree-specific analyses, which may provide important insights into these closely related groups.

Speciation history of Antirrhinum

Rothmaler (1956) and Webb (1972) proposed a model of isolation–contact–isolation based on the distribution of important morphological characters in Antirrhinum. This model evokes the idea that the Kickxiella and Antirrhinum morphologies first evolved during periods of isolation. Periods of secondary contact would then allow the introgression of alleles underlying morphological traits, with this introgression potentially facilitating the colonisation of new habitats. In our study, we found indirect evidence for such a model, with extensive hybridisation between species from across the genus not just in well characterised hybrid zones but in numerous sympatric taxa as well as species in allopatry. Vargas et al. (2004) and Vargas et al. (2009) reported recent putative hybridisation based on nuclear ribosomal ITS sequences, while Vargas et al. (2009) found evidence for species within the same broad geographic area sharing chloroplast haplotypes. Similarly, Wilson & Hudson (2011) found a mismatch between chloroplast lineages and morphology in species with overlapping distributions. Several Antirrhinum hybrid zones exist and have been well characterised, particularly between Antirrhinum majus pseudomajus and A. majus striatum, subspecies that differ primarily by flower colour (Bradley et al., 2017; Tavares et al., 2018). These subspecies showed no evidence of genome-wide barriers to gene flow (Ringbauer et al., 2018). This suggests that the genomes may be largely exchanged following secondary contact (Tavares et al., 2018) despite the presence of local barriers to gene flow at flower colour genes. Given the recent origin and rapid radiation of the genus, like many other study systems for investigating speciation, this pattern of adaptive divergence in the presence of gene flow might be typical of incipient lineages undergoing speciation. Similar patterns of phenotypic evolution in response to harsh environments have been found in the Senecio laetus species complex, where different populations have adapted recurrently to dune, headland and alpine environments and show evidence of recent gene flow (Roda et al., 2013). Similarly, in Stickleback fish, recurrent colonisation of freshwater habitats was accompanied by the evolution of similar phenotypic traits, including changes in body shape, skeletal armour and pigmentation (Jones et al., 2012). Overall, our results show how natural selection can promote and maintain suites of phenotypic differences, even in the presence of gene flow, and place Antirrhinum as a promising system for future studies of adaptive divergence.

Acknowledgements

Research by ADT is supported by NERC grants NE/R010699/1 and NE/L011336/1 and AH by the Biotechnology and Biological Sciences Research council (BBSRC; grant number BB/D552089/1) and through the SynTax award scheme supported by BBSRC and NERC. The Royal Botanic Garden Edinburgh is supported by the Scottish Government’s Rural and Environmental Science and Analytical Services Division. MDC was supported through a National Council of Science and Technology of Mexico (CONACYT) graduate scholarship and YW by a BBSRC postgraduate studentship. We thank Antirrhinum collection holders and field botanists for providing additional samples, including: I. Mateu-Andres, T. Gübitz, E. Coen, M. Burrus and C. Thébault, and M. Els for help with figures. We thank Lynda Delph, Peter Morrell and an anonymous reviewer for useful comments that improved the manuscript. Additional fieldwork was supported by a research grant from the Carnegie Trust for the Universities of Scotland and the Genetics Society.

Author contributions

AH and YW collected samples and performed laboratory work. MCD processed the sequence data and carried out the analyses with the help of ADT. MCD and ADT led the writing of the manuscript with support from DF and AH.

ORCID

Mario Durán-Castillo https://orcid.org/0000-0001-9928-2359

David L. Field https://orcid.org/0000-0002-4014-8478

Andrew Hudson https://orcid.org/0000-0001-9049-0100

Alex D. Twyford https://orcid.org/0000-0002-8746-6617

Data availability

Raw sequence reads are available in the Sequence Read Archive ( Biosamples SAMN18237715–SAMN18237800). Phylogenetic trees and morphological data are available on Dryad (https://doi.org/10.5061/dryad.xgxd254gr).

References

Beaulieu JM, O’Meara BC, Donoghue MJ. 2013. Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. Systematic Biology 62: 725–737.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimming tool for Illumina sequence data. Bioinformatics 30: 2114–2120.

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.

Brandrud MK, Paun O, Lorenzo MT, Nordal I, Brysting AK. 2017. RADseq provides evidence for parallel ecotypic divergence in the autotetraploid Cochlearia officinalis in Northern Norway. Scientific Reports 7: 5573.

Broad Institute. 2018. Picard tools. [WWW document] URL http://broad institute.github.io/picard/ [accessed 1 July 2018].

Butlin RK, Saura M, Charrier G, Jackson B, André C, Caballero A, Coyne JA, Galindo J, Grahame JW, Hollander J et al. 2014. Parallel evolution of local adaptation and reproductive isolation in the face of gene flow. Evolution 68: 935–949.

Catchen J, Hohenlohe PA, Basham S, Amores A, Cresko WA. 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology 22: 3124–3140.

Chifman J, Kubatko L. 2015. Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-
specific rate variation, and invariant sites. *Journal of Theoretical Biology* 374: 35–47.

Chou J, Gupta A, Yaduvanshi S, Davidson R, Nute M, Mirabar S, Warnow T. 2015. A comparative study of SVDquartets and other coalescent-based species tree estimation methods. *BMC Genomics* 16: 11.

Coen ES, Carpenter R, Martin C. 1986. Transposable elements generate novel spatial patterns of gene expression in *Antirrhinum majus*. *Cell* 47: 285–296.

Coen ES, Meyerowitz EM. 1991. The war of the whorls: Genetic interactions controlling flower development. *Nature* 353: 31–37.

Daneeck P, Auten A, Abeasis G, Albers CA, Banks E, DePristato Ma, Handsaker RE, Lunter G, Marth GT, Sherry ST et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.

Doley JJ, Doyle JL. 1987. *Leach*. *New Phytologist* 105: 357–365.

Doran-Castillo M. 2019. Genomic pattern of gene flow and divergence in *Antirrhinum*. Unpublished PhD thesis, The University of Edinburgh, UK.

Ellis TJ, Field DL. 2016. Repeated gains in yellow and anthocyanin pigmentation in flower colour transitions in the *Antirrhinaceae*. *Annals of Botany* 117: 1133–1140.

Elmer KR, Fan S, Kusche H, Luise Spreitzer M, Kautt AF, Franchini P, Meyer Forrest A, Escudero M, Heuertz M, Wilson Y, Cano E, Vargas P. 2017. Genomic patterns of gene flow and divergence in *Antirrhinum*. *New Phytologist* 209: 1022–1032.

Malinsky M, Matschiner M, Svardal H. 2020. Dsuite – Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources* 21: 584–595.

Mateu-Andrés I, De Paco L. 2005. Allozymic differentiation of the *Antirrhinum majus* and *A. siculum* species groups. *Annals of Botany* 95: 465–473.

Miller MR, Dunham JP, Amores A, Cresko WA, Johnson SA. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research* 17: 240–248.

Nicotra AB, Leigh A, Boyce CK, Jones CS, Niklas KJ, Rothery DL, Tsukaya H. 2011. The evolution and functional significance of leaf shape in the angiosperms. *Functional Plant Biology* 38: 535–552.

Oke KB, Rolshausen G, LeBlond C, Hendry AP. 2017. How parallel is parallel evolution? A comparative analysis in fishes. *American Naturalist* 190: 1–6.

Parkhurst DF, Loucks OL. 1972. Optimal leaf size in relation to environment. *Journal of Ecology* 50: 537.

Pezz JE, Brown JW, Walker JF, Hinchcliffe CE, Smith SA. 2018. Quartet Sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany* 105: 385–403.

Rambaut A. 2016. *FigTree v.1.4.3*. [WWW document] URL http://tree.bio.ed.ac.uk/software/figtree/

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

Ringbauer H, Kolesnikov A, Field DL, Barton NH. 2018. Estimating barriers to gene flow from distorted isolation-by-distance patterns. *Genetics* 208: 1231–1245.

Rodà F, Ambrolo L, Walter GM, Liu HL, Schaud A, Lowe A, Pelser PB, Prentis P, Rieseberg LH, Ortiz-Barrientos D. 2013. Genomic evidence for the parallel evolution of coastal forms in the *Senecio laetus* complex. *Molecular Ecology* 22: 2941–2952.

Rosenblum EB, Parent CE, Brandi EE. 2014. The molecular basis of phenotypic convergence. *Annual Review of Ecology, Evolution, and Systematics* 45: 203–226.

Rothmaier W. 1956. *Taxonomische monographie Antirrhinum majus*. Berlin: Akademie Verlag.

Salzburger W. 2009. The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Molecular Ecology* 18: 169–185.

Schwarz-Sommer Z, Davies B, Hudson A. 2003. An everlasting pioneer: the story of *Antirrhinum* research. *Nature Review Genetics* 4: 655–664.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.

Tan Y, Barnbrook M, Wilson Y, Molnár A, Bukys A, Hudson A. 2020. Shared mutations in a novel glutaredoxin repressor of multicellular trichome fate underlie parallel evolution of *Antirrhinum* species. *Current Biology* 30: 1357–1366.

Tavares H, Whibley A, Field DL, Bradley D, Couchman M, Copsey L, Elloeot J, Burren M, Andalo C, Li M et al. 2018. Selection and gene flow shape genomic islands that control floral guides. *Proceedings of the National Academy of Sciences, USA* 115: 1106–11011.

Truccoli E, Frajman B, Haverkamp THA, Schonwetter P, Paun O. 2017. Genomic analyses suggest parallel ecological divergence in *Heliopera pusillum* (caryophyllaceae). *New Phytologist* 216: 267–278.

Twfford AD, Ennos RA. 2012. Next-generation hybridization and introgression. *Heredity* 108: 179–189.

Twfford AD, Friedman J. 2015. Adaptive divergence in the monkey flower *Mimulus guttatus* is maintained by a chromosomal inversion. *Evolution* 69: 1476–1486.
Vachaspati P, Warnow T. 2018. SVDquest: Improving SVDquartets species tree estimation using exact optimization within a constrained search space. *Molecular Phylogenetics and Evolution* 124: 122–136.

Vargas P, Carrió E, Guzmán B, Amat E, Giemes 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, Ornosa C, Ortiz-Sánchez FJ, Arroyo J. 2010. Is the occluded corolla of *Antirrhinum* bee-specialized? *Journal of Natural History* 44: 1427–1443.

Vargas P, Rosselló J, Oyama R, Giemes 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.

Warnow T. 2015. Concatenation analyses in the presence of incomplete lineage sorting, *PLoS Currents* 7.

Whibley AC, Langlade NB, Andalo C, Hanna AI, Bangham A, Thébaud C, Coen E. 2006. Evolutionary paths underlying flower color variation in *Antirrhinum*. *Science* 18: 963–966.

Wilding CS, Weetman D, Rippon EJ, Steen K, Maweje HD, Barsukov I, Donnelly MJ. 2014. Parallel evolution or purifying selection, not introgression, explains similarity in the pyrethroid detoxification linked GSTE4 of *Anopheles gambiae* and *A. arabiensis*. *Molecular Genetics and Genomics* 290: 201–215.

Wilson Y. 2010. Evolutionary relationships and morphological variation between *Antirrhinum* species. Unpublished PhD Thesis, The University of Edinburgh, UK.

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

Xue Y, Carpenter R, Dickinson HG, Coen ES. 1996. Origin of allelic diversity in antirrhinum S locus RNases. *Plant Cell* 8: 804–814.

Zwettler D, Vieira CP, Schlötterer C. 2002. Polymorphic microsatellites in *Antirrhinum* (Scrophulariaceae), a genus with low levels of nuclear sequence variability. *Journal of Heredity* 93: 217–221.

**Supporting Information**

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

**Fig. S1** Heatmap showing a conservative estimate of hybridisation across *Antirrhinum* samples, based on the *Dmin* statistic.

**Fig. S2** Heatmap showing an alternative estimate of the hybridisation statistic D across *Antirrhinum* samples, based on the BBAA trio arrangement.

**Table S1** Collection information for *Antirrhinum* samples.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.