Botanical Journal of the Linnean Society, 2022, **200**, 15–38. With 6 figures.

A timeframe for mint evolution: towards a better understanding of trait evolution and historical biogeography in Lamiaceae

JEFFREY P. ROSE[1,*], CHUN-LEI XIANG[2], KENNETH J. SYTSMA[3] and BRYAN T. DREW[1,*]

[1]*Department of Biology, University of Nebraska-Kearney, Kearney, NE 68849, USA*
[2]*CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China*
[3]*Department of Botany, University of Wisconsin-Madison, Madison, WI 53706, USA*

Received 9 April 2021; revised 8 September 2021; accepted for publication 2 December 2021

Lamiaceae are one of the largest and most economically important families of flowering plants. Despite focused study on relationships within subclades, higher-level relationships have been under-studied. Moreover, the herbaceous habit of much of the family has resulted in a poor fossil record and has hampered estimates of divergence times. Using a new dataset of five plastid loci from 178 members of Lamiaceae representing all subfamilies and nearly all tribes, we clarify major infrafamilial relationships and present a robust set of divergence times. We use this phylogenetic hypothesis as a platform to test previous hypotheses regarding the historical biogeography and evolution of major traits in the family. We confirm the placement of subfamily Nepetoideae, show continued uncertainty in the placement of subfamilies Ajugoideae and Premnoideae and highlight extreme discordance with recent results from nuclear data. Lamiaceae originated during the Late Cretaceous as woody plants with nutlet fruits and four stamens, probably in South-East Asia. Most subfamilies diverged during the Eocene, perhaps facilitated by climatic cooling. Our results provide a valuable set of secondary dates for Lamiaceae and highlight the need for focused study of subfamilies Callicarpoideae and Viticoideae. Our results also provide several hypotheses regarding trait or range-dependent diversification.

ADDITIONAL KEYWORDS: Eocene – habit – K-Pg – Lamiales – nutlets – *Salvia* – stamen number.

INTRODUCTION

An accurate estimate of temporal history is requisite for properly articulating organismal evolution. The DNA sequencing revolution in conjunction with the notion of molecular clocks (Zuckerkandl & Pauling, 1965) provided a broad platform to apply dates to phylogenetic trees. Subsequently, relaxed molecular clock models allowed for divergence time estimations that are ostensibly accurate and stable (Thorne, Kishino & Painter, 1998; Huelsenbeck, Larget & Swofford, 2000; Douzery *et al.*, 2004; Drummond *et al.*, 2006). There had been resistance to molecular dating techniques (e.g. Ayala, 1986, 1999; Graur & Martin, 2004), but molecular dating has become a routine exercise in phylogenetic studies, although pitfalls still exist (e.g. Sytsma, Spalink & Berger, 2014; Beaulieu *et al.*, 2015; Sauquet & Magallón, 2018). Indeed, phylogenetic trees calibrated to absolute or relative time (ultrametric) are recommended, if not required, for many phylogenetic comparative analyses including tracing character evolution, measuring diversification rates and reconstructing historical biogeography. Despite the importance of time-calibrated phylogenetic trees, many angiosperm clades do not yet have robust estimates of divergence times, especially in groups in which their morphological traits and/or biogeography are not conducive to fossilization (e.g. herbaceous, tropical and/or xerophytic lineages).

Lamiaceae (the mint family) are one of the most easily recognized families of flowering plants. They are nearly worldwide in distribution and are of economic importance for timber (teak: *Tectona* L.f., white teak: *Gmelina arborea* Roxb.), culinary uses (e.g. mint: *Mentha* L., basil: *Ocimum* L.), horticulture (e.g.

---

*Corresponding authors. E-mail: rosej@unk.edu; bdrewfb@yahoo.com

beautyberry: *Callicarpa* L., lavender: *Lavandula* L., sage: *Salvia* L.) and weeds (e.g. ground-ivy: *Glechoma* L., dead-nettle: *Lamium* L., pignut: *Mesosphaerum suaveolens* (L.) Kuntze). With *c.* 7000 currently recognized species (Harley *et al.* 2004; Zhao *et al.*, 2021), Lamiaceae are the largest family of Lamiales and the sixth largest plant family overall. Based on both morphological and molecular phylogenetic results, the family has recently expanded in size as most taxa traditionally placed in Verbenaceae are now treated in Lamiaceae (Cantino, 1992a, 1992b; Cantino, Harley & Wagstaff, 1992; Wagstaff & Olmstead, 1997; Wagstaff *et al.*, 1998; Harley *et al.*, 2004). In Lamiales, Lamiaceae are probably sister to a clade composed of Mazaceae + Orobanchaceae + Paulowniaceae + Phrymaceae (Schäferhoff *et al.*, 2010; Refulio-Rodriguez & Olmstead, 2014; Xu *et al.*, 2019; Zhang *et al.*, 2020).

PHYLOGENY AND CLASSIFICATION OF LAMIACEAE

Most studies of Lamiaceae have focused on relationships at or below the subfamilial level (e.g. Paton *et al.*, 2004; Bramley, Forest & de Kok, 2009; Conn *et al.*, 2009; Bräuchler, Meimberg & Heubl, 2010; Scheen *et al.*, 2010; Bendiksby *et al.*, 2011; Drew & Sytsma, 2012; Roy & Lindqvist, 2015; Xiang *et al.*, 2018; Zhao *et al.*, 2020), whereas studies of subfamilial relationships have been largely lacking, especially with comprehensive sampling of major clades. However, investigations into major relationships with comprehensive sampling have been conducted by Li *et al.* (2016) using five plastid loci and Zhao *et al.* (2021) using 79 protein-coding plastid genes. These studies support the recognition of 12 subfamilies in Lamiaceae (Li *et al.*, 2016; Li & Olmstead, 2017; Zhao *et al.*, 2021). The monophyly of all proposed subfamilies except Cymarioideae has been affirmed from these studies, but relationships among subfamilies are still unclear, with Bayesian analysis recovering high posterior probabilities for most relationships and parsimony and maximum likelihood analyses finding low (or no) support for some relationships.

Nuclear evidence for relationships in Lamiaceae was presented by the Mint Evolutionary Genomics Consortium (MEGC, 2018) using data from 520 single-copy nuclear exons and representatives of all subfamilies except Cymarioideae. These data were used, in part, to infer species trees for the family based on concatenation and coalescent analyses. Trees from these two approaches are largely congruent, but they differ topologically in two ways: (1) Premnoideae were not recovered as monophyletic in the coalescent analysis; and (2) Ajugoideae were sister to Lamioideae in the coalescent species tree but sister to Lamioideae + Peronematoideae in the concatenated species tree. Overall, nuclear data appear congruent with plastid data (Li *et al.*, 2016; Zhao *et al.*, 2021) in recovering Callicarpoideae + Prostantheroideae as sister to the rest of Lamiaceae, followed by Nepetoideae. However, the two genomes present considerably different relationships among the remaining clades, although given the conflict between the coalescent and concatenated analyses of MEGC (2018) it is not exactly clear why.

TIMEFRAME FOR DIVERSIFICATION OF LAMIACEAE

Despite the importance of and interest in Lamiaceae, no comprehensive analysis of divergence times in the family has been conducted, with most studies focusing on groups below the subfamily level and including limited outgroup sampling. Estimates of the stem age of Lamiaceae from angiosperm-wide analyses of divergence times have been typically estimated at < 40 Mya (Martínez-Millán, 2010; Magallón *et al.*, 2015; Wikström *et al.*, 2015; Barba-Montoya *et al.*, 2018; Li *et al.*, 2019), but Zhang *et al.* (2020) estimated a Cretaceous age for stem Lamiaceae. Inferring divergence times for Lamiaceae is challenging due to a paucity of fossils that can be used to constrain nodes within the family (Harley *et al.*, 2004), but some fossils that do exist can be confidently, albeit conservatively, placed (e.g. Drew & Sytsma, 2012). Particularly widely used have been two fossils used to constrain the stem of *Melissa* L. and the crown of Nepetoideae (Drew & Sytsma, 2012; Yu *et al.*, 2014; Deng *et al.*, 2015; Yao *et al.*, 2016; Li *et al.*, 2017) and a fossil that can be used to constrain *Stachys* L. (Roy *et al.*, 2013, 2016; Roy & Lindqvist, 2015; Yao *et al.*, 2016; Kriebel *et al.*, 2019). Chronograms calibrated with the aforementioned fossils have differed markedly from age estimates of Lamiaceae as inferred from angiosperm-wide studies. For example, crown Nepetoideae have been dated to *c.* 57.0 Mya (Drew & Sytsma, 2012) or slightly older at 63.4 Mya (Li *et al.*, 2017), whereas Roy & Lindqvist (2015) dated crown Lamioideae to at least 20 Mya. In their investigation of *Pogostemon* Desf., Yao *et al.* (2016) provided more representative family-wide sampling with widespread sampling and calibrations in Lamiaceae and across Lamiales, dating crown Lamiaceae to *c.* 65 Mya. This large disparity in ages estimated based on angiosperm-wide studies versus those with good sampling of Lamiaceae and Lamiales has yet to be fully addressed.

HISTORICAL BIOGEOGRAPHY OF LAMIACEAE

Hedge (1992) and Harley *et al.* (2004) reviewed the biogeography of Lamiaceae, pointing to six or seven areas where the family shows high species diversity, with Harley *et al.* (2004) adding the Indo–Malesian region due to the inclusion of genera from Verbenaceae and Symphoremataceae in a broader concept of Lamiaceae (Cantino *et al.*, 1992). Building on previous work, Harley *et al.* (2004) proposed a Late Cretaceous or Early Tertiary Gondwanan origin of the family, with a rapid radiation to other areas, and they also hypothesized an Asian origin of Nepetoideae. Studies investigating the historical biogeography of the entire Lamiaceae using molecular phylogenetic trees have been lacking, with studies focusing on relationships at or below the subfamilial level (e.g. Paton *et al.*, 2004; Scheen & Albert, 2009; Mathiesen, Scheen & Lindqvist, 2011; Drew & Sytsma, 2012, 2013; Roy *et al.*, 2013; Roy & Lindqvist, 2015; Yao *et al.*, 2016; Drew *et al.*, 2017a; Li *et al.*, 2017), and many of these studies have used parsimony character mapping. Model-based analyses of historical biogeography at the subfamilial and tribal levels have suggested a strong signal for an Asian origin of Lamioideae (Roy & Lindqvist, 2015), a Mediterranean/South-West Asian origin of Nepetoideae: Mentheae (Drew & Sytsma, 2012) and a South-East Asian origin of Nepetoideae: Elsholtzieae (Li *et al.*, 2017). So far, these are the best available biogeographical reconstructions for deep nodes in Lamiaceae, so the hypotheses of Harley *et al.* (2004) have yet to be rigorously tested with molecular data.

Trait evolution in Lamiaceae

One of the most conspicuous features of temperate Lamiaceae, especially subfamilies Lamioideae and Nepetoideae, is the distinctive nutlet fruit (essentially a schizocarp) that splits into four mericarps at maturity and is usually, but not always, associated with a gynobasic style. This fruit type was used as a defining feature for Lamiaceae prior to molecular data (Ryding, 1995). However, a large number of tropical Lamiaceae (former Verbenaceae) have drupaceous or nut-like fruits [e.g. Ajugoideae: *Clerodendrum* L. *s.l.* (Steane, de Kok & Olmstead, 2004; Yuan *et al.*, 2010); Callicarpoideae: *Callicarpa*, Prostantheroideae: Chloantheae (Conn *et al.*, 2009]); Viticoideae: *Vitex* L. *s.s.*], or even capsular fruits [Viticoideae: *Teijsmanniodendron* Koord. (de Kok, Go & Latiff, 2009)]. No recent hypothesis has yet been provided for when nutlets evolved, or what the ancestral fruit type of Lamiaceae might have been. Based on parsimony reconstructions on a suboptimal Lamiales phylogenetic tree that placed Myoporaceae as sister to Lamiaceae (as well as questionable coding in which nutlets were treated as indehiscent fruits), Wagstaff & Olmstead (1997) reconstructed the most recent common ancestor for all Lamiaceae as bearing an ovary with four uniovulate locules and a fleshy fruit. More recently, Li *et al.* (2016) suggested that nutlets evolved independently from other Lamiaceae in Prostantheroideae. The presence of drupaceous fruits in the most recent common ancestor in all Lamiaceae is therefore a reasonable hypothesis, but needs to be more rigorously scrutinized.

Likewise, although most of Lamiaceae are herbaceous, the drupaceous members are also often woody (Harley *et al.*, 2004). Transition between woody and herbaceous habit in Lamiales appears complex, confounded by uncertainty regarding relationships in the order. As such, the ancestral habit of the family has never been estimated although, as with fruit type, a woody habit seems to be a reasonable starting hypothesis in need of testing.

Perhaps the most evolutionary noteworthy morphological trait in Lamiaceae is the occurrence of two fertile stamens, brought about by either reduction to staminodes or, in most cases, the loss of one pair of stamens. The presence and evolutionary consequence of this reduction has been the topic of much study, especially in *Salvia* (Claßen-Bockhoff, Wester & Tweraser, 2003; *et al.*, 2004; Walker & Sytsma, 2007; Celep *et al.*, 2020; Kriebel *et al.*, 2020). However, several genera of New World Menthinae also have two stamens (Harley *et al.*, 2004; Drew & Sytsma, 2012; Drew *et al.*, 2017a). In Menthinae (Nepetoideae), Drew & Sytsma (2012) suggested four or five origins of two stamens from ancestors with four stamens, with four to six reversions to four stamens. When other Nepetoideae are considered, there appear to be an additional three or four origins of two stamens, with possibly one or two more reversions to four stamens (Drew & Sytsma, 2012). Uncertainty in the number of losses/gains in stamen number is a result of both uncertainty in relationships in New World Menthinae (Drew & Sytsma, 2012; Drew *et al.*, 2017a), and uncertainty in the ancestral reconstructions in New World Menthinae and the crown nodes of Salviinae and Lycopinae + Menthinae + Nepetinae. Conversely, some lineages of Lamiaceae have more than four stamens (up to 16). Such conditions are found in several (sub)tropical clades (Harley *et al.*, 2004) such as *Lachnostachys* Hook. (five to eight), *Symphorema* Roxb. (six), and *Tectona* (five or six), although when stamens are greater than four the genus is generally polymorphic regarding stamen number. In their phylogenetic tree of Lamiales, Schäferhoff *et al.* (2010) reconstructed stamen number in Lamiales. Lamiaceae were represented as a single tip coded as fixed for four stamens, but the most recent common ancestor of Lamiaceae and its sister clade was reconstructed as bearing four stamens.

Given a current lack of consensus on relationships in Lamiaceae, particularly with representative taxonomic sampling, a poorly understood timeframe for the diversification of the family and untested or tenuously supported hypotheses regarding the historical biogeography and fruit/fertile stamen evolution in the family, we sought to address these issues. Specifically, our goals were to: (1) provide a hypothesis of relationships in Lamiaceae based on rapidly evolving plastid data and with representative sampling in the family; (2) provide a robust estimate of subfamilial and inter-subfamilial divergence times; (3) clarify the historical biogeography of the deepest nodes of Lamiaceae, specifically testing the hypothesis of a K-Pg, Gondwanan origin of the family; (4) assess whether ancestral Lamiaceae were woody or herbaceous; (5) investigate the evolution of fruit type in the family to better place the origin of nutlets and examine the lability of fruit type and (6) investigate fertile stamen number evolution in the family, including clarifying fertile stamen number at the crown of subtribe Salviinae.

MATERIAL AND METHODS

Taxonomic sampling in Lamiales and Lamiaceae

Our analyses included two datasets, a Lamiales-wide dataset and a second dataset that focused on Lamiaceae. The Lamiales-wide dataset contained 196 taxa and included 63 accessions from Lamiaceae (Supporting Information, File S1). Sequences for the Lamiales-wide data were downloaded from GenBank and were largely from Schäferhoff *et al.* (2010), Bendiksby *et al.* (2011) and Refulio-Rodriguez & Olmstead (2014), but were augmented with our own data and sequences from GenBank. The Lamiales-wide dataset contained representatives from all families of Lamiales, with the exception of the newly described Wightiaceae (Liu *et al.*, 2020), and included taxa from all 12 subfamilies of Lamiaceae (Li & Olmstead 2017). Two species of Solanaceae (Solanales) and one species each from Montiniaceae (Solanales), Apocynaceae (Gentianales) and Rubiaceae (Gentianales) served as an outgroup, with the two taxa from Gentianales used to root the phylogenetic tree (Refulio-Rodriguez & Olmstead 2014).

The second dataset focused on Lamiaceae and included 195 species. The Lamiaceae dataset included data from Drew & Sytsma (2012), GenBank and newly sequenced data (Supporting Information, File S1). Of these 195 species, 178 species were from Lamiaceae and represented all 12 subfamilies of Lamiaceae, and 17 species were from closely related families [Acanthaceae (two species), Bignoniaceae (one species), Gesneriaceae (two species), Linderniaceae (one species), Mazaceae (one species), Orobanchaceae (four species), Paulowniaceae (one species), Pedaliaceae (one species), Phrymaceae (two species), Plantaginaceae (one species) and Scrophulariaceae (one species)] and served as an outgroup. Gesneriaceae were used to root the phylogenetic tree based on Refulio-Rodriguez & Olmstead (2014).

DNA extraction, gene regions and sequencing

We used the DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) to extract DNA from silica-dried leaves and herbarium collection specimens. Thermal cycler conditions for PCR were as in Drew *et al.* (2017a), and we usedTaKaRa Ex Taq (Otsu, Shiga, Japan) PCR products. PCRs were diluted 30× in water and subsequently cycle sequenced, electrophoresed and analysed at the University of Arizona Genetics Core.

The Lamiales-wide analyses used the *matK*, *ndhF*, *rps16* and *trnL-trnF* plastid gene regions. For phylogenetic analyses of the Lamiaceae dataset we used the entire *ycf1* gene (also a small portion of *trnN-GUU* and the *trnN-ycf1* spacer) and the *ycf1-rps15* spacer, *trnL-F* and *rpl32-trnL* regions. The *trnN-ycf1* spacer and the first *c.* 900 nucleotides of *ycf1* were sequenced and amplified for most taxa using the primers *trnNGUU*-f (TAA CAG CCG ACC GCT CTA CC) and *ycf1*-921r (CAT TCA CTC GGA TTT CTT CSS), but for some more recalcitrant taxa we used the primers *ycf1*-167f (TCA GCA ACR ACT GGT TTT ATT ACG) and *ycf1*-936r (ATT CAT TTT CCT TTC CAT TCA CTC G). Primers for the remainder of *ycf1*, the *ycf1-rps15* spacer and the *trnL-F* and *rpl32-trnL* gene regions are detailed in Drew & Sytsma (2011, 2012).

Sequence analyses, phylogenetic inference and divergence time estimation

Contigs of the plastid DNA sequences (for the Lamiaceae dataset) were constructed and edited using Sequencher v.4.7 (Gene Codes, Ann Arbor, MI, USA) and Geneious v.9.1.8 (Kearse *et al.*, 2012). Sequences for both the Lamiales-wide and Lamiaceae datasets were aligned and edited using Mesquite v.3.61 (Maddison & Maddison, 2019). For both datasets, phylogenetic and divergence time analyses were conducted simultaneously using BEAST v.1.10.4 (Suchard *et al.*, 2018).

In BEAST analyses of both Lamiales and Lamiaceae, datasets were partitioned to account for sequence rate heterogeneity between different gene regions. The Lamiales dataset had four partitions corresponding to the *matK*, *ndhF*, *rps16* and *trnL-trnF* plastid regions. The Lamiaceae dataset had five partitions corresponding to the *trnN-ycf1* spacer, *ycf1*, *ycf1-rps15* spacer, *trnL-F* and *rpl32-trnL* regions. For each of our data partitions, we used a model of evolution as determined by the Akaike information criterion in jModelTest2 (Darriba *et al.*, 2012). For the Lamiales dataset jModelTest2 suggested the TVM + Γ + I (*matK*) and GTR + Γ + I (*ndhF*, *rps16*, *trnL-trnF*) models, and for the Lamiaceae dataset we used the GTR + Γ (*trnN-ycf1* spacer and *rpl32-trnL*) and GTR + Γ + I (*trnL-trnF*, *ycf1*, *ycf1-rps15* spacer) models.

To estimate divergence times in Lamiaceae we used a two-step approach. The first dataset consisted of a Lamiales-wide data matrix composed of *matK*, *ndhF*, *rps16* and *trnL-F*. The alignment initially consisted of 196 taxa and 7451 nucleotides (*matK* = 1804, *ndhF* = 2202, *rps16* = 1742 and *trnL-F* = 1703). Subsequently, after deletion of ambiguously aligned characters and long single-taxon insertions the Lamiales data matrix we used for analyses consisted of 6154 nucleotides (*matK* = 1732, *ndhF* = 2172, *rps16* = 1128, and *trnL-F* = 1122). The root of the tree was constrained with a uniform prior with ages ranging from 81 to 121 My, and the crown of Lamiales + Solanales was constrained with a uniform prior ranging from 79 to 115 My based on Barba-Montoya *et al.* (2018; nodes 1052 and 1053, respectively). The crown of Lamiales was constrained with a uniform prior ranging from 67 to 107 My based on several analyses that previously estimated dates for the crown node of Lamiales (Bremer, Friis & Bremer, 2004; Janssens *et al.*, 2009; Magallón & Castillo, 2009; Magallón *et al.*, 2015; Barba-Montoya *et al.*, 2018). We used the following five fossil calibration points within Lamiales. (1) For the most recent common ancestor (MRCA) of Acanthaceae *s.s.* and *Avicennia* L. we used a lognormal distribution (LD) with a minimum age of 37.8 My, a mean (mu) of 1.5 and a standard deviation (sigma; SD) of 1.0 based on Cavagnetto & Anadón (1996), Tripp & McDade (2014) and Woodcock *et al.* (2020). (2) Stem Bignoniaceae were constrained using a LD with a minimum age of 56 My, a mean of 2.0 and a SD of 1.0 based on Palaeocene fossils from western North America (Brown, 1962; Manchester, 2014; Manchester, Grímsson & Zetter, 2015) and Japan (Horiuchi & Manchester, 2011). (3) Crown Oleaceae were constrained with a LD, a minimum age of 65.4 My, a mean of 1.5 and a SD of 1.0 based on Late Cretaceous wood fossils described in Srivastava *et al.* (2015). (4) Crown Lamiaceae were constrained using a LD, a minimum age of 65.4 My, a mean of 1.5 and a SD of 1.0 based on Late Cretaceous wood fossils described in Wheeler *et al.* (2017). These fossils from Wheeler *et al.* (2017) were not attributable to a specific lineage, with affinities to *Gmelina* L. and *Vitex*, so we conservatively constrained the crown of Lamiaceae. (5) The crown of Nepetoideae were constrained with a LD, a minimum age of 47.8 My, a mean of 2.6 and a SD of 0.5 based on early Eocene *Ocimum* pollen fossils (Kar 1996; Sarkar & Prasad 2002; see Drew & Sytsma 2012). Additional fossil calibrations, such as *Melissa* (Nepetoideae) and *Ajuginucula* E.Reid (Ajugoideae) identified as early-Oligocene by Reid & Chandler (1926), have been previously employed to estimate divergence times in Lamiaceae (Drew & Sytsma, 2012, 2013; Yu *et al.*, 2014; Deng *et al.*, 2015; Yao *et al.*, 2016; Drew *et al.*, 2017a; Li *et al.*, 2017; Zhang *et al.*, 2017), but the identity of these fossils was questioned by Manchester, Grímsson & Zetter (2015). Additionally, Boltenhagen (1976a, 1976b) described Lamiaceae (Nepetoideae) pollen from the Coniacian, but Muller (1981) considered this finding as 'pending'.

Since a major goal of this paper is to estimate divergence times within Lamiaceae, we specifically tested how removing fossil calibrations in Lamiaceae would affect dating estimates within the family. To this end, we also experimented with several other calibration strategies to test the robustness of the calibrations outlined above. In concert with the aforementioned calibrations we: (1) constrained the crown of Oleaceae with a LD, a minimum age of 41.2 My (instead of 65.4 My as before; mean and SD as described previously) based on mid-Eocene fruit fossils described by Call & Dilcher (1992); (2) constrained the crown of Oleaceae with a LD and a minimum age of 41.2 My and removed constraints on the crown nodes of both Lamiaceae and Nepetoideae; (3) constrained the crown of Oleaceae with a LD, a minimum age of 83.6 My (mean and SD as described before) based on early Campanian pollen fossils described in Manchester, Grímsson & Zetter (2015); (4) constrained the crown of Oleaceae with a minimum age of 65.4 Myr (other parameters as described before) and removed constraints on the crown Lamiaceae; (5) constrained the crown of Oleaceae with a minimum age of 65.4 My and removed constraints on the crown Nepetoideae and (6) constrained the crown of Oleaceae with a minimum age of 65.4 My and removed constraints on the crown nodes of both Lamiaceae and Nepetoideae.

The second step of our divergence time estimation approach focused on Lamiaceae and contained the *trnN-ycf1* spacer, *ycf1*, *ycf1-rps15* spacer, *trnL-F* and *rpl32-trnL* gene regions. The data in this alignment was obtained from Drew & Sytsma (2012) and newly produced sequences, augmented with data from GenBank (particularly for outgroup taxa). The new sequences in this study include 26 new *trnL-F*, *rpl32-trnL* and *trnN-ycf1* spacer-*ycf1*-*ycf1-rps15* spacer region sequences and *c.* 1200 nucleotides (*trnN-ycf1* spacer and *c.* 900 nucleotides of *ycf1*) from 55 additional species. The initial alignment contained 195 taxa and 11 251 nucleotides (*trnN-ycf1* = 317, *ycf1* = 7146, *ycf1-rps15* spacer = 886, *trnL-F* = 1365, *rpl32-trnL* = 1537). After removing sites due to ambiguous alignment and/or long single-taxon insertions, the Lamiaceae alignment we used for analyses contained 9635 nucleotides (*trnN-ycf1* = 304, *ycf1* = 6606, *ycf1-rps15* spacer = 660, *trnL-F* = 1099, *rpl32-trnL* = 966). The Lamiaceae alignment included all 12 subfamilies and 17 of 22 tribes (all missing tribes were from Lamioideae). The largest three subfamilies were the most densely sampled with 105 Nepetoideae (*c*. 3500 total species), 28 Lamioideae (*c*. 1260 species) and 12 Ajugoideae (*c*. 770 species) species represented. This sampling was somewhat biased towards Nepetoideae, but we accounted for missing taxa in our coding schemes (see next), and the unsampled tribes of Lamioideae should not impact results or conclusions in this study. We used *Rehmannia* Libosch. ex Fisch. & C.A.Mey, (Orobanchaceae), *Triaenophora* (Hook.f.) Soler. (Orobanchaceae), *Pedicularis* L. (Orobanchaceae), *Lindenbergia* Doweld (Orobanchaceae), *Paulownia* Siebold & Zucc. (Paulowniaceae), *Phryma* L. (Phrymaceae), *Erythranthe* Spach (Phrymaceae), *Mazus* Lour. (Mazaceae), *Echinacanthus* Nees (Acanthaceae), *Aphelandra* R.Br. (Acanthaceae), *Sesamum* L. (Pedaliaceae), *Adenocalymma* Benth. (Bignoniaceae), *Digitalis* L. (Plantaginaceae), *Scrophularia* L. (Scrophulariaceae), *Lindernia* All. (Linderniaceae), *Boea* Comm. ex Lam. (Gesneriaceae) and *Primulina* Hance (Gesneriaceae) as outgroups. Gesneriaceae were used to root the tree following Refulio-Rodriguez & Olmstead (2014).

We constrained the root of this tree with a uniform prior and lower and upper ages of 80.8 and 94.1 My based on our initial Lamiales-wide analysis. We also constrained the MRCA of Lamiaceae and Orobanchaceae (lower = 69.6, upper = 80.3), crown Lamiaceae (lower = 65.6, upper = 71.9) and crown Nepetoideae (lower = 51.3, upper = 61.3) based on ages obtained from the Lamiales-wide analysis. We also constrained the crown of tribe Stachydeae with a LD, an offset of 13.8 My, a mean of 1.5 and a SD of 0.5 based on fossils described in Mai (2001) and used in previous studies of Lamioideae (e.g. Roy *et al.*, 2013). To explore the effect that constraining Lamiaceae and Nepetoideae had on dates obtained for Lamiaceae we performed two additional analyses. (1) We conducted an analysis as before, but without constraining Lamiaceae or Nepetoideae. (2) We conducted an analysis as before, but used the fossil calibrations for crown Lamiaceae and crown Nepetoideae as described in the Lamiales-wide BEAST analysis instead of using uniform priors.

For our BEAST Lamiales-wide analysis we conducted seven separate analyses of 150 000 000 generations of Markov chain Monte Carlo (MCMC), sampling every 5000 generations. The specific burn-in percentages for each individual Lamiales analysis were assessed using Tracer v.1.7.1 (Rambaut *et al.*, 2018), and varied between 10 and 15%. We conducted six BEAST runs for the Lamiaceae analyses of 120 000 000 MCMC generations with samples taken every 5000 generations. After analysing the log files in Tracer v.1.7.1, we discarded the first 10% of trees as burn-in for each run. For both the Lamiales-wide and Lamiaceae analyses, the MCMC output analyses were combined using logCombiner v.1.10.4 and the maximum clade credibility (MCC) tree was subsequently produced using TreeAnnotator v.1.10.4; both programs are included in the BEAST package.

PHYLOGENETIC PLACEHOLDERS

Comparative analyses, including analyses of historical biogeography and character evolution, are most accurate when the character states of unsampled tips are accounted for. As our dataset contains representative but incomplete sampling of tribes and especially genera, we selected tips as placeholders, attempting to group as finely as possible to minimize the total number of dropped tips. Placeholders were selected to represent monophyletic groups while ensuring that all known genera of Lamiaceae were properly accounted for in our comparative analyses. If multiple tips existed for a particular placeholder (e.g. multiple samples of *Stachys* for Stachydeae), all tips but one were dropped to ensure monophyly. Detailed composition and justification of taxonomic placeholders are fully documented in the Supporting Information (Supplementary Methods), but a summary of the process follows. Using a list of all tips in our phylogenetic tree, we selected the greatest number of monophyletic groups represented by our tips while considering the distribution of unsampled genera. The placement of unsampled genera was inferred using information from previous studies of Lamiaceae, especially (but not exclusively) subfamilial and tribal level studies (e.g. Paton *et al.*, 2004; Bramley *et al.*, 2009; Conn *et al.*, 2009; Bräuchler *et al.*, 2010; Yuan *et al.*, 2010; Bendiksby *et al.*, 2011; Pastore *et al.*, 2011; Drew & Sytsma, 2012; Salmaki *et al.*, 2013; Li *et al.*, 2016; Roy *et al.*, 2016; Li *et al.*, 2017; Paton, Mwanyambo & Culham, 2018; Xiang *et al.*, 2018) and checking our list of genera against Harley *et al.* (2004), adding genera as necessary and accounting for changes in circumscription.

Using known relationships among genera as a guide, we pruned our MCC tree to 115 tips in Lamiaceae, maximizing the number of terminals represented without presenting an inaccurate summary of generic relationships. Due to uncertainty about interfamilial relationships in Lamiales and sparse sampling among other families of Lamiales, we excluded outgroups. We were able to ensure monophyletic and representative placeholders for all subfamilies with the exception of Lamioideae. In Lamioideae, we were missing data for tribes Paraphlomideae and Synandreae. We completely excluded Paraphlomoideae to avoid combining this small clade of *c*. 23 species with the larger tribes Lamieae, Leucadeae, Leonureae, Marrubieae and Phlomoideae (Bendiksby *et al.*, 2011). Because our plastid topology of relationships among tribes Galeopsideae and Stachydeae differed from Zhao *et al.* (2021), we used previous nuclear results to place Synandreae in a clade with Colquhounieae + Galeopsideae + Stachydeae for placeholder purposes (Roy & Lindqvist, 2015; Roy *et al.*, 2016). We were unable to account for eight small genera that do not have any molecular data available to determine their exact phylogenetic position (Bendiksby *et al.*, 2011; Xiang *et al.*, 2018): *Benguellia* G.Tayl. (Nepetoideae: Ocimeae, one species), *Eriothymus* (Benth.) Schmidt (Nepetoideae: Mentheae, one species), *Madlabium* Hedge (Lamioideae, one species) *Metastachydium* Airy Shaw ex C.Y.Wu & H.W.Li (Lamioideae, one species), *Monochilus* Fisch. & C.A.Mey (Ajugoideae, two species), *Pseudomarrubium* Popov (Lamioideae, one species) and *Renschia* Vatke (Scutellarioideae, one species). Last, because of rampant non-monophyly in New World Menthinae (Bräuchler *et al.*, 2010; Drew & Sytsma, 2012; Drew *et al.*, 2017a), we coded all tips in this clade directly rather than using placeholders.

Ancestral range estimation

We scored our taxonomic placeholders for presence/absence in eight regions, delimited on the location of present geographical barriers and past tectonic activity and with a particular focus on the areas of high mint diversity recognized by Hedge (1992) and refined by Harley *et al.* (2004), largely corresponding to the area delimitation of Drew & Sytsma (2012) and Kriebel *et al.* (2019): (1) Africa south of the Sahel including the southern Arabian Peninsula and Indian Ocean islands; (2) Australasia including Australia, New Zealand, New Guinea, and south Pacific islands; (3) Eurasia excluding the Mediterranean north of the Central Asian deserts and west of the Altai Mountains; (4) North America south to Chiapas in Mexico; (5) North-East Asia east of the Altai Mountains and north of the Himalayas, and including temperate China, Japan and the Korean peninsula; (6) South and Central America south of Chiapas including the Caribbean; (7) South-East Asia from India, east to tropical China and south to Indonesia and (8) the Mediterranean and South-West Asia from the Canary Islands east to Pakistan, including North Africa, the Middle East and the Caucasus.

We determined the geography of extant tips using a multi-step process. First, we downloaded all records of Lamiaceae from GBIF (https://www.gbif.org/), applying a filter to remove coordinates located in heavily urban areas, country centroids or in the ocean using the function *clean_coordinates* in the R package 'CoordinateCleaner' (Zizka *et al.*, 2019). We then converted country and state/province/department to one of the eight areas delimited previously and scored each genus for presence/absence in these areas. Last, we checked genus distributions against reported ranges in Harley *et al.* (2004) and *Flora of China* (http://www.efloras.org/). We then combined the range of each genus into a 'super range' for each placeholder clade. Geography of each genus and placeholder is available in the Supporting Information (File S2).

In several cases in which placeholder tips were wide-ranging (seven or eight of all eight possible ranges), we were concerned that the broad extant distribution would add noise to reconstructions when a priori knowledge from published analyses at lower taxonomic levels allows objective removal of ostensibly derived ranges. We therefore removed some ranges from these placeholders to incorporate this knowledge. We only used studies in which ranges were nearly equally or more finely partitioned than our own: [*Elsholtzia* Willd. *s.s.*: Li *et al.*, 2017; *Isodon* (Schrad. ex Benth.) Spach: Yu *et al.*, 2014; subgenera of *Salvia*: Kriebel *et al.*, 2019; and *Pogostemon*: Yao *et al.*, 2016] or based on clear phylogenetic signal that temperate areas are derived in Scutellarioideae (Li *et al.*, 2016; Zhao *et al.*, 2020).

For four clades in which phylogenetic data exist but were either without or with only a parsimony-based ancestral range reconstruction (*Teucrium* L. *s.l.*: Salmaki *et al.*, 2016; Leucadeae: Scheen & Albert, 2009; *Mentha*: Bunsawat *et al.*, 2004; and Stachydeae: Roy *et al.*, 2013), we reanalysed the available aligned DNA matrices from these studies, assigning priority to nuclear DNA matrices, when available, and ignoring plastid data if both were available. Matrices were analysed using RAxML (Stamatakis, 2014) under GTR + Γ with 20 search replicates for the best tree. For each clade, the most likely phylogram across all search replicates was made ultrametric to relative time using penalized likelihood with the *chronos* function in 'ape' v.5.0 (Paradis & Schliep, 2018) with a smoothing parameter of 0.1. Each tip was scored for presence/absence in the eight ranges defined for the entire family. Ancestral range estimation (ARE) was conducted in BioGeoBEARS v.1.1.2 (Matzke, 2012; 2013) using the dispersal-extinction-cladogenesis (DEC) model (Ree & Smith, 2008) without any additional assumptions. Note that the primary aim of this analysis was to obtain a reasonable idea of the ARE for the crown of each of the four placeholder lineages that we analysed here, not a definitive analysis of their historical biogeography. As such, we incorporated uncertainty about the ARE into our refined coding by including all possible ranges inferred at each crown, irrespective of their probabilities. Phylogenetic trees, tip coding and results for these clade-specific analyses are available in the Dryad repository accompanying this paper.

Using the refined presence/absence matrix for our placeholder clades in each of the eight areas, we conducted ARE for Lamiaceae using BioGeoBEARS, testing between the DEC model with and without the additional 'j' parameter, which models jump dispersal/founder events. The significance of the fit of the DEC and DECj models were tested using a likelihood ratio test. We allowed a maximum range size of up to seven areas (the maximum found in any extant tip post cleaning). We conducted a stratified analysis, with dispersal probabilities between pairs of areas specified for four time slices (0–35, 35–65, 65–90 and > 90 Mya) with dispersal multipliers modified in each time slice based on the adjacency of landmasses during each period of time. The justification for these time slices and dispersal probabilities have been detailed in other large groups with wide geographical distributions, as in Lamiaceae (e.g. Buerki *et al.*, 2011: Sapindaceae; Berger *et al.*, 2016: Myrtales; Cardinal-McTeague, Sytsma & Hall, 2016: Brassicales; Spalink *et al.*, 2016: Cyperaceae; Rose *et al.*, 2018: Ericales; Kriebel *et al.*, 2019: *Salvia*) and are based on known geological events affecting geographical distributions in such widely dispersed groups. These time slices allowed for testing the importance of continental vicariance and collision as well as the role of possible Northern Hemisphere land bridges (Tiffney, 1985, 2000; Tiffney & Manchester, 2001; Graham, 2011). The wide time range between 0 and 35 Mya was used to allow the potential for relatively easy dispersal between areas given our relatively sparse sampling towards the tips. All input files are provided in the Dryad repository accompanying this paper.

ANCESTRAL STATE RECONSTRUCTION OF TRAITS

We scored habit, fruit type and fertile stamen number for all genera of Lamiaceae, using the same list as for the ARE analysis previously and scoring 'supergenera' based on the trait combinations found in their constituent genera. We scored habit based on descriptions in Harley *et al.* (2004), treating genera as either woody or herbaceous, with 'undershrubs' grouped as woody and the ambiguous state of 'subshrubs' scored as either woody or herbaceous based on the prevailing habit in the rest of the genus. We scored fruit type as one of three categories (capsule, drupe/drupaceous/nut or schizocarp/nutlet) based on fruit descriptions in Harley *et al.* (2004). Last, we scored fertile stamen number as one of four categories: 2, 4, 5 or > 5, again using Harley *et al.* (2004). We allowed placeholders to be polymorphic. Ancestral states were reconstructed by fitting a Markov-k (Mk) model using maximum likelihood as implemented in the *rayDISC* function in the R package 'corrHMM' v.1.22 (Beaulieu, Oliver & O'Meara, 2017). We reconstructed traits using an equal rate (ER), symmetric rates (SYM) and all rates different (ARD) model of trait evolution, with the root state estimated using the procedure of FitzJohn, Maddison & Otto (2009). We selected the best model for each trait using Akaike weights derived from the AICc score of each model. Trait scoring for each genus and placeholder is available in the Supporting Information (File S3).

RESULTS

LAMIACEAE IN LAMIALES AND SUBFAMILIAL RELATIONSHIPS

Our analysis of the aligned matrix of 6154 bp of 196 species of Lamiales resulted in a largely well-supported hypothesis of relationships in Lamiaceae (Fig. 1; Supplementary Information, Fig. S1). Lamiaceae were recovered as sister to Mazaceae + Orobanchaceae + Paulowniaceae + Phrymaceae with a posterior probability (PP) = 1.0. In Lamiaceae, all subfamilies, when represented by multiple tips, were supported by PP = 1.0 except Lamioideae (PP > 0.99). In Lamiaceae, Callicarpoideae + Prostantheroideae formed a clade (PP > 0.99) strongly supported as sister to the rest of the family (PP = 1.0). Nepetoideae were in turn sister to all remaining subfamilies (PP > 0.99). In Nepetoideae, we recovered Mentheae as sister to Elsholtzieae + Ocimeae, but with low support (PP = 0.77). Relationships among the clade formed by the remaining subfamilies (PP > 0.99) were largely well-supported, with the clade Symphorematoideae + Viticoideae (PP = 1.0) sister to the remainder of the clade (PP > 0.99), and Tectonoideae subsequently sister to remainder of the clade, which comprised Ajugoideae + Cymarioideae + Lamioideae + Peronematoideae + Premnoideae + Scutellarioideae (PP = 1.0). In this last clade, relationships were less clear, with Ajugoideae weakly supported as sister to the remaining families in the clade (PP = 0.82). Although weakly supported, Peronematoideae were sister to a strongly supported clade of Cymarioideae + Lamioideae + Scutellarioideae (PP = 1.0), with Scutellarioideae in turn sister to Cymarioideae + Lamioideae (PP = 1.0).

LAMIACEAE DIVERGENCE TIMES

Our various calibration strategies yielded largely congruent divergence time estimates (Supplementary Information, Table S1). The results that follow are based on our BEAST analyses that used five fossil calibration points (including two in Lamiaceae) and Oleaceae constrained at 65.4 Mya. Major divergence

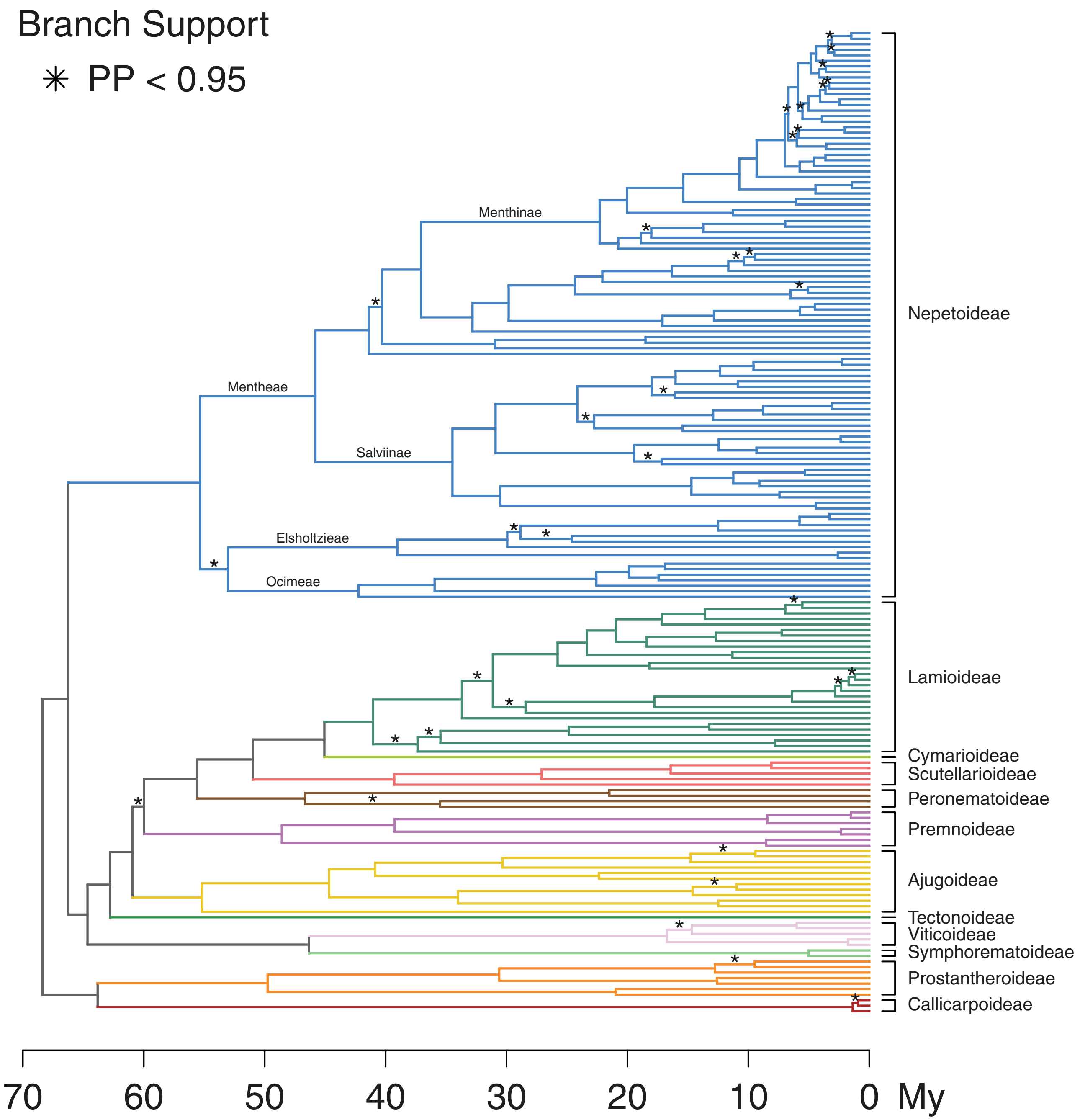

Figure 1. BEAST chronogram of Lamiaceae with outgroups removed. Branches are coloured based on subfamily, which are labelled to the right. Tribes and subtribes in Nepetoideae as discussed in the text are indicated on internal branches. All branches are supported by posterior probabilities (PP) > 0.95, unless indicated by an asterisk (*).

times in Lamiaceae are presented in Table 1. The analyses that explored different calibration strategies regarding the Lamiaceae and Nepetoideae crowns yielded similar ages for major clades in Lamiaceae, with dates slightly older when fossil calibrations for the Lamiaceae and Nepetoideae crowns were used (data not shown, but see supplementary table and tree files in Dryad). Lamiaceae diverged from other Lamiales in the Late Cretaceous [stem age 74.4 My, 95% highest posterior density (HPD) = 69.8–78.9 My] and began to diversify soon after (crown age 68.4 My, 95% HPD = 65.6–71.3 My) (Fig. 1; Table 1). Stem lineages of most extant subfamilies diverged within a narrow window during the Late Cretaceous and

Table 1. Summary of crown and stem divergence times (in My) and 95% highest posterior density (HPD) for major clades (family, subfamily, and select tribes) of Lamiaceae as outlined in Figure 1. An asterisk (*) indicates clades for which, due to sampling, have their deepest divergences not represented in this dataset. In the case of Callicarpoideae (and probably also Viticoideae), infra-subfamilial relationships are not well known such that the crown age based on extant tips is uncertain. In the other cases, genera around the crown radiation are missing for Prostantheroideae (*Brachysola* Rye), Scutellarioideae (*Wenchengia* C.Y.Wu & S.Chow) and Symphorematoideae (*Sphenodesme* Jack), so the crown ages presented here are under-estimated. A dagger (†) indicates subfamilies crown ages of which cannot be reported as they are represented by a single tip in this dataset

Clade	Crown mean	Crown 95% HPD	Stem mean	Stem 95% HPD
Lamiaceae	68.4	65.6–71.3	74.4	69.8–78.9
Ajugoideae	55.2	48.0–61.8	60.9	55.8–66.0
Callicarpoideae*	1.4	0.3–2.8	63.8	55.8–70.3
Cymarioideae†	–	–	45.1	38.2–52.0
Lamioideae	41.0	34.4–47.6	45.1	38.2–52.0
Nepetoideae	55.3	51.3–59.7	66.2	62.8–70.3
Nepetoideae: Elsholtzieae	39.0	28.5–49.1	53.0	47.4–59.2
Nepetoideae: Mentheae	45.8	38.9–52.6	55.3	51.3–59.7
Nepetoideae: Ocimeae	42.2	31.0–52.8	53.0	47.4–59.2
Peronematoideae	46.7	32.1–58.2	55.6	49.3–61.7
Premnoideae	48.6	33.5–60.5	60.0	54.7–65.2
Prostantheroideae*	49.8	36.6–61.7	63.8	55.8–70.3
Scutellarioideae*	39.3	28.0–50.0	51.0	44.2–57.6
Symphorematoideae*	5.0	1.2–9.9	46.3	29.4–61.5
Tectonoideae†	–	–	62.8	58.1–67.6
Viticoideae	16.7	8.3–26.3	46.3	29.4–61.5

Palaeocene, especially *c.* 55–65 Mya, and stem lineages of all extant subfamilies originated by the mid-Eocene *c.* 45 Mya, with the divergence between Cymarioideae-Lamioideae occurring most recently (Fig. 1; Table 1).

ANCESTRAL RANGE ESTIMATION

Reduced ranges of widespread clades

DEC on widespread clades led us to reduce the distribution of Leucadeae to Africa/South-East Asia/South-West Asia (Supplementary Information, Fig. S2), *Mentha* to Australasia/South-West Asia (Supplementary Information, Fig. S3), *Stachys s.l.* to Eurasia/South-West Asia (Supplementary Information, Fig. S4) and *Teucrium s.l.* to Australasia/South America/South-East Asia/South-West Asia (Supplementary Information, Fig. S5).

Lamiaceae biogeography

The DEC model without any additional parameters (LnL = −348.11, $d = 1.02 \times 10^{-2}$, $e = 1.00 \times 10^{-12}$; Supplementary Information, Fig. S6) was a worse fit for the data than DEC with a model with jump dispersal added (LnL = −340.44, $d = 9.6 \times 10^{-3}$, $e = 1.00 \times 10^{-12}$, $j = 0.020$; Fig. 2; Supplementary Information, Fig. S7) as indicated by a likelihood ratio test ($D = 15.34$, d.f. = 1, $P = 8.99 \times 10^{-5}$), and we therefore focus our discussion on the DECj results. Crown Lamiaceae were ambiguously reconstructed, with the highest probability belonging to a joint Australasian + South-East Asian origin ($P = 0.16$). Crown Callicarpoideae + Prostantheroideae were ambiguously reconstructed as Australasian in origin ($P = 0.30$), whereas the crown of the remainder of the family was reconstructed as South-East Asian ($P = 0.76$). Crown Nepetoideae originated in South-East Asia, although this reconstruction is not fully clear ($P = 0.57$), but the crown of Elsholtzieae + Ocimeae more clearly originated in this region ($P = 0.92$). Crown Mentheae may have originated in South-West Asia ($P = 0.58$), with a strong signal for this area along most of the backbone of Mentheae. There was strong South-East Asian signal among other subfamilies, especially for inter-subfamilial nodes and subfamilial nodes (when multiple placeholders are present), with the exception of Symphorematoideae + Viticoideae and Premnoideae, which were ambiguously reconstructed as South-East Asian in origin ($P = 0.54$ and 0.41, respectively). As with Nepetoideae: Mentheae, crown Lamioideae excluding tribe Pogostemoneae were reconstructed as originating in South-West Asia, albeit ambiguously ($P = 0.36$).

ANCESTRAL STATE RECONSTRUCTION OF HABIT

Akaike model weights did not clearly favour one model of trait evolution, so we selected the simplest

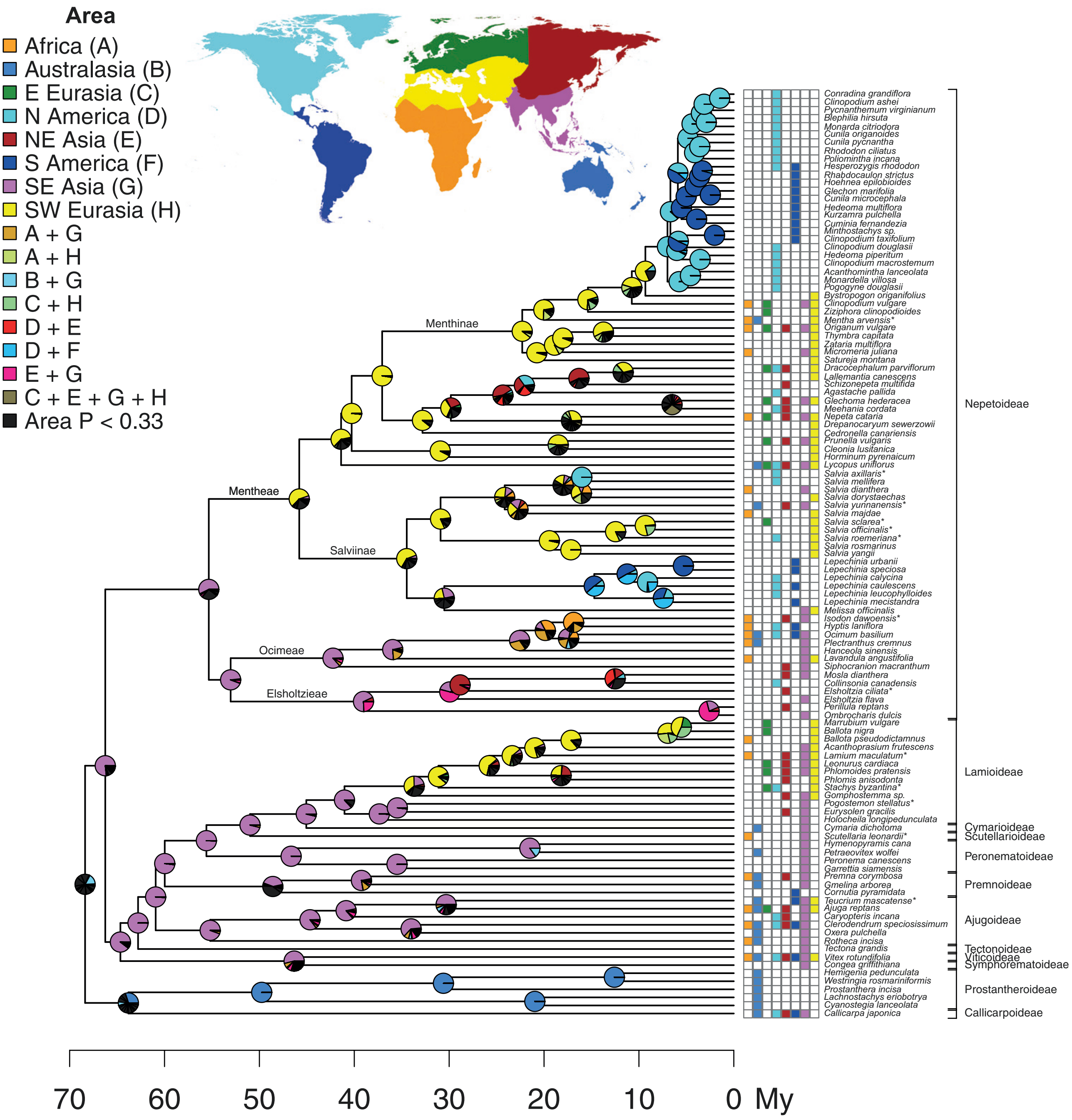

Figure 2. Ancestral range estimation of Lamiaceae under the DECj model. Node pies represent the relative probability of each range/joint range. Areas with a global maximum probability < 0.33 are coloured black. Tip boxes show presence or absence of each placeholder in each of the eight areas. Placeholders marked with an asterisk (*) have had their ranges reduced based on a priori knowledge of their ancestral range (see Materials and Methods). Subfamilies are indicated to the right. Tribes and subtribes in Nepetoideae as discussed in the text are indicated on internal branches.

(ER) model of evolution (weight = 0.31; Fig. 3). Lamiaceae were reconstructed as ancestrally woody ($P = 0.99$), and most subfamilies were reconstructed as ancestrally woody but with major shifts to herbaceous habit by crown Nepetoideae ($P = 0.79$), and at least crown Lamioideae excluding tribe Pogostemoneae ($P = 0.72$), but possibly also crown Lamioideae ($P = 0.55$). There were numerous transitions back

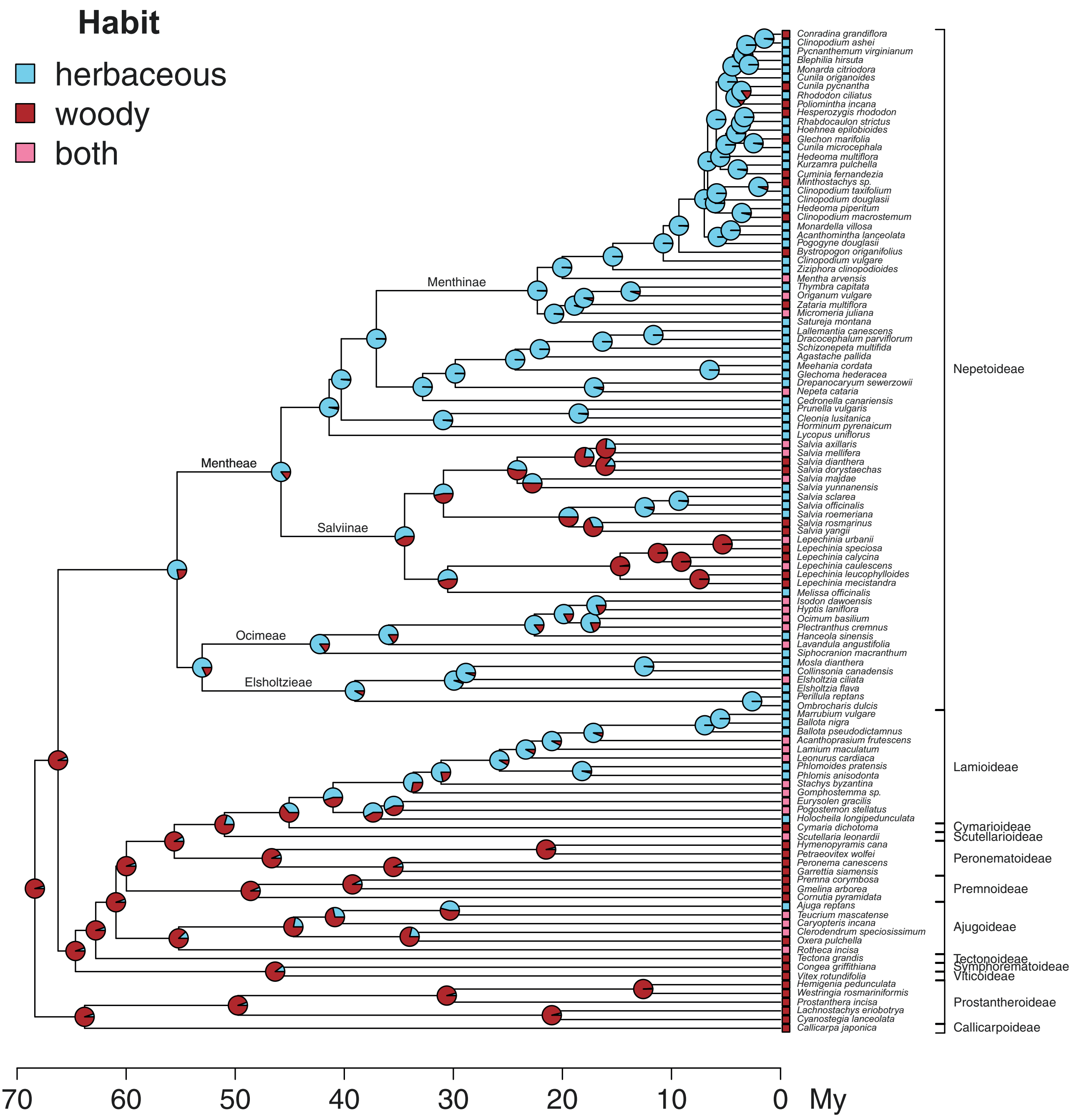

Figure 3. Ancestral state reconstruction of habit in Lamiaceae under an equal rates model of trait evolution. Subfamilies are indicated to the right. Tribes and subtribes in Nepetoideae as discussed in the text are indicated on internal branches.

to woody habit in Nepetoideae and Lamioideae in addition to transition that occurred along the stems to placeholder tips, notably in Nepetoideae subtribe Salviinae. However, it was unclear whether woodiness evolved independently in *Salvia* and *Lepechinia* Willd. or was ancestral in the tribe and was lost in clades of *Lepechinia*, *Melissa* and *Salvia*, as crown Salviinae were reconstructed as nearly equally probably woody or herbaceous ($P_{\text{herbaceous}} = 0.57$).

ANCESTRAL STATE RECONSTRUCTION OF FRUIT

Akaike model weights favoured the symmetric rates model of trait evolution (weight = 0.80; Fig. 4). Nutlets

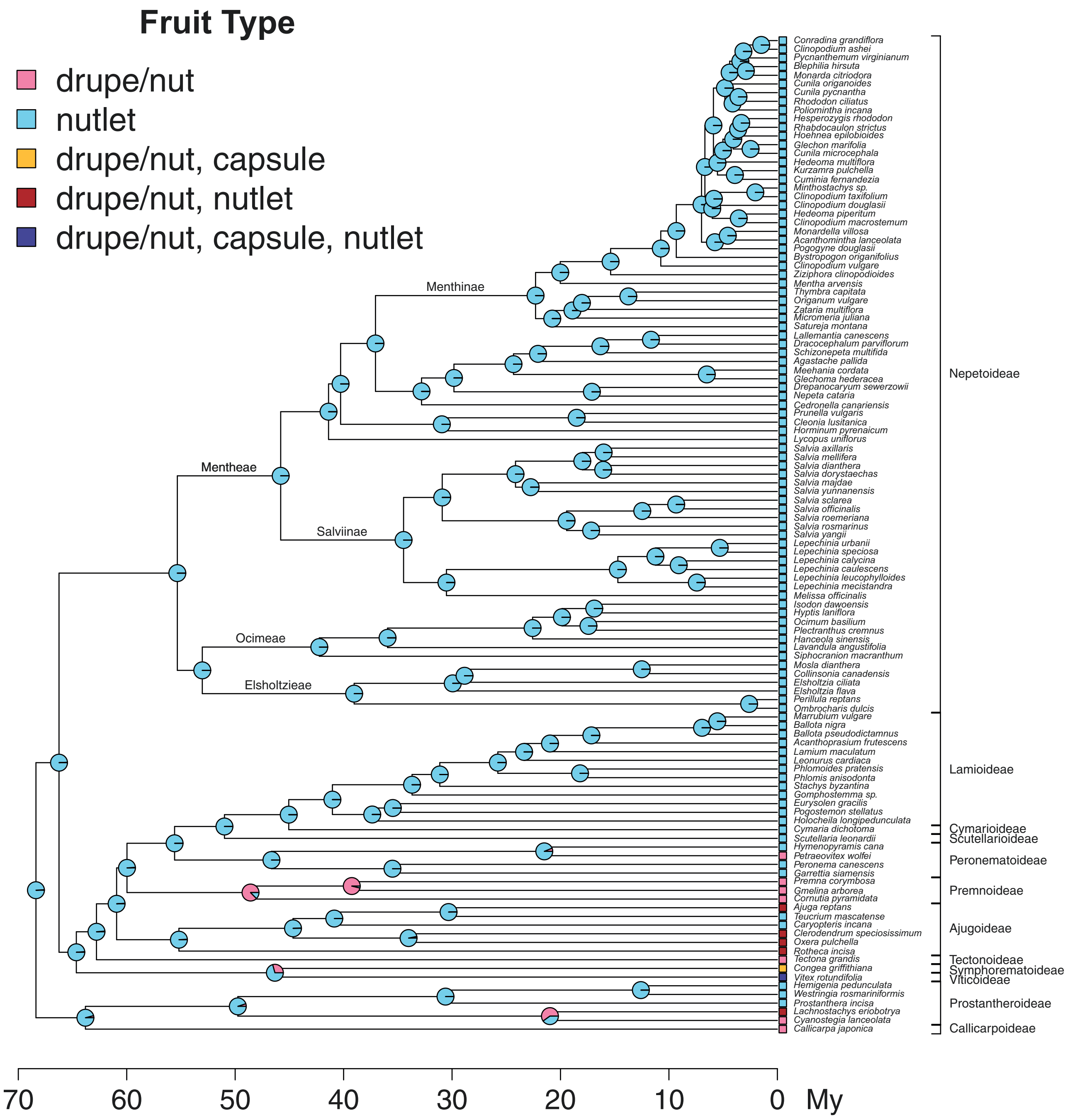

Figure 4. Ancestral state reconstruction of fruit type in Lamiaceae under a symmetric rates model of trait evolution. Subfamilies are indicated to the right. Tribes and subtribes in Nepetoideae as discussed in the text are indicated on internal branches.

were strongly favoured as the ancestral fruit type of Lamiaceae ($P > 0.99$) and as the ancestral fruit type for nearly all deeper nodes of the family. Capsular fruits evolved independently in Symphorematoideae and Viticoideae, and drupes evolved at least seven times, with the common ancestor of Premnoideae and Prostantheroideae tribe Chloantheae having drupaceous fruits ($P = 0.87$ and $P = 0.60$, respectively).

ANCESTRAL STATE RECONSTRUCTION OF STAMENS

Akaike model weights favoured the symmetric rates model of fertile stamen number evolution (weight = 0.91; Fig. 5). Ancestral Lamiaceae were inferred to have had four fertile stamens ($P > 0.99$). Transitions inferred at internal nodes were to two fertile stamens in Prostantheroideae, possibly at the crown of Westringeae ($P_{\text{2stamens}} = 0.43$) or more likely at the common ancestor of *Hemigenia* R.Br. and *Westringia* Sm. ($P = 0.76$), at crown *Salvia* ($P = 0.92$) and in New World Menthinae ($P = 0.70$). Taking into account evolution on branches leading to terminals, two stamens evolved at least 15 times in Lamiaceae, including independently in the closely related *Collinsonia* L. and *Mosla* (Benth.) Buch.-Ham. ex Maxim. (Nepetoideae: Elsholtzieae; Fig. 5). There were at least three reversions to four stamens from

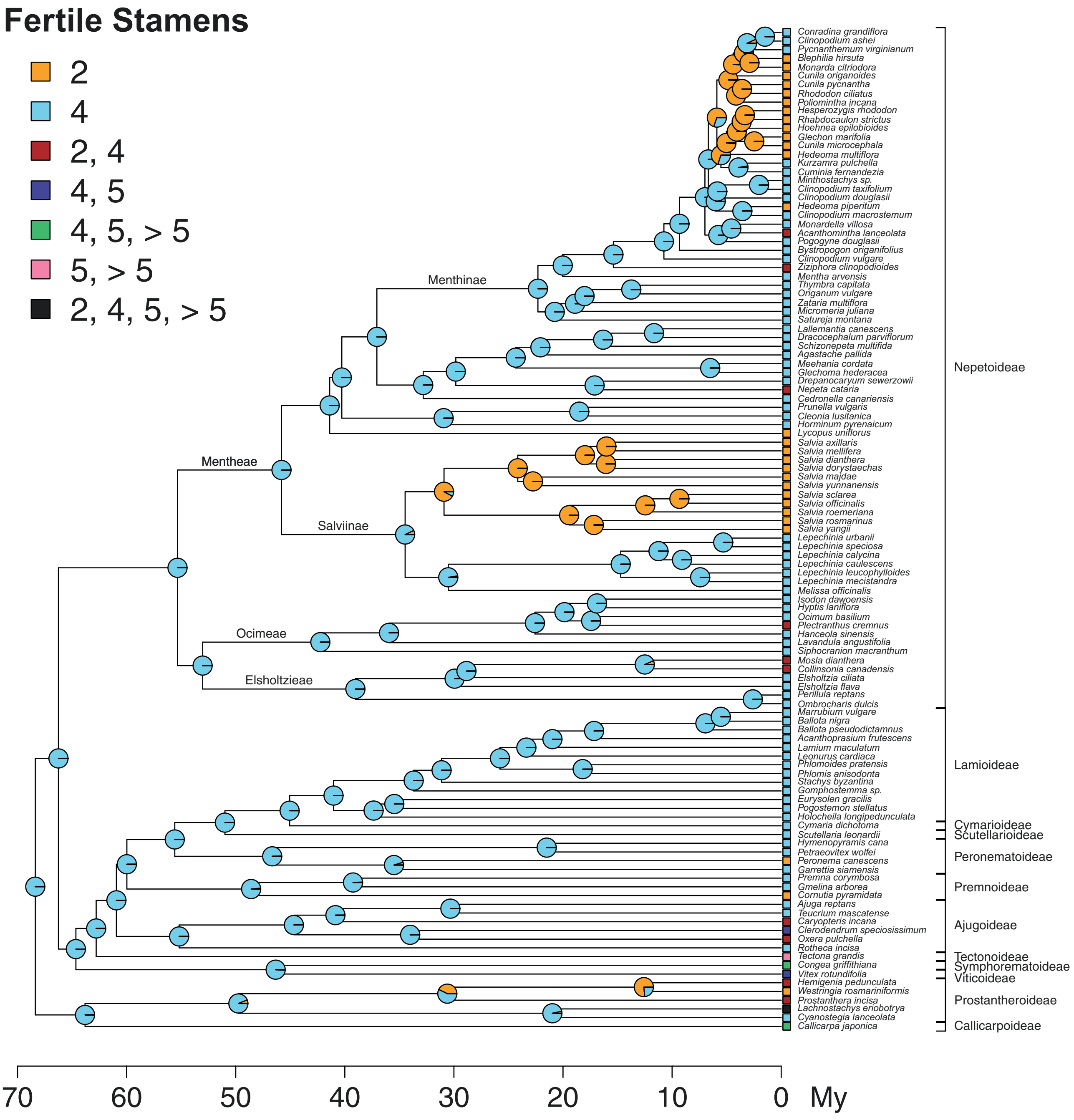

Figure 5. Ancestral state reconstruction of fertile stamen number in Lamiaceae under a symmetric rates model of trait evolution. Subfamilies are indicated to the right. Tribes and subtribes in Nepetoideae as discussed in the text are indicated on internal branches.

two stamens: in the *Hemigenia* clade and twice in New World Menthinae. On the other hand, stamen proliferation only occurred in smaller clades, especially in Prostantheroideae tribe Chloantheae. Few common ancestors in Lamiaceae were reconstructed as possessing a large number of stamens.

DISCUSSION

Over the past several decades, slow but substantial progress has been made towards understanding relationships in Lamiaceae. Despite this progress, much of the phylogenetic, biogeographical and morphological information has not been combined into a holistic view of the family, and most energy has been devoted towards understanding relationships rather than evolutionary phenomena. Elucidating relationships is a necessary first step, but a full understanding of evolution of the mint family requires time-calibrated trees combined with a robust series of comparative analyses. Such analyses are more routinely being done in the community of Lamiaceae systematists. In this way, our study helps to provide a framework for future studies by providing a robust family-wide context and a series of secondary fossil calibrations and refining hypotheses about the historical biogeography and trait evolution in the family. Here, we continue the process of clarifying the evolution of Lamiaceae and provide a framework for future studies by presenting a well-sampled and supported phylogenetic hypothesis that includes the first family-wide divergence estimates for major lineages in Lamiaceae and elucidate habit, fruit and stamen evolution within the family.

Progress towards a phylogenetic hypothesis for Lamiaceae

During the last five years, relationships among subfamilies have been proposed based on plastid and nuclear data (Li *et al.*, 2016; MEGC, 2018; Zhao *et al.*, 2021). Although resolved in some areas, relationships between some subfamilies remain unclear, and the monophyly of some subfamilies is also not fully elucidated (e.g. Cymarioideae, Viticoideae). Given that our molecular dataset is also derived from the plastid genome, but reflects a different set of loci, our topology largely matches the topology of Li *et al.* (2016) and Zhao *et al.* (2021) and is fully congruent with the summary topology of the former (Fig. 1). Moreover, our use of more rapidly evolving plastid loci clarifies most unresolved relationships in Li *et al.* (2016), including placing Nepetoideae as sister to the remainder of the family excluding Callicarpoideae + Prostantheroideae, positions also strongly supported by an analysis of 79 more conserved plastid loci (Zhao *et al.*, 2021) and nuclear data (MEGC, 2018).

Although our study still fails to confidently resolve the placement of Ajugoideae and Premnoideae, our MCC tree suggests that Premnoideae are more closely related to Lamioideae than are Ajugoideae (Fig. 1; PP = 0.82). This contrasts markedly with the relationships resolved by Zhao *et al.* (2021), who suggested that Ajugoideae are more closely related to Lamioideae than are Premnoideae, albeit without full support. Zhang *et al.* (2020), using whole plastomes from representatives of six ingroup subfamilies, suggested a third, albeit variably supported topology showing a sister relationship of Premnoideae and Tectonoideae. The relative placement of Ajugoideae and Premnoideae appears to be the most recalcitrant inter-subfamilial relationship in the plastid genealogy, and neither our sampling of a few relatively rapidly evolving loci nor the large number of more conserved loci appears sufficient to resolve this. A third, still untested strategy to robustly resolve the plastid genealogy may be to analyse complete plastome data from all subfamilies with multiple samples for all subfamilies.

More importantly, the plastid topology regarding relationships among Ajugoideae, Lamioideae, Peronematoideae, Premnoideae and Scutellarioideae (even ignoring the uncertain placement of Ajugoideae) is in strong conflict with relationships among these subfamilies inferred from 520 nuclear genes (MEGC, 2018; Godden *et al.*, 2019; Fig. 6), with Cymarioideae not represented by nuclear data. Based on these nuclear data, Scutellarioideae are strongly (but not fully) supported as sister to the remaining subfamilies, in contrast to being sister to Lamioideae as found based on plastid data. Additionally, nuclear data place Ajugoideae as essentially unresolved in relation to Lamioideae and Peronematoideae, in strong conflict with their placement in any previous study. One obvious caveat is that the nuclear data does not convincingly resolve this part of the tree. Furthermore, there is strong conflict between concatenation and coalescent species tree approaches when using these nuclear data. It is not immediately clear what the cause of conflict between genomes is, but at least some of topological conflict is probably real. Biologically, such conflict may be the result of horizontal gene flow in the evolutionary history of Lamiaceae, as has been demonstrated in *Salvia* (Rose *et al.*, 2021). Alternatively, some conflict may be artificially caused by error in gene tree estimation in MEGC (2018), which in turn introduces error into species tree estimation. This may arise from estimating gene trees for loci that contain little or no phylogenetic information about relationships at these nodes (cf. fig. S2b of MEGC, 2018).

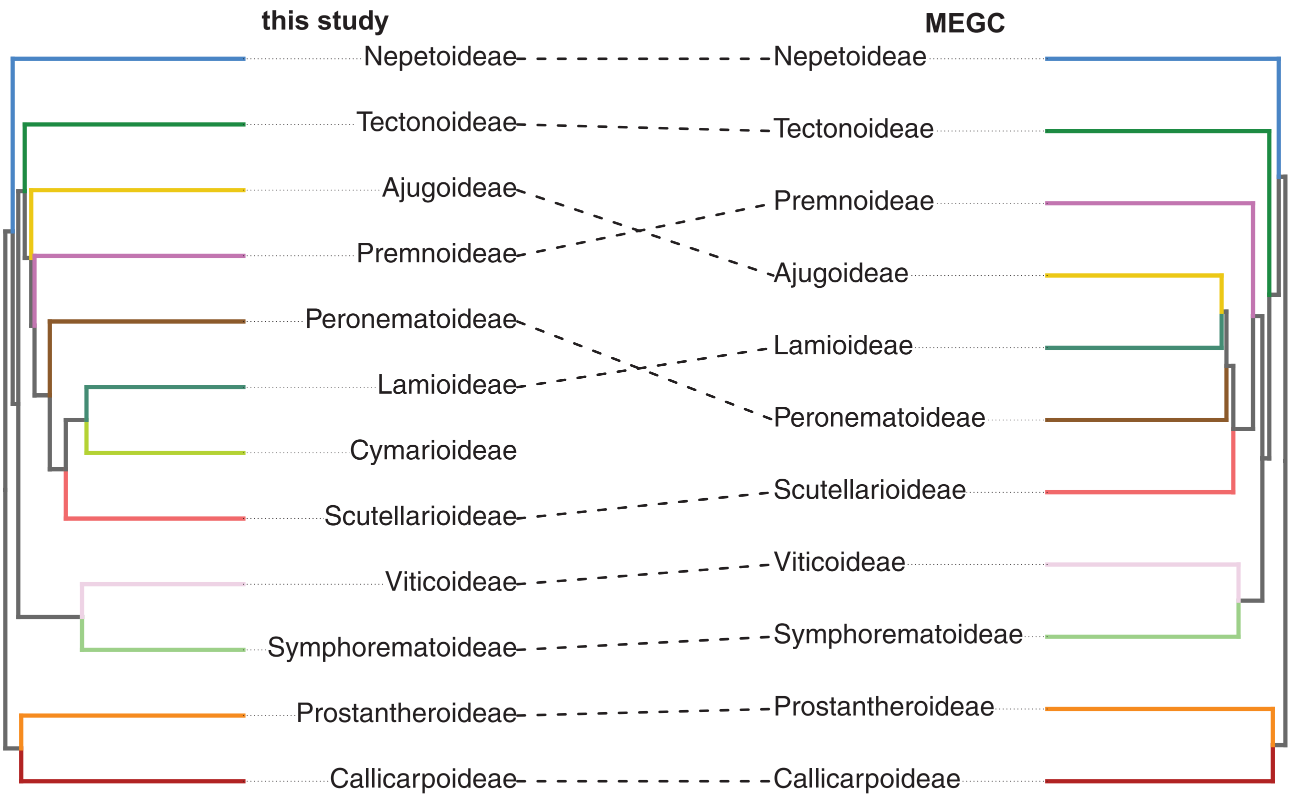

Figure 6. Tanglegram showing concurrence and conflict between plastid and nuclear relationships in Lamiaceae as currently understood. Time-calibrated plastid relationships are shown on left and are as found in this study, and nuclear relationships are shown at right based on the chronogram of Godden *et al.* (2019) using the tree of MEGC (2018). Links connect subfamilies and nodes are rotated to minimize link overlap. Note the conflicting placement of subfamilies Ajugoideae and Scutellarioideae.

A POST-K-PG DIVERSIFICATION OF LAMIACEAE

Our study finally provides a comprehensive, time-calibrated hypothesis for relationships throughout Lamiaceae. Our finding of a Late Cretaceous origin of stem and crown Lamiaceae clearly contradicts nearly all previous estimates of the divergence time of stem Lamiaceae at < 40 Mya (Fig. 1; Table 1). However, our ages are consistent with previous estimates of the age of crown Lamiaceae in studies focused within the family and using one or more primary internal calibrations (Drew & Sytsma, 2012; Yao *et al.*, 2016). Subfamily divergence times are nearly identical for those of Nepetoideae and tribes Mentheae and Ocimeae (Drew & Sytsma, 2012; Yu *et al.*, 2014; Li *et al.*, 2017), but the age of crown Elsholtzieae is younger in this study that estimated by Li *et al.* (2017) (39.0 vs. 50.1 My), perhaps a result of fairly sparse outgroup sampling in Li *et al.* (2017). Outside of Nepetoideae, our estimated divergence time of *c*. 41 Mya for crown Lamioideae is consistent with the estimate of Yao *et al.* (2016), but is much older than the estimate of *c.* 20–26 Mya by Roy & Lindqvist (2015). The latter young age is almost certainly due to constraining the age of Lamiaceae based on the young stem age for the family given in Martínez-Millán (2010). In Ajugoideae, Salmaki *et al.* (2016) estimated the crown age of Clades II–IV of Xiang *et al.* (2018) to be *c*. 16 Myr, considerably younger than our estimate of 45 Myr for this node. Although the methods of Salmaki *et al.* (2016) are somewhat unclear, they may have estimated these divergence times based on secondary calibrations from Roy & Lindqvist (2015). Given the apparent disparity in ages recovered across previous studies, we hope that our list of vetted and conservatively assigned primary calibrations will aid in choosing a common set of fossils with which to calibrate phylogenetic trees. This, in combination with a comprehensive set of secondary dates throughout the major clades of Lamiaceae, will ideally result in consistent set of divergence times across studies investigating questions at a diversity of taxonomic scales in Lamiaceae.

The clustering of divergence times of major Lamiaceae clades just after the K-Pg boundary, especially in the 40–50 Mya timeframe, suggest that the warm temperatures during the Eocene climatic optimum and, in particular, the subsequent rapid cooling had profound impacts on present day diversity of Lamiaceae. Rapid cladogenesis during this timeframe is also seen in other angiosperm families, particularly in the huge radiation of Asteraceae (Panero & Crozier, 2016; Mandel *et al.*, 2019), but also in smaller clades (Loranthaceae: Liu *et al.*, 2018; Ranunculaceae: Zhai *et al.*, 2019). Future studies should examine whether rates of speciation in Lamiaceae also increased during this timeframe.

Biogeographical origins: Out of South-East Asia and around the world?

In expanding on previous work, Harley *et al.* (2004) hypothesized that Lamiaceae originated during the Late Cretaceous on landmasses formerly part of Gondwana and that subfamily Nepetoideae originated in South-East Asia. Our ARE provides support for a Late Cretaceous origin for Lamiaceae and also supports the hypothesis that Nepetoideae originated in South-East Asia (Fig. 2). However, our study does not resolve if the family originated in the former parts of Gondwana or Laurasia, although our results indicate that the latter is more likely.

Our ARE suggests that crown Lamiaceae originated in either Australasia (former Gondwana) or South-East Asia (former Laurasia), with the possibility of a joint Australasian + South-East Asian origin, although the landmasses were far more distant then than at present. An origin involving one or both of these landmasses is evident from both the high probability of a South-East Asian origin of most subfamilies, as well as the strong Australasian signal in Prostantheroideae, which is endemic to Australia. Noise in our ARE is two-fold: first, in excluding outgroups from our analyses, and second because of the wide-ranging distribution of Callicarpoideae. The first problem awaits a better supported and widely sampled estimate of the phylogenetic tree of both Lamiaceae and close relatives in Lamiales. The second problem awaits a better understanding of relationships in Callicarpoideae, of which only 19/170 species have ever been sampled in a single phylogenetic tree and in which relationships are largely unresolved (Bramley, 2009; Li *et al.*, 2016). However, most species of Callicarpoideae occur in South-East Asia.

Although the crown ancestral range of Lamiaceae remains murky, South-East Asia has nevertheless clearly been important in the biogeographical history of the family. South-East Asia is reconstructed as the most likely area of origin for nearly all super-subfamilial nodes and the crowns of all subfamilies when multiple placeholders exist for those subfamilies (Fig. 2). The somewhat ambiguous reconstruction of the ancestral range of Symphorematoideae + Viticoideae is probably an artefact stemming from uncertainty given that we coded nearly cosmopolitan Viticoideae from a single tip. As with Callicarpoideae, Viticoideae are in critical need of a better sampled and supported phylogenetic tree incorporating a more representative portion of the species diversity.

One of the more intriguing biogeographical patterns in Lamiaceae is found in Premnoideae, in which the Central/South American genus *Cornutia* L. (*c.* 12 species) is sister to the rest of the subfamily, all restricted to the Old World, and almost exclusively in the tropics. Inferring the ancestral range of crown of Premnoideae is inherently difficult, given the unusual distribution and an old crown age (48.6 Myr; Table 1). Our reconstruction of the crown of the subfamily is ambiguous, although possibly of South-East Asian origin with subsequent dispersal to South America (Fig. 2), and a reliable understanding of the origin of *Cornutia* awaits, at the very least, a better taxonomic sampling of *Gmelina* and *Premna* L. However, arrival in South America via Antarctica remains a likely and testable hypothesis.

All previous biogeographical analyses of Lamiaceae above the generic level have focused on Lamioideae and Nepetoideae, and particularly on the latter. Despite using different biogeographical models and/or programs, our results largely corroborate previous findings. Using plastid data, Roy & Lindqvist (2015) reconstructed Lamioideae as South-East Asian in origin, as we suggest here, although their nuclear dataset suggested a joint South-East Asian and rather broadly defined 'temperate Asian' origin for the subfamily. Drew & Sytsma (2012) reconstructed the historical biogeography of Mentheae (Nepetoideae) and proposed a Mediterranean origin for most backbone nodes in the tribe. Their Mediterranean region largely corresponds to what we have delimited in South-West Asia.

Again, our analysis is largely concurrent in finding a strong South-West Asian signal along backbone Mentheae but is less certain of the ancestral range of the ancestor that marks the divergence of the nearly cosmopolitan subtribe Lycopinae (*Lycopus* L.). One area in which we disagree with previous reconstructions is that Li *et al.* (2017) reconstructed the historical biogeography of Elsholtzieae (Nepetoideae) as having a North-West Asian origin of the tribe, while we suggest a more south-eastern or widespread Asian origin for the tribe.

The movement of Mentheae (Nepetoideae) and Lamioideae from South-East Asia to South-West Asia during Eocene cooling appears likely to be correlated with a shift in bioclimatic niche. Future studies should examine if this is the case and, given that these two clades are particularly species rich (Harley *et al.*, 2004; Scheen *et al.*, 2010; Bendiksby *et al.*, 2011; Drew & Sytsma, 2012), if any bioclimatic niche shifts are associated with increased rates of speciation.

TRAIT EVOLUTION AND THE UNEXPECTED ORIGIN OF NUTLETS

Our analyses clarify the evolution of key traits in Lamiaceae by examining the evolution of fruit and habit using a broad sampling of the family and re-assess the evolution of fertile stamen number with much denser sampling outside of Nepetoideae compared to previous work (Drew & Sytsma, 2012). We reconstruct the ancestor of all Lamiaceae as being woody with four fertile stamens and having nutlet fruits (Figs 3–5). A woody habit for the ancestor of the family is not unexpected given the large number of woody clades around the deepest nodes. Although woodiness is reconstructed with a probability of nearly 1.0 along the backbone of the clade formed by the MRCA of Tectonoideae and Lamioideae, better taxonomic sampling in this clade would clarify some of the more ambiguously reconstructed nodes (Fig. 3). In particular, a comprehensively sampled phylogenetic tree would probably result in a higher probability of woodiness at the crown of Ajugoideae Clade I (*Rotheca* Raf. as the placeholder) given that relationships in this clade suggest that the herbaceous habit is derived (Xiang *et al.*, 2018). The same is probably true for crown Scutellarioideae + Cymarioideae/Lamioideae given relationships in Scutellarioideae (Li *et al.*, 2012; Zhang *et al.*, 2020). Likewise, the origin of herbaceous habit in Lamioideae is ambiguous but probably originated at least after the divergence of tribe Pogostemoneae. However, the ancestral habit of Pogostemoneae is not intuitive based on a consideration of the phylogenetic tree of the clade, and a clarification of the origin of woodiness in Lamioideae awaits future study. It is curious that the two largest subfamilies (Nepetoideae and Lamioideae), which combined account for roughly two-thirds of total species diversity in mints are probably herbaceous in origin. Future study should explicitly test for the effect of the origin of herbaceous habit on increased rates of diversification (Soltis *et al.*, 2013; Boucher *et al.*, 2017), perhaps testing the effect of presence in South-West Asia on the evolution of herbaceous habit (see above). Additionally, although we do not test this association here, many of the woody mint lineages are also restricted to tropical biomes in Australasia and South-East Asia. The interplay between, woody habit, occurrence in the tropics and rates of diversification should be finely teased apart in the future.

Unexpectedly, we reconstruct nutlets as the ancestral fruit type of all Lamiaceae (Fig. 4). This fruit type has generally been thought of as an apomorphy for Lamiaceae in the traditional, narrow sense (Cantino, 1992a; Ryding, 1995), but appears to be ancestral for the entire family. Scoring fruit type in Lamiaceae is complicated both by ovule number and abortion (Cantino, 1992a; Harley *et al.*, 2004), and by the presence of drupaceous fruit in at least some members of all subfamilies excepting Lamioideae and Nepetoideae, and nutlets or nutlet-like fruits appearing in at least some members of many of these other subfamilies. The ancestral nutlet pattern may be driven in part by our relatively sparse sampling of Prostantheroideae and Viticoideae, which are polymorphic for nutlet and non-nutlet fruits. Furthermore, Li *et al.* (2016) suggested that nutlets evolved independently in Prostantheroideae, although they hypothesized that it was shared by the common ancestor of the entire subfamily. In Viticoideae, only Neotropical *Pseudocarpidium* Millsp. has nutlets. *Pseudocarpidium* was not included in Bramley, Forest & de Kok (2009) or Zhao *et al.* (2021), and although Li *et al.* (2016) found it to be embedded in Viticoideae in some analyses, the genus formed a polytomy with the rest of the subfamily in other analyses. Therefore, given our refined phylogenetic hypothesis it appears that the nutlet/schizocarp fruit of Prostantheroideae is inherited from the common ancestor of mints, rather than derived independently. Based on our analyses, drupaceous fruits have arisen independently in Callicarpoideae, Prostantheroideae tribe Chloantheae, Viticoideae, Symphorematoideae and Tectonoideae, among other clades. More comprehensive sampling of these clades and better support for the placement of *Pseudocarpidium* will resolve the exact number of times in which drupes have evolved from nutlets. Fruits have evolved in a similar way in the closely related Verbenaceae, in which drupaceous fruits have originated multiple times from ancestors with schizocarps (O'Leary *et al.*, 2012). In Boraginaceae (Boraginales) nutlets have evolved twice from capsular fruited ancestors (Weigend *et al.*, 2014). Although not tested here, such a transition from capsules to nutlets is also probable in Lamiaceae given the predominance of capsule fruits in the clade sister to Lamiaceae, as well as in other closely related Lamiales.

Last, we have clarified the evolution of fertile stamen number in Lamiaceae. We strongly, and largely unsurprisingly, reconstruct the ancestor of Lamiaceae as having four fertile stamens (Fig. 5). Stamen proliferation above four appears to have little, if any, effect on the macroevolutionary dynamics of Lamiaceae except for Tectonoideae (androecium penta- or hexamerous), and it is unclear to what extent stamen numbers in excess of five are found in any interspecific common ancestors in Lamiaceae, apart from the *c.* 17 species in the *Sphenodesme* Jack + *Symphorema* clade in Symphorematoideae (Li *et al.*, 2016), or if they have originated in individual species. On the other hand, reduction in fertile stamen number has had important evolutionary consequences, at least in the particularly well-documented case of its impact on the diversification of the *c*. 1000 *Salvia* spp. (Claßen-Bockhoff *et al*., 2003, 2004; Walker & Sytsma, 2007; Drew *et al.*, 2017b; Kriebel *et al.*, 2020), in which the reduction in fertile stamen number may be a preadaptation to the elongation of the filament connective and therefore the evolution of the staminal lever mechanism. However, it is unclear if a shift in stamen number has led to an increased speciation rate in *Salvia* as a whole, or if other processes have led to increased rates of speciation in subclades of *Salvia* (Kriebel *et al*., 2019, 2020). At the least, reduction in stamen number may lead to increased floral specialization and therefore increased opportunities for speciation (although perhaps not an abrupt shift in rates). This hypothesis remains to be tested and, whereas the most species-rich clade in Lamiaceae consists of species with two fertile stamens (*Salvia*, New World Menthinae), most other lineages of Lamiaceae with two fertile stamens are relatively depauperate.

CONCLUSIONS

This study provides the first comprehensive and time-calibrated phylogenetic hypothesis for Lamiaceae, which samples all currently recognized subfamilies and samples most key nodes in these subfamilies. Our results reinforce an emerging consensus of plastid relationships with the family, and there is great promise that whole plastome sequences, with non-coding regions included, will resolve and provide support for the most recalcitrant relationships. At the same time, large nuclear datasets are increasingly being employed to elucidate broad scale relationships in the family. Although species trees derived from organellar and nuclear genomic compartments are concordant regarding many relationships, several major relationships are still unresolved and possibly discordant. It remains to be seen if relationships inferred from these genomes remain discordant as additional data are accumulated and unresolved relationships in both datasets are clarified and, if so, what processes have led to such discordance. Given the strong signal for identical ancestral ranges and character states in most inter-subfamilial nodes, it seems unlikely that topological differences between the plastid and nuclear genomes (as presently understood, with an unknown placement of Cymarioideae based on nuclear data) would have an impact on our reconstructions.

From a woody, four-stamened, nutlet-bearing ancestor that arose in the Late Cretaceous, possibly in what is now South-East Asia, mints have diversified into one of the largest plant families, possibly aided by climatic cooling and the invasion of xeric habitats since the Eocene. Future work should target taxonomic and genetic sampling in Callicarpoideae and Viticoideae to continue to clarify the historical biogeography of the family and test hypotheses regarding the role of reduction of stamen number and evolution of herbaceous habit on diversification rates in the family.

ACKNOWLEDGEMENTS

The authors thank Holly Forbes from the UC-Berkeley Botanical Gardens and Cindy Newlander and Mike Kintgen from the Denver Botanical Gardens for granting permission and helping us to collect specimens. We also thank Jay Walker and N. Ivalú Cacho for help collecting specimens. This paper was funded in part by the University of Wisconsin Botany Department Hofmeister Endowment, an NSF-DOB grant to Sytsma (DOB-1046355), an NSF-DEB collaborative grant to Sytsma and Drew (DEB-1655606 and DEB-1655611) and a Yunnan Fundamental Research Project Grant to Xiang (2019FI009).

DATA AVAILABILITY

All alignments, trees, input files, data matrices, and R scripts needed to reproduce the analyses have been deposited in Dryad (https://doi.org/10.5061/dryad.zw3r2288z).

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

REFERENCES

Ayala FJ. 1986. On the virtues and pitfalls of the molecular evolutionary clock. *Journal of Heredity* **77:** 226–235.

Ayala FJ. 1999. Molecular clock mirages. *Bioessays* **21:** 71–75.

Barba-Montoya J, dos Reis M, Schneider H, Donoghue PC, Yang Z. 2018. Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous terrestrial revolution. *New Phytologist* **218:** 819–834.

Beaulieu JM, Oliver JC, O'Meara B. 2017. *corHMM: analysis of binary character evolution. R package v. 1.22*. Available at: https://CRAN.R-project.org/package=corHMM.

Beaulieu JM, O'Meara BC, Crane P, Donoghue MJ. 2015. Heterogeneous rates of molecular evolution and diversification could explain the Triassic age estimate for angiosperms. *Systematic Biology* **64:** 869–878.

Bendiksby M, Thorbek L, Scheen AC, Lindqvist C, Ryding O. 2011. An updated phylogeny and classification of Lamiaceae subfamily Lamioideae. *Taxon* **60:** 471–484.

Berger BA, **Kriebel R**, **Spalink, D Sytsma, KJ. 2016.** Divergence times, historical biogeography, and shifts in speciation rates of Myrtales. *Molecular Phylogenetics and Evolution* **95:** 116–136.

Boltenhagen E. 1976a. Pollens et spores Sénoniennes du Gabon. *Cahiers de Micropaléontologie* **3:** 1–21.

Boltenhagen E. 1976b. La microflore Sénonienne du Gabon. *Revue de Micropaléontologie* **18:** 191–199.

Boucher FC, **Verboom GA**, **Musker S**, **Ellis AG. 2017.** Plant size: a key determinant of diversification? *New Phytologist* **216:** 24–31.

Bramley GLC. 2009. The genus *Callicarpa* (Lamiaceae) on Borneo. *Botanical Journal of the Linnean Society* **159:** 416–455.

Bramley GLC, **Forest F**, **de Kok RPJ. 2009.** Troublesome tropical mints: re-examining generic limits of *Vitex* and relations (Lamiaceae) in South East Asia. *Taxon* **58:** 500–510.

Bräuchler C, **Meimberg H**, **Heubl, G. 2010.** Molecular phylogeny of Menthinae (Lamiaceae, Nepetoideae, Mentheae)—taxonomy, biogeography and conflicts. *Molecular Phylogenetics and Evolution* **55:** 501–523.

Bremer K, **Friis EM**, **Bremer B. 2004.** Molecular phylogenetic dating of asterid flowering plants shows early Cretaceous diversification. *Systematic Biology* **53:** 496–505.

Brown RW. 1962. Paleocene flora of the Rocky Mountains and Great Plains. *U. S. Geological Survey Professional Paper* **375:** 1–119.

Buerki S, **Forest F**, **Alvarez N**, **Nylander JAA**, **Arrigo N**, **Sanmartín I. 2011.** An evaluation of new parsimony-based versus parametric inference methods in biogeography: a case study using the globally distributed plant family Sapindaceae. *Journal of Biogeography* **38:** 531–550.

Bunsawat J, **Elliott NE**, **Hertweck KL**, **Sproles E**, **Alice LA. 2004.** Phylogenetics of *Mentha* (Lamiaceae): evidence from chloroplast DNA sequences. *Systematic Botany* **29:** 959–964.

Call VB, **Dilcher DL. 1992.** Investigations of angiosperms from the Eocene of southeastern North America: samaras of *Fraxinus wilcoxiana* Berry. *Review of Palaeobotany and Palynology* **74:** 249–266.

Cantino P, **Harley RM**, **Wagstaff SJ. 1992.** Genera of Labiatae: status and classification. In: Harley RM, Reynolds T, eds. *Advances in Labiatae science.* Kew: Royal Botanic Gardens, 511–522.

Cantino PD. 1992a. Evidence for a polyphyletic origin of the Labiatae. *Annals of the Missouri Botanical Garden* **79:** 361–379.

Cantino PD. 1992b. Toward a phylogenetic classification of the Labiatae. In: Harley RM, Reynolds T, eds. *Advances in Labiatae science.* Kew: Royal Botanic Gardens, 27–37.

Cardinal-McTeague WM, **Sytsma KJ**, **Hall JC. 2016.** Biogeography and diversification of Brassicales: a 103 million year tale. *Molecular Phylogenetics and Evolution* **99:** 204–224.

Cavagnetto C, **Anadón P. 1996.** Preliminary palynological data on floristic and climatic changes during the Middle Eocene-Early Oligocene of the eastern Ebro Basin, northeast Spain. *Review of Palaeobotany and Palynology* **92:** 281–305.

Celep F, **Atalay Z**, **Dikmen F**, **Doğan M**, **Sytsma KJ**, **Claßen-Bockhoff R. 2020.** Pollination ecology, specialization, and genetic isolation in sympatric bee pollinated *Salvia* (Lamiaceae). *International Journal of Plant Sciences* **181:** 800–811.

Claßen-Bockhoff R, **Speck T**, **Tweraser E**, **Wester P**, **Thimm S**, **Reith M. 2004.** The staminal lever mechanism in *Salvia* L. (Lamiaceae): a key innovation for adaptive radiation? *Organisms Diversity & Evolution* **4:** 189–205.

Claßen-Bockhoff R, **Wester P**, **Tweraser E. 2003.** The staminal lever mechanism in *Salvia* L. (Lamiaceae)—a review. *Plant Biology* **5:** 33–41.

Conn BJ, **Streiber N**, **Brown EA**, **Henwood MJ**, **Olmstead RG. 2009.** Infrageneric phylogeny of Chloantheae (Lamiaceae) based on chloroplast *ndhF* and nuclear ITS sequence data. *Australian Systematic Botany* **22:** 243–256.

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

Deng T, **Nie ZL**, **Drew BT**, **Volis S**, **Kim C**, **Xiang CL**, **Zhang JW**, **Wang YH**, **Sun H. 2015.** Does the Arcto–Tertiary biogeographic hypothesis explain the disjunct distribution of Northern Hemisphere herbaceous plants? The case of *Meehania* (Lamiaceae). *PLoS One* **10:** e0117171.

Douzery EJ, **Snell EA**, **Bapteste E**, **Delsuc F**, **Philippe H. 2004.** The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? *Proceedings of the National Academy of Sciences, USA* **101:** 15386–15391.

Drew BT, **González-Gallegos JG**, **Xiang C-L**, **Kriebel R**, **Drummond CP**, **Walker JB**, **Sytsma KJ. 2017b.** Salvia united: the greatest good for the greatest number. *Taxon* **66:** 133–145.

Drew BT, **Liu S**, **Bonifacino JM**, **Sytsma KJ. 2017a.** Amphitropical disjunctions in New World Menthinae: three Pliocene dispersals to South America following late Miocene dispersal to North America from the Old World. *American Journal of Botany* **104:** 1695–1707.

Drew BT, **Sytsma KJ. 2011.** Testing the monophyly and placement of *Lepechinia* in the tribe Mentheae (Lamiaceae). *Systematic Botany* **36:** 1038–1049.

Drew BT, **Sytsma KJ. 2012.** Phylogenetics, biogeography, and staminal evolution in the tribe Mentheae (Lamiaceae). *American Journal of Botany* **99:** 933–953.

Drew BT, **Sytsma KJ. 2013.** The South American radiation of *Lepechinia* (Lamiaceae): phylogenetics, divergence times and evolution of dioecy. *Botanical Journal of the Linnean Society* **171:** 171–190.

Drummond AJ, **Ho SY**, **Phillips MJ**, **Rambaut A. 2006.** Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4:** e88.

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

Godden GT, **Kinser TJ**, **Soltis PS**, **Soltis DE. 2019.** Phylotranscriptomic analyses reveal asymmetrical gene duplication dynamics and signatures of ancient polyploidy in mints. *Genome Biology and Evolution* **11:** 3393–3408.

Graham A. 2011. The age and diversification of terrestrial New World ecosystems through Cretaceous and Cenozoic time. *American Journal of Botany* **98:** 336–351.

Graur D, **Martin W. 2004.** Reading the entrails of chickens: molecular timescales of evolution and the illusion of precision. *Trends in Genetics* **20:** 80–86.

Harley RM, **Atkins S**, **Budantsev AL**, **Cantino PD**, **Conn BJ**, **Grayer R**, **Harley MM**, **de Kok RPJ**, **Krestovskaya TV**, **Morales R**, **Paton AJ**, **Ryding O**, **Upson T. 2004.** Labiatae. In: Kadereit JW, Jeffrey C, eds. *The families and genera of vascular plants, Vol. 7*. Berlin, Heidelberg: Springer, 167–275.

Hedge C. 1992. A global survey of the biogeography of the Labiatae. In: Harley RM, Reynolds T, eds. *Advances in Labiatae science*. Kew: Royal Botanic Gardens, 7–17.

Horiuchi J, **Manchester S. 2011.** The Paleocene Yachinaka flora of the Minato Formation from Kuji, northeastern Honshu, Japan. Abstract. Botany 2011 Conference; St. Louis, Missouri.

Huelsenbeck JP, **Larget B**, **Swofford D. 2000.** A compound Poisson process for relaxing the molecular clock. *Genetics* **154:** 1879–1892.

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

Kar RK. 1996. On the Indian origin of *Ocimum* (Lamiaceae): a palynological approach. *Palaeobotanist* **43:** 45–50.

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

de Kok RPJ, **Go R**, **Latiff A. 2009.** The genus *Teijsmanniodendron* Koord. (Lamiaceae). *Kew Bulletin* **64:** 587–625.

Kriebel R, **Drew B**, **González-Gallegos JG**, **Celep F**, **Heeg L**, **Mahdjoub MM**, **Sytsma KJ. 2020.** Pollinator shifts, contingent evolution, and evolutionary constraint drive floral disparity in *Salvia* (Lamiaceae): evidence from morphometrics and phylogenetic comparative methods. *Evolution* **74:** 1335–1355.

Kriebel R, **Drew BT**, **Drummond CP**, **González-Gallegos JG**, **Celep F**, **Mahdjoub MM**, **Rose JP**, **Xiang CL**, **Hu GX**, **Walker JB**, **Lemmon EM**, **Lemmon AR**, **Sytsma KJ. 2019.** Tracking temporal shifts in area, biomes, and pollinators in the radiation of *Salvia* (sages) across continents: leveraging anchored hybrid enrichment and targeted sequence data. *American Journal of Botany* **106:** 573–597.

Li B, **Cantino PD**, **Olmstead RG**, **Bramley GL**, **Xiang CL**, **Ma ZH**, **Tan YH**, **Zhang DX. 2016.** A large-scale chloroplast phylogeny of the Lamiaceae sheds new light on its subfamilial classification. *Scientific Reports* **6:** 1–18.

Li B, **Xu W**, **Tu T**, **Wang Z**, **Olmstead RG**, **Peng H**, **Francisco-Ortega J**, **Cantino PD**, **Zhang D. 2012.** Phylogenetic position of *Wenchengia* (Lamiaceae): a taxonomically enigmatic and critically endangered genus. *Taxon* **61:** 392–401.

Li BO, **Olmstead RG. 2017.** Two new subfamilies in Lamiaceae. *Phytotaxa* **313:** 222–226.

Li HT, **Yi TS**, **Gao LM**, **Ma PF**, **Zhang T**, **Yang JB**, **Gitzendanner MA**, **Fritsch PW**, **Cai J**, **Luo Y**, **Wang H. 2019.** Origin of angiosperms and the puzzle of the Jurassic gap. *Nature Plants* **5:** 461–470.

Li P, **Qi ZC**, **Liu LX**, **Ohi-Toma T**, **Lee J**, **Hsieh TH**, **Fu CX**, **Cameron KM**, **Qiu YX. 2017.** Molecular phylogenetics and biogeography of the mint tribe Elsholtzieae (Nepetoideae, Lamiaceae), with an emphasis on its diversification in East Asia. *Scientific Reports* **7:** 2057.

Liu B, **Le CT**, **Barrett RL**, **Nickrent DL**, **Chen Z**, **Lu L**, **Vidal-Russell R. 2018.** Historical biogeography of Loranthaceae (Santalales): diversification agrees with emergence of tropical forests and radiation of songbirds. *Molecular Phylogenetics and Evolution* **124:** 199–212.

Liu B, **Tan YH**, **Liu S**, **Olmstead RG**, **Min DZ**, **Chen ZD**, **Joshee N**, **Vaidya BN**, **Chung RC**, **Li B. 2020.** Phylogenetic relationships of *Cyrtandromoea* and *Wightia* revisited: a new tribe in Phrymaceae and a new family in Lamiales. *Journal of Systematics and Evolution* **58:** 1–17.

Maddison WP, **Maddison DR. 2019.** *Mesquite: a modular system for evolutionary analysis. Version 3.61*. Available at: http://www.mesquiteproject.org.

Magallón S, **Castillo A. 2009.** Angiosperm diversification through time. *American Journal of Botany* **96:** 349–365.

Magallón S, **Gómez-Acevedo S**, **Sánchez-Reyes LL**, **Hernández-Hernández T. 2015.** A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist* **207:** 437–453.

Mai DH. 2001. Die mittelmiozaenen und obermiozaenen Floren aus der Meuroer und Raunoer Folge in der Lausitz: Teil II: Dicotyledonen. *Palaeontographica Abteilung B Palaeophytologie* **257:** 35–174.

Manchester SR. 2014. Revisions to Roland Brown's North American Paleocene flora. *Acta Musei Nationalis Pragae* **70:** 153–210.

Manchester SR, **Grímsson F**, **Zetter R. 2015.** Assessing the fossil record of asterids in the context of our current phylogenetic framework. *Annals of the Missouri Botanical Garden* **100:** 329–363.

Mandel JR, **Dikow RB**, **Siniscalchi CM**, **Thapa R**, **Watson LE**, **Funk VA. 2019.** A fully resolved backbone phylogeny reveals numerous dispersals and explosive diversifications throughout the history of Asteraceae. *Proceedings of the National Academy of Sciences, USA* **116:** 14083–14088.

Martínez-Millán M. 2010. Fossil record and age of the Asteridae. *The Botanical Review* **76:** 83–135.

Mathiesen C, **Scheen AC**, **Lindqvist C. 2011.** Phylogeny and biogeography of the lamioid genus *Phlomis* (Lamiaceae). *Kew Bulletin* **66:** 83–99.

Matzke NJ. 2012. Founder-event speciation in BioGeoBEARS package dramatically improves likelihoods and alters parameter inference in dispersal-extinction-cladogenesis (DEC) analyses. *Frontiers in Biogeography* **4:** 210.

Matzke NJ. 2013. *BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R Scripts*. Vienna:

CRAN: The Comprehensive R Archive Network. Available at: http://cran.r-project.org/package=BioGeoBEARS.

Mint Evolutionary Genomics Consortium (MEGC). 2018. Phylogenomic mining of the mints reveals multiple mechanisms contributing to the evolution of chemical diversity in Lamiaceae. *Molecular Plant* **11:** 1084–1096.

Muller J. 1981. Fossil pollen records of extant angiosperms. *Botanical Review* **47:** 1–142.

O'Leary N, **Calviño CI**, **Martínez S**, **Lu-Irving P**, **Olmstead RG**, **Múlgura ME. 2012.** Evolution of morphological traits in Verbenaceae. *American Journal of Botany* **99:** 1778–1792.

Panero JL, **Crozier BS. 2016.** Macroevolutionary dynamics in the early diversification of Asteraceae. *Molecular Phylogenetics and Evolution* **99:** 116–132.

Paradis E, **Schliep K. 2018.** ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35:** 526–528.

Pastore JFB, **Harley RM**, **Forest F**, **Paton A**, **van den Berg C. 2011.** Phylogeny of the subtribe Hyptidinae (Lamiaceae tribe Ocimeae) as inferred from nuclear and plastid DNA. *Taxon* **60:** 1317–1329.

Paton A, **Mwanyambo M**, **Culham A. 2018.** Phylogenetic study of *Plectranthus*, *Coleus* and allies (Lamiaceae): taxonomy, distribution and medicinal use. *Botanical Journal of the Linnean Society* **188:** 355–376.

Paton AJ, **Springate D**, **Suddee S**, **Otieno D**, **Grayer RJ**, **Harley MM**, **Willis F**, **Simmonds MSJ**, **Powell MP**, **Savolainen V. 2004.** Phylogeny and evolution of basils and allies (Ocimeae, Labiatae) based on three plastid DNA regions. *Molecular Phylogenetics and Evolution* **31:** 277–299.

Rambaut A, **Drummond AJ**, **Xie D**, **Baele G**, **Suchard MA. 2018.** Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67:** 901.

Ree RH, **Smith SA. 2008.** Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* **57:** 4–14.

Refulio-Rodriguez NF, **Olmstead RG. 2014.** Phylogeny of Lamiidae. *American Journal of Botany* **101:** 287–299.

Reid EM, **Chandler MEJ. 1926.** *Catalogue of Cainozoic plants in the Department of Geology. Vol. 1. The Brembridge Flora*. London: British Museum (Natural History).

Rose JP, **Kleist TJ**, **Löfstrand SD**, **Drew BT**, **Schönenberger J**, **Sytsma KJ. 2018.** Phylogeny, historical biogeography, and diversification of angiosperm order Ericales suggest ancient Neotropical and East Asian connections. *Molecular Phylogenetics and Evolution* **122:** 59–79.

Rose JP, **Kriebel R**, **Kahan L**, **DiNicola A**, **González-Gallegos JG**, **Celep F**, **Lemmon EM**, **Lemmon AR**, **Sytsma KJ**, **Drew BT. 2021.** Sage insights into the phylogeny of *Salvia*: dealing with sources of discordance within and across genomes. *Frontiers in Plant Sciences* **12:** 767478.

Roy T, **Catlin NS**, **Garner DM**, **Cantino PD**, **Scheen AC**, **Lindqvist C. 2016.** Evolutionary relationships within the lamioid tribe Synandreae (Lamiaceae) based on multiple low-copy nuclear loci. *PeerJ* **4:** e2220.

Roy T, **Chang TH**, **Lan T**, **Lindqvist C. 2013.** Phylogeny and biogeography of New World Stachydeae (Lamiaceae) with emphasis on the origin and diversification of Hawaiian and South American taxa. *Molecular Phylogenetics and Evolution* **69:** 218–238.

Roy T, **Lindqvist C. 2015.** New insights into evolutionary relationships within the subfamily Lamioideae (Lamiaceae) based on pentatricopeptide repeat (PPR) nuclear DNA sequences. *American Journal of Botany* **102:** 1721–1735.

Ryding O. 1995. Pericarp structure and phylogeny of the Lamiaceae-Verbenaceae-complex. *Plant Systematics and Evolution* **198:** 101–141.

Salmaki Y, **Kattari S**, **Heubl G**, **Bräuchler C. 2016.** Phylogeny of non-monophyletic *Teucrium* (Lamiaceae: Ajugoideae): implications for character evolution and taxonomy. *Taxon* **65:** 805–822.

Salmaki Y, **Zarre S**, **Ryding O**, **Lindqvist C**, **Bräuchler C**, **Heubl G**, **Barber J**, **Bendiksby M. 2013.** Molecular phylogeny of tribe Stachydeae (Lamiaceae subfamily Lamioideae). *Molecular Phylogenetics and Evolution* **69:** 535–551.

Sarkar S, **Prasad V. 2002.** *Ocimum* pollen grains from the Subathu Formation (Late Ypresian) of Shimla Hills, Himachal Pradesh, India. *Palaeobotanist* **51:** 165–167.

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

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

Scheen AC, **Albert VA. 2009.** Molecular phylogenetics of the *Leucas* group (Lamioideae; Lamiaceae). *Systematic Botany* **34:** 173–181.

Scheen AC, **Bendiksby M**, **Ryding O**, **Mathiesen C**, **Albert VA**, **Lindqvist C. 2010.** Molecular phylogenetics, character evolution and suprageneric classification of Lamioideae (Lamiaceae). *Annals of the Missouri Botanical Garden* **97:** 191–219.

Soltis DE, **Mort ME**, **Latvis M**, **Mavrodiev EV**, **O'Meara BC**, **Soltis PS**, **Burleigh JG**, **Rubio de Casas R. 2013.** Phylogenetic relationships and character evolution analysis of Saxifragales using a supermatrix approach. *American Journal of Botany* **100:** 916–929.

Spalink D, **Drew BT**, **Pace MC**, **Zaborsky JG**, **Starr JR**, **Cameron KM**, **Givnish TJ**, **Sytsma KJ. 2016.** Biogeography of the cosmopolitan sedges (Cyperaceae) and the area-richness correlation in plants. *Journal of Biogeography* **43:** 1893–1904.

Srivastava R, **Wheeler EA**, **Manchester SR**, **Baas P. 2015.** Wood of Oleaceae from the latest Cretaceous of India–the earliest olive branch? *IAWA Journal* **36:** 443–451.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 312–313.

Steane DA, **de Kok RPJ**, **Olmstead RG. 2004.** Phylogenetic relationships between *Clerodendrum* (Lamiaceae) and other ajugoid genera inferred from nuclear and chloroplast DNA sequence data. *Molecular Phylogenetics and Evolution* **32:** 39–45.

Suchard MA, **Lemey P**, **Baele G**, **Ayres DL**, **Drummond AJ**, **Rambaut A. 2018.** Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* **4:** vey016.

Sytsma KJ, **Spalink D**, **Berger B. 2014.** Calibrated chronograms, fossils, outgroup relationships, and root priors: Re-examining the historical biogeography of Geraniales. *Biological Journal of the Linnean Society* **113:** 29–49.

Thorne JL, **Kishino H**, **Painter IS. 1998.** Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* **15:** 1647–1657.

Tiffney BH. 1985. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phytogeography of the Northern Hemisphere. *Journal of the Arnold Arboretum* **66:** 243–273.

Tiffney BH. 2000. Geographic and climatic influence on the Cretaceous and Tertiary history of Euramerican floristic similarity. *Acta Universitatis Carolinae, Geologica* **44:** 5–16.

Tiffney BH, **Manchester SR. 2001.** The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere Tertiary. *International Journal of Plant Sciences* **162:** S3–17.

Tripp EA, **McDade LA. 2014.** A rich fossil record yields calibrated phylogeny for Acanthaceae (Lamiales) and evidence for marked biases in timing and directionality of intercontinental disjunctions. *Systematic Biology* **63:** 660–684.

Wagstaff SJ, **Hickerson L**, **Spangler R**, **Reeves PA**, **Olmstead RG. 1998.** Phylogeny in Labiatae *s.l.*, inferred from cpDNA sequences. *Plant Systematics and Evolution* **209:** 265–274.

Wagstaff SJ, **Olmstead RG. 1997.** Phylogeny of Labiatae and Verbenaceae inferred from *rbcL* sequences. *Systematic Botany* **22:** 165–179.

Walker JB, **Sytsma KJ. 2007.** Staminal evolution in the genus *Salvia* (Lamiaceae): molecular phylogenetic evidence for multiple origins of the staminal lever. *Annals of Botany* **100:** 375–391.

Weigend M, **Luebert F**, **Gottschling M**, **Couvreur TLP**, **Hilger HH**, **Miller JS. 2014.** From capsules to nutlets—phylogenetic relationships in the Boraginales. *Cladistics* **30:** 508–518.

Wheeler EA, **Srivastava R**, **Manchester SR**, **Baas P. 2017.** Surprisingly modern latest Cretaceous–earliest Paleocene woods of India. *IAWA Journal* **38:** 456–542.

Wikström N, **Kainulainen K**, **Razafimandimbison SG**, **Smedmark JE**, **Bremer B. 2015.** A revised time tree of the asterids: establishing a temporal framework for evolutionary studies of the coffee family (Rubiaceae). *PLoS One* **10:** e0126690.

Woodcock D, **Meyer HW**, **Prado Y**, **Rufasto E**, **Terry DO. 2020.** The Piedra Chamana Fossil Forest: a scientific resource of global importance. *Revista Peruana de Biología* **27:** 517–528.

Xiang CL, **Zhao F**, **Cantino PD**, **Drew BT**, **Li B**, **Liu ED**, **Soltis DE**, **Soltis PS**, **Peng H. 2018.** Molecular systematics of *Caryopteris* (Lamiaceae) and its allies with reference to the molecular phylogeny of subfamily Ajugoideae. *Taxon* **67:** 376–394.

Xu WQ, **Losh J**, **Chen C**, **Li P**, **Wang RH**, **Zhao YP**, **Qiu YX**, **Fu CX. 2019.** Comparative genomics of figworts (*Scrophularia,* Scrophulariaceae), with implications for the evolution of *Scrophularia* and Lamiales. *Journal of Systematics and Evolution* **57:** 55–65.

Yao G, **Drew BT**, **Yi TS**, **Yan HF**, **Yuan YM**, **Ge XJ. 2016.** Phylogenetic relationships, character evolution and biogeographic diversification of *Pogostemon s.l.* (Lamiaceae). *Molecular Phylogenetics and Evolution* **98:** 184–200.

Yu XQ, **Maki M**, **Drew BT**, **Paton AJ**, **Li HW**, **Zhao JL**, **Conran JG**, **Li J. 2014.** Phylogeny and historical biogeography of *Isodon* (Lamiaceae): rapid radiation in south-west China and Miocene overland dispersal into Africa. *Molecular Phylogenetics and Evolution* **77:** 183–194.

Yuan YW, **Mabberley DJ**, **Steane DA**, **Olmstead RG. 2010.** Further disintegration and redefinition of *Clerodendrum* (Lamiaceae): implications for the understanding of the evolution of an intriguing breeding strategy. *Taxon* **59:** 125–133.

Zhai W, **Duan X**, **Zhang R**, **Guo C**, **Li L**, **Xu G**, **Shan H**, **Kong H**, **Ren Y. 2019.** Chloroplast genomic data provide new and robust insights into the phylogeny and evolution of the Ranunculaceae. *Molecular Phylogenetics and Evolution* **135:** 12–21.

Zhang C, **Zhang T**, **Luebert F**, **Xiang Y**, **Huang CH**, **Hu Y**, **Rees M**, **Frohlich MW**, **Qi J**, **Weigend M**, **Ma H. 2020.** Asterid phylogenomics/phylotranscriptomics uncover morphological evolutionary histories and support phylogenetic placement for numerous whole-genome duplications. *Molecular Biology and Evolution* **37:** 3188–3210.

Zhang ML, **Zeng XQ**, **Sanderson SC**, **Byalt VV**, **Sukhorukov AP. 2017.** Insight into Central Asian flora from the Cenozoic Tianshan montane origin and radiation of *Lagochilus* (Lamiaceae). *PLoS One* **12:** e0178389.

Zhao F, **Chen YP**, **Salmaki Y**, **Drew BT**, **Wilson TC**, **Scheen AC**, **Celep F**, **Bräuchler C**, **Bendiksby M**, **Wang Q**, **Min DZ**, **Peng H**, **Olmstead RG**, **Li B**, **Xiang CL. 2021.** An updated tribal phylogeny of Lamiaceae: evidence from complete plastome sequences. *BMC Biology* **19:** 1–27.

Zhao F, **Li B**, **Drew BT**, **Chen YP**, **Wang Q**, **Yu WB**, **Liu ED**, **Salmaki Y**, **Peng H**, **Xiang CL. 2020.** Leveraging plastomes for comparative analysis and phylogenomic inference within Scutellarioideae (Lamiaceae). *PLoS One* **15:** e0232602.

Zizka A, **Silvestro D**, **Andermann T**, **Azevedo J**, **Duarte Ritter C**, **Edler D**, **Farooq H**, **Herdean A**, **Ariza M**, **Scharn R**, **Svantesson S. 2019.** CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution* **10:** 744–751.

Zuckerkandl E, **Pauling L. 1965.** Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ, eds. *Evolving genes and proteins.* Cambridge: Academic Press, 97–166.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Supplementary Methods. Word document detailing the constituent genera in each taxonomic placeholder and justification for these groupings.

Supplementary File S1. Excel workbook showing the taxa used in the Lamiales-wide and Lamiaceae analyses, and their GenBank accession numbers.

Supplementary File S2. Excel workbook showing the coding of taxonomic placeholders for the biogeographical analysis.

Supplementary File S3. Excel workbook showing the coding of taxonomic placeholders for the morphological analyses.