Exploring the potential of gametic reconstruction of parental genotypes by F₁ hybrids as a bridge for rapid introgression

Journal:	<i>Genome</i>
Manuscript ID	gen-2016-0181.R1
Manuscript Type:	Article
Date Submitted by the Author:	25-Apr-2017
Complete List of Authors:	Cannon, Charles; The Morton Arboretum, Center for Tree Science; The Morton Arboretum Scher, C.; The Morton Arboretum, Center for Tree Science
Is the invited manuscript for consideration in a Special Issue? :	Evolution of Tree Diversity
Keyword:	conspecific pollen precedence, crossing-over, F ₁ hybrids, random assortment, introgression

SCHOLARONE™
Manuscripts

1 Exploring the potential of gametic reconstruction of parental genotypes by F₁ hybrids as a bridge
2 for rapid introgression

3

4 Charles H. Cannon* and C. Lane Scher

5

6 Center for Tree Science, The Morton Arboretum, Lisle, IL 60532

7 *corresponding author: ccannon@mortonarb.org

Draft

Abstract

Interspecific hybridization and genetic introgression are commonly observed in natural populations of many species, especially trees. Among oaks, gene flow between closely-related species has been well documented. And yet, hybridization does not lead to a ‘melting pot’, i.e. the homogenization of phenotypic traits. Here, we explore how the combination of several common reproductive traits and genomic biology could create an avenue for interspecific gene flow that partially explains this apparent paradox. During meiosis, F_1 hybrids will produce approximately $(\frac{1}{2})^n$ ‘reconstructed’ parental gametes, where n equals the number of chromosomes. Crossing-over would introduce a small amount of introgressive material. The resulting parental-type gametophytes would probably possess a similar fertilization advantage as conspecific pollen. The resulting ‘backcross’ would actually be the genetic equivalent of a conspecific out-cross, with a small amount of heterospecific DNA captured through crossing-over. Even with detailed genomic analysis, the resulting offspring would not appear to be a backcross. This avenue for rapid introgression between species through the F_1 hybrid will be viable for organisms that meet certain conditions: low base chromosome number, conserved genomic structure and size, production of billions of gametes/gametophytes during each reproductive event, and conspecific fertilization advantage.

Keywords

Conspecific pollen precedence; crossing-over; F_1 hybrids; random assortment; meiosis; introgression

Introduction

Oaks are famous for their ability to hybridize and yet they maintain their phenotypic and ecological integrity (Cottam et al. 1959; Hardin 1975). Trees in this genus span a broad ecological and phenotypic range and have diversified on several different continents (Nixon 1997; Wu et al. 1999; Manos and Stanford 2001; Menitsky 2005; Kremer et al. 2007; Hipp et al. 2014). While oaks are often considered unusual in their promiscuity, the ability to produce viable hybrid offspring among closely related species appears to be a characteristic of many trees (Sutton et al. 1991; Martinsen et al. 2001; Field et al. 2009). Viable crosses have been performed between tree species which have been separated for hundreds of thousands of years (Santamour 1972; Bacon and Spellenberg 1996). In the case of *Liriodendron*, the North American and Chinese species of tulip poplars produce viable crosses despite a separation of >14 million years (Nie et al. 2008).

Our rather poor state of knowledge about tree reproduction is strongly biased towards temperate species. Very little is known about gene flow and reproductive biology of tropical tree species (Ward et al. 2005). Given the ecological setting of tropical trees in mega-diverse communities where sympatry is a common feature of these ecosystems, abundant opportunity for interspecific gene flow exists. The ability to remain inter-fertile also enhances the stochastic ability of rare species to avoid local extinction (Cannon and Lerdau 2015). Given these basic characteristics of tree reproductive biology, the syngameon (Lotsy 1925) appears to be an appropriate evolutionary model for tree evolution. Here, we define the term syngameon as a suite of partially inter-fertile species that exchange genes, either directly or indirectly, under certain evolutionary and biogeographic conditions and where inter-specific gene flow is episodic and spatially variable in nature.

One condition required for participation in a syngameon is a conserved genomic structure and ploidy level among the inter-fertile species. Little polyploidization has been observed in most woody plants (Soltis and Soltis 2009; Chen et al. 2014), with natural examples occurring in low frequencies. Genome architecture is also deeply conserved in trees. For example, the peach reference genome served as a valuable resource for the assembly and alignment of an oak genome (Plomion et al. 2016), despite >90 million years of evolutionary divergence (Hedges et al. 2015). Widespread synteny has been observed in large diverse and inter-fertile genera like *Eucalyptus* (Hudson et al. 2011). Genomic structure is more closely related to growth form than phylogeny, indicating that being woody and long-lived can have a strong impact on genomic structure (Staton et al. 2015). Most angiosperm tree genomes are small and have few chromosomes (Table 1, Ohri 2005; Rice et al. 2015), which probably enhances the conservation of genomic structure. These characteristics would facilitate the flow of genetic material among tree species participating in a syngameon.

These syngameon are not ‘melting pots’, where species slowly merge through time (Stanford 1995). Divergence with gene flow is an increasingly documented mode of diversification and many examples have been observed and studied (Arnold 2015). Even in oak populations with many hybrid offspring and evidence of introgression, there is no phenotypic melting pot (Moran et al. 2012). The three most species rich genera of tree in the world all belong to the Myrtaceae family (Beech et al. 2017), which includes *Eucalyptus*, widely known to hybridize, and two tropical groups, neither one of which has been the subject of a taxonomic monograph, possibly indicating the complexity of the evolutionary patterns in the group. Many genetic studies of natural tree populations observe F₁ hybrids but few of the subsequent backcross or F₂ hybrids (Nason et al. 1992; Milne et al. 2003; Milne and Abbott 2008; Pellegrino

et al. 2009) and yet introgression of the nuclear genome has been widely documented (Dodd and Afzal-Rafii 2004, Eaton et al. 2015) with introgressed material found in 1.7-18.83% of individuals (Curtu et al. 2007, Owusu et al. 2015). The degree of introgression was often related to geographic distance between species (Valbuena-Carabaña et al. 2005; Ortego et al. 2015). Rapid recovery of parental type oaks from hybrids and lack of morphologically intermediate individuals in mixed populations of *Q. petraea* and *Q. robur* was attributed to the idea that few genes were necessary to distinguish oak species (Stebbins 1950, Kremer et al. 2002), but recent genomic evidence shows that there are actually a large number of QTLs involved in distinguishing these two species, spread over at least 9 of the 12 chromosomes (Saintagne et al. 2003).

Here, we provide an alternative explanation for these observations. We explore the likelihood that Mendelian segregation and pollen competition would allow F_1 hybrids to act as a direct genetic bridge among partially interfertile tree species in a syngameon. Through recombination, small portions of the genome could be introgressed in a single backcross. Finally, we discuss the evolutionary conditions and probabilities associated with this genetic bridge. During meiosis, random assortment of chromosomes in the F_1 gametes will produce a proportion of gametes that are largely indistinguishable from the parental genotype at a probability of $(\frac{1}{2})^n$, where n equals the number of linkage units or chromosomes (Fig. 1). Given expected crossing-over rates between the two parental types in the F_1 hybrid, the resulting “parental genotype” gametophyte would contain small fragments of the other parental species. We propose that introgression without the presence of an intermediate series of backcrosses is a logical outcome from a particular set of evolutionary (low chromosome number, enormous gamete production, conserved genomic structure, and conspecific gamete advantage).

Trees, particularly wind-pollinated species, produce enormous quantities of pollen (Table 2, Molina et al. 1996) and considerably fewer ovules, which causes increased competition among pollen (Friedman and Barrett 2009). This remarkable production of gametophytes enables the mechanism explored here. In this study, we use computer simulation studies to estimate the proportion of parental type gametophytes expected from an F_1 hybrid during a single reproductive event, given the number of chromosomes and the frequency of crossing-over. We also propose that in the competition for the limited number of ovules, conspecific pollen precedence (Howard 1999) will give parental-type gametophytes greater fertilization success on the stigma of the parental species. The assumption of unequally competitive pollen is based on evidence of conspecific pollen precedence across flowering plants (Carney et al. 1994; Carney et al. 1996; Klips 1999; Chapman et al. 2005), including a variety of trees (Williams et al. 1999; Williams et al. 2001; Vanden Broeck et al. 2006; Lepais et al. 2013). The style and its regulation of pollen tube growth play a critical role in the outcome of mixed open-pollination and ultimate fertilization, providing the mother plant considerable control over mate choice (Lora et al. 2016).

Here, we ask several basic questions in relation to these ideas: 1) how does number of chromosomes and gametes produced during each reproductive event affect the production of 'reconstructed' parental gametophytes from an F_1 hybrid? 2) how does an empirical probability of crossing over affect the proportion of parental gametophytes? and 3) what proportion of parental type individuals would be observed in the subsequent F_2 backcross generation, assuming some degree of fertilization advantage for parent-type gametophytes.

Methods

We simulated gamete production, given various numbers of linkage groups or chromosomes with and without crossing over, assuming basic random assortment of the resulting

chromatids into the gametes, using Python 3.4 (Python Software Foundation). We applied a probability of crossing over that would result in a genomic average of 2 chiasmata per bivalent. This crossover rate is probably an overestimate, given estimates calculated from genetic maps for several species that range from 0.6 to 1.79 (Supp. Table 1) but was based on direct observations of two chiasmata per bivalent in several plant species (Bell 1982) and precise estimates of 1-2 chiasmata per bivalent in *Arabidopsis* (Mézard 2006; Drouaud et al. 2006). Crossing over is known to be highly variable within and across linkage groups (Choi and Henderson 2015; Melamed-Bessudo et al. 2016) but our simplified model of equal probability of crossing over across the genome adequately addresses our questions at the genome level.

For each combination of chromosome number and crossover rate, we simulated between one million and ten million F_1 hybrid parent cells. For each replicate, we created a base set of chromosomes, given n which ranged from 8 to 23, and each pair of chromosomes was independently and randomly assigned between 10 and 50 possible cross-over points. The probability of crossing over was constant across all units across all chromosomes, meaning that chromosomes with more possible cross-over points were more likely to be recombinant. A simple 0.035 probability of crossover between each unit was used to produce an average of 2 crossovers per chromosome across all chromosome lengths. After recombination, four gametes were randomly generated from each parent cell, assuming random assortment of chromatids.

To simulate the subsequent backcross generation, each F_1 hybrid gametophyte was crossed back to the parent species. The probability that each gametophyte would successfully fertilize an ovule was equal to its percent identity to the parent species, for an overall fertilization success rate of ~50%.

Results

With an average of 2 chiasmata per chromosome, 0.006% of the gametophytes from an oak ($n=12$) F_1 hybrid would be more than 90% identical to a gametophyte produced by one of the parental species, given both tails of the distribution, while the proportion of gametophytes that are 80% similar to one of the parental types increases to 0.42%, ~70 times greater (Fig. 2, Table 2). Because an oak tree can produce hundreds of billions of pollen grains in a single reproductive event (Molina et al. 1996), one F_1 hybrid tree would produce 9.3 and 659 million pollen grains that are >90% or >80% identical to each parental genotype, respectively, each year (Table 2). The proportion of parental type gametophytes produced is negatively correlated to chromosome number (Fig. 2) and pollen grains which are 90% parental-type are no longer produced once chromosome number reaches 23 (Table 2), presenting an upper threshold for chromosome number and the direct effect of this mechanism. Even in trees with 23 chromosomes, like ash (*Fraxinus*) and olive (*Olea*), the number of >80% parental genotype pollen grains produced would still be in the millions, indicating that while the effect of this mechanism could be greatly diminished, it may not be entirely absent.

As would be expected, crossing over diminishes the production of parental genotype gametophytes (Fig. 2) but it is interesting to note that the slope of the relationship is almost identical between >90% parental type with no crossing over and >80% parental type with two chiasmata per bivalent. The impact of increasing chromosome number is substantially stronger for >90% parental type gametophytes than >80% parental type gametophytes, as the relationships become increasingly divergent. As our crossover rate is probably an overestimate, our results probably represent a lower bound for the effect of this mechanism.

When the entire distribution of gametophytes produced by a 12-chromosome F_1 hybrid with a crossover rate of 0.035 are backcrossed to a parental type, given a stochastic success rate

of fertilization equal to its identity to that parent, roughly 0.006% of the fertilized ovules would be >95% identical to that parent while 0.37% would be >90% identical to that parent. Due to the traditional assumption that all backcrossed individuals are 75% identical to the parent, these individuals would not fall into the backcross category, despite in reality belonging to that category. Furthermore, we make the simple assumption that the vigor of a pollen grain would be directly related to its similarity to the parental stigma tissue, resulting in a higher proportion of parent-like F_2 backcrosses than would be expected if each gametophyte had equal fitness (Fig. 3). Given the assumption that fitness is related to genotype and that parental genotypes are more fit in a particular environment, these ‘reconstructed’ F_2 offspring, almost identical to a parental type, would also enjoy greater success and fitness and increase in proportion as the plants develop and mature.

Discussion

Some of the basic mechanisms discussed here are well understood while others are poorly known. For example, the independent assortment of chromatids into the gametes is a ‘law’ of Mendelian genetics, although segregation distortion could play an underappreciated role in the meiotic process (Taylor and Ingvarsson 2003). Perhaps surprisingly, we know relatively little about the distribution and frequency of crossing over during meiosis for many organisms. For trees in particular, we know almost nothing about the total reproductive effort of an individual during a single flowering event, e.g. the number of gametophytes produced. As introduced above, tree reproduction represents an extreme combination of these basic reproductive and genomic traits, where individuals can produce huge numbers of gametophytes (Molina et al. 1996), genomes are generally small (Ohri 2005; Chen et al. 2014; Ng et al. 2016) with few chromosomes (Table 1), polyploidization is rare (Soltis and Soltis 2009), and the

overall structure of the genome is deeply conserved (Plomion et al. 2016; Staton et al. 2015). Under these unique circumstances, rapid introgression between species through an F_1 hybrid intermediary becomes possible through the reconstruction of parental genotypes, carrying a proportion of introgressive material directly related to crossover probabilities. This mechanism would allow gene flow among several species while limiting the erosion of overall genetic integrity of each species.

This avenue of gene flow would be strongly dependent on many other factors and would probably only play a significant role during times of rapid change in the selective environment. We want to emphasize that it opens a new alternative for organisms with the appropriate combination of reproductive traits and would be most effective when conspecific fertilization advantage is strong (Howard 1999). Despite the much lower abundance of ‘reconstructed’ parental-type gametophytes, they could potentially outcompete the far more abundant ‘hybrid’ gametophytes because of their greater genetic compatibility with an individual of the same parental type (Fig. 3) and thus achieve fertilization. In controlled studies of pollen competition, conspecific pollen has been shown to be more successful than heterospecific pollen (Williams et al. 1999; Vanden Broeck et al. 2006). Conspecific pollen advantage lowers hybridization rates (Lepais et al. 2013) and reduces hybrid seed production, given the relative proportions of gametophytes in the pollen load (Carney et al. 1994; Chapman et al. 2005). Additionally, seeds fertilized by heterospecific pollen are aborted more often than those fertilized by conspecific pollen (Williams et al. 2001). Genetic studies of natural populations have suggested a similar pattern, demonstrating that trees produce more fruit and seedlings from conspecific and backcross fertilizations than expected (Lepais and Gerber 2011).

The combination of these traits together would suggest subsequent backcross generations

would be rarer in the population than expected, given the presence of F_1 hybrids and equivalent fitness. A study comparing mixed-species communities of three different plants (oaks, manzanitas, and irises: chromosome numbers=12, 13, and 22, respectively), identified only F_1 hybrids among the oaks and no backcrosses while the opposite was found in the manzanita and iris communities (Nason et al. 1992). Studies of two populations of mixed *Rhododendron* ($n=12$) populations found that 80 and 100% of hybrids were F_1 (Milne and Abbott 2008; Milne et al. 2003). With 12 chromosomes, as in oaks and rhododendrons, gametophytes of each parental type make up about 0.2% of total pollen produced (Fig. 2). Mean pollen capture by stigmas among 19 anemophilous species is 0.32% (Friedman and Barrett 2009), meaning that, statistically, at least one parental type pollen grain could land on most stigmas. If conspecific fertilization advantage exists, all of the F_2 backcrosses would be basically identical to a conspecific outcrossing event, despite the fact that the involvement of an F_1 hybrid. Shorter-lived, less reproductively prolific plants like Iris probably face different selective pressures than trees, so other processes like differential flowering times, pollinator fidelity, or polyploidization may arise as reproductive barriers. While backcrosses are frequently found in mixed iris populations, F_1 hybrids are considerably more rare (Nason et al. 1992), suggesting that this bottleneck against the formation of F_1 hybrids may prevent the rapid meltdown of species (Arnold 2015).

Support for our proposed mechanism from existing empirical evidence is difficult to discern, particularly as it proposes an absence of genetic evidence instead of any obvious proof. In our model, the F_1 hybrid cross with a parental species through a reconstructed parental gametophyte with some introgressive material would be indistinguishable from a simple outcrossing event between conspecific individuals of a single species that already possessed

some introgressive material. Introgression of genetic material across species boundaries is common in natural populations of oaks (Eaton et al. 2015) and the proportion of phenotypically ‘pure’ individuals with introgressive material ranges from 1.7 to 18.83% (Curtu et al. 2007; Owusu et al. 2015). Phenotype and genotype are often not correlated. In some mixed populations of oaks, the species assigned to an individual based on morphology may not match the species assigned based on genetic evidence (Hipp and Weber 2008, Cavender-Bares and Pahlich 2009), and in general there are fewer morphologically intermediate individuals than genetically intermediate individuals (González-Rodríguez et al. 2004). We have little evidence about the evolutionary origin of this introgressed material or how dynamic it might be.

Among tree species that do not possess the entire suite of reproductive traits most conducive to this mechanism, like *Populus*, the melting pot effect should be more obvious, unless other barriers to gene flow have evolved. Poplars have 19 chromosomes and produce considerably less pollen than *Quercus* (~4% in abundance, see Table 3), thus reducing the production of parental type pollen by at least two orders of magnitude in comparison to oaks (Table 2). About 0.01% of all pollen produced by *Populus* would be parental type (Figure 2), which is far below the level of probability that at least one parental type gametophyte would land on each stigma. In mixed *Populus* populations, fewer of the hybrids (56-85%) were strictly F_1 (Christe et al. 2016) compared to oaks. In a poplar plantation containing a pure species and large numbers of its F_1 hybrid, there was no evidence of introgression back to the pure species (Vanden Broeck et al. 2006), suggesting the hybrids truly were an evolutionary dead end. Similarly, *Fraxinus* species, with a base number of 23 chromosomes, produce “swarms” of intermediate hybrids with less genetic and morphological distinctness (Hinsinger et al. 2013). The high number of chromosomes in ash trees greatly reduces their potential to produce F_1

hybrids (Figure 2). The common occurrence of polyploids in ash trees, which are generally uncommon in trees, may indicate that other mechanisms have evolved to help avoid the melting pot.

Here, we propose that F_1 hybrid trees can serve as a genetic bridge between partially interfertile species that allows rapid introgression between species while recovering the parental genotype in a single backcross. Current models of introgression implicitly assume that F_1 hybrids produce mixed gametophytes that are intermediate between the parental types. We have shown through quantitative analysis that organisms with similar reproductive traits to an oak tree should produce a considerable fraction of gametophytes that largely reconstruct a parental genotype. While no direct evidence of this result exists, hybrids can be expected to produce gametophytes with a range of genotypes (Figure 1) and fertilization fitness on either parental individual (Figure 3). Ultimately, techniques based on single-pollen genotyping may provide an empirical test in the future (Suyama 2010). This mechanism could be even further enhanced by the possibility of meiotic drivers during the assortment process and potential biases could exist towards the parental gametophyte types in F_1 hybrids. We do not believe that this process would be mutually exclusive of other better known processes and would not affect other forms of hybrid speciation, including homoploid hybrid speciation. We argue that the outlined tree-like traits enable a new pathway that does not replace others.

The results of our simple exercise could help solve the apparent paradox of frequent hybridization in trees like oaks, with the common production of F_1 hybrids and genetic introgression contrasted against the maintenance of phenotypic identity and ecological specialization within a species. The maintenance of a genetic bridge among species through the F_1 hybrid offspring creates a powerful avenue for capturing advantageous alleles and testing

novel alleles, evolved in another species, against a compatible but divergent genomic background. Hybridization will play an increasingly important role in the Anthropocene as populations and communities shift and migrate through both direct human mediation and indirect climate drivers (Chunco 2014). The mechanism described here helps explain how species can remain coherent despite this gene flow and why genetic swamping and phenotypic intermediacy may not be such a threat (Banks et al. 2016).

Acknowledgements

This research was funded by the Center for Tree Science at the Morton Arboretum. Previous versions of the manuscript was greatly improved by comments and advice from A. Hipp and S. Hoban.

Draft

299 References

- 300 Arnold, M. L. 2015. Divergence with Genetic Exchange. Oxford University Press.
- 301 Backs, J. R., Terry, M., and Ashley, M. V. 2016. Using Genetic Analysis to Evaluate
302 Hybridization as a Conservation Concern for the Threatened Species *Quercus hinckleyi*
303 C.H. Muller (Fagaceae). Int. J. Plant Sci., **177**(2): 122–131.
- 304 Bacon, J. R., and Spellenberg, R. 1996. Hybridization in two distantly related Mexican black
305 oaks *Quercus conzattii* and *Quercus eduardii* (Fagaceae: Quercus: Section Lobatae).
306 Sida, **17**(1): 17–41.
- 307 Barreneche, T., Bodenes, C., Lexer, C., Trontin, J.-F., Fluch, S., Streiff, R., Plomion, C.,
308 Roussel, G., Steinkellner, H., Burg, K., Favre, J.-M., Glössl, J., and Kremer, A. 1998. A
309 genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR,
310 microsatellite, minisatellite, isozyme and 5S rDNA markers. Theor. Appl. Genet., **97**(7):
311 1090–1103.
- 312 Cai, Q., Guy, C. L., and Moore, G. A. 1994. Extension of the linkage map in Citrus using
313 random amplified polymorphic DNA (RAPD) markers and RFLP mapping of cold-
314 acclimation-responsive loci. Theor. Appl. Genet., **89**(5): 606–614.
- 315 Cannon, C. H., and Ler dau, M. 2015. Variable mating behaviors and the maintenance of tropical
316 biodiversity. Front. Genet., **6**: 183.
- 317 Carney, S. E., Cruzan, M. B., and Arnold, M. L. 1994. Reproductive Interactions Between
318 Hybridizing Irises: Analyses of Pollen-Tube Growth and Fertilization Success. Am. J.
319 Bot., **81**(9): 1169–1175.
- 320 Carney, S. E., Hodges, S. A., and Arnold, M. L. 1996. Effects of Differential Pollen-Tube
321 Growth on Hybridization in the Louisiana Irises. Evolution, **50**(5): 1871–1878.

- 322 Casasoli, M., Mattioni, C., Cherubini, M., and Villani, F. 2001. A genetic linkage map of
323 European chestnut (*Castanea sativa* Mill.) based on RAPD, ISSR and isozyme markers.
324 Theor. Appl. Genet., **102**(8): 1190–1199.
- 325 Causse, M., Santoni, S., Damerval, C., Maurice, A., Charcosset, A., Deatrick, J., and Vienne, D.
326 1996. A composite map of expressed sequences in maize. Genome, **39**(2): 418–432.
- 327 Cavender-Bares, J., and Pahlich, A. 2009. Molecular, morphological, and ecological niche
328 differentiation of sympatric sister oak species, *Quercus virginiana* and *Q. geminata*
329 (Fagaceae). Am. J. Bot., **96**(9): 1690–1702.
- 330 Cervera, M. T., Storme, V., Ivens, B., Gusmão, J., Liu, B. H., Hostyn, V., Van Slycken, J., Van
331 Montagu, M., and Boerjan, W. 2001. Dense genetic linkage maps of three *Populus*
332 species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and
333 microsatellite markers. Genetics, **158**(2): 787–809.
- 334 Chapman, M. A., Forbes, D. G., and Abbott, R. J. 2005. Pollen competition among two species
335 of *Senecio* (Asteraceae) that form a hybrid zone on Mt. Etna, Sicily. Am. J. Bot., **92**(4):
336 730–735.
- 337 Chen, S.-C., Cannon, C. H., Kua, C.-S., Liu, J.-J., and Galbraith, D. W. 2014. Genome size
338 variation in the Fagaceae and its implications for trees. Tree Genet. Genomes, **10**(4):
339 977–988.
- 340 Choi, K., and Henderson, I. R. 2015. Meiotic recombination hotspots - a comparative view. Plant
341 J., **83**(1): 52–61.
- 342 Christe, C., Stölting, K. N., Bresadola, L., Fussi, B., Heinze, B., Wegmann, D., and Lexer, C.
343 2016. Selection against recombinant hybrids maintains reproductive isolation in
344 hybridizing *Populus* species despite F₁ fertility and recurrent gene flow. Mol. Ecol.,

- 345   **25**(11): 2482–2498.
- 346   Chunco, A. J. 2014. Hybridization in a warmer world. *Ecol. Evol.*, **4**(10): 2019–2031.
- 347   Cottam, W. P., Tucker, J. M., and Drobnick, R. 1959. Some Clues to Great Basin Postglacial
- 348   Climates Provided by Oak Distributions. *Ecology*, **40**(3): 361–377.
- 349   Curtu, A. L., Gailing, O., and Finkeldey, R. 2007. Evidence for hybridization and introgression
- 350   within a species-rich oak (*Quercus* spp.) community. *BMC Evol. Biol.*, **7**: 218.
- 351   Dirlewanger, E., Pronier, V., Parvery, C., Rothan, C., Guye, A., and Monet, R. 1998. Genetic
- 352   linkage map of peach [*Prunus persica* (L.) Batsch] using morphological and molecular
- 353   markers. *Theor. Appl. Genet.*, **97**(5-6): 888–895.
- 354   Dodd, R. S., and Afzal-Rafii, Z. 2004. Selection and dispersal in a multispecies oak hybrid zone.
- 355   *Evolution*, **58**(2): 261–269.
- 356   Drouaud, J., Camilleri, C., Bourguignon, P.-Y., Canaguier, A., Bérard, A., Vezon, D., Giancola,
- 357   S., Brunel, D., Colot, V., Prum, B., Quesneville, H., and Mézard, C. 2006. Variation in
- 358   crossing-over rates across chromosome 4 of *Arabidopsis thaliana* reveals the presence of
- 359   meiotic recombination “hot spots.” *Genome Res.*, **16**(1): 106–114.
- 360   Eaton, D. A. R., Hipp, A. L., González-Rodríguez, A., and Cavender-Bares, J. 2015. Historical
- 361   introgression among the American live oaks and the comparative nature of tests for
- 362   introgression. *Evolution*, **69**(10): 2587–2601.
- 363   Field, D. L., Ayre, D. J., Whelan, R. J., and Young, A. G. 2009. Molecular and morphological
- 364   evidence of natural interspecific hybridization between the uncommon *Eucalyptus*
- 365   *aggregata* and the widespread *E. rubida* and *E. viminalis*. *Conserv. Genet.*, **10**(4): 881–
- 366   896.
- 367   Friedman, J., and Barrett, S. C. H. 2009. Wind of change: new insights on the ecology and

- 368 evolution of pollination and mating in wind-pollinated plants. *Ann. Bot.*, **103**(9): 1515–
369 1527.
- 370 González-Rodríguez, A., Arias, D. M., Valencia, S., and Oyama, K. 2004. Morphological and
371 RAPD analysis of hybridization between *Quercus affinis* and *Q. laurina* (fagaceae), two
372 Mexican red oaks. *Am. J. Bot.*, **91**(3): 401–409.
- 373 Grattapaglia, D., and Sederoff, R. 1994. Genetic linkage maps of *Eucalyptus grandis* and
374 *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers.
375 *Genetics*, **137**(4): 1121–1137.
- 376 Hardin, J. W. 1975. Hybridization and Introgression in *Quercus alba*. *J. Arnold Arbor.*, **56**(3):
377 336–363.
- 378 Hedges SB, Marin J, Suleski M, Paymer M & Kumar S (2015) Tree of Life Reveals Clock-Like
379 Speciation and Diversification. *Mol Biol Evol* 32: 835-845.
- 380 Hinsinger, D. D., Basak, J., Gaudeul, M., Cruaud, C., Bertolino, P., Frascaria-Lacoste, N., and
381 Bousquet, J. 2013. The phylogeny and biogeographic history of ashes (fraxinus,
382 oleaceae) highlight the roles of migration and vicariance in the diversification of
383 temperate trees. *PLoS One*, **8**(11): e80431.
- 384 Hipp, A. L., Eaton, D. A. R., Cavender-Bares, J., Fitzek, E., Nipper, R., and Manos, P. S. 2014.
385 A framework phylogeny of the American oak clade based on sequenced RAD data. *PLoS*
386 *One*, **9**(4): e93975.
- 387 Hipp, A. L., and Weber, J. A. 2008. Taxonomy of Hill's Oak (*Quercus ellipsoidalis*: Fagaceae):
388 Evidence from AFLP Data. *Syst. Bot.*, **33**(1): 148–158.
- 389 Howard, D. J. 1999. Conspecific Sperm and Pollen Precedence and Speciation. *Annu. Rev. Ecol.*
390 *Syst.*, **30**: 109–132.

- 391 Hudson, C. J., Kullán, A. R. K., Freeman, J. S., Faria, D. A., Grattapaglia, D., Kilian, A.,
 392 Myburg, A. A., Potts, B. M., and Vaillancourt, R. E. 2011. High synteny and colinearity
 393 among Eucalyptus genomes revealed by high-density comparative genetic mapping. *Tree*
 394 *Genet. Genomes*, **8**(2): 339–352.
- 395 Kleinhofs, A., Kilian, A., Saghai Maroof, M. A., Biyashev, R. M., Hayes, P., Chen, F. Q.,
 396 Lapitan, N., Fenwick, A., Blake, T. K., Kanazin, V., Ananiev, E., Dahleen, L., Kudrna,
 397 D., Bollinger, J., Knapp, S. J., Liu, B., Sorrells, M., Heun, M., Franckowiak, J. D.,
 398 Hoffman, D., Skadsen, R., and Steffenson, B. J. 1993. A molecular, isozyme and
 399 morphological map of the barley (*Hordeum vulgare*) genome. *Theor. Appl. Genet.*, **86**(6):
 400 705–712.
- 401 Klips, R. A. 1999. Pollen competition as a reproductive isolating mechanism between two
 402 sympatric Hibiscus species (Malvaceae). *Am. J. Bot.*, **86**(2): 269–272.
- 403 Kremer, A., Casasoli, M., Barreneche, T., Bodenes, C., Sisco, P., Kubisiak, T., Scalfi, M.,
 404 Leonardi, S., Bakker, E., Buiteveld, J., Romero-Severson, J., Arumuganathan, K.,
 405 Derory, J., Scotti-Saintagne, C., Roussel, G., Bertocchi, M. E., Lexer, C., Porth, I.,
 406 Hebard, F., Clark, C., Carlson, J., Plomion, C., Koelewijn, H.-P., and Villani, F. 2007.
 407 Fagaceae. In C. Kole (Ed.), *Genome Mapping and Molecular Breeding in Plants: Forest*
 408 *Trees* (pp. 161–187). Springer-Verlag.
- 409 Kremer, A., Dupouey, J. L., Douglas Deans, J., Cottrell, J., Csaikl, U., Finkeldey, R., Espinel, S.,
 410 Jensen, J., Kleinschmit, J., Van Dam, B., Ducousso, A., Forrest, I., de Heredia, U. L.,
 411 Lowe, A. J., Tutkova, M., Munro, R. C., Steinhoff, S., and Badeau, V. 2002. Leaf
 412 morphological differentiation between *Quercus robur* and *Quercus petraea* is stable
 413 across western European mixed oak stands. *Ann. For. Sci.*, **59**(7): 777–787.

- 414 Lepais, O., and Gerber, S. 2011. Reproductive patterns shape introgression dynamics and species
415 succession within the European white oak species complex. *Evolution*, **65**(1): 156–170.
- 416 Lepais, O., Roussel, G., Hubert, F., Kremer, A., and Gerber, S. 2013. Strength and variability of
417 postmating reproductive isolating barriers between four European white oak species. *Tree*
418 *Genet. Genomes*, **9**(3): 841–853.
- 419 Lora, J., Hormaza, J. I., and Herrero, M. 2016. The Diversity of the Pollen Tube Pathway in
420 Plants: Toward an Increasing Control by the Sporophyte. *Front. Plant Sci.*, **7**: 107.
- 421 Lotsy, J. P. 1925. Species or Linneon. *Genetica* **7** (5-6): 487–506.
- 422 Manos, P. S., and Stanford, A. M. 2001. The historical biogeography of Fagaceae: tracking the
423 Tertiary history of temperate and subtropical forests of the Northern hemisphere. *Int. J.*
424 *Plant Sci.*, **162**(6 Suppl.): S77–S93.
- 425 Martinsen, G. D., Whitham, T. G., Turek, R. J., and Keim, P. 2001. Hybrid populations
426 selectively filter gene introgression between species. *Evolution*, **55**(7): 1325–1335.
- 427 Melamed-Bessudo, C., Shilo, S., and Levy, A. A. 2016. Meiotic recombination and genome
428 evolution in plants. *Curr. Opin. Plant Biol.*, **30**: 82–87.
- 429 Mézard, C. 2006. Meiotic recombination hotspots in plants. *Biochem. Soc. Trans.*, **34**(Pt 4):
430 531–534.
- 431 Menitsky, Y. L. 2005. Oaks of Asia. Science Publishers, New Hampshire.
- 432 Milne, R. I., and Abbott, R. J. 2008. Reproductive isolation among two interfertile
433 *Rhododendron* species: low frequency of post-F₁ hybrid genotypes in alpine hybrid
434 zones. *Mol. Ecol.*, **17**(4): 1108–1121.
- 435 Milne, R. I., Terzioglu, S., and Abbott, R. J. 2003. A hybrid zone dominated by fertile F₁s:
436 maintenance of species barriers in *Rhododendron*. *Mol. Ecol.*, **12**(10): 2719–2729.

- 437 Molina, R. T., Rodríguez, A. M., Palaciso, I. S., and López, F. G. 1996. Pollen production in
438 anemophilous trees. *Grana*, **35**(1): 38–46.
- 439 Moran, E. V., Willis, J., and Clark, J. S. 2012. Genetic evidence for hybridization in red oaks
440 (*Quercus* sect. *Lobatae*, *Fagaceae*). *Am. J. Bot.*, **99**(1): 92–100.
- 441 Nason, J. D., Ellstrand, N. C., and Arnold, M. L. 1992. Patterns of Hybridization and
442 Introgression in Populations of Oaks, Manzanitas, and Irises. *Am. J. Bot.*, **79**(1): 101–
443 111.
- 444 Ng, C. H., Lee, S. L., Tnah, L. H., Ng, K. K. S., Lee, C. T., and Madon, M. 2016. Genome size
445 variation and evolution in *Dipterocarpaceae*. *Plant Ecol. Divers.*, **9**(5-6): 437–446.
- 446 Nie, Z.-L., Wen, J., Azuma, H., Qiu, Y.-L., Sun, H., Meng, Y., Sun, W.-B., and Zimmer, E. A.
447 2008. Phylogenetic and biogeographic complexity of *Magnoliaceae* in the Northern
448 Hemisphere inferred from three nuclear data sets. *Mol. Phylogenet. Evol.*, **48**(3): 1027–
449 1040.
- 450 Nixon, K. C. 1997. In *Flora of North America* (pp. 436–506). Oxford University Press.
- 451 Ohri, D. 2005. Climate and growth form: the consequences for genome size in plants. *Plant Biol.*
452 , **7**(5): 449–458.
- 453 Ortego, J., Nogueras, V., Gugger, P. F., and Sork, V. L. 2015. Evolutionary and demographic
454 history of the Californian scrub white oak species complex: an integrative approach. *Mol.*
455 *Ecol.*, **24**(24): 6188–6208.
- 456 Owusu, S. A., Sullivan, A. R., Weber, J. A., Hipp, A. L., and Gailing, O. 2015. Taxonomic
457 Relationships and Gene Flow in Four North American *Quercus* Species (*Quercus* section
458 *Lobatae*). *Syst. Bot.*, **40**(2): 510–521.
- 459 Pellegrino, G., Bellusci, F., and Musacchio, A. 2009. Genetic integrity of sympatric hybridising

460 plant species: the case of *Orchis italica* and *O. anthropophora*. *Plant Biol.* , **11**(3): 434–
461 441.

462 Plomion, C., Aury, J.-M., Amselem, J., Alaeitabar, T., Barbe, V., Belser, C., Bergès, H.,
463 Bodénès, C., Boudet, N., Boury, C., Canaguier, A., Couloux, A., Da Silva, C., Duplessis,
464 S., Ehrenmann, F., Estrada-Mairey, B., Fouteau, S., Francillonne, N., Gaspin, C.,
465 Guichard, C., Klopp, C., Labadie, K., Lalanne, C., Le Clainche, I., Leplé, J.-C., Le
466 Provost, G., Leroy, T., Lesur, I., Martin, F., Mercier, J., Michotey, C., Murat, F., Salin,
467 F., Steinbach, D., Faivre-Rampant, P., Wincker, P., Salse, J., Quesneville, H., and
468 Kremer, A. 2016. Decoding the oak genome: public release of sequence data, assembly,
469 annotation and publication strategies. *Mol. Ecol. Resour.*, **16**(1): 254–265.

470 Python Software Foundation. Python Language Reference, version 3.4. Available at
471 <https://www.python.org/>

472 Reiter, R. S., Williams, J. G., Feldmann, K. A., Rafalski, J. A., Tingey, S. V., and Scolnik, P. A.
473 1992. Global and local genome mapping in *Arabidopsis thaliana* by using recombinant
474 inbred lines and random amplified polymorphic DNAs. *Proc. Natl. Acad. Sci. U. S. A.*,
475 **89**(4): 1477–1481.

476 Rice et al. 2015. The Chromosome Counts Database (CCDB) – a community resource of plant
477 chromosome numbers. *New Phytol.* **206**(1): 19-26

478 Saintagne, C., Bodénès, C., Barreneche, T., Pot, D., Plomion, C., and Kremer, A. 2004.
479 Distribution of genomic regions differentiating oak species assessed by QTL detection.
480 *Heredity* , **92**(1): 20–30.

481 Santamour, F. S. 1972. Interspecific Hybrids in *Liriodendron* and Their Chemical Verification.
482 *For. Sci.*, **18**(3): 233–236.

- 483 Soltis, P. S., and Soltis, D. E. 2009. The role of hybridization in plant speciation. *Annu. Rev.*
 484 *Plant Biol.*, **60**: 561–588.
- 485 Stanford, P. K. 1995. For Pluralism and against Realism about Species. *Philos. Sci.*, **62**(1): 70–
 486 91.
- 487 Staton, M., Zhebentyayeva, T., Olukolu, B., Fang, G. C., Nelson, D., Carlson, J. E., and Abbott,
 488 A. G. 2015. Substantial genome synteny preservation among woody angiosperm species:
 489 comparative genomics of Chinese chestnut (*Castanea mollissima*) and plant reference
 490 genomes. *BMC Genomics*, **16**: 744.
- 491 Stebbins, G. L. 1950. Variation and evolution in plants. Columbia University Press.
- 492 Sutton, B. C. S., Flanagan, D. J., Gawley, J. R., Newton, C. H., Lester, D. T., and El-Kassaby, Y.
 493 A. 1991. Inheritance of chloroplast and mitochondrial DNA in *Picea* and composition of
 494 hybrids from introgression zones. *Theor. Appl. Genet.*, **82**(2): 242–248.
- 495 Suyama, Y. 2010. Procedure for Single-Pollen Genotyping. In *Single-Pollen Genotyping* (pp. 7–
 496 15).
- 497 Tanksley, S. D., Ganai, M. W., Prince, J. P., de Vicente, M. C., Bonierbale, M. W., Broun, P.,
 498 Fulton, T. M., Giovannoni, J. J., Grandillo, S., and Martin, G. B. 1992. High density
 499 molecular linkage maps of the tomato and potato genomes. *Genetics*, **132**(4): 1141–1160.
- 500 Taylor, D. R., and Ingvarsson, P. K. 2003. Common features of segregation distortion in plants
 501 and animals. *Genetica*, **117**(1): 27–35.
- 502 Valbuena-Carabaña, M., González-Martínez, S. C., Sork, V. L., Collada, C., Soto, A.,
 503 Goicoechea, P. G., and Gil, L. 2005. Gene flow and hybridisation in a mixed oak forest
 504 (*Quercus pyrenaica* Willd. and *Quercus petraea* (Matts.) Liebl.) in central Spain.
 505 *Heredity*, **95**(6): 457–465.

- 506 Vanden Broeck, A., Cottrell, J., Quataert, P., Breyne, P., Storme, V., Boerjan, W., and Van
507 Slycken, J. 2006. Paternity analysis of *Populus nigra* L. offspring in a Belgian plantation
508 of native and exotic poplars. *Ann. For. Sci.*, **63**(7): 783–790.
- 509 Verhaegen, D., and Plomion, C. 1996. Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus*
510 *grandis* using RAPD markers. *Genome*, **39**(6): 1051–1061.
- 511 Ward, M., Dick, C. W., Gribel, R., and Lowe, A. J. 2005. To self, or not to self... a review of
512 outcrossing and pollen-mediated gene flow in neotropical trees. *Heredity* , **95**(4): 246–
513 254.
- 514 Williams, J. H., Boecklen, W. J., and Howard, D. J. 2001. Reproductive processes in two oak
515 (*Quercus*) contact zones with different levels of hybridization. *Heredity* , **87**(Pt 6): 680–
516 690.
- 517 Williams, J. H., Jr, Friedman, W. E., and Arnold, M. L. 1999. Developmental selection within
518 the angiosperm style: using gamete DNA to visualize interspecific pollen competition.
519 *Proc. Natl. Acad. Sci. U. S. A.*, **96**(16): 9201–9206.
- 520 Wu, Z., Raven, P. H. and Hong, D. 1999. *Flora of China*. Science Press/Missouri Botanical
521 Garden, Beijing/St. Louis.

522   Table 1. Median, minimum, and maximum number of chromosomes in each family and some
523   genera (Rice et al. 2015)

Family	Median (Min - max)
Genus	
Annonaceae	8 (7 - 24)
<i>Polyalthia</i>	9 (7 - 28)
Dilleniaceae	13 (8 - 32)
Dipterocarpaceae	7 (6 - 15)
Ebenaceae	15 (12 - 57)
Ericaceae	13 (2 - 78)
Fabaceae	10 (1 - 134)
<i>Robinia</i>	10 (10 - 15)
<i>Acacia</i>	13 (12 - 104)
<i>Mimosa</i>	13 (8 - 32)
<i>Inga</i>	13 (13 - 26)
Fagaceae	12 (6 - 45)
<i>Quercus</i>	12 (6 - 14)
Juglandaceae	16 (11 - 32)
<i>Juglans</i>	16 (6 - 32)
Lauraceae	12 (6 - 48)
Magnoiaceae	19 (7 - 95)
<i>Magnolia</i>	19 (7 - 95)
Malvaceae	15 (4 - 164)
Meliaceae	25 (4 - 180)
Moraceae	13 (2 - 154)
<i>Ficus</i>	13 (11 - 26)
Myristicaceae	22 (19 - 26)
Myrtaceae	11 (5 - 45)
Oleaceae	23 (10 - 138)
<i>Fraxinus</i>	23 (23 - 138)
Rosaceae	15 (1 - 112)
<i>Prunus</i>	8 (8 - 90)
Sapindaceae	13 (7 - 105)
<i>Acer</i>	13 (6 - 52)
Theaceae	15 (5 - 75)

524

525

Table 2. Estimates of total pollen produced by each species in one year and the number of parental type pollen grains of 10 anemophilous trees, as well as ragweed (*Ambrosia artimisiifolia*) and maize (*Zea mays*) for comparison between herbaceous and woody plants (Molina et al. 1996; Fumanal et al. 2007; Jarosz et al. 2003). Parental type pollen estimates are based on total pollen production estimate, chromosome number, and simulation results without crossing over.

	Pollen produced per year ($\times 10^9$)	Number of chromosomes	Number of pollen grains 90% identical to parent ($\times 10^6$)	Number of pollen grains 80% identical to parent ($\times 10^6$)
<i>Zea mays</i>	0.014	10	0.002	0.07
<i>Pinus pinaster</i>	25.1	12	0.8	56.8
<i>Quercus rotundifolia</i>	291.1	12	9.3	658.9
<i>Acer negundo</i>	10.1	13	0.2	15.7
<i>Ulmus minor</i>	6.2	14	0.05	6.5
<i>Juglans regia</i>	2.3	16	0.005	1.1
<i>Ambrosia artimisiifolia</i> L.	1.19	18	0.0003	0.3
<i>Populus nigra</i>	41.7	19	0.01	6.9
<i>Platanus hispanica</i>	164	21	0.04	11.6
<i>Olea europaea</i>	59.4	23	0	2.2
<i>Fraxinus angustifolia</i>	76.1	23	0	2.8
<i>Salix atrocinerea</i>	12.5	38	0	0

Figure captions

Figure 1. Distribution of potential gametophyte types produced by two parental species and their F_1 hybrid. Each parental type will produce 100% of their species type while a F_1 hybrid would produce $(\frac{1}{2})^n$ gametes of each parental type, where n = number of linkage groups. Crossing-over will produce “introgressed” gametophytes that are largely of one parental type but carry a small proportion of the other type, depending on crossing-over rate.

Figure 2. Relationship between chromosome number and proportion of parental-type gametophytes produced from an F_1 hybrid. For no crossing-over (purple) and 3.5% crossing-over (green), the frequency of gametophytes which are >90% (light) and >80% (dark) similar to a parental type are shown.

Figure 3. The gametophyte with that matches the genotype of the style tissue would probably have a greater vigor than gametophytes with a mixture of chromosomes. During the competition for fertilization, even if the matching gametophyte is comparatively rare, it would still have an advantage for fertilizing the ovule.

Figure 1. Distribution of potential gamete types produced by two parental species and their F1 hybrid. Each parental type will produce 100% of their species type while a F1 hybrid would produce $(\frac{1}{2})^n$ gametes of each parental type, where n = number of linkage groups. Crossing-over will produce "introgressed" gametes that are largely of one parental type but carry a small proportion of the other type, depending on crossing-over rate.

26x17mm (300 x 300 DPI)

Figure 2. Relationship between chromosome number and proportion of parental-type gametes produced from a F1 hybrid. For no crossing-over (purple) and 3.5% crossing-over (green), the frequency of gametes which are >90% (light) and >80% (dark) similar to a parental type are shown.

24x18mm (300 x 300 DPI)

Draft

Figure 3. The pure gamete with matching chromosomes as the mother tissue of the stigma and style would probably have a greater vigor than mixed gametes. During the competition for fertilization, even if the pure gamete is comparatively rare, it would still have an advantage for fertilizing the ovule.

41x47mm (300 x 300 DPI)