CrossMark

ORIGINAL ARTICLE

# Recombination provides evidence for ancient hybridisation in the *Silene aegyptiaca* (Caryophyllaceae) complex

**Bernard E. Pfeil**[1] · **Zeynep Toprak**[1,2] · **Bengt Oxelman**[1]

**Abstract** Recombination events among distinct alleles complicate phylogenetic estimation. Various in vivo and in vitro processes can bring distinct alleles into the same genome to then undergo recombination, which may subsequently mislead phylogenetic inference if not assessed properly. Among the processes bringing divergent alleles together, hybridisation is perhaps the simplest and most likely, but alternatives need to be considered before hybridisation can be accepted as the underlying cause. Such alternatives include the presence of paralogues or deeply coalescing alleles, as well as amplification artefacts. Here, we document a recombination event that apparently took place between two divergent lineages of the *Silene aegyptiaca* complex in the flowering plant family Caryophyllaceae. We evaluate several possible mechanisms that might be responsible for the observed pattern. An ancient introgressive hybridisation event was the simplest explanation for the observations, compatible with geographic proximity of the affected lineages, whereas paralogy and deep coalescence are difficult to reconcile with the evidence obtained from a species tree of the group based on six different, non-recombinant genes and gene trees inferred using two partitions of the recombinant locus.

Bernard E. Pfeil and Zeynep Toprak are equal contributors.

**Electronic supplementary material** The online version of this article (doi:10.1007/s13127-017-0331-9) contains supplementary material, which is available to authorized users.

✉ Bernard E. Pfeil
bernard.pfeil@bioenv.gu.se

[1] Department of Biological and Environmental Sciences, University of Gothenburg, 40530 Gothenburg, Sweden

[2] Department of Biology, Faculty of Sciences, University of Dicle, 21280 Diyarbakir, Turkey

## Introduction

Recombination among distinct allele lineages can pose a challenge for phylogenetic inference of gene trees (Sanderson and Doyle 1992; Posada and Crandall 2002). In vivo recombination occurs during meiosis when a chromatid of each homologous chromosome crosses over and is exchanged, thus potentially forming chimaeric alleles. The presence of chimaeric alleles has the potential to introduce additional homoplasy into a dataset if non-monophyletic alleles contributed to the chimaera (Sanderson and Doyle 1992). The result can be the disruption of the historical signal, which in turn can lead to incorrect gene tree inference (Sanderson and Doyle 1992; Posada and Crandall 2002), biased branch length (Penny et al. 2008) and node age (Pfeil 2009) estimation.

Although recombination can be seen simply as a confounding process, it can also be an opportunity to detect the footprint of ancient hybridisation (Wendel and Doyle 1998; Linder and Rieseberg 2004; Rautenberg et al. 2008; Kelly et al. 2010). Hybridisation has the potential to bring together divergent alleles into the same genome, where they can then be recombined. If a recombined allele becomes fixed and persists to be sampled, it may be possible to infer which lineages contributed to this allele, for example, by separate phylogenetic analysis of non-recombined parts of the locus (Rautenberg et al. 2008; Kelly et al. 2010; Schwartz et al. 2015) combined with multi-labelling approaches (Visser et al. 2012). However, competing explanations for the origins of chimaeric alleles need to be considered before hybridisation can be inferred (e.g. Rautenberg et al. 2008; Frajman et al. 2009).

🖉 Springer

One alternative explanation is that deeply coalescing allele lineages have contributed to the formation of a chimaera. Coalescent theory predicts that allele lineages that differ in selectively neutral ways can be maintained for some time after speciation in descendent species, depending on the effective population size of the locus (e.g. Rosenberg 2003). If the allele lineages persist through a second speciation event, with one allele lineage being passed to both descendent species, but the other surviving only in one descendent (the deeply coalescing allele lineage), then recombination between the surviving lineages in the one species carrying both lineages can create chimaeric alleles. One part of such a chimaeric allele may indicate a closer relationship of the chimaera to alleles from the closely related species (the non-deeply coalescing lineage), but the other part could present a different relationship, as deeply coalescing alleles can coalesce with allele lineages from different species when polymorphism occurs in a common ancestral species (Rosenberg 2003). Thus, different topologies may occur in the separate trees of recombined portions of alleles, even though hybridisation has not occurred.

Another explanation is that alleles residing at paralogous loci may occasionally pair during meiosis (e.g. through unequal recombination among proximal copies) and also contribute to the formation of recombined alleles (Jelesko et al. 2004; Yandeau-Nelson et al. 2006). If the sampling of alleles at each locus is incomplete, or some descendent loci have been lost, the origins of the chimaeric allele may mimic the topological expectations of deep coalescence.

A third alternative explanation is recombination during the data acquisition steps, such as might be caused by template switching during PCR amplification (Cronn et al. 2002). If paralogous loci reside in the genome, or contaminating DNA from a related species is present, suitable templates for template switching could be available to generate chimaeric alleles. In the case of contamination, the recombined alleles would be expected to be identical (or nearly identical, within the sequencing error margin) to different alleles in the dataset at different parts of the sequence (e.g. Pfeil et al. 2004). This could mimic the pattern expected for recombination that followed a very recent hybridisation event where the introgressed alleles have not diverged from the originating alleles.

In an initial survey of seven loci selected to study species delimitation in the *Silene aegyptiaca* complex (below), we found indications of recombination in one of the low copy nuclear (LCN) loci. This locus was of necessity excluded from the earlier study (Toprak et al. 2016), because it violates the assumptions of the methods employed there (Heled and Drummond 2010). Here we examine the sequence patterns within this locus further and test (1) whether the evidence for recombination in the LCN locus is compelling, (2) how many recombination events might have taken place, (3) which lineages have contributed to the formation of chimaeric alleles

and (4) which alternative explanations can be ruled out in order to infer the origin(s) of the chimaeric sequences. We also provide a tentative evaluation of which modes of speciation are more compatible with the cause of recombination and discuss the implications this has with respect to the porous nature of the barriers to gene flow between these species in the *S. aegyptiaca* complex.

### The *S. aegyptiaca* (Caryophyllaceae) complex

The *S. aegyptiaca* (L.) L. fil. 1782 species complex includes a number of highly similar annual species that are found in the Eastern Mediterranean Region, especially southern Anatolia (Coode and Cullen 1967; Erixon and Oxelman 2008; Aydin et al. 2014a). A recent study based on six nuclear loci revealed several cryptic species in the group (Toprak et al. 2016). The analysis was based on a Bayesian multispecies coalescent analysis that tests the assignment of samples to species while co-estimating the species tree (DISSECT; Jones et al. 2015). A conservative interpretation of the results suggested that the observed species diversity was partitioned into three main clades, containing a total of nine putative taxonomic species (colour coded) that are partly geographically structured (Fig. 1; modified from Toprak et al. 2016). The evidence for these nine species, exhibited by monophyletic gene trees, is strong. If samples come from the same species, then recombination among them is expected for sexually reproducing organisms. However, recombination among species can provide evidence for rare unexpected gene flow.

## Material and methods

### Data generation

DNA sequences of the LCN locus EST09 were generated from dried material of members of the *S. aegyptiaca* group (see Petri et al. 2013, for details on how the locus was discovered, assessed to be low copy and PCR primers designed). Protocols of molecular methods including DNA extraction, PCR amplification, sequencing and multiple alignment can be found in Aydin et al. (2014b), and vouchers are presented in Toprak et al. (2016). For the cases where the initial sequences were polymorphic, sequence-specific primers were used to separate allele copies (Scheen et al. 2012).

### Data analysis

The alignment was tested for recombination using RDP4 (Martin et al. 2010). An initial *p* value of 0.1 was used with three detection methods (RDP, MaxChi and Chimaera) to screen for recombination events at a relaxed stringency (high
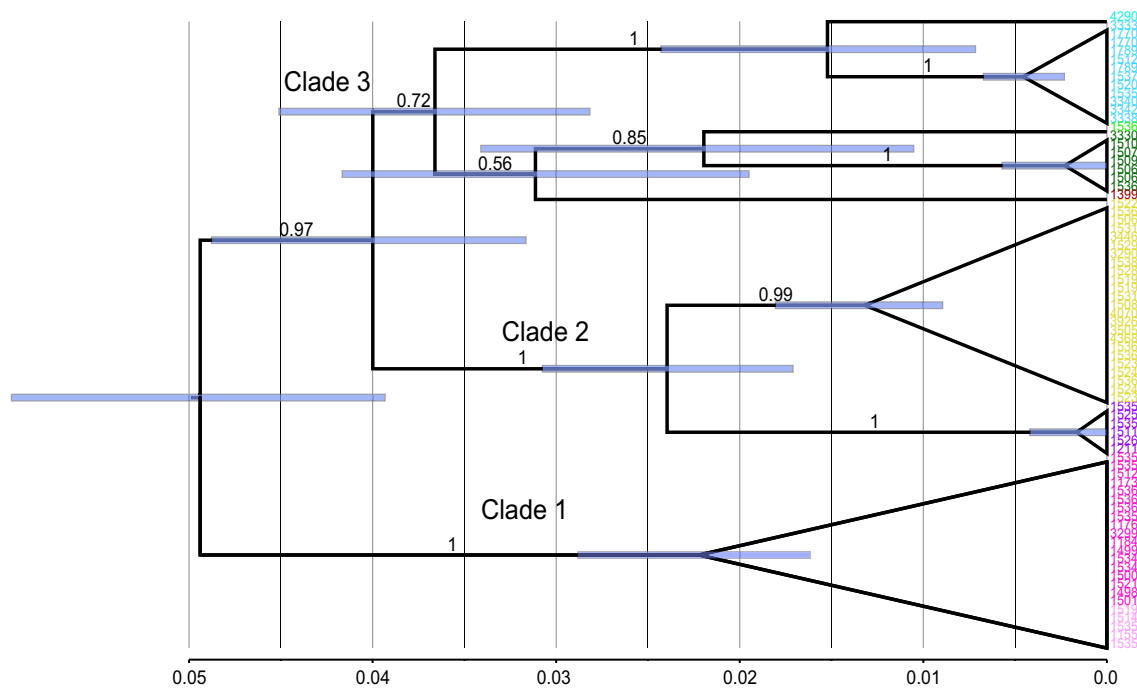
**Fig. 1** Species tree inferred by the DISSECT module of *BEAST based on six nuclear loci. Clades containing conspecific samples have been shown as *triangles*. Six clades and three singletons comprise the nine species discussed in the text. *Numbers above branches* are the clade posterior probabilities (PP); only PP > 0.5 are shown on tree. *Bars overlaying nodes* show the 95% credibility interval of node heights. The timescale is in substitutions per site

type 1 error). Any putative recombination events found by any of these methods were then re-checked with all methods. Only events with $p \leq 0.05$ in at least two methods, and with corresponding phylogenetic evidence consistent with recombination, were considered to be real. Although several events were inferred within species and excluded from further analyses (not shown), we focused only on the one event that we report here that is a case of recombination between species.

We checked the phylogenetic pattern of the putative recombination event first within RDP4, using UPGMA trees (not shown), but then by running a Bayesian analysis (using MrBayes v3.2; Ronquist and Huelsenbeck 2003) run for 500,000 generations in two parallel runs with two chains each (sampling every 1000 generations), with mixed substitution models and the addition of a gamma parameter (otherwise default settings). The alignment was partitioned into two parts, those either side of the recombination breakpoint estimated by RDP4 (positions 1–813 and 814–1573) and run separately.

We then ran the putatively recombined part of the alignment (positions 814–1573) in BEAST v1.8.0 (Drummond et al. 2012) for 10,000,000 generations (sampling every 5000 generations), using an uncorrelated log-normal relaxed clock model, to derive the relative age of the recombination event with respect to speciation events. We calibrated the gene tree using the 95% HPD (highest posterior density) interval of the species tree root height from the previous DISSECT analysis (Toprak et al. 2016) as the 95% central interval of a normal distribution as the prior on the gene tree root height.

We used the HKY substitution model, which was the most visited in the MrBayes mixed model analysis, with the addition of a gamma among-site rate heterogeneity parameter, as in the previous analyses. We changed the relaxed clock mean prior (ucld.mean) to a broad uniform prior ($1 \times 10^{-6}$ to $1 \times 10^{2}$) and subsequently checked that these limits were not encountered as a boundary by inspection in Tracer after the analysis.

We did not calibrate the BEAST tree in absolute time units, because the evidence for a specific calibration is scanty (e.g. fossil record, mutation rate). For the purpose of the present study, absolute ages are not needed, as we are just interested in whether the recombination event took place before or after a speciation event. Therefore, we use relative time units (scaled as substitutions per site) and a prior density given by a species tree based on six other loci.

## Results

### Recombination results

We found one recombination event using RDP4 that was supported by four methods with the following $p$ values: GeneConv $p = 0.02$; BootScan $p = 0.01$; MaxChi $p = 0.03$; Chimaera $p < 0.01$. Three sequences belonging to the individuals 15,247, 3290 and 15,363 (all from the *S. aegyptiaca*

(yellow) group in Fig. 1) shared this event. The recombination breakpoint was suggested to be around position 813.

The MrBayes analyses of the 5′ partition of the alignment (1–813) converged quickly and produced a consensus tree with two of these three sequences in a supported clade with a member of the same group based on the DISSECT species tree, individual 15,223 (Fig. 2). All three sequences are members of a large clade consisting of individuals from distinct lineages (yellow, green, blue and lilac) of *S. aegyptiaca* complex, with little supported resolution within it but clearly distinct from the *S. atocioides* Boiss. (pink) group sequences (posterior probability [PP] = 1.0; the colour coding for species is shown in Fig. 1). The light pink group sequences could be sister to the *S. atocioides* (pink) group, if one unresolved node is favourably resolved (to shift the clade with individuals 15,350 and 15,146 to be sister to the *S. atocioides* (pink) group, thus making this partition of the gene somewhat compatible with the species tree in Fig. 1.

Phylogenies inferred using the 3′ partition (814–1573), however, produced a tree with the putative recombinant sequences instead in a single clade and strongly supported as sister to the clade of the *S. atocioides* (pink) group (Fig. 3). This result is strongly in conflict with the species tree inferred by DISSECT (Fig. 1), *BEAST (shown in Toprak et al. 2016) and the 5′ partition of the same locus (Fig. 2). Three strongly supported branches (each PP = 1.0) exclude the placement of these three sequences with the sequence two of them are most closely related to in the 5′ partition (individual 15,223).

## BEAST dating results

The dated tree of positions 814–1573 inferred by BEAST also converged quickly. The coefficient of variation did not abut zero (mean = 0.37), indicating that a relaxed clock was appropriate and that the rate variation was not excessive. The relative age of the recombination event (i.e. the divergence time of the recombinant clade from the sister group, 0.026) was approximately half the age of the species tree root height (0.049) and did not overlap with the latter age (95% HPD 0.016–0.036 and 0.039–0.060, respectively, Fig. 4).

## Discussion

### Evidence for recombination

We found clear evidence for recombination supported by several methods. Phylogenetic analysis of the partitions from either side of the inferred breakpoint produced supported conflicting topologies consistent with the inferred recombination. These analyses implicate the *S. atocioides* (pink) group as the closest extant relative of the donor lineage of the 3′ partition of EST09. The six genes analysed previously (Toprak et al. 2016) place the individuals carrying the recombined EST09 sequences in a background consistent with other members of the *S. aegyptiaca* (yellow) group (Fig. 1). This suggests that a small proportion of the genome has successfully introgressed



**Fig. 2** Consensus phylogram inferred from the Bayesian analysis of the 5′ partition of EST09. *Numbers above branches* are the clade posterior probabilities. The root position was placed along the long branch that separated the pink group individuals, which is also consistent with mid-point rooting. The putatively recombined sequences are marked by *red branches*. The *scale bar* is in units of substitutions per site
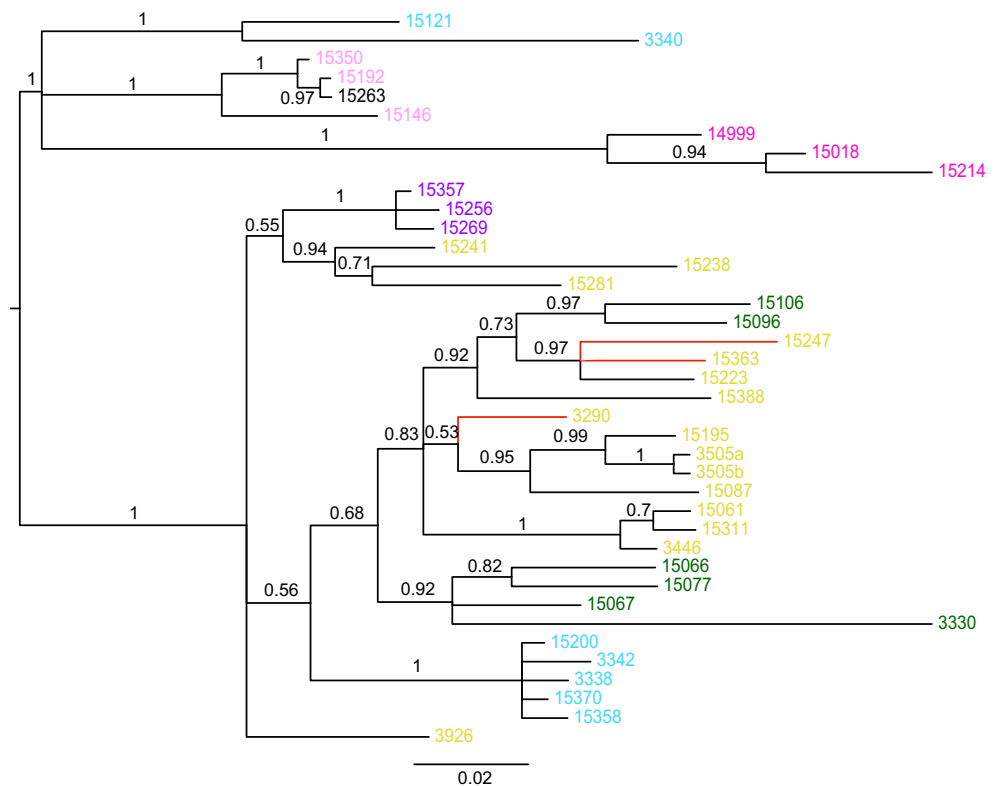
**Fig. 3** Consensus phylogram inferred using MrBayes of the 3′ partition of EST09. *Numbers above branches* are the clade posterior probabilities. The root position was placed along the long branch that separated the *S. atocioides* (*pink*) group individuals and close relatives from the majority of other individuals, which is also consistent with mid-point rooting. The putatively recombined sequences are marked by *red branches*. The *scale bar* is in units of substitutions per site
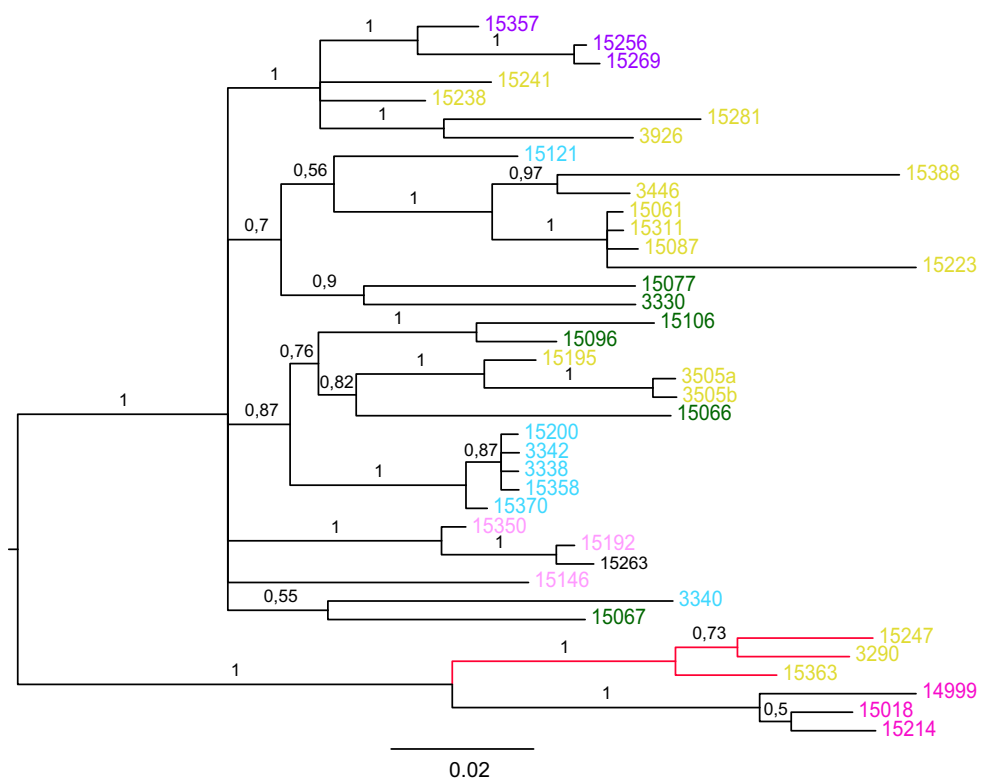


from the *S. atocioides* (pink) group into the *S. aegyptiaca* (yellow) group.

The monophyly of the three sequences in the 3′ part of EST09 suggests that a single successful introgression event took place, with subsequent divergence of allele lineages after the event. The lack of monophyly of these sequences in the 5′ part of the gene is compatible with this scenario, if the 3′ donor sequence was independently recombined at least twice (to explain the two supported different positions in Fig. 2) with different alleles present in the host species, *S. aegyptiaca* (yellow), after the introgression event (Fig. 5). Recombination hot spots are well known (Lichten and Goldman 1995; Hey 2004), so recombination in approximately the same position is not unexpected.

## Origin of recombination via introgressive hybridisation

If our calibration is correctly placed, the donor lineage of the recombination *S. atocioides* group (pink) had already diverged from the remaining species, including the *S. aegyptiaca* (yellow) group, by the time the recombination occurred. The *S. aegyptiaca* (yellow) group, on the other hand, diverged from its sister taxon *S. aegyptiaca* subsp. *ruderalis* (lilac) (Fig. 1) at 0.017–0.031, which is largely overlapping with the time of recombination. However, the crown age of the *S. aegyptiaca* (yellow) group only barely overlaps with the recombination time (0.009–0.018). This suggests that the recombination event most likely took place in the stem of the *S. aegyptiaca* (yellow)

group, presumably shortly after speciation, because no trace of the event has yet been found in the *S. a.* subsp. *ruderalis* (lilac) group. However, caution should be taken in this respect, because more extensive sampling within the *S. a* subsp. *ruderalis* (lilac) group may yet reveal recombinant sequences within that lineage.

The timing of the recombination event is consistent with ancient hybridisation that has brought together disparate alleles into the same genome and allowed meiotic recombination to occur (Fig. 5). In contrast, the timing we observe for the divergence between the position of the recombined sequences in the 5′ and 3′ parts of the gene is not consistent with paralogy or deeply coalescing alleles as the source. This is because the divergence between the source lineage and the recombinant partition is around the same age as the divergence between the yellow and lilac groups, rather than earlier (i.e. earlier than the speciation of all species here) as would be expected for a paralogue or a deeply coalescing allele drawn from the yellow group that has its closest relationship to the pink group).

To explain the source of the recombined alleles by either paralogous loci or deeply coalescing alleles requires several additional ad hoc assumptions. For paralogy, a gene duplication (one event), loss of copies to leave a single copy per individual (several events) and finally two unequal recombinations between the paralogous copies (two events) would be required to explain the observations. A similar number of events are also required for recombination involving a deeply
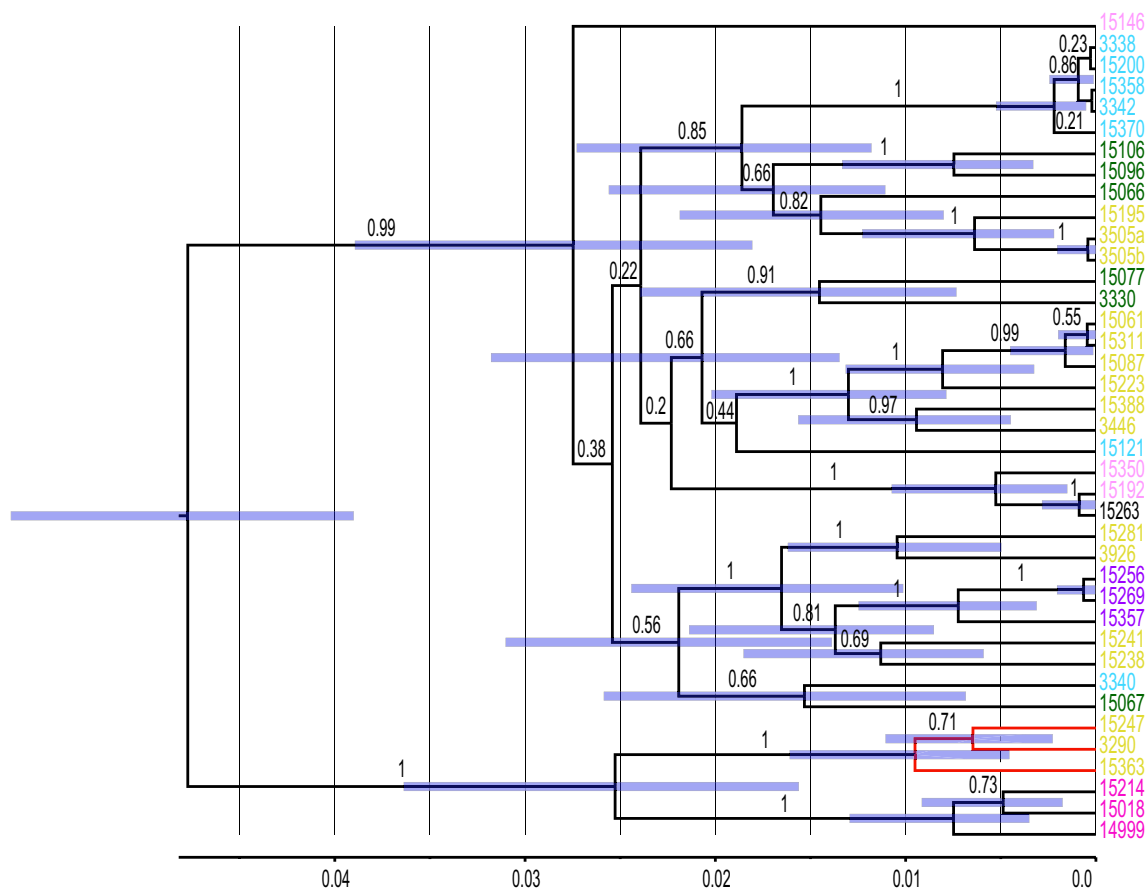
**Fig. 4** Maximum clade credibility chronogram inferred using BEAST of the 3′ partition of EST09. *Numbers above branches* are the clade posterior probabilities. The putatively recombined sequences are marked by *red branches*. *Bars overlaying nodes* show the 95% credibility interval of node ages. The timescale is in units of substitutions per site

coalescing allele, but especially important is the maintenance of multiple allele lineages through several speciation events despite genetic drift. This is considered unlikely, because the relevant branches in the species tree (Fig. 1) are also observed in most genes trees when independently analysed (not shown).

These explanations are less parsimonious than introgressive hybridisation, which requires the introgression itself (one event) and two ordinary meiotic recombinations (two events). The hybridisation is also compatible with the geographic proximity of the *S. atocioides* (pink) and *S. aegyptiaca* (yellow) groups, with the individuals bearing the recombined sequences found along the westernmost part of the distribution of the *S. aegyptiaca* (yellow) group, closest to the *S. atocioides* (pink) group individuals (Fig. 6).

We also rule out template switching during amplification (where the templates in question are derived from either the paralogy or deep coalescence scenarios above) as a viable explanation of the origin of the recombined sequences. This is for the simple reason that template switching during amplification requires the presence of the templates that the

polymerase jumps between, in the same DNA extract (Pääbo et al. 1990). This means that both templates must have been in the genome to start with, which generates the same problems with paralogy and deeply coalescing alleles as before. If the ad hoc assumptions are considered too problematic to believe that these templates were in the same genome ready for recombination, then they are also not available for template switching. However, template switching involving distinct allele lineages brought into the population via hybridisation may have occurred. Whether the recombination occurred in vivo or during amplification does not affect our conclusion that hybridisation was the cause of the distinct sequences, but it does change when and where the recombination occurred.

Contamination with other samples in this study can also be ruled out, because the branch lengths of the recombined piece in the 3′ partition analysis shows many changes since recombination, which is clearly at odds with this explanation. Contamination should result in identical or near-identical sequences in the recombined part to the source of the contamination (which is most likely from the same study, given PCR primer specificities).
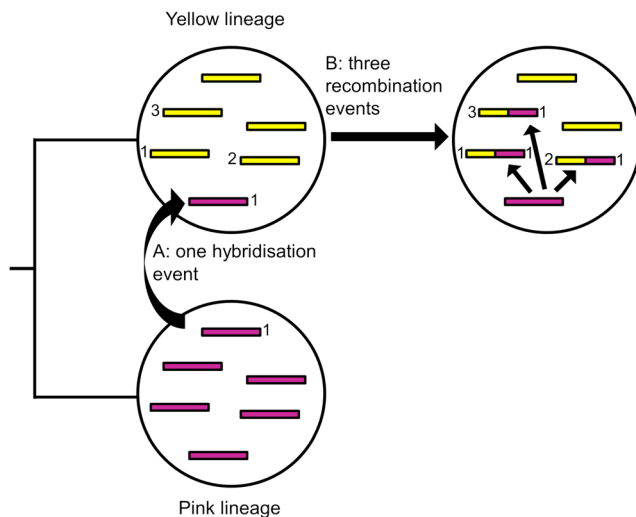
**Fig. 5** Proposed model of the introgression and recombination between the *S. atocioides* (pink) and *S. aegyptiaca* (yellow) *S. aegyptiaca* lineages. Event A is a hybridisation episode leading to the introgression of a *S. atocioides* (pink) allele of EST09 into the *S. aegyptiaca* (yellow) population. This would have occurred at ca. 0.026 relative time units. Event B is recombination among alleles in the *S. aegyptiaca* (yellow) population over time since introgression to produce several combinations between the single *S. atocioides* (pink) allele lineage from the source population and several *S. aegyptiaca* (yellow) allele lineages present in the *S. atocioides* (pink) population. We presuppose that a recombination hotspot is present to generate approx. the same breakpoint in each case. This results in alelle combinations with monophyletic *S. atocioides* (pink) alleles (all derived from pink allele 1) in the recombinants, but more diverse *S. aegyptiaca* (yellow) alleles (derived from yellow alleles 1, 2 and 3) in the recombinants, consistent with the gene tree observations. Alternatively, two independent recombinations (rather than three) and subsequent divergence of one *S. aegyptiaca* (yellow) allele lineage in the recombinants could explain the gene trees

## Implications of hybridisation on barriers to gene flow and mode of speciation

Our results imply rare hybridisation between lineages that have clearly diverged. This may offer some insight into their mode of speciation. Barrier formation can reduce or cut-off gene flow, simply due to isolation by distance, but generally only genetic drift acts to differentiate the now separated gene pools (Seehausen et al. 2014). Secondary contact can often be expected to result in some gene flow, limited by the amount of time since the barrier was formed.

In contrast, other modes of speciation not involving physical barriers to gene flow include chromosomal, ecological or ploidy-driven speciation. These are expected to result in intrinsic genetic isolation fairly early in the speciation continuum, such that secondary contact is unlikely to result in gene flow (Coyne and Orr 2004; Seehausen et al. 2014; Nosil et al. 2008; Servedio et al. 2011; Muir et al. 2012; Guirao-Rico et al. 2017). Given that hybridisation appears to be the cause of recombination (either in vivo or in vitro) in these *Silene* individuals, and that the species involved are distinct (Toprak et al.

2016), it appears that secondary contact has resulted in some but limited gene flow. The "*S. atocioides*" (pink) and "*S. aegyptiaca*" (yellow) clades occupy vicariant geographical areas, adding to the probability that speciation by geographical isolation is the most likely mode of speciation between the ancestors of these species.

## Biogeographic implications of hybridisation across the Anatolian Diagonal

Anatolia is biogeographically interesting because it is a hotspot of biological diversity (Ansell et al. 2011) and because it stands as a bridge between Europe and the Near East, allowing a mixture of the biotas from these areas (Davis 1965–1985). Stretching across central-southern to northeastern Anatolia is the Anatolian Diagonal (AD), a chain of mountains known to be a barrier to biological exchange between the European and Near Eastern sides of the region (Davis 1965–1985; Çiplak et al. 1993), although only for some taxa (Gündüz et al. 2007; Bilgin 2011).

In some taxa, the AD is thought to have separated eastern and western refugia during the last glacial maximum (LGM) (Bilgin 2011), when mountains above 2000 m were glaciated (Atalay 1996). However, in the taxa examined here, the clade on the western side of the AD and the clade on the eastern side (and the mountains of the Diagonal itself) include two and seven species, respectively. This much greater diversity (compared to phylogeographic partitioning of haplotypes within a single species) suggests that the processes that drove speciation between these clades was much earlier than the LGM. However, given the very high sequence diversity observed (and therefore probably high local rates) coupled with the absence of a reliable local calibration point for this group, we cannot pinpoint divergence times any more precisely than this. In the absence of more precise divergence times, it becomes rather speculative to link abiotic factors to clade divergences, but a few possibilities can at least be discussed.

The first of these possibilities is the refilling of the Mediterranean basin at the end Messinian, ca. 5.3 Ma (Garcia-Castellanos and Villaseñor 2011). As the sea level was much lower before the refilling took place, for about 600,000 years (Garcia-Castellanos and Villaseñor 2011), there would have been more land at lower altitudes available as habitat. This could have allowed a more widespread and possibly continuous distribution of the ancestral species between what is now southern Anatolia and the Levant. With the catastrophic refilling of the basin, the ancestor may have been split between these two endpoints and diverged as a consequence of the isolation (i.e. vicariant speciation).

Another process with an overlapping time frame is the folding, uplifting and increased volcanism in the eastern Anatolian plateau that raised the height of the plateau and of the AD in particular, thought to have taken place around the

**Fig. 6** Distribution of individuals in southern Anatolia and western Middle East. Western clade individuals are in *pink* and *light pink*. The three individuals with recombined EST09 sequences from the yellow group are marked with *arrows*. The *figure on the lower right* shows the approximate position of the Anatolian Diagonal (*dashed red lines*)
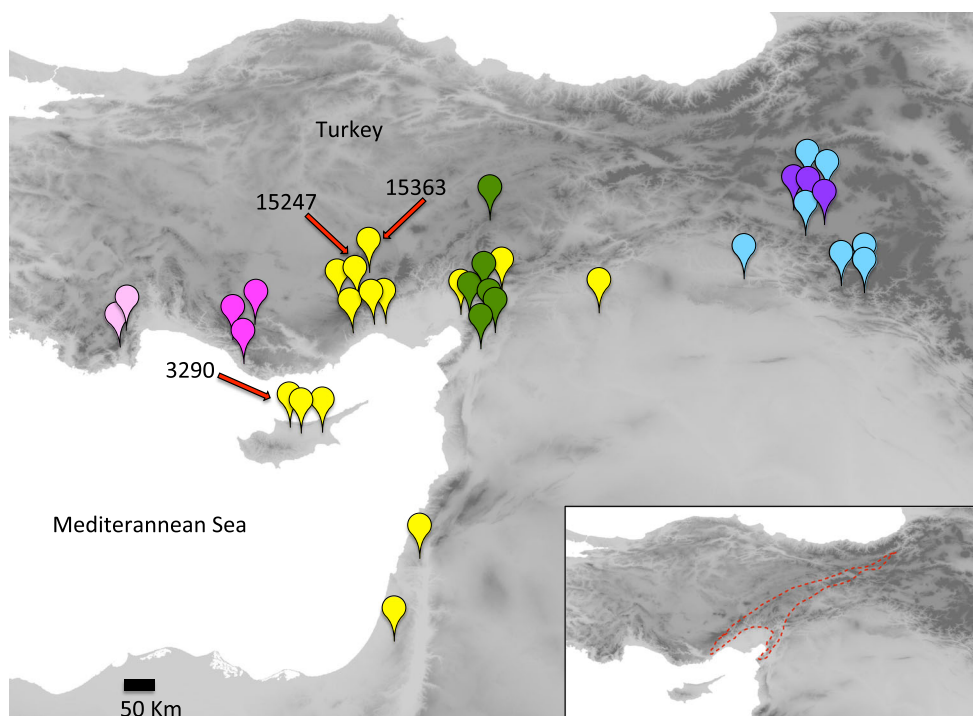


mid- to late-Miocene (Şaroğlu and Yilmaz 1986). If the ancestral species was more restricted to mountain slopes, rather than able to thrive at near sea level altitudes, then the increasing height of the Anatolian Diagonal may at first have increased the available habitat area by increasing the area of mountainous regions. Then later, as the peaks became too high, the presence of glaciers on the highest peaks during glacial times (e.g. from the early Pleistocene, c. 2.5 Ma; Ansell et al. 2011) could have caused the mountain chain itself to become a barrier between the western and eastern populations, again leading to speciation.

Any of the glacial cycles up to the LGM could have led to the production of a barrier along the Anatolian Diagonal, but perhaps the initial creation of a barrier would have produced the isolation necessary to cause speciation, if the ancestral species was already in position on either side of the mountains at that time.

Changes in sea level during the Pleistocene as a result of glacial cycles could also have opened up more habitats away from the glaciated mountains more recently than the end Messinian. Thus, a low-altitude corridor may have repeatedly opened and closed from 2.5 Ma onwards, forming a connection at times and a barrier at times more recently than the initial refilling of the basin. Therefore, an ancestral species adapted to lower altitudes could have been split between the western and eastern regions discussed here at many different times.

In summary, there have been several opportunities for ancestral populations adapted to either lower or higher altitudes to have been split between regions to the west and east of the AD since the end of the Miocene. Indeed, initial divergence without

physical barriers is possible, but perhaps less plausible, given the relatively strong geographical pattern. We are therefore left with many reasonable hypotheses to test, if the means to achieve more precise dating becomes available, but no clear indication with the data at hand as to which of these might be more plausible. But in any case, the evidence for ancient hybridisation that we have found, via recombination in EST09, shows that whichever barrier may have driven speciation, it was a semi-porous barrier that has allowed for at least some exchange at some time(s) after the initial divergence among the *S. atocioides* (pink) and *S. aegyptiaca* (yellow) lineages.

### An example of gradual speciation?

Hybridisation among close relatives with similar morphologies can complicate the distinctiveness of affected species. Evidence for hybridisation among closely related plants has been steadily accumulating in various groups of angiosperms (Rieseberg 1997; Mallet 2005). Where speciation has been caused by abiotic barriers that are permeable, secondary contact that results in some gene flow is perhaps to be expected. Our results, with many fine-grained but conflicting clusters suggested by the DISSECT analysis, within each of the nine that are subtended by highly supported branches (Toprak et al. 2016), are suggestive of gradual speciation. The signal of an isolated ancient introgression event that we present here, well after the primary divergence of the species involved, may represent the "last gasp" of gene flow between these gradually diverging gene pools.

# References

Ansell, S. W., Stenøien, H. K., Grundmann, M., Russell, S. J., Koch, M. A., Schneider, H., & Vogel, J. C. (2011). The importance of Anatolian mountains as the cradle of global diversity in *Arabis alpina*, a key arctic–alpine species. *Annals of Botany, 108*, 241–252.

Atalay, I. (1996). Palaeosols as indicators of the climatic changes during quaternary period in S. Anatolia. *Journal of Arid Environments, 32*, 23–35.

Aydin, Z., Ertekin, A. S., Långström, E., & Oxelman, B. (2014a). A new section of Silene (Caryophyllaceae) including a new species from south Anatolia, Turkey. *Phytotaxa, 178*, 98–112.

Aydin, Z., Marcussen, T., Ertekin, A. S., & Oxelman, B. (2014b). Marginal likelihood estimate comparisons to obtain optimal species delimitations in Silene sect. Cryptoneurae (Caryophyllaceae). *PloS One, 9*, e106990.

Bilgin, R. (2011). Back to the suture: the distribution of intraspecific genetic diversity in and around Anatolia. *International Journal of Molecular Sciences, 12*, 4080–4103.

Çiplak, B., Demirsoy, A., & Bozcuk, A. N. (1993). Distribution of Orthoptera in relation to the Anatolian Diagonal in Turkey. *Articulata, 8*, 1–20.

Coode, M. J. E., & Cullen, J. (1967). Silene L. In P. H. Davis (Ed.), *Flora of Turkey and the East Aegean islands* (Vol. 2, pp. 179–242). Edinburgh: Edinburgh University Press.

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland: Sinauer Associates.

Cronn, R. C., Cedroni, M., Haselkorn, T., Grover, C., & Wendel, J. F. (2002). PCR–mediated recombination in amplification products derived from polyploid cotton. *Theoretical and Applied Genetics, 104*, 482–489.

Davis, P. H. (1965–1985). *Flora of Turkey and the East Aegean islands* (Vol. 1–9). Edinburgh: Edinburgh University Press.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution, 29*, 1969–1973.

Erixon, P., & Oxelman, B. (2008). Whole–gene positive selection, elevated synonymous substitution rates, duplication, and indel evolution of the chloroplast clpP1 gene. *PloS One, 3*, e1386.

Frajman, B., Eggens, F., & Oxelman, B. (2009). Hybrid origins and homoploid reticulate evolution within Heliosperma (Sileneae, Caryophyllaceae)—a multigene phylogenetic approach with relative dating. *Systematic Biology, 58*, 328–345.

Garcia–Castellanos, D., & Villaseñor, A. (2011). Messinian salinity crisis regulated by competing tectonics and erosion at the Gibraltar arc. *Nature, 480*, 359–363.

Guirao–Rico, S., Sánchez-Gracia, A., & Charlesworth, D. (2017). Sequence diversity patterns suggesting balancing selection in partially sex-linked genes of the plant *Silene latifolia* are not generated by demographic history or gene flow. *Molecular Ecology, 26*, 1357–1370. doi:10.1111/mec.13969.

Gündüz, I., Jaarola, M., Tez, C., Yeniyurt, C., Polly, P. D., & Searle, J. B. (2007). Multigenic and morphometric differentiation of ground squirrels (Spermophilus, Sciuridae, Rodentia) in Turkey, with a description of a new species. *Molecular Phylogenetics and Evolution, 43*, 916–935.

Heled, J., & Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution, 27*, 570–580.

Hey, J. (2004). What's so hot about recombination hotspots? *PLoS Biology, 2*, e190.

Jelesko, J. G., Carter, K., Thompson, W., Kinoshita, Y., & Gruissem, W. (2004). Meiotic recombination between paralogous RBCSB genes on sister chromatids of *Arabidopsis thaliana*. *Genetics, 166*, 947–957.

Jones, G., Aydin, Z., & Oxelman, B. (2015). DISSECT: an assignment–free Bayesian discovery method for species delimitation under the multispecies coalescent. *Bioinformatics, 31*, 991–998.

Kelly, L. J., Leitch, A. R., Clarkson, J. J., Hunter, R. B., Knapp, S., & Chase, M. W. (2010). Intragenic recombination events and evidence for hybrid speciation in Nicotiana (Solanaceae). *Molecular Biology and Evolution, 27*, 781–799.

Lichten, M., & Goldman, A. S. H. (1995). Meiotic recombination hotspots. *Annual Review of Genetics, 29*, 423–444.

Linder, C. R., & Rieseberg, L. H. (2004). Reconstructing patterns of reticulate evolution in plants. *American Journal of Botany, 91*, 1700–1708.

Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology Evolution, 20*, 229–237.

Martin, D. P., Lemey, P., Lott, M., Moulton, V., Posada, D., & Lefeuvre, P. (2010). RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics, 26*, 2462–2463.

Muir, G., Dixon, C. J., Harper, A. J., & Filatov, D. A. (2012). Dynamics of drift, gene flow and selection during speciation in Silene. *Evolution, 66*, 1447–1458.

Nosil, P., Egan, S. P., & Funk, D. J. (2008). Heterogeneous genomic differentiation between walking–stick ecotypes: "isolation by adaptation" and multiple roles for divergent selection. *Evolution, 62*, 316–336. doi:10.1111/j.1558-5646.2007.00299.x.

Pääbo, S., Irwin, D. M., & Wilson, A. C. (1990). DNA damage promotes jumping between templates during enzymatic amplification. *Journal of Biological Chemistry, 265*, 4718–4721.

Penny, D., White, W. T., Hendy, M. D., & Phillips, M. J. (2008). A bias in ML estimates of branch lengths in the presence of multiple signals. *Molecular Biology and Evolution, 25*, 239–242.

Petri, A., Pfeil, B. E., & Oxelman, B. (2013). Introgressive hybridization between anciently diverged lineages of Silene (Caryophyllaceae). *PloS One, 8*, e67729.

Pfeil, B. E. (2009). The effect of incongruence on molecular dates. *Taxon, 58*, 511–518.

Pfeil, B. E., Brubaker, C. L., Craven, L. A., & Crisp, M. D. (2004). Paralogy and orthology in the Malvaceae rpb2 gene family: investigation of gene duplication in Hibiscus. *Molecular Biology and Evolution, 21*, 1428–1437.

Posada, D., & Crandall, K. A. (2002). The effect of recombination on the accuracy of phylogenetic estimation. *Journal of Molecular Evolution, 54*, 396–402.

Rautenberg, A., Filatov, D., Svennblad, B., Heidari, N., & Oxelman, B. (2008). Conflicting phylogenetic signals in the SlX1/Y1 gene in Silene. *BMC Evolutionary Biology, 8*, 299.

Rieseberg, L. H. (1997). Hybrid origins of plant species. *Annual Review of Ecology and Systematics, 28*, 359–389.

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics, 19*, 1572–1574.

Rosenberg, N. A. (2003). The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly, and polyphyly in a coalescent model. *Evolution, 57*, 1465–1477.

Sanderson, M. J., & Doyle, J. J. (1992). Reconstruction of organismal and gene phylogenies from data on multigene families: concerted evolution, homoplasy, and confidence. *Systematic Biology, 41*, 4–17.

Şaroğlu, F., & Yilmaz, Y. (1986). Geological evolution and basin models during neotectonic episode in the eastern Anatolia. http://www.mta.gov.tr/v2.0/eng/.

Scheen, A.-C., Pfeil, B. E., Petri, A., Heidari, N., Nylinder, S., & Oxelman, B. (2012). Use of allele–specific sequencing primers is an efficient alternative to PCR subcloning of low–copy nuclear genes. *Molecular Ecology Resources, 12*, 128–135.

Schwartz, T., Nylinder, S., Ramadugu, C., Antonelli, A., & Pfeil, B. E. (2015). The origins of oranges: a multi–locus phylogeny of Rutaceae subfamily Aurantioideae. *Systematic Botany, 40*, 1053–1062. doi:10.1600/036364415X690067.

Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., Peichel, C. L., Saetre, G.-P., C, B., Brannstrom, A., Brelsford, A., Clarkson, C. S., Eroukhmanoff, F., Feder, J. L., Fischer, M. C., Foote, A. D., Franchini, P., Jiggins, C. D., Jones, F. C., Lindholm, A. K., Lucek, K., Maan, M. E., Marques, D. A., Martin, S. H., Matthews, B., Meier, J. I., Most, M., Nachman, M. W., Nonaka, E., Rennison, D. J., Schwarzer, J., Watson, E. T., Westram, A. M., & Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics, 15*, 176–192.

Servedio, M. R., Doorn, G. S. V., Kopp, M., Frame, A. M., & Nosil, P. (2011). Magic traits in speciation: 'magic' but not rare? *Trends in Ecology and Evolution, 26*, 389–397.

Toprak, Z., Pfeil, B. E., Jones, G., Marcussen, T., Ertekin, A. S., & Oxelman, B. (2016). Species delimitation without prior knowledge: DISSECT reveals extensive cryptic speciation in the Silene aegyptiaca complex (Caryophyllaceae). *Molecular Phylogenetics and Evolution, 102*(2016), 1–8. doi:10.1016/j.ympev.2016.05.024.

Visser, J. C., Bellstedt, D. U., & Pirie, M. D. (2012). The recent recombinant evolution of a major crop pathogen, potato virus Y. *PloS One, 7*, e50631.

Wendel, J. F., & Doyle, J. J. (1998). Phylogenetic incongruence: window into genome history and molecular evolution. In D. E. Soltis, P. S. Soltis, & J. J. Doyle (Eds.), *Molecular systematics of plants II: DNA sequencing* (pp. 265–296). Norwell: Kluwer Academic Publishers.

Yandeau–Nelson, M. D., Xia, Y., Li, J., Neuffer, M. G., & Schnable, P. S. (2006). Unequal sister chromatid and homolog recombination at a tandem duplication of the a1 locus in maize. *Genetics, 173*, 2211–2226.