"Pharaoh's Dance": the Oak Genomic Mosaic

Andrew L. Hipp
The Morton Arboretum
4100 Illinois Route 53
Lisle, IL 60532-1293, USA
ahipp@mortonarb.org

I listened to all the session reels. There were some low moments, some starts and stops. But with all that music, what we know as Bitches Brew *could have been assembled twenty different ways.*
—Bob Belden, producer of the 1999 *Bitches Brew* reissue (Szwed 2004).

Thus, there is not one history with which we can describe the relationship of our genome to the genomes of the African apes, but instead different histories for different segments of our genome. In this respect, our genome is a mosaic, where each segment has its own relationship to that of the African apes.
—Svante Pääbo, Director of The Department of Genetics at the Max Planck Institute for Evolutionary Anthropology, Leipzig (Pääbo 2003).

The past two weeks, I have listened about a dozen times to Joe Zawinul's "Pharaoh's Dance," the first track on Miles Davis' album *Bitches Brew*. I encourage you to pull the album from your shelf if you have it or play it online. The song starts out with a driving vamp on the electric piano, drum steady, pushing you forward… but at 1 minute 40 seconds, it may strike you that the song has changed tracks, starting over from the beginning, abruptly, as though the tape had been cut and a second take of the opening spliced in. It had: the song was strongly edited in post-production to execute Miles' vision. Hours of studio time between August 19 and 21, 1969 were distilled and run through effects, looped, spliced together by producer Teo Macero and Miles over the months that followed to make the final recording of 20 minutes, 5 seconds in length. "Pharaoh's Dance" as we know it now is a mosaic created of 19 different edits. The final piece – the album as a whole – was not even immediately recognizable to some of the musicians who recorded it.[1] Yet I must have listened to it scores of times over the course of about 20 years before I gave any thought to where each section came from. When I listen to this song and to the title track of the album, which is almost as heavily edited, it works as a whole.

Oaks are similar: when we look at an oak, we see one organism, despite the fact that each of the estimated 25,808 genes in the oak genome has a unique history (Plomion et al. 2018). The history of those genes may be described as the history of gene duplications, divergences, origins and extinctions of alleles, some of which differ from each other functionally – some contributing to a longer petiole, some contributing to earlier flowering – but most of which will contribute little or not at all to adaptive differences between individuals (Ohta 1992). Because of the recombination that happens at meiosis, the many genes on a chromosome, even adjacent ones, may have different histories.[2] Gene histories give us some of the best information we have about

---

[1] "I didn't really like the sessions at the time," Zawinul reminisced. "I didn't think they were exciting enough. But a short while later I was at the CBS offices, and a secretary was playing this incredible music. It was really smoking. So I asked her, "Who the hell is this?" And she replied, "It's that Bitches Brew thing." "I thought, Damn, that's great." (Tingen 2017)

[2] This is true even at the base of the "Tree of Life," where bacteria, eukaryotes, and archaea may have traded genes extensively during the early divergences of life's three domains… and continue to today (Pennisi 1998; Quammen 2018).

the history of speciation, but that relationship between the evolutionary history of alleles and the origins of populations and species is not straightforward (Maddison 1997; Knowles 2009). To the contrary, every individual's genome comprises genes whose alleles may have arisen within the species to which that individual seems to belong, others that were born long before that species arose and are shared by descent among numerous species, and still others that shuttled into the population from a different species by hybridization (Wendel and Doyle 1998). An oak's genome – as any species' genome – is a mosaic spliced together from disparate histories of genes and genome regions (Pääbo 2003; Baum and Smith 2013).

When an oak population first becomes geographically isolated from others of its species, it is little if at all distinguishable from the populations from which it arose (Darwin 1859). It might harbor some uncommon alleles that help it adapt to a new environment and that consequently have the potential to increase rapidly in frequency in the population. If present, these alleles may belong to genes that inhabit inverted chromosomal segments, dragging a range of genetic variation along with them, protecting a whole region of the genome from mixing with other populations (Ayala and Coluzzi 2005; Rieseberg 2001)[3] or perhaps allowing that variation to accumulate, providing fodder for future speciation (Fuller et al. 2018; Han et al. 2017). Whatever novelties there are in this isolated population are likely to be minor, at least in oaks, because oak populations each contain a large proportion of the genetic variation in their species as a whole (Michaud et al. 1992; Muir and Schlötter 2005; but cf. Lexer et al. 2006). Indeed, trees as a group tend to be highly genetically variable within compared to among populations (Hampe and Petit 2006), and we therefore expect tree speciation to be messy. A large ancestral population with high genetic variation gives rise to populations with high genetic variation, which because of trees' large effective population sizes will only slowly lose shared alleles (Hudson and Coyne 2002). As a consequence, many individual genes we look at will fail to distinguish species from each other.

In the past few decades, it has become clear that so long as we have sampled a large number of nucleotides from across a wide range of the genome, we can generally detect species boundaries in oaks using population genetic and conventional phylogenetic approaches (Muir et al. 2000; Hipp and Weber 2008; Cavender-Bares and Pahlich 2009; Guichoux et al. 2011; Fitzek et al. 2018) or phylogenetic approaches that explicitly account for random sorting of ancestral genetic variation (Ortego et al. 2017; McVay et al. 2017a, b; Eaton et al. 2015). Most of the time, sophisticated analytical methods are not needed: when we analyze large volumes of data using conventional methods, the history of population divergence generally swamps the signals of random allele sorting and hybridization (Hipp 2015; Rokas et al. 2003). We recover species boundaries and phylogenetic histories that are consistent with morphology, ecology, and biogeography (Fitz-Gibbon et al. 2017; Cavender-Bares et al. 2015; Hipp et al. 2018). Occasionally, however, our efforts to reconstruct relationships among species and populations will be fouled up even with large numbers of genes, and these cases serve as a reminder that oak genomes are stitched together from histories that may differ dramatically from one another. The seams between the genes are sometimes obvious ecologically or morphologically, like the transition at 1:40 on "Pharaoh's Dance." You have probably at least once stood in an oak population and puzzled over the imprint of genes introgressed from another species you knew well. Such moments of discordance draw our attention to the myriad splices that comprise the oak genome. They serve as a point of entry into the editing room, where we can unravel the

---

[3] These regions must be small if present at all in oaks, as the oak genome is remarkably conserved, exhibiting only very fine scale rearrangements across the entire family (Bodénès et al. 2012; Barreneche 2004; Cannon et al. 2018; Kremer et al. 2012).

varying histories from the bits of recording tape strewn across the table. Each alternate history that we infer gives us a clearer understanding of what an oak species is and where its constituent bits arose.

      The origin of the Eurasian White Oaks (McVay et al. 2017a), the Roburoids, is just such a case study.[4] When my colleague Ian Pearse visited our lab in 2009 to generate the first well-resolved, species-level phylogeny for White and Red Oaks (Pearse and Hipp 2009), we found the Roburoids to form a distinct lineage (Fig. 1), despite the fact that they had not in previous studies been genetically distinguishable as a group from the North American oaks (Denk and Grimm 2010; Oh and Manos 2008).[5] The Roburoids were placed sister to all other White Oaks except for the live oaks of section *Virentes*, which matched previous findings based on analysis of chloroplast and nuclear ribosomal DNA (Manos et al. 1999) and was satisfying from a biogeographical standpoint. This result was, however, not robustly supported, and it seemed sensitive to which subset of our data we analyzed. Moreover, this result was based on amplified fragment length polymorphism data, which provide less precise information about genetic similarity and relatedness than sequence data do. As a consequence, when, in our next published phylogeny (Hipp et al. 2014), based on much more and higher quality genomic data, the Roburoid oaks popped out sister to the Eastern North American *Quercus alba* and *Q. michauxii* (Fig. 2), we were fully prepared to accept the change in the story. This was exciting, though not entirely new: the placement had been found previously using a single-copy nuclear gene, *CRABSCLAW*, albeit with low support (Oh and Manos 2008), and a relationship between the Eastern North American and East Asian White Oaks that approximated the groups we recovered had been hypothesized in the 1980s based on morphological data (Axelrod 1983). We thought that the problem was solved: with high-quality genomic data, the origins of the Roburoid oaks were safely pinned to Eastern North American ancestors.

      Problems arose, however, as we added more species to the dataset. In different analyses on the same data, we found two fascinating things. First, two chestnut-leaved White Oaks growing approximately 6,000 miles from each other – *Q. sadleriana* of California, *Q. pontica* of the Caucasus – consistently found each other as phylogenetic sisters and fell, together, sister to all of the remaining White Oaks, including the *Virentes*. No matter what we did or how we attacked the data, this result stood. These two species are now treated as *Quercus* section *Ponticae*, making them the smallest section in the genus (Denk et al. 2017). Second, the Roburoids were phylogenetically unstable, moving between the two positions that had shown up in our two earlier papers depending on how we analyzed the data (Fig. 3). As we varied taxon sampling, sampling of individuals within taxa, genomic clustering parameters, and sampling of subsets of loci, section *Ponticae* either fell sister to the remaining White Oaks, forming the second lineage to diverge after *Q. sadleriana* + *Q. pontica*; or sister to a clade of three Eastern North American species (*Q. alba* and *Q. michauxii* as shown before, plus *Q. montana*, which we had not previously sampled). The instability was unsettling.

      With meticulous work, however, the postdoctoral researcher on this project, John McVay, discerned that there were genuinely two stories embedded in the genome. The first was a sister relationship between the Roburoids and *Q. alba* and company; the second was a close relationship between *Q. pontica* and the Roburoids, but not between *Q. sadleriana* and the Roburoids. This was peculiar. By necessity, all species in any single clade are equally closely

---

[4] For another equally fascinating case study, see John McVay's exciting and beautifully illustrated (2017b) study of *Quercus gambelii* and its relationship to *Q. macrocarpa* and *Q. lobata* (McVay et al. 2017a).

[5] Presciently, Denk and Grimm (2010) found that *Q. pontica*, traditionally treated as a Roburoid, *was* genetically distinguishable from the other white oaks (Denk and Grimm 2010). At that time, we did not yet have confirmation of the relationship between *Q. pontica* and *Q. sadleriana*.

related to all species of any other clade you chose, provided that all genes are inherited along ancestry lines defined by the phylogeny. Thus *Q. pontica* could only be more closely related to the Roburoid White Oaks if genes were moving along some path other than the dominant phylogeny. The most straightforward explanation for this is introgressive hybridization in Europe between *Q. pontica* and the Roburoid oaks. A history of crossing and backcrossing between *Q. pontica* and the Roburoids had shuttled genes from *Q. pontica* into the Roburoids, dragging the Roburoids out to the edge of the White Oak phylogeny while leaving little if any obvious morphological evidence of the gene flow.[6] Eliminating genes that appeared to be subject to introgression between *Q. pontica* and the Roburoid White Oaks, McVay recovered what we believe to be the correct placement of the Roburoid White Oaks sister to Eastern North American *Q. alba* and its kin, returning us to Axelrod's (1983) prediction, but with refined definitions of the lineages involved.

All organisms are genomic mosaics. Modern humans have alleles that arose in chimpanzee and Neanderthal populations, some of which appear likely to have been under selection in anatomically modern humans as they migrated into Asia (Simonti et al. 2016). More dramatically, a gene crucial to building the mammalian placenta arose in mammals from a series of parallel retroviral infections (Quammen 2018). In other words, chimp, Neanderthal, and viral genes have all shaped the fitness of modern humans. In the same way, advantageous alleles appear in some cases to flow between oak species in ways that facilitate shared adaptations (Cannon et al. 2018; Khodwekar and Gailing 2017) and might play an important role in the evolution and persistence of species (Anderson 1953; Cannon and Lerdau 2015; Hampe and Petit 2006) in other cases, selected genes appear to differentiate species even in the face of ongoing gene flow across the remainder of the genome (Oney-Birol et al. 2018). We are only beginning to get a sense of which of these stories is more generally the case: selection as a barrier to gene flow versus gene flow as mechanism by which adaptations are shared among species. With two good oak genomes online now (Plomion et al. 2016, 2018; Sork et al. 2016) we are finally in a position to start focusing on the seams between genes, on the discrete and variable histories that make up oaks. We can start to ask a previously unanswerable question: what do we mean by "species" and "phylogenies" when we are talking about oaks?

The answer to this question, like the meaning of "Pharaoh's Dance," probably does not rest in any single set of genes or edits. While it appears there are genomic islands of differentiation among closely related oaks (Scotti-Saintagne et al. 2004), at least some of which are conserved among major clades Sullivan et al. 2016), it may turn out that the genes that distinguish *Q. macrocarpa* from *Q. bicolor* in Wisconsin are not the genes that distinguish it from *Q. stellata* in Missouri.[7] Integrating over *all* the stories in the genome will require us to build on phylogenomic approaches that explicitly model divergence with gene flow, as reviewed in the earlier paragraphs. But such an integration risks painting a simplistic single story over the

---

[6] Edgar Anderson, in his seminal *Introgressive Hybridization* (Anderson 1949), made the point that hybridization is probably most important in organismal evolution when it is least obvious. He writes, "If introgression proves to be a primary factor in evolution it will be because it so greatly enriches variation in the participating species. As raw material for evolution, the bizarre hybrid swarms described in Chapter 1 are not so important as the *Asclepias* introgression described by Woodson (1947), which was barely noticeable in any one locality and extended as a trend through a long intermediate zone. By the time of the third backcross of the original hybrid to one of the parental species, there would be little or no external indication of hybridity in the mongrel progeny. Yet in terms of gene frequencies, the effects of introgression in such mongrels would far outweigh the immediate effects of gene mutation" (pp. 61-62). The Roburoid oaks may be such a case. As a side-note, it is remarkable to an oak enthusiast to find that the genus *Quercus* did not make it into this (1949) book, though two references to introgression in the genus showed up in Anderson's 1953 *Biological Reviews* article of the same title (Anderson 1953).

[7] The reader will probably recognize Leigh van Valen here, who wrote, "It may well be that *Quercus macrocarpa* in Quebec exchanges many more genes with local *Q. bicolor* than it does with *Q. macrocarpa* in Texas" (Van Valen 1976). The idea is similar, though I consider that van Valen's concern has largely been put to rest by molecular studies of the past 20 years, beginning most resoundly with Muir et al.(2000) and reviewed in two previous articles (Hipp 2015; Hipp 2016) in this journal.

more than 25,000 splices that make up the genome of the *Q. bicolor* in my front yard. The most interesting story underlying the oak phylogeny may reside in the seams themselves, the sutures between the genes, the changes from one story to another.

"Pharaoh's Dance" could have been assembled 20 different ways. Oaks are similar, assembled from the scraps of tape that make up the oak genome. But unlike "Pharaoh's Dance," of which there is only one realization, there have been millions upon millions of *Q. bicolor*, all assembled differently.[8] This fall, as I kick up the *Q. alba* and *Q. macrocarpa* and *Q. rubra* litter on my bike ride into work, I see alleles spliced in from everywhere. Are all these genes interchangeable? Are there a select few that have the power to change one species to another? What are the limits to gene flow and recombination? How disparate are the histories that make up the individual *Q. alba* growing in The Morton Arboretum's East Woods? Those are questions for the coming decade.

Acknowledgments

---

[8] These are of course not unconstrained assemblies: for an eloquent discussion of this, please see Anderson (1949), ch. 3, in which he discusses at length the limits to recombination in organisms, why hybrids don't fill all of the possible recombination space but rather tend to fall along what he terms a "recombination spindle" between the poles of the two parents.

Works cited

Anderson, E. 1949. *Introgressive hybridization*. New Jersey: Wiley.

Anderson, E. 1953. Introgressive Hybridization. *Biol Rev* 28(3): 280-307.

Axelrod, D.I. 1983. Biogeography of oak in the arcto-tertiary province. *Ann Mo Bot Gard* 70: 629-657.

Ayala, F.J., and M. Coluzzi. 2005. Chromosome speciation: Humans, Drosophila, and mosquitoes. *Proc Natl Acad Sci U S A* 102(suppl_1): 6535-6542.

Barreneche, T., M. Casasoli, K. Russell, A. Akkak, H. Meddour, C. Plomion, F. Villani, and A. Kremer. 2004. Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs). *Theoretical and Applied Genetics* 108(3): 558-566.

Baskett, M.L., and R. Gomulkiewicz. 2011. Introgressive hybridization as a mechanism for species rescue. *Theoretical Ecology* 4(2): 223-239.

Baum, D.A., and S.D. Smith. 2013. *Tree Thinking: An Introduction to Phylogenetic Biology* Greenwood Village, CO: Roberts and Company Publishers, Inc.

Bodénès, C., E. Chancerel, O. Gailing, G.G. Vendramin, F. Bagnoli, J. Durand, P.G. Goicoechea, C. Soliani, F. Villani, C. Mattioni, H.P. Koelewijn, F. Murat, J. Salse, G. Roussel, C. Boury, F. Alberto, A. Kremer, and C. Plomion. 2012. Comparative mapping in the Fagaceae and beyond with EST-SSRs. *BMC plant biology* 12(1): 153.

Cannon, C.H., and M. Lerdau. 2015. Variable mating behaviors and the maintenance of tropical biodiversity. *Frontiers in Genetics* 6. doi:10.3389/fgene.2015.00183.

Cannon, C.H., O. Brendel, M. Deng, A.L. Hipp, A. Kremer, C.-S. Kua, C. Plomion, J. Romero-Severson, and V. Sork. 2018. Gaining a global perspective on Fagaceae genomic diversification and adaptation. *New Phytologist* 218(3): 894-897.

Cavender-Bares, J., and A. Pahlich. 2009. Molecular, morphological, and ecological niche differentiation of sympatric sister oak species, *Quercus virginiana* and *Q. geminata* (Fagaceae). *American Journal of Botany* 96(9): 1690-1702.

Cavender-Bares, J., A. González-Rodriguez, D.A.R. Eaton, A.L. Hipp, A. Beulke, and P.S. Manos. 2015. Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): A genomic and population genetics approach. *Molecular Ecology* 24: 3668-3687.

Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. First Edition. London: John Murray.

Denk, T., and G.W. Grimm. 2010. The oaks of western Eurasia: Traditional classifications and evidence from two nuclear markers. *Taxon* 59: 351-366.

Denk, T., G.W. Grimm, P.S. Manos, M. Deng, and A.L. Hipp. 2017. An updated infrageneric classification of the oaks: review of previous taxonomic schemes and synthesis of evolutionary patterns. In *Oaks Physiological Ecology, Exploring the Functional Diversity of Genus* Quercus *L.*, edited by E. Gil-Pelegrin, J.J. Peguero-Pina, and D. Sancho-Knapik, pp. 13-38. Cham, Switzerland: Springer International Publishing AG.

Eaton, D.A.R., A.L. Hipp, A. González-Rodríguez, and J. Cavender-Bares. 2015. Historical introgression among the American live oaks and the comparative nature of tests for introgression. *Evolution* 69(10): 2587-2601.

Fitz-Gibbon, S., A.L. Hipp, K.K Pham, P.S. Manos, and V. Sork. 2017. Phylogenomic inferences from reference-mapped and de novo assembled short-read sequence data using RADseq sequencing of California white oaks (*Quercus* subgenus *Quercus*). *Genome* 60: 743-755.

Fitzek, E., A. Delcamp, E. Guichoux, M. Hahn, M. Lobdell and A.L. Hipp. 2018. A nuclear DNA barcode for eastern North American oaks and application to a study of hybridization in an Arboretum setting. *Ecology and Evolution* 8(11): 5837-5851.

Fuller, Z.L., C.J. Leonard, R.E. Young, S.W. Schaeffer, and N. Phadnis. 2018. Ancestral polymorphisms explain the role of chromosomal inversions in speciation. *PLOS Genet* 14(7): e1007526.

Guichoux, E., L. Lagache, S. Wagner, P. Léger, and R.J. Petit. 2011. Two highly validated multiplexes (12-plex and 8-plex) for species delimitation and parentage analysis in oaks (*Quercus* spp.). *Molecular Ecology Resources* 11(3): 578-585.

Hampe, A., and R.J. Petit. 2006. Some Evolutionary Consequences of Being a Tree. *Annu. Rev. Ecol. Evol. Syst* 37: 187-214.

Han, F., S. Larmichhaney, B.R. Grant, P.R. Grant, L. Andersson, and M.T. Webster. 2017. Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome research* 27(6): 1004-1015.

Hipp, A.L. 2015. Should hybridization make us skeptical of the oak phylogeny? *International Oaks* 26: 9-18.

Hipp, A.L. 2016. Oak Research in 2015: a Snapshot from the IOS Conference. *International Oaks* 27: 15-22.

Hipp, A.L., D.A.R. Eaton, J. Cavender-Bares, E. Fitzek, R. Nipper, and P.S. Manos. 2014. A framework phylogeny of the American oak clade based on sequenced RAD data. *PLoS ONE* 9: e93975.

Hipp, A.L., P.S. Manos, A. González-Rodriguez, M. Hahn, M. Kaproth, J.D. McVay, S. Valencia Avalos, and J. Cavender-Bares. 2018. Sympatric parallel diversification of major oak clades in the Americas and the origins of Mexican species diversity. *New Phytologist* 217(1): 439-452.

Hipp, A.L., and J.A. Weber. 2008. Taxonomy of Hill's Oak (*Quercus ellipsoidalis*: Fagaceae): Evidence from AFLP Data. *Systematic Botany* 33: 148-158.

Hudson, R.R., and J.A. Coyne. 2002. Mathematical Consequences of the Genealogical Species Concept. *Evolution* 56(8): 1557-1565.

Khodwekar, S., and O. Gailing. 2017. Evidence for environment-dependent introgression of adaptive genes between two red oak species with different drought adaptations. *American Journal of Botany* 104(7): 1088-1098.

Knowles, L.L. 2009. Estimating Species Trees: Methods of Phylogenetic Analysis When There Is Incongruence across Genes. *Systematic Biology* 58(5): 463-467.

Kremer, A., A.G. Abbott, J.E. Carlson, P.S. Manos, C. Plomion, P. Sisco, M.E. Staton, S. Ueno, and G.G. Vendramin. 2012. Genomics of Fagaceae. *Tree Genetics & Genomes* 8: 583-610.

Lexer, C., A. Kremer, and R.J. Petit. 2006. Shared alleles in sympatric oaks: recurrent gene flow is a more parsimonious explanation than ancestral polymorphism. *Molecular Ecology* 15: 2007-2012.

Maddison, W.P. 1997. Gene trees in species trees. *Systematic Biology* 46: 523-536.

Manos, P.S., J.J. Doyle, and K.C. Nixon. 1999. Phylogeny, Biogeography, and Processes of Molecular Differentiation in *Quercus* Subgenus *Quercus* (Fagaceae). *Molecular Phylogenetics and Evolution* 12(3): 333-349.

McVay, J.D., A.L. Hipp, and P.S. Manos. 2017a. A genetic legacy of introgression confounds phylogeny and biogeography in oaks. *Proc. R. Soc. B.* 284(1854): 20170300.

McVay, J.D., D. Hauser, A.L. Hipp, P.S. Manos. P.S. 2017b. Phylogenomics reveals a complex evolutionary history of lobed-leaf white oaks in western North America. *Genome* 60(9): 733-742.

Michaud, H., R. Lumaret, and F. Romane. 1992. Variation in the genetic structure and reproductive biology of holm oak populations. *Vegetation* 99-100: 107-113.

Muir, G., C.C. Fleming, and C. Schlötterer. 2000. Taxonomy: species status of hybridizing oaks. *Nature* 405(6790): 1016.

Muir, G., and C. Schlötterer. 2005. Evidence for shared ancestral polymorphism rather than recurrent gene flow at microsatellite loci differentiating two hybridizing oaks (*Quercus* spp.). *Molecular ecology* 14(2): 549-561.

Oh, S.-H., and P.S. Manos. 2008. Molecular phylogenetics and cupule evolution in Fagaceae as inferred from nuclear CRABS CLAW sequences. *Taxon* 57: 434-451.

Ohta, T. 1992. The Nearly Neutral Theory of Molecular Evolution. *Annual Review of Ecology and Systematics* 23(1): 263-286.

Oney-Birol, S., S. Fitz-Gibbon, J.-M. Chen, P.F. Gugger, and V.L. Sork. 2018. Assessment of shared alleles in drought-associated candidate genes among southern California white oak species (*Quercus* sect. *Quercus*). *BMC genetics* 19(1): 88.

Ortego, J., P.F. Gugger, and V.L. Sork. 2017. Genomic data reveal cryptic lineage diversification and introgression in Californian golden cup oaks (section *Protobalanus*). *New Phytologist* doi:10.1111/nph.14951.

Pääbo, S. 2003. The mosaic that is our genome. *Nature* 421(6921): 409-412.

Pearse, I.S., and A.L. Hipp. 2009. Phylogenetic and trait similarity to a native species predict herbivory on non-native oaks. *Proceedings of the National Academy of Sciences* 106(43): 18097-18102.

Pennisi, E. 1998. Genome data shake tree of life. *Science* 280(5364): 672-674.

Plomion, C., J.M. Aury, J. Amselem, T. Alaeitabar, V. Barbe, C. Belser, H. Bergès, C. Bodénès, N. Boudet, C. Boury, A. Canaguier, A. Couloux, C. Da Silva, S. Duplessis, F. Ehrenmann, B. Estrada-Mairey, S. Fouteau, N. Francillonne, C. Gaspin, C. Guichard, C. Klopp, K. Labadie, C. Lalanne, I. Le Clainche, J.C. Leplé, G. Le Provost, T. Leroy, I. Lesur, F. Martin, J. Mercier, C. Michotey, F. Murat, F. Salin, D. Steinbach, P. Faivre-Rampant, P. Wincker, S. Salse, H. Quesneville, and A. Kremer. 2016. Decoding the oak genome: public release of sequence data, assembly, annotation and publication strategies. *Molecular Ecology Resources doi: 10.1111/1755-0998.12425*

Plomion, C., J.-M. Aury, J. Amselem, T. Leroy, F. Murat, S. Duplessis, S. Faye, N. Francillonne, K. Labadie, G. Le Provost, I. Lesur, J. Bartholomé, P. Faivre-Rampant, A. Kohler, J.-C. Leplé, N. Chantret, J. Chen, A. Diévart, T. Alaeitabar, V. Barbe, C. Belser, H. Bergès, C. Bodénès, M.-B. Bogeat-Triboulot, M.-L. Bouffaud, B. Brachi, E. Chancerel, D. Cohen, A. Couloux, C. Da Silva, C. Dossat, F. Ehrenmann, C. Gaspin, J. Grima-Pettenati, E. Guichoux, A. Hecker, S. Herrmann, P. Hugueney, I. Hummel, C. Klopp, C. Lalanne, M. Lascoux, E. Lasserre, A.Lemainque, M.-L. Desprez-Loustau, I. Luyten, M.-A. Madoui, S. Mangenot, C. Marchal, F. Maumus, J. Mercier, C. Michotey, O. Panaud, N. Picault, N. Rouhier, O. Rué, C. Rustenholz, F. Salin, M. Soler, M. Tarkka, A. Velt, A.E. Zanne, F. Martin, P. Wincker, H. Quesneville, A. Kremer, and Jérôme Salse. 2018. Oak genome reveals facets of long lifespan. *Nature Plants* 4(7): 440-452.

Quammen, D. 2018. *The Tangled Tree: A Radical New History of Life*. New York, NY: Simon and Schuster.

Rieseberg, L.H. 2001. Chromosomal rearrangements and speciation. *Trends in Ecology and Evolution* 16(7): 351-358.

Rokas, A., Williams, B.L., King, N., Carroll, S.B. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425(6960): 798-804.

Scotti-Saintagne, C., S. Mariette, I. Porth, P.G. Goicoechea, T. Barreneche, C. Bodénès, K. Burg, and A. Kremer. 2004. Genome Scanning for Interspecific Differentiation Between Two Closely Related Oak Species [*Quercus robur* L. and *Q. petraea* (Matt.) Liebl.]. *Genetics* 168(3): 1615-1626.

Simonti, C.N., B. Vernot, L. Bastarache, E. Bottinger, D.S Carrell, R.L. Chisholm, D.R. Crosslin, S.J. Hebbring, G.P. Jarvik, I.J. Kullo, R. Li, J. Pathak, M.D. Ritchie, D.M. Roden, SS. Verma, G. Tromp, J.D. Prato, W.S. Bush, J.M. Akey, J.C. Denny, J.A. Capra. 2016. The phenotypic legacy of admixture between modern humans and Neandertals. *Science* 351(6274): 737-741.

Sork, V.L., S.T. Fitz-Gibbon, D. Puiu, M. Crepeau, P.F. Gugger, R. Sherman, K. Stevens, C.H. Langley, M. Pellegrini and S.L. Salzberg. 2016. First Draft Assembly and Annotation of the Genome of a California Endemic Oak *Quercus lobata* Née (Fagaceae). *G3: Genes, Genomes, Genetics* 6(11): 3485-3495.

Sullivan, A.R., S.A. Owusu, J.A. Weber, A.L. Hipp, and O. Gailing. 2016. Hybridization and divergence in multi-species oak (*Quercus*) communities. *Botanical Journal of the Linnean Society* 181(1): 99-114.

Szwed, J. 2004. *So What: The Life of Miles Davis*. New York, NY: Simon and Schuster. Available at: http://www.simonandschuster.com/books/So-What/John-Szwed/9780684859835 [Accessed November 8, 2018].

Tingen, P. 2017. "Miles Davis and the Making of Bitches Brew." *JazzTimes*, July, 10. Available at: https://jazztimes.com/features/miles-davis-and-the-making-of-bitches-brew-sorcerers-brew. Accessed November 18, 2018.

Van Valen, L. 1976. Ecological species, multispecies, and oaks. *Taxon* 25: 233-239.

Wendel, J.F., and J.J. Doyle. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. In *Molecular Systematics of Plants II: DNA Sequencing*, edited by D.E. Soltis, P.S. Soltis, and J.J. Doyle, pp. 265-296. Norwell, MA: Kluwer Academic Publishers.
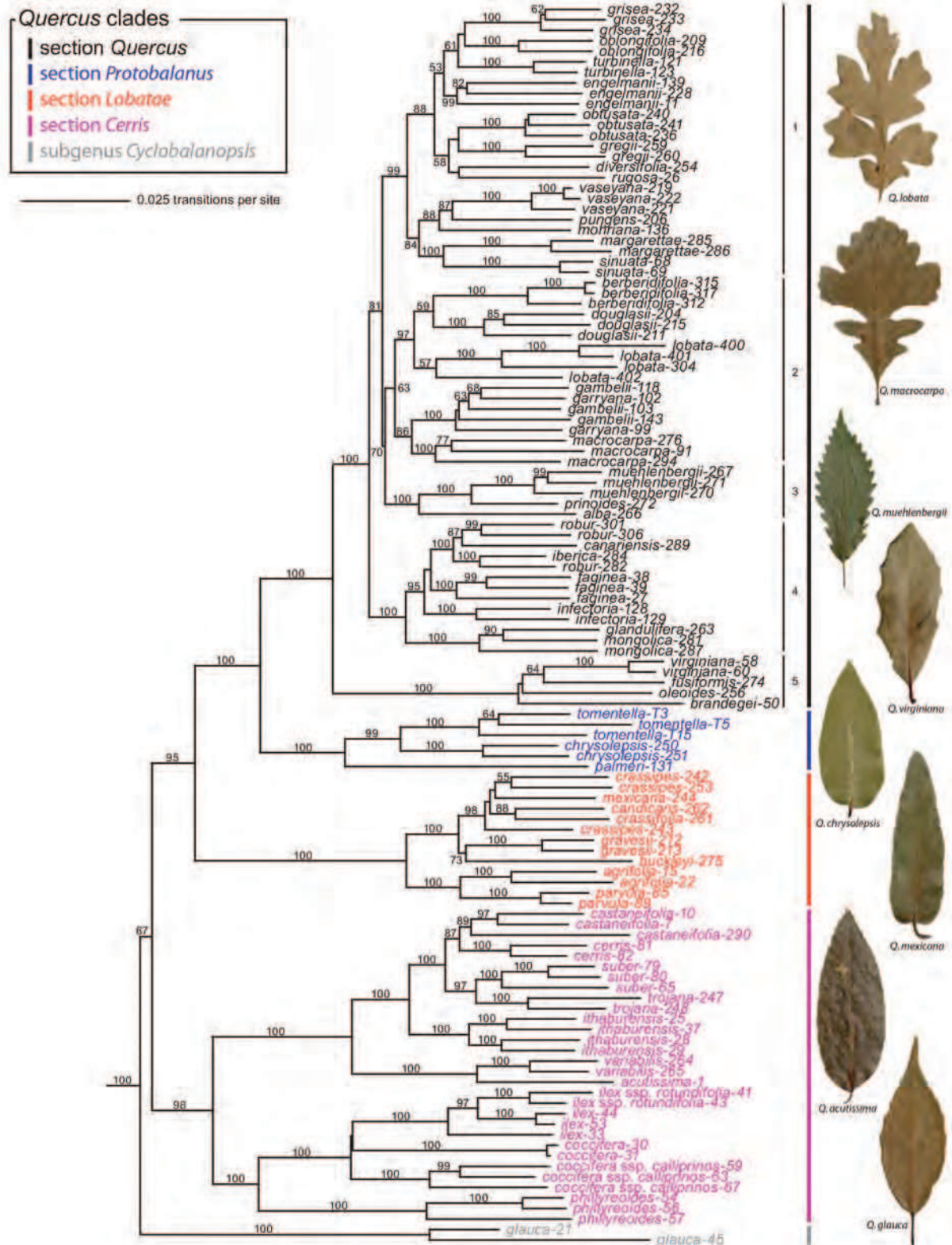
Figure captions

Figure 1. AFLP phylogeny, showing Roburoids sister to the remaining white oaks. From Pearse and Hipp 2009.

Figure 2. First oak RAD-seq phylogeny, showing Roburoids sister to *Quercus alba* and allies. From Hipp et al. 2004.

Figure 3. The resolution: the two placements of the Roburoids that resolve from alternative analyses of RAD-seq data. There are two stories in the genome, one resulting from ancient introgression (left panel), one from speciation and population divergence history (right panel). From McVay et al. 2017a.

Figure 1. AFLP phylogeny, showing Roburoids sister to the remaining white oaks. From Pearse and Hipp 2009.
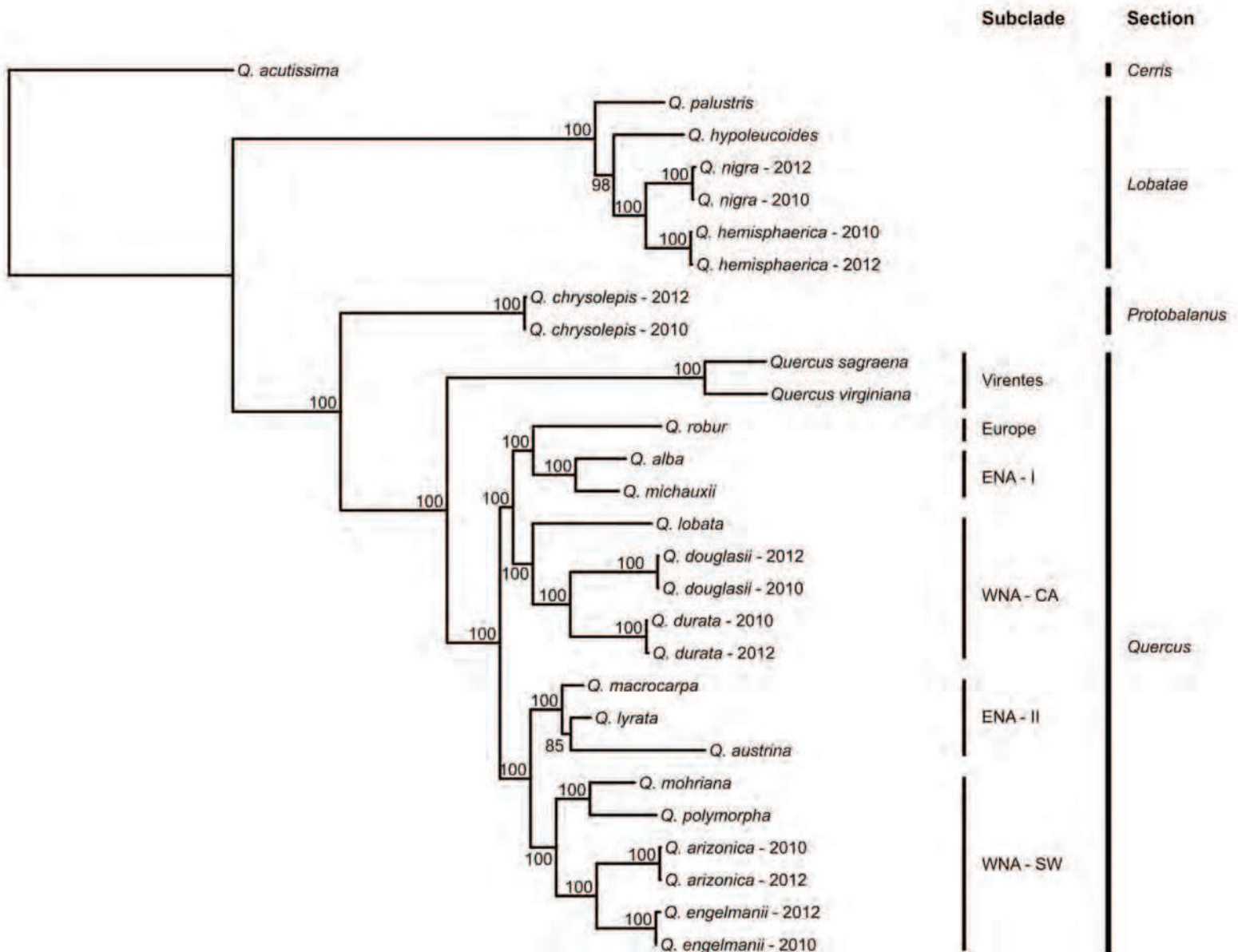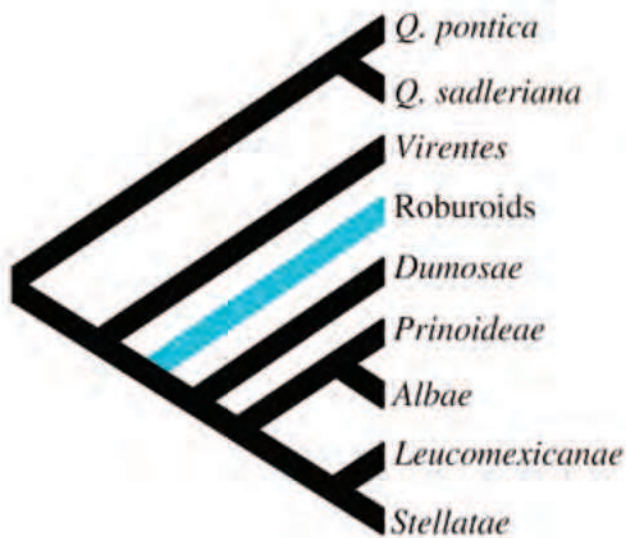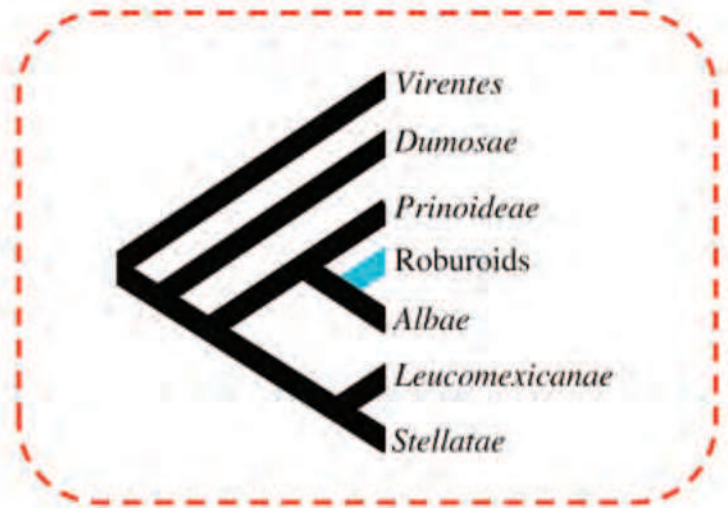
Figure 2. First oak RAD-seq phylogeny, showing Roburoids sister to *Quercus alba* and allies. From Hipp et al. 2004.

Phylogenetic reconstructions showing the Roburoids outside the core American white oak lineage are due to ancient hybridization between the Roburoids and *Quercus pontica*.

The corrected history, identified by pulling out loci that reflect a history of ancient gene flow between lineages, places the Eurasian white oaks sister to *Quercus alba* and friends.

Figure 3. The resolution: the two placements of the Roburoids that resolve from alternative analyses of RAD-seq data. There are two stories in the genome, one resulting from ancient introgression (left panel), one from speciation and population divergence history (right panel). From McVay et al. 2017a.