Targeted amplicon sequencing of 40 nuclear genes supports a single introduction and rapid radiation of Hawaiian *Metrosideros* (Myrtaceae)

Julian R. Dupuis1,2,\*, Yohan Pillon3,4, Tomoko Sakishima4,5, Chrissen E.C. Gemmill6, Srikar Chamala7,8, W. Brad Barbazuk7, Scott M. Geib2, Elizabeth A. Stacy4,5,\*

1U.S. Department of Agriculture-Agricultural Research Service, Daniel K. Inouye U.S. Pacific Basin Agricultural Research Center, Hilo, HI 96720, USA

2Department of Plant and Environmental Protection Services, University of Hawaiʻi at Mānoa, Honolulu, HI 96822, USA

3LSTM, IRD, INRA, CIRAD, Montpellier SupAgro, University of Montpellier, Montpellier, France

4Department of Biology and Tropical Conservation Biology and Environmental Science Program, University of Hawaiʻi, Hilo, HI 96720, USA

5School of Life Sciences, University of Nevada, Las Vegas, Nevada 89154, USA

6School of Science, University of Waikato, Hamilton, Aotearoa/New Zealand

7Department of Biology, University of Florida, Gainesville, FL 32611, USA

8Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, FL 32611, USA

\*corresponding authors: JRD: jrdupuis@hawaii.edu, phone: 808-959-4335, fax: 808-959-5470; EAS: elizabeth.stacy@unlv.edu, phone: 702-895-4461, fax: 702-895-3956

Running title: Forty nuclear genes support monophyly of Hawaiian *Metrosideros*

ABSTRACT

Compared to traditionally used plastid or ribosomal markers, nuclear gene markers provide many advantages for molecular systematics of plants, and high-throughput sequencing technologies are making large nuclear datasets available at an unprecedented rate. We used targeted amplicon sequencing of 44 nuclear genes to construct a time-calibrated phylogeny of genus *Metrosideros* (Myrtaceae), evaluate recent systematic revisions, and assess whether phylogenetic signal within the Hawaiian Archipelago is correlated to island biogeography or morphological diversification. We generated a final data set of 40 nuclear genes for 187 specimens, used multiple search heuristics and species-tree analysis to estimate a phylogeny, and incorporated new fossils for the genus to estimate divergence times across the dataset. All analyses supported the monophyly of *Metrosideros*, including *Carpolepis* and *Tepualia*. Hawaiian *Metrosideros* were monophyletic and dated to 3.1 MYA using new fossils for the genus, which is intermediate to previous estimates based on nuclear ribosomal/chloroplast loci and calibrated with island ages. Within the Hawaiian *Metrosideros* clade, we observed short branch lengths and unresolved relationships, and phylogenetic patterns were not concordant with biogeographic hypotheses of island progression, or the delineation of taxa or morphotypes. Average nucleotide diversity was relatively consistent across the Hawaiian Islands with the exception of slightly lower diversity on Kauaʻi.These results provide a data-rich estimate of the timing of a single introduction of *Metrosideros* to Hawai‘i and highlight the need for molecular markers with higher evolutionary rates for resolution of relationships within this recent radiation.

KEY WORDS island biogeography; molecular dating; next-generation sequencing; Pacific Islands; phylogenomics

INTRODUCTION

Plant molecular systematics has historically been dominated by the use of relatively few plastid (e.g., *mat*K, *ndh*F, *rbc*L *trn*H*-psb*A, *trnL-*F) or nuclear ribosomal (e.g., ETS, ITS, 18S, 26S) markers (Hollingsworth et al. 2009; Kress 2017; Olmstead and Palmer 1994; Soltis and Soltis 1998). The utility of universal primers, relatively simple dynamics of inheritance and evolution, and expected orthology have facilitated the wide use of these markers (Kelchner 2000; Olmstead and Palmer 1994; Soltis and Soltis 1998). Nuclear genes, on the other hand, have been less widely adopted in plant molecular systematics, mostly due to the difficulty in isolating and identifying orthologous genes, as well as more complicated evolutionary dynamics (Pillon et al. 2013b; Sang 2002; Small et al. 2004; Zimmer and Wen 2012) and longer coalescence times relative to cytoplasmic genes (Moore 1995; Pillon et al. 2013a). Despite these difficulties, nuclear genes often exhibit faster overall rates of evolution than plastid or ribosomal loci and provide a wealth of unlinked loci for phylogenetic inference (Duarte et al. 2010; Small et al. 2004). When properly characterized, nuclear genes can provide great insight into phylogenetic relationships (Bailey and Doyle 1999; Carstens and Satler 2013; Oh and Potter 2003; Sang 2002; Syring et al. 2005), and provide powerful means to explore introgression, hybridization, and ancient allopolyploidization events (Duarte et al. 2010; Sang 2002; Wendel and Doyle 1998). Additionally, high-throughput sequencing technologies have facilitated monumental leaps in generating nuclear gene data, which is greatly increasing capacity for genomic-scale phylogenetic inference in plants (de Sousa et al. 2014; Duarte et al. 2010; Straub et al. 2012; Uribe-Convers et al. 2016; Weitemier et al. 2014; Zimmer and Wen 2015).

*Metrosideros* Banks ex Gaertn. (Myrtaceae) is a medium-sized genus (ca. 60 species) of woody plants comprising trees, shrubs, vines and mat plants that are distributed predominantly across the islands of the Pacific (Govaerts et al. 2008). Recent work by Pillon et al. (2015) based on ITS and five single-copy nuclear genes expanded *Metrosideros* to include *Carpolepis* J.W. Dawson, endemic to New Caledonia, and the monotypic *Tepualia* Griseb. of Chile and Argentina. This broad circumscription of the genus *Metrosideros* is equivalent to tribe Metrosidereae sensu Wilson (2011), however the basal relationships remain unresolved. *Metrosideros* belongs to the BKMMST clade (Biffin et al. 2010), which includes the tribes Backhousieae, Kanieae, Metrosidereae, Myrteae, Syzygieae, Tristanieae, and the unplaced genus *Cloezia* Brongn. & Gris (for various treatments, see Biffin et al. 2010; Wilson 2011), but the relationships within this group are unresolved. Pillon et al. (2015) concluded that resolving phylogenetic relationships within *Metrosideros* will require a large number of nuclear genes.

*Metrosideros* spp*.* are present on most high islands of the Pacific (except Micronesia), where they are a common component of ridge vegetative communities, and certain early-successional taxa are pioneers on new lava flows on volcanic islands. Of particular interest to biogeographers is determining when and how the genus reached these remote islands, particularly the Hawaiian Islands, which form the most isolated archipelago on Earth. The genus is ubiquitous across the archipelago, where it is one of only four lineages that likely reached the archipelago through air flotation (i.e., wind), the other three being orchids with minute seeds (Price and Wagner 2018). Five endemic species are currently recognized, including *Metrosideros waialealae* Rock with two varieties and the morphologically diverse *Metrosideros polymorpha* Gaudich. with eight varieties (Wagner et al. 2005, but see below).

Using nuclear ribosomal ITS, Wright et al. (2000) (Wright et al. 2001)suggested that two dispersal events to the remote islands of Polynesia occurred from New Zealand during the Pleistocene, possibly associated with altered wind currents during periods of maximum glaciation. Using a minimum mutation rate of 1 bp per 1.5 million years based on the distribution of sequences across islands of known age, the authors estimated colonization of Hawaiʻi at < 2 MYA. However, in their parsimony-based analysis, they included only five species from New Zealand and Kermadec Island, all of which are trees, and few clades were supported. Subsequent attempts to date the arrival of *Metrosideros* to the Hawaiian Islands using ITS and ETS, and 10 noncoding chloroplast DNA (cpDNA) loci, respectively, yielded estimates of 0.5-1.0 MYA based on limited ETS evolution within the archipelago (Wright et al. 2001) and 3.9 (1.4-6.3) MYA based on the ages of volcanic islands (Percy et al. 2008). In both analyses, *Metrosideros collina* (J.R. Forst. & G. Forst.) A. Gray from the Marquesas Islands was nested within the Hawaiian *Metrosideros* clade, but both studies concluded that the Marquesas likely served as a stepping stone to the Hawaiian archipelago in part because of the intermediate location of these islands between New Zealand and the Hawaiian Islands (Percy et al. 2008; Wright et al. 2001). The cpDNA-based phylogeny, which included more intensive sampling within Hawaiʻi, suggested a monophyletic origin for Hawaiian *Metrosideros* (Percy et al. 2008).

Within the Hawaiian Islands, *Metrosideros* has diversified into a large number of vegetatively distinct, predominantly single-island-endemic forms that are nonrandomly distributed across an exceptionally broad range of environments. Collectively, these forms dominate the Hawaiian Islands’ native mesic and wet forests, occurring in large, continuous communities within each island. Molecular-based attempts to uncover relationships among forms within and among islands have yielded some insights into the evolutionary history of the group. Percy et al. (2008) found that both nucleotide and haplotype diversity at cpDNA loci were greatest on the oldest island of Kauaʻi and with a single exception decreased with island age, and the authors suggested that this pattern may be due to initial colonization of Kauaʻi with a temporal gap prior to the colonization of the younger islands following the biogeographical progression rule (Wagner and Funk 1995). The exception to this trend was the intermediately positioned island of Molokaʻi, which showed surprisingly high haplotype diversity for its age, consistent with colonization from both older and younger islands. More recent observations in the field have revealed a greater diversity of morphotypes within Hawaiian *Metrosideros* than previously recorded (e.g. Stacy and Sakishima 2019; Sur et al. 2018). An analysis of nuclear microsatellite (SSR) variation across 1,486 individuals of 23 taxa or morphotypes from across the archipelago plus additional populations from other Pacific islands (Stacy and Sakishima 2019) revealed that populations loosely followed the progression rule with Tahitian populations clustering most strongly with those on Kauaʻi and Oʻahu. The population-genetics approach also revealed significant isolation by distance across the 530-km archipelago and greater private allelic richness on Kauaʻi, consistent with limited gene flow across islands, as well as peak genetic diversity on Oʻahu, concordant with peak morphological diversity on that island. Both the cpDNA- and SSR-based studies found a predominant pattern of clustering of populations within islands and multiple colonizations of each of Oʻahu and Hawaiʻi Island, though Stacy and Sakishima (2019) revealed very weak separation between *Metrosideros* taxa on the younger islands of Molokaʻi and Hawaiʻi and other exceptions to island-based clustering. Collectively, results from both studies may be consistent with colonization of new Hawaiian Islands from older islands by just two hybridizing taxa that are adapted to the conditions on new islands (early- and late-successional *Metrosideros polymorpha* Gaudich. var. *incana* (Levl.) Skottsb. and *Metrosideros polymorpha* var. *glaberrima* (Levl.) St., respectfully), followed by predominantly within-island diversification and restricted gene flow among islands (Stacy and Sakishima 2019). Lastly, whereas cpDNA haplotypes were not aligned with boundaries among taxa or morphotypes or even pubescent versus glabrous forms (Percy et al. 2008), population-level analyses revealed a broad distribution of strengths of isolation among taxa from weak to strong, including weak but significant differentiation between pubescent and glabrous taxa on two islands (DeBoer and Stacy 2013; Stacy et al. 2014; Stacy and Sakishima 2019).

Here, we use targeted amplicon sequencing (Bybee et al. 2011; O'Neill et al. 2013) to generate a dataset of 44 nuclear genes for 187 specimens, focusing on 23 taxa (including unnamed morphotypes (“races”) of Hawaiian *Metrosideros*. We place the Hawaiian taxa in phylogenetic context by including 18 species and varieties from elsewhere (predominantly taxa from Pacific islands) and use this nuclear dataset to construct a fossil-calibrated, dated phylogeny of *Metrosideros* to address two main questions. First, is *Metrosideros* strongly supported as monophyletic? If so, what is the sister taxon to *Metrosideros*? Second, when did *Metrosideros* colonize the Hawaiian Islands and where did this lineage originate? We then explore fine-scale phylogenetic patterns of this genus across the Hawaiian Islands to address the following questions. First, are phylogenetic patterns consistent with the progression rule, including colonization of new islands by just two widespread varieties of *M. polymorpha* followed by within-island diversification and limited repeat- and back-colonizations of islands? Second, does nucleotide diversity vary across islands consistent with observed patterns for either cpDNA loci or SSR loci? And third, does the partitioning of nucleotide diversity align with current taxonomic boundaries (including races) or with a split between glabrous and pubescent taxa?

MATERIALS AND METHODS

***Specimen sampling***

We selected 187 specimens, including 163 *Metrosideros,* 123 of which represented all five Hawaiian species, including the two varieties of *M. waialealae* and 18 varieties or morphotypes (“races”) of *M. polymorpha*, as well as species previously included in *Carpolepis* (*Metrosideros laurifolia* Brogn. & Gris. from New Caledonia) and *Tepualia* (*Metrosideros stipularis* (Hook. & Arn.) Hook.f. from South America), along with *Metrosideros angustifolia* (L.) Sm. from South Africa. The remaining 24 specimens comprised other Myrtaceae to provide phylogenetic context to *Metrosideros* and serve as outgroups, including *Backhousia* Hook. & Harv., *Cloezia* Brogn. & Gris, *Ecualyptus* L’Hér., *Eugenia* L., *Melaleuca* L., *Pimenta* Lindl., *Psidium* L., *Purpureostemon* Gugerli, *Rhodomyrtus* (DC.) Rchb., *Syzygium* R.Br. ex Gaertn., *Tristaniopsis* Brongn. & Griseb., *Xanthomyrtus* Diels, and *Xanthostemon* F.Muell.. Summarized specimen information is provided in Table 1, and detailed information is provided in Online Resource 1.

***Sequence data***

We selected 44 low-copy genes for targeted amplicon sequencing (Online Resource 1), expanding on the datasets used by Pillon et al. (2015). Primers were designed as in Pillon et al. (2014), and amplification and sequencing followed Pillon et al. (2015). Briefly, these markers were designed from a *Metrosideros* transcriptome that was mined for phylogenetically-informative low-copy genes identified by Duarte et al. (2010). Primer sets for these markers were validated on a small test set of specimens of Hawaiian *Metrosideros*. The amplicons were typically 200-500 bp long and included a single exon or two exons separated by one intron. Forty-eight of the best-performing markers (i.e., most consistent amplification) were selected for amplification using the Fluidigm Access Array (Fluidigm Corporation, South San Francisco, CA) with a 48 x 48 Integrated Fluidic Circuit (IFC) cartridge, following manufacturer’s recommendations. Libraries were generated for four IFCs, and each was sequenced on approximately a quarter lane (although not necessarily at the same time) of an Illumina MiSeq v3, 2x300 bp sequencing run (Illumina, San Diego, CA).

***Data processing and filtering***

We used default parameters and options for all data processing and analysis steps unless otherwise mentioned. Raw sequencing files were de-multiplexed by specimen and converted to FASTQ using the FASTQ generation analysis v1.0.0 in BaseSpace (Illumina, San Diego, CA). We removed Illumina adapters using Cutadapt (Martin 2011), merged paired-end reads with FLASh (Magoc and Salzberg 2011), and used Cutadapt again to demultiplex merged reads by locus using the primer sequences with a minimum overlap (-O) of 10 bp. On a per-gene basis, we calculated read depth of each unique sequence per individual and removed any sequence with <20 reads as well as any containing Ns. From these filtered sequences, we created a concatenated multi-FASTA file for each gene, including potentially multiple sequences per individual, and aligned these sequences using the L-INS-I algorithm in MAFFT 7.182 (Katoh and Standley 2013). We then assessed the quality of these alignments by eye using AliView v1.18 (Larsson 2014) and manually removed sequences that were obviously off-target (e.g. completely different sequence composition or long strings of missing data in a single sequence that caused blocks of gaps in the alignment). While obviously subjective, this manual filtering effectively removed outlier sequences that were likely caused by PCR error or off-target amplification. Additionally, we evaluated the unique sequences present per specimen and their corresponding read depth and identified four genes that were likely not single-copy (highly variable sequences with high depth per individual that were found in many individuals); these were removed from further analyses. We then generated a consensus sequence per individual per gene with heterozygous positions represented by IUPAC ambiguity codes (i.e. non-phased sequence data). From these 40 filtered gene alignments, we created final individual gene alignments and a concatenated dataset and removed individuals with poor coverage from each dataset. We removed any individuals with >75% missing data (i.e. <10 genes resolved out of 40), calculated alignment-specific statistics using Alignment Assessment v2 (Portik et al. 2016), and used the Pegas library v0.11 (Paradis 2010) in R v3.4.4 (R Core Team 2018) to calculate nucleotide diversity (Nei 1987) with 10,000 bootstrap replicates (with subsampling replacement) to measure standard error.

***Phylogenetic analyses***

We used a combination of phylogenetic approaches for the concatenated multiple sequence alignment. We conducted maximum likelihood (ML) tree searches using IQ-Tree v1.6.7 (Nguyen et al. 2015). The analyses were partitioned by gene, allowing each partition to have its own evolutionary rate (option -spp), and we used IQ-Tree’s ModelFinder (Kalyaanamoorthy et al. 2017) to determine the best-fit substitution model for each partition and to merge partitions with the same best model (option –m MFP+MERGE). We assessed node support with 10,000 replicates each of ultrafast bootstrapping (ufBS, Hoang et al. 2018a) and the Shimodaira-Hasegawa approximate likelihood-ratio test (SH-aLRT, Guindon et al. 2010), and resampling for the ufBS was conducted within partitions. Maximum Parsimony (MP) tree searches were conducted in MPBoot (Hoang et al. 2018b) using 10,000 replicates of ufBS to assess node support.

We then used ASTRAL-III v5.6.2 (Zhang et al. 2018) to estimate a species tree from individual gene trees under the multi-species coalescent model, which has been shown to be effective when there are high levels of incomplete lineage sorting (Chou et al. 2015). We used consensus trees from IQ-Tree (generated as above, but with no partition information) as input gene trees, did not force individuals of the same species to be together in the species tree, and assessed support using local posterior probability (locPP, Sayyari and Mirarab 2016) as implemented in ASTRAL-III. Phylogenies were visualized using FigTree v1.4.4 (Rambaut and Drummond 2010).

***Divergence time estimation***

We used BEAST v2.5.1 (Bouckaert et al. 2014) to co-estimate divergence times and relationships in a Bayesian framework. We used partitions identified by IQ-Tree’s ModelFinder, applied a birth-death process for the branching process prior, and used a separate relaxed lognormal distribution clock model for each partition (Drummond et al. 2006). We used several fossils to calibrate the molecular clock, and predominantly followed Thornhill et al. (2015) to select these fossils. We used *Myrtaceidites mesonesu* Cookson & Pike to calibrate all genera except *Xanthostemon* (Couper 1960) (calibration #1), *Eucalyptus* fruits to calibrate all genera except *Xanthostemon* and *Szyzgium* (Gandolfo et al. 2011) (calibration #2, see below), *Paleomyrtinaea princetonensis* Stockey & Maxwell to calibrate Myrteae (Pigg et al. 1993) (calibration #3), and *Myrtaceidites verrucosus* Partridge to calibrate *Psidium* + *Pimenta* + *Eugenia* (Graham 1985; Romero and Zamaloa 1985) (calibration #4). We used the prior exponential distributions of Thornhill et al. (2015) for these calibrations: means = 1.03, 1.844, 4.066, and 3.09, and offsets = 61.7, 51.2, 50, and 37.2, respectively. Several new fossil *Metrosideros* have been documented recently (Tarran et al. 2016; Tarran et al. 2017) that increase the minimum age of this genus from the Miocene (Pole et al. 2008) to the Oligocene-Eocene (30-40 MYA) (Tarran et al. 2016). Again, we preferred to be conservative with these dating priors, so we applied a normally distributed calibration to the *Metrosideros* clade with a mean = 30 and sigma = 10; this established a 95% confidence interval for this calibration from 13.6 to 46.4 MYA, thus broadly encompassing all fossil data for the genus (calibration #5). As proposed by Pillon et al. (2015), previous analyses generally placed *Carpolepis* and *Tepualia* within *Metrosideros*; to avoid conflicting topologies across analyses, we included these genera in this calibration clade. We ran four independent BEAST analyses, each with 100-million Markov Chain Monte Carlo replicates that were sampled every 10,000 generations. Convergence (effective sample sizes > 200) was assessed using TRACER v1.7.1, and independent analyses were combined with LOGCOMBINER, where we discarded the first 25% of generations as burn-in. We used TREEANNOTATOR to generate a maximum clade credibility tree with median ages and 95% credibility intervals (CI) for all nodes. We used FigTree v1.4.4 (Rambaut and Drummond 2010) to visualize trees.

RESULTS

Combining all sequencing results yielded 11.13 million reads, including 2.43 million reads successfully demultiplexed by both locus and individual. After removing sequences with Ns and those with a read depth <20 on a per individual+locus basis, we used the remaining 2.38 million reads to call the raw consensus sequences (average per gene = 54,243 reads, average per individual = 14,916 reads). From our initial 187 individuals and 44 genes, 23 individuals failed outright (no gene recovery) likely due to poor or variable quality DNA input and errors in sample handling and preparation. After removing the four genes that were likely not single copy within the study lineage, and an additional 14 individuals with >75% missing data (i.e., missing genes), the resulting concatenated alignment comprised 151 individuals and 12,702 bp (40 genes). Fifty-one individuals had 100% recovery of genes, and the average number of genes recovered per individual was 32.3 (minimum = 10, maximum = 40). There was an average of 122.8 individuals per gene (minimum = 62, maximum = 151), average gene length was 317.5 bp (minimum = 203, maximum = 509), and average percentage of variable sites was 12.2% (minimum = 7%, maximum = 25.7%). Alignment and partition information are provided in Online Resource 2 and Online Resource 3, respectively.

We assessed phylogenetic relationships in this dataset with three alternative search heuristics, partitioned ML analysis, non-partitioned MP, and gene-tree species-tree analysis using ASTRAL-III. The inferred phylogenies from these approaches were virtually identical when branches lacking strong node support were collapsed. Following the authors’ recommendations for each approach, we considered a node to have strong support with support values of >0.95 ufBS for MP, >0.8/0.95 SH-aLRT/ufBS for ML, and >0.95 locPP for ASTRAL-III. We present the ML consensus tree, with support values from all analyses, in Fig. 1, and all trees are provided in Online Resources 4. Overall, we find strong support for the monophyly of most of the BKMMST clade, represented here by *Backhousia* (tribe Backhousieae), *Tristaniopsis* (Kanieae), *Metrosideros* (Metrosidereae), *Eugenia*, *Pimenta*, *Psidium*, *Rhodomyrtus* (Myrteae), *Xanthomyrtus* (Tristanieae), and *Clozeia* (unplaced to tribe). However, *Syzygium* (Syzygieae) was placed outside of this clade and sister to the other taxa in our tree, although the node support in the basal part of the tree was poor. In fact, we found no strong support for any of the deeper relationships between tribes and genera, except for the monophyly of tribe Myrteae. *Xanthomyrtus* was resolved as sister to the clade of *Cloezia* and *Metrosideros*. *Cloezia* was strongly supported as monophyletic and was recovered as sister to *Metrosideros* with almost significant node support in ML (90.8/94 SH-aLRT/ufBS), but much lower support in other analyses (0.51 ufBS in MP, 0.48 locPP in ASTRAL-III). ASTRAL-III recovered more nodes at this intergeneric level of the phylogeny with high support (Fig. 1, inset), but many of the relationships were still unsupported. ASTRAL-III input provided in Online Resource 5, and gene trees provided in Online Resources 6 and 7.

In the ingroup, all analyses strongly supported the monophyly of *Metrosideros* including *M. laurifolia* (previously *Carpolepis laurifolia*) and *M. stipularis* (previously *Tepualia stipularis*) (Fig. 1). *Metrosideros laurifolia* tended to fall as sister to other *Metrosideros*, although without strong support. *Metrosideros stipularis* and *M. angustifolia* from South Africa were well supported sister taxa in both ML and MP. Within *Metrosideros*, we consistently observed some finer grouping of species but many of the interspecific relationships were not well supported. *Metrosideros colensoi* Hook.f. + *Metrosideros diffusa* Sol. Ex Gaertn. were strongly resolved as sister taxa and sister to *Metrosideros fulgens* Sol. Ex Gaertn., at least in the ML analysis. *Metrosideros boninensis* (Hayata ex Koidz) Tuyama (Japan) was resolved as sister to *Metrosideros tetrasticha* Guillaumin (New Caledonia), and these in turn were sister to *Metrosideros nervulosa* C.Moore & F.Muell.(Lord Howe Island). The large New Zealand tree species *Metrosideros excelsa* Sol. Ex Gaertn. and *Metrosideros robusta* A.Cunn. also formed a clade. *Metrosideros collina* *s.l.* formed two well-supported clades, *M. collina s.s.* (J.R.Forst & G.Forst.) A.Gray and *M. vitiensis* (A.Gray) Pillon, with the latter also being resolved as two strongly supported, geographic clades*.* Hawaiian *Metrosideros* spp. formed a well-supported monophyletic group in all analyses except for ASTRAL-III, and was sister to *M. collina s.s*.. Throughout the clade of Hawaiian *Metrosideros*, branch lengths were very short. Except for one well supported clade in ML and MP that roughly separated the Hawaiian sampling in half, only a few small clades were well supported and only in ML analysis. There was no apparent phylogenetic structure associated with islands, species or subspecific taxa within the Hawaiian clade, nor was there a division between glabrous and pubescent taxa. Average gene diversity within each island ranged from 0.00125 in Kauaʻi to 0.00262 in Molokaʻi (Table 2).

Using fossil calibrations for five positions on this phylogeny, we estimated divergence times of the partitioned dataset and overall found ages that were similar to those obtained in previous studies (Berger et al. 2016; Thornhill et al. 2015). The overall crown age of our tree was estimated in the Paleocene to the late Cretaceous (72.8 MYA), while the crown age of *Metrosideros* fell in the Oligocene to early Miocene (29.9 MYA) (Fig. 2). Lastly, the Hawaiian *Metrosideros* clade had an estimated crown age in the Pliocene (3.1 MYA) (Fig. 2). BEAST input provided in Online Resource 8.

DISCUSSION

***Broad phylogenetic patterns***

Using 40 nuclear markers, we find support for a monophyletic BKMMST clade with the exception of *Syzygium*, although node support for its placement was poor (Fig. 1). Given the low node support in the basal part of the tree, the unexpected placement of *Syzygium* (represented by a single accession) could be an artifact of poor taxon sampling, particularly in comparison to Biffin et al. (2010) and Thornhill et al. (2015), who sampled Syzygieae more comprehensively. Inadequate rooting might also explain the unexpected placement of *Syzygium*, considering that we did not include any non-Myrtaceae outgroup taxa and only 12 loci were successfully sequenced for *Xanthostemon*. The conflict among studies over the placement of *Syzygium* relative to the other tribes could be real considering that our study relies entirely on nuclear genes, while previous Myrtaceae phylogenetic studies relied heavily on plastid loci (Berger et al. 2016; Biffin et al. 2010; Thornhill et al. 2015; Vasconcelos et al. 2017). Incongruence among plastid, mitochondrial, and nuclear genomes has been recently reported at a deep phylogenetic level (Sun et al. 2015). In the current study, the relationships between the tribes of the BKMMST clade were weakly supported, however, the New Caledonia-endemic genus *Cloezia* (unplaced to tribe) was strongly supported as monophyletic and sister to a strongly supported monophyletic *Metrosideros* rather than to tribe Kanieae as was observed by Biffin et al. (2010). Future work combining comprehensive sampling of both taxa and the genome (i.e. phylogenomic approaches) will undoubtedly provide greater resolution of intertribal relationships within the Myrtaceae.

We find support for the inclusion of *Carpolepis* (New Caledonia)and *Tepualia* (South America)within *Metrosideros*, as proposed by Pillon et al. (2015), although our sampling did not include some of the most divergent species within these lineages, including those from subsection *Crassinervis* from New Caledonia, New Guinea, New Zealand, the Solomon Islands, and the Philippines. The New Zealand *Metrosideros* taxa were resolved in multiple clades, indicating that this diverse group of species is clearly not monophyletic and has colonized New Zealand multiple times (and refuting the hypothesis of Wright et al. (2000) and Wright et al. (2001)). Interestingly, within New Zealand, we observed close relationships between species of similar habit (e.g. three of the vine species, *M. diffusa*, *M. colensoi*, and *M. fulgens*, and the two large tree species, *M. robusta* and *M. excelsa*). Two distinct clades are recovered in the paraphyletic *M. collina s.l.*, one from French Polynesia (Marquesas, Tahiti) from where the species was described, and one from Fiji, Samoa, Vanuatu, which is now recognized as *M. vitiensis* (Pillon et al. 2015). Although our sampling of *M. collina* in French Polynesia is limited and does not include specimens from the Austral Islands, we recovered the group as monophyletic. This result does not support the hypothesis of a dispersal from Hawaiʻi to the Marquesas, a possibility suggested previously (Percy et al. 2008). Instead, the resulting tree topology is consistent with French Polynesia serving as a stepping stone between the Western Pacific and the Hawaiian Islands for *Metrosideros* (Percy et al. 2008; Wright et al. 2000; Wright et al. 2001). Our data do not distinguish between the islands of French Polynesia, however, so we are unable to pinpoint their last stop before crossing the equator. Furthermore, the close relationships between *M. excelsa*/*M. robusta* (New Zealand) and species in eastern Melanesia/Polynesia/Hawaiʻi may suggest a New Zealand origin for the latter lineages, consistent with the conclusions of Wright et al. (2000) and Wright et al. (2001). *Metrosideros vitiensis* was resolved into two strongly supported distinct clades, which could represent two different species, although morphological variation in this species is relatively limited (Pillon 2018). Within the Hawaiian Islands, we found no phylogenetic resolution among islands, taxa, or leaf types (glabrous vs. pubescent) consistent with recent diversification of, and hybridization among, taxa. Weak resolution at phylogenetically useful loci despite remarkable morphological diversification among taxa is common among Hawaii’s young plant lineages (e.g. *Tetramolopium*: Lowrey et al. (2001), *Pittosporum*: Gemmill et al. (2002), *Psychotria*: Nepokroeff et al. (2003) and Stallman et al. (2019)).

***Dating Hawaiian* Metrosideros**

Here, we provide the most data-rich estimate of the timing of *Metrosideros* colonization of the Hawaiian Islands, based on fossil calibration including new discoveries of *Metrosideros* fossils in Tasmania (Tarran et al. 2016; Tarran et al. 2017). Using minimum mutation rates of ribosomal markers, Wright et al. (2000) and Wright et al. (2001) estimated an age of <2 and 0.5-1 MYA, respectively, for this colonization. Percy et al. (2008) presented an older estimate (2.4-3.9 MYA) based on cpDNA markers and island ages with wide confidence intervals (1.4-6.3 MYA). Our estimate supports the older estimates of Percy et al. (2008), but narrows the range of this estimate substantially. Unlike Percy et al. (2008), we did not find any phylogenetic structure corresponding to island progression, so cannot make iterative estimates of island colonization times within the archipelago. However, the presence of both Kauaʻi (5.1 MYA) and Oʻahu roughly 3 MYA (Clague et al. 2009; Price and Clague 2002) suggests that multiple islands were available for colonization when *Metrosideros* arrived. Given the exceptional dispersability of *Metrosideros* seeds (Drake 1992), spread to the second island was likely rapid following establishment on the first. Such a scenario may help to explain the high variation observed on these islands relative to that on the younger islands in terms of cpDNA haplotype diversity (Kauaʻi) (Percy et al. 2008), morphology (Oʻahu) (Sur et al. 2018), and nuclear microsatellites (Oʻahu) (Stacy and Sakishima 2019). However, these observations conflict with observation in the present study of slightly lower nucleotide diversity on Kauaʻi relative to all other islands. Kauai’s lower diversity at low-copy nuclear genes may stem from a lower number of taxa/morphotypes sampled from this island compared to other islands (Table 2), but Maui also had relatively few taxa sampled and did not have reduced diversity, supporting the trend of higher genetic diversity on the intermediate-aged Hawaiian Islands (Gillespie and Baldwin 2010; Stacy and Sakishima 2019). If colonization of Kauaʻi and Oʻahu was near simultaneous, the disparate pattern of diversity at these markers on these islands may be associated with the contrasting environmental conditions around the time of colonization (post-shield-building phase on Kauaʻi vs. shield-building phase on Oʻahu); i.e. Kauai’s relatively higher chloroplastic diversity and isolation by distance may reflect slower evolution of these markers and a longer history of *Metrosideros* on a highly dissected landscape there (somewhat limiting seed dispersal).

Within the Hawaiian Islands, we did not find any phylogenetic patterns consistent with the progression rule (Wagner and Funk 1995), or any indication of repeat- or back-colonizations of islands (Bess et al. 2014). Further, within each of the two widespread varieties of *M. polymorpha* there was no pattern of clustering across islands consistent with initial island colonization by these taxa. Although nuclear markers are generally faster-evolving than chloroplast or ribosomal markers, perhaps they are still too slow to reflect the rapid diversification of this group (e.g. Pillon et al. 2013a). Additionally, it is important to note that hybridization between species/varieties of *Metrosideros* is common (e.g. Corn and Hiesey 1973; Gardner et al. 2004; Stacy et al. 2016), so any signature of multiple introductions to individual islands may be lost due to hybridization (Johnson et al. 2019). The wind dispersal of *Metrosideros* seeds likely exacerbates the homogenizing effect of hybridization predominantly within islands (Stacy and Sakishima 2019), again speaking to the need for even faster-evolving molecular markers and population genetic analyses (discussed below).

***Hawaiian* Metrosideros *diversity***

Using reduced specimen sampling and just five of the 40 nuclear markers reported here, Pillon et al. (2015) suggested that a large number of genes would be necessary to resolve the phylogenetic relationships within *Metrosideros*. Given our results based on 40 nuclear markers and extensive specimen sampling across the Hawaiian Islands, however, it is clear that adding more genes derived from the same approach will not resolve the relationships in this group. Within the archipelago, we observed no well-supported phylogenetic structure concordant with taxonomic or morphotype boundaries, island of origin, or any other axes of variation (e.g. glabrous vs. pubescent leaves). Observation of genetic differentiation across islands (Percy et al. 2008) and taxa (Stacy and Sakishima 2019) through other approaches suggests that these low-copy nuclear markers apparently lack the resolution needed for fine-scale population-level analysis. The utility of these low-copy markers at broad phylogenetic scales (Duarte et al. 2010) may inherently limit their utility at finer phylogenetic scales; testing this idea is beyond the scope of this study, however, other studies have shown the utility of bioinformatically selecting loci that are informative for the particular group in question (e.g. Dupuis et al. 2018). Additionally, given the success of microsatellite-based analyses of fine-scale genetic structure in *Metrosideros* (DeBoer and Stacy 2013; Stacy et al. 2014; Stacy and Sakishima 2019), and the high resolution of population genomic approaches in general (Restriction-site Associated DNA sequencing, RADseq, e.g., Andrews et al. 2016), future work in this group using genome-wide SNPs is promising.

CONCLUSIONS

We used targeted amplicon sequencing to generate a large nuclear dataset for *Metrosideros*, focusing on the Hawaiian Islands. We found strong support for the inclusion of *Carpolepis* and *Tepualia* in *Metrosideros*, supporting previous studies. Hawaiian *Metrosideros* formed a strongly supported monophyletic group, and using recently discovered *Metrosideros* fossils, we dated the colonization of the Hawaiian archipelago to ~3.1 MYA, which is intermediate to previous estimates. The Hawaiian clade was sister to *M. collina* from French Polynesia. In turn these shared a common ancestor with taxa from Fiji, Vanuatu, and Samoa, indicating a westerly migration across the Pacific. Within the monophyletic Hawaiian clade, however, we found no phylogenetic pattern associated with islands or taxonomic units, and no signatures of progressive colonization of the islands or repeat or back-colonizations of individual islands. Given the recent findings of genetic differentiation within Hawaiian *Metrosideros* based on microsatellite markers, it is clear that very fast-evolving markers are required to investigate relationships in this group. Our estimate of *Metrosideros*’arrival in the Hawaiian Islands at 3.1 MYA suggests that evolution of the group’s exceptional phenotypic diversity has been rapid.

ACKNOWLEDGEMENTS

The authors thank J. Johansen and A. Williams for assistance with sample collection and the Hawaiʻi Division of Forestry and Wildlife for permission to collect from state forests. They also thank Laure Barrabé, Abby Cuttriss, Gildas Gâteblé, Melissa Johnson, Jean-Yves Meyer, Greg Plunkett (“Plant and People of Vanuatu” Project), Laurence Ramon, Ravahere Taputuarai, the staff of the national Park of American Samoa, the herbarium NOU, Eve Lucas and the Kew DNA bank for providing DNA samples, and Mark Simmons for comments on an earlier draft of this manuscript. Fluidigm Access Array amplification and MiSeq sequencing were conducted at the Institute for Bioinformatics and Evolutionary Studies (IBEST) at the University of Idaho. Data processing and analysis were conducted on the moana HPC cluster at the United Stated Department of Agriculture-Agriculture Research Service, Daniel K. Inouye Pacific Basin Agricultural Research Center. USDA is an equal opportunity employer. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. Funding was provided by NSF DEB 0954274 and NSF HRD 0833211. Figures were prepared using FigTree v1.4.4 (Rambaut and Drummond 2010) and Inkscape v0.91 (The Inkscape Team 2017).

COMPLIANCE WITH ETHICAL STANDARDS

DISCLOSURE OF POTENTICAL CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

ELECTRONIC SUPPLEMENTARY MATERIAL

Note, specimen identifiers might differ between supplementary files, but DNA numbers are consistent. For detailed specimen information, see Online Resource 1.

**Online Resource 1**. Detailed specimen information, and summarized gene information in final dataset. **Online Resource 2**. Main DNA alignment. **Online Resource 3**. Partitions for main DNA alignment. **Online Resource 4**. Maximum likelihood consensus tree with SH-aLRT/ufBS values for node support, maximum parsimony consensus tree with ufBS values for node support, and ASTRAL-III species tree with locPP values for node support. **Online Resource 5**. Input tree file for ASTRAL-III analysis. **Online Resource 6**. Gene trees in newick format**. Online Resource 7**. Gene trees in pdf format, rooted with best available outgroup or else rooted mid-tree. **Online Resource 8**. Input XML file for BEAST analysis.

LITERATURE CITED

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. Nat Rev Genet 17:81-92 doi:10.1038/nrg.2015.28

Bailey CD, Doyle JJ (1999) Potential phylogenetic utility of the low-copy nuclear gene *pistillata* in Docotyledonous plants: comparison to nrDNA ITS and *trnL* intron in *Sphaeocardamum* and other Brassicaceae. Mol Phylogenet Evol 13:20-30

Berger BA, Kriebel R, Spalink D, Sytsma KJ (2016) Divergence times, historical biogeography, and shifts in speciation rates of Myrtales. Mol Phylogenet Evol 95:116-136 doi:10.1016/j.ympev.2015.10.001

Bess EC, Catanach TA, Johnson KP, Fernández-Palacios JM (2014) The importance of molecular dating analyses for inferring Hawaiian biogeographical history: a case study with bark lice (Psocidae:Ptycta). J Biogeogr 41:158-167 doi:10.1111/jbi.12191

Biffin E, Lucas EJ, Craven LA, Ribeiro da Costa I, Harrington MG, Crisp MD (2010) Evolution of exceptional species richness among lineages of fleshy-fruited Myrtaceae. Ann Bot 106:79-93 doi:10.1093/aob/mcq088

Bouckaert R et al. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput Biol 10:e1003537 doi:10.1371/journal.pcbi.1003537

Bybee SM et al. (2011) Targeted amplicon sequencing (TAS): a scalable next-gen approach to multilocus, multitaxa phylogenetics. Genome Biol Evol 3:1312-1323 doi:10.1093/gbe/evr106

Carstens BC, Satler JD (2013) The carniverous plant described as *Sarracenia alata* contains two cryptic species. Biol J Linn Soc 109:737-746

Chou J, Gupta A, Yaduvanshi S, Davidson R, Nute M, Mirarab S, Warnow T (2015) A comparative study of SVDquartets and other coalescent-based species tree estimation methods. BMC Genomics 16:S2

Clague DA, Braga JC, Bassi D, Fullagar PD, Renema W, Webster JM (2009) The maximum age of Hawaiian terrestrial lineages: geological constraints from Kōko Seamount. J Biogeogr 37:1022-1033 doi:10.1111/j.1365-2699.2009.02235.x

Corn CA, Hiesey WM (1973) Altitudinal variation in Hawaiian *Metrosideros*. Am J Bot 60:991-1002

Couper RA (1960) New Zealand Mesozoic and Cainozoic Plant Microfossils. Government Printer, Wellington, New Zealand

de Sousa F, Bertrand YJ, Nylinder S, Oxelman B, Eriksson JS, Pfeil BE (2014) Phylogenetic properties of 50 nuclear loci in *Medicago* (Leguminosae) generated using multiplexed sequence capture and next-generation sequencing. PLoS One 9:e109704 doi:10.1371/journal.pone.0109704

DeBoer N, Stacy EA (2013) Divergence within and among 3 Varieties of the Endemic Tree, 'Ohi'a Lehua (*Metrosideros polymorpha*) on the Eastern Slope of Hawai'i Island. J Hered 104:449-458 doi:10.1093/jhered/est027

Drake DR (1992) Seed dispersal of *Metrosideros polymorpha* (Myrtaceae): a pioneer tree of Hawaiian lava flows. Am J Bot 79:1224-1228

Drummond AJ, Ho SY, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. PLoS Biol 4:e88 doi:10.1371/journal.pbio.0040088

Duarte JM et al. (2010) Identification of shared single copy nuclear genes in *Arabadopsis*, *Populus*, *Vitis* and *Oryza* and their phylogenetic utility across various taxonomic levels. BMC Evol Biol 10:61

Dupuis JR, Bremer FT, Kauwe A, San Jose M, Leblanc L, Rubinoff D, Geib SM (2018) HiMAP: Robust phylogenomics from highly multiplexed amplicon sequencing. Mol Ecol Res 18:1000-1019 doi:10.1111/1755-0998.12783

Gandolfo MA et al. (2011) Oldest known *Eucalyptus* macrofossils are from South America. PLoS One 6:e21084 doi:10.1371/journal.pone.0021084

Gardner RC, De Lange PJ, Keeling DJ, Bowala T, Brown HA, Wright SD (2004) A late Quaternary phylogeography for *Metrosideros* (Myrtaceae) in New Zealand inferred from chloroplast DNA haplotypes. Biol J Linn Soc 83:399-412

Gemmill CE, Allan GJ, Wagner WL, Zimmer EA (2002) Evolution of insular Pacific *Pittosporum* (Pittosporaceae): origin of the Hawaiian radiation. Mol Phylogenet Evol 22:31-42 doi:10.1006/mpev.2001.1019

Gillespie RG, Baldwin BG (2010) Island biogeography of remote archipelagoes. In: Losos JB, Ricklefs RE (eds) The Theory of Island Biogeography revisited. Princeton University Press, Princeton,

Govaerts R et al. (2008) World Checklist of Myrtaceae: 1-455. Kew Publishing, Royal Botanical Gardens, Kew

Graham A (1985) Studies in Neotropical Paleobotany IV. The Eocene communities of Panama. Ann Missouri Bot Gard 72:504-534

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307-321 doi:10.1093/sysbio/syq010

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018a) UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol Biol Evol 35:518-522 doi:10.1093/molbev/msx281

Hoang DT, Vinh LS, Flouri T, Stamatakis A, von Haeseler A, Minh BQ (2018b) MPBoot: fast phylogenetic maximum parsimony tree inference and bootstrap approximation. BMC Evol Biol 18:11 doi:10.1186/s12862-018-1131-3

Hollingsworth PM et al. (2009) A DNA barcode for land plants. Proc Natl Acad Sci U S A 106:12794-12797 doi:[www.pnas.org/cgi/doi/10.1073/pnas.0905845106](file:///Users/juliandupuis/Documents/Geib_Lab/Metrosideros_phylogeny/submission2_PSE2/www.pnas.org/cgi/doi/10.1073/pnas.0905845106)

Johnson MA, Pillon Y, Sakishima T, Price DK, Stacy EA (2019) Multiple colonizations, hybridization and uneven diversification in *Cyrtandra* (Gesneriaceae) lineages on Hawai'i Island. J Biogeogr 46:1178-1196 doi:10.1111/jbi.13567

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587-589 doi:10.1038/nmeth.4285

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772-780 doi:10.1093/molbev/mst010

Kelchner SA (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. Ann Missouri Bot Gard 87:482-498

Kress WJ (2017) Plant DNA barcodes: Applications today and in the future. J Syst Evol 55:291-307 doi:10.1111/jse.12254

Larsson A (2014) AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30:3276-3278 doi:10.1093/bioinformatics/btu531

Lowrey TK, Quinn CJ, Taylor RK, Chan R, Kimball RT, De Nardi JC (2001) Molecular and morphological reassessment of relationships within the *Vittadinia* group of Astereae (Asteraceae). Am J Bot 88:1279-1289

Magoc T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27:2957-2963 doi:10.1093/bioinformatics/btr507

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet Journal 17:10-12

Moore WS (1995) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution 49:718-726 doi:<https://www.jstor.org/stable/2410325>

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nepokroeff M, Sytsma KJ, Wagner WL, Zimmer EA (2003) Reconstructing Ancestral Patterns of Colonization and Dispersal in the Hawaiian Understory Tree Genus *Psychotria* (Rubiaceae): A Comparison of Parsimony and Likelihood Approaches. Syst Biol 52:820-838 doi:10.1080/10635150390251072

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268-274 doi:10.1093/molbev/msu300

O'Neill EM et al. (2013) Parallel tagged amplicon sequencing reveals major lineages and phylogenetic structure in the North American tiger salamander (*Ambystoma tigrinum*) species complex. Mol Ecol 22:111-129 doi:10.1111/mec.12049

Oh S-H, Potter D (2003) Phylogenetic utility of the second intron of LEAFY in *Neillia* and *Stephanandra* (Rosaceae) and implications for the origin of *Stephanandra*. Mol Phylogenet Evol 29:203-215 doi:10.1016/s1055-7903(03)00093-9

Olmstead RG, Palmer JD (1994) Chloroplast DNA systematics: a review of methods and data analysis. Am J Bot 81:1205-1224

Paradis E (2010) pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics 26:419-420 doi:10.1093/bioinformatics/btp696

Percy DM, Garver AM, Wagner WL, James HF, Cunningham CW, Miller SE, Fleischer RC (2008) Progressive island colonization and ancient origin of Hawaiian *Metrosideros* (Myrtaceae). Proc Royal Soc B 275:1479-1490 doi:10.1098/rspb.2008.0191

Pigg KB, Stockey RA, Maxwell SL (1993) *Paleomyrtinaea*, a new genus of permineralized myrtaceous fruits and seeds from the Eocene of British Columbia and Paleocene of North Dakota. Can J Bot 71:1-9

Pillon Y (2018) A new species of *Metrosideros* (Myrtaceae) from Vanuatu and notes on the genus. Phytotaxa 347 doi:10.11646/phytotaxa.347.2.10

Pillon Y et al. (2013a) Potential use of low-copy nuclear genes in DNA barcoding: a comparison with plastid genes in two Hawaiian plant radiations. BMC Evol Biol 13:35

Pillon Y, Johansen J, Sakishima T, Chamala S, Barbazuk WB, Stacy EA (2014) Primers for low-copy nuclear genes in *Metrosideros* and cross-amplification in Myrtaceae. Appl Plant Sci 2 doi:10.3732/apps.1400049

Pillon Y, Johansen JB, Sakishima T, Roalson EH, Price DK, Stacy EA (2013b) Gene discordance in phylogenomics of recent plant radiations, an example from Hawaiian *Cyrtandra* (Gesneriaceae). Mol Phylogenet Evol 69:293-298 doi:10.1016/j.ympev.2013.05.003

Pillon Y, Lucas E, Johansen JB, Sakishima T, Hall B, Geib SM, Stacy EA (2015) An expanded *Metrosideros* (Myrtaceae) to include *Carpolepis* and *Tepualia* based on nuclear genes. Syst Bot 40:782-790 doi:10.1600/036364415x689249

Pole M, Dawson J, Denton T (2008) Fossil Myrtaceae from the Early Miocene of southern New Zealand. Aust J Bot 56:67 doi:10.1071/bt07032

Portik DM, Smith LL, Bi K (2016) An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura). Mol Ecol Res 16:1069-1083 doi:10.1111/1755-0998.12541

Price JP, Clague DA (2002) How old is the Hawaiian biota? Geology and phylogeny suggest recent divergence. Proc Royal Soc B 269:2429-2435 doi:10.1098/rspb.2002.2175

Price JP, Wagner WL (2018) Origins of the Hawaiian flora: Phylogenies and biogeography reveal patterns of long‐distance dispersal. J Syst Evol 56:600-620 doi:10.1111/jse.12465

R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria URL <https://wwwR-projectorg/>

Rambaut A, Drummond AJ (2010) FigTree v1.4.4. Institute of Evolutionary Biology, University of Edinburgh Available at: <http://treebioedacuk/software/figtree>

Romero EJ, Zamaloa MC (1985) Polen de Angiospermas de la formacion Rio Turbio (Eoceno), Provincia de Santa Cruz, Republica Argentina. Ameghiniana 22:43-51

Sang T (2002) Utility of low-copy nuclear gene sequences in plant phylogenetics. Crit Rev Biochem Mol Biol 37:121-147 doi:10.1080/10409230290771474

Sayyari E, Mirarab S (2016) Fast Coalescent-Based Computation of Local Branch Support from Quartet Frequencies. Mol Biol Evol 33:1654-1668 doi:10.1093/molbev/msw079

Small RL, Cronn RC, Wendel JF (2004) Use of nuclear genes for phylogeny reconstruction in plants. Aust Syst Bot 17:145-170

Soltis DE, Soltis PS (1998) Choosing an approach and an appropriate gene for phylogenetic analysis. In: Soltis DE, Soltis PS, Doyle JJ (eds) Molecular Systematics of Plants II: DNA Sequencing. Kluwer Academic Publishers, Boston, MA,

Stacy EA, Johansen JB, Sakishima T, Price DK (2016) Genetic analysis of an ephemeral intraspecific hybrid zone in the hypervariable tree, *Metrosideros polymorpha*, on Hawai‘i Island. Heredity 117:173-183 doi:10.1038/hdy.2016.40

Stacy EA, Johansen JB, Sakishima T, Price DK, Pillon Y (2014) Incipient radiation within the dominant Hawaiian tree *Metrosideros polymorpha*. Heredity 113:334-342 doi:10.1038/hdy.2014.47

Stacy EA, Sakishima T (2019) Phylogeography of the highly dispersible landscape‐dominant woody species complex, *Metrosideros*, in Hawaii. J Biogeogr doi:10.1111/jbi.13650

Stallman JK, Funk VA, Price JP, Knope ML (2019) DNA barcodes fail to accurately differentiate species in Hawaiian plant lineages. Bot J Linn Soc 190:374-388 doi:10.1093/botlinnean/boz024

Straub SC, Parks M, Weitemier K, Fishbein M, Cronn RC, Liston A (2012) Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. Am J Bot 99:349-364 doi:10.3732/ajb.1100335

Sun M, Soltis DE, Soltis PS, Zhu X, Burleigh JG, Chen Z (2015) Deep phylogenetic incongruence in the angiosperm clade Rosidae. Mol Phylogenet Evol 83:156-166 doi:10.1016/j.ympev.2014.11.003

Sur GL, Keating R, Snow N, Stacy EA (2018) Leaf Micromorphology Aids Taxonomic Delineation within the Hypervariable Genus *Metrosideros* (Myrtaceae) on O‘ahu. Pac Sci 72:345-361 doi:10.2984/72.3.6

Syring J, Willyard A, Cronn R, Liston A (2005) Evolutionary relationships among *Pinus* (Pinaceae) subsections inferred from multiple low-copy nuclear loci. Am J Bot 92:2086-2100

Tarran M, Wilson PG, Hill RS (2016) Oldest record of *Metrosideros* (Myrtaceae): Fossil flowers, fruits, and leaves from Australia. Am J Bot 103:754-768 doi:10.3732/ajb.1500469

Tarran M, Wilson PG, Macphail MK, Jordan GJ, Hill RS (2017) Two fossil species of *Metrosideros* (Myrtaceae) from the Oligo-Miocene Golden Fleece locality in Tasmania, Australia. Am J Bot 104:891-904 doi:10.3732/ajb.1700095

The Inkscape Team (2017) Inkscape v0.91. Available at: <https://inkscapeorg>

Thornhill AH, Ho SY, Kulheim C, Crisp MD (2015) Interpreting the modern distribution of Myrtaceae using a dated molecular phylogeny. Mol Phylogenet Evol 93:29-43 doi:10.1016/j.ympev.2015.07.007

Uribe-Convers S, Settles ML, Tank DC (2016) A Phylogenomic Approach Based on PCR Target Enrichment and High Throughput Sequencing: Resolving the Diversity within the South American Species of *Bartsia* L. (Orobanchaceae). PLoS One 11:e0148203 doi:10.1371/journal.pone.0148203

Vasconcelos TNC et al. (2017) Myrteae phylogeny, calibration, biogeography and diversification patterns: Increased understanding in the most species rich tribe of Myrtaceae. Mol Phylogenet Evol 109:113-137 doi:10.1016/j.ympev.2017.01.002

Wagner WL, Funk VA (1995) Hawaiian biogeography: evolution on a hot spot archipelago. Smithsonian Institution Press, Washington DC

Wagner WL, Herbst DR, Lorence DH (2005) Flora of the Hawaiian Islands website. <http://botany.si.edu/pacificislandbiodiversity/hawaiianflora/index.htm> [accessed February 2019].

Weitemier K, Straub SC, Cronn RC, Fishbein M, Schmickl R, McDonnell A, Liston A (2014) Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. Appl Plant Sci 2 doi:10.3732/apps.1400042

Wendel JF, Doyle JJ (1998) Phylogenetic incongruence: window into genome history and molecular evolution. In: Soltis PS, Soltis DE (eds) Molecular Systematics of Plants II. Kluwer Academic, Dordrecht, Netherlands, pp 265-296

Wilson PG (2011) Myrtaceae. In: Kubitzki K (ed) The families and genera of vascular plants. Volume X. Sapindales, Cucurbitales, Myrtaceae. Springer-Verlag, Heidelberg, pp 212-271

Wright SD, Yong CG, Dawson JW, Whittaker DJ, Gardner RC (2000) Riding the ice age El Nino? Pacific biogeography and evolution of *Metrosideros* subg. *Metrosideros* (Myrtaceae) inferred from nuclear ribosomal DNA. Proc Natl Acad Sci U S A 97:4118-4123 doi:10.1073/pnas.050351197

Wright SD, Yong CG, Wichman SR, Dawson JW, Gardner RC (2001) Stepping stones to Hawaii: a trans-equatorial dispersal pathway for *Metrosideros* (Myrtaceae) inferred from nrDNA (ITS+ETS). J Biogeogr 28:769-774

Zhang C, Rabiee M, Sayyari E, Mirarab S (2018) ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19:153 doi:10.1186/s12859-018-2129-y

Zimmer EA, Wen J (2012) Using nuclear gene data for plant phylogenetics: progress and prospects. Mol Phylogenet Evol 65:774-785 doi:10.1016/j.ympev.2013.01.005

Zimmer EA, Wen J (2015) Using nuclear gene data for plant phylogenetics: Progress and prospects II. Next-gen approaches. J Syst Evol 53:371-379 doi:10.1111/jse.12174

DATA AVAILABILITY STATEMENT

Raw reads are provided as a NCBI Sequence Reads Archive (BioProject PRJNA526920, SRA accessions: SRR8737094-SRR8737280), and data input and parameter files are available in the Electronic Supplementary Material.

Table 1. Summarized specimen information, including sample size (N). For Hawaiian *Metrosideros*, General distribution also contains occurrence and general ecology/habitat characteristics, followed by whether the taxon has glabrous or pubescent leaves. For detailed individual information, see Appendix S1.

|  |  |  |  |
| --- | --- | --- | --- |
| Genus | Species epithet | N | General distribution |
| *Metrosideros* | *macropus* | 6 | Oʻahu; montane; glabrous |
| *Metrosideros* | *polymorpha* race B1 | 2 | Oʻahu; Koʻolau backbone, cloud forest; glabrous |
| *Metrosideros* | *polymorpha* race C1 | 2 | Oʻahu; montane; pubescent |
| *Metrosideros* | *polymorpha* var. *dieterii* | 4 | Kauaʻi; ridge; glabrous |
| *Metrosideros* | *polymorpha* race F1 | 6 | Oʻahu; montane, cloud forest; pubescent |
| *Metrosideros* | *polymorpha* var. *glaberrima*2 | 34 | Hawaiʻi, Kauaʻi, Maui, Molokaʻi, Oʻahu; wet forest, bog; glabrous |
| *Metrosideros* | *polymorpha* var. *incana*2 | 17 | Hawaiʻi, Maui, Molokaʻi, Oʻahu; new lava flows, dry, low elevation; pubescent |
| *Metrosideros* | *polymorpha* race L1 | 4 | Oʻahu; montane; glabrous |
| *Metrosideros* | *polymorpha* var. *macrophylla* | 3 | Hawaiʻi; wet forest; glabrous |
| *Metrosideros* | *polymorpha* race MB1 | 1 | Molokaʻi; montane; glabrous |
| *Metrosideros* | *polymorpha* race MM1 | 2 | Molokaʻi; ridge; glabrous |
| *Metrosideros* | *polymorpha* var. *newellii* | 4 | Hawaiʻi; riparian; glabrous |
| *Metrosideros* | *polymorpha* var. *nuda* | 1 | Hawaiʻi; subalpine; glabrous |
| *Metrosideros* | *polymorpha* var. *polymorpha* | 7 | Hawaiʻi, Maui, Molokaʻi; subalpine; pubescent |
| *Metrosideros* | *polymorpha* var. *prostrata* | 4 | Oʻahu; extreme wind; pubescent |
| *Metrosideros* | *polymorpha* var. *pumila* | 5 | Kauaʻi, Molokaʻi, Oʻahu; bog; glabrous and pubescent |
| *Metrosideros* | *polymorpha* race S1 | 2 | Oʻahu; Kaʻala Bog; glabrous |
| *Metrosideros* | *polymorpha* race SM1 | 2 | Kauaʻi; wet forest; glabrous |
| *Metrosideros* | *polymorpha* race Z1 | 2 | Oʻahu; montane, cloud forest; glabrous |
| *Metrosideros* | *rugosa* | 5 | Oʻahu, Koʻolau backbone, cloud forest; pubescent |
| *Metrosideros* | *tremuloides* | 5 | Oʻahu; steep windward slopes; glabrous |
| *Metrosideros* | *waialealae var. fauriei* | 3 | Molokaʻi; ridge; glabrous |
| *Metrosideros* | *waialealae var. waialealae* | 4 | Kauaʻi; riparian; glabrous |
| *Metrosideros* | *Angustifolia* | 1 | South Africa |
| *Metrosideros* | *bartlettii* | 1 | New Zealand |
| *Metrosideros* | *boninensis* | 1 | Japan |
| *Metrosideros* | *bacuminum* | 1 | New Caledonia |
| *Metrosideros* | *carminea* | 2 | New Zealand |
| *Metrosideros* | *colensoi* | 1 | New Zealand |
| *Metrosideros* | *collina s.s.* | 5 | Marquesas, Tahiti |
| *Metrosideros* | *cordata* | 1 | New Guinea |
| *Metrosideros* | *diffusa* | 1 | New Zealand |
| *Metrosideros* | *excelsa* | 4 | New Zealand |
| *Metrosideros* | *fulgens* | 1 | New Zealand |
| *Metrosideros* | *nervulosa* | 1 | Lord Howe Island |
| *Metrosideros* | *nitida* | 1 | New Caledonia |
| *Metrosideros* | *operculata* var. *francii* | 2 | New Caledonia |
| *Metrosideros* | *operculata* var. *operculata* | 1 | New Caledonia |
| *Metrosideros* | *perforata* | 2 | New Zealand |
| *Metrosideros* | *robusta* | 1 | New Zealand |
| *Metrosideros* | *tetrasticha* | 1 | New Caledonia |
| *Metrosideros* | *vitiensis* | 8 | Fiji, Vanuatu, Tutuila island |
| *Metrosideros (Carpolepis)* | *laurifolia* | 2 | New Caledonia |
| *Metrosideros (Tepualia)* | *stipularis* | 2 | South America |
| *Backhousia* | *hughesii* | 1 | Australia |
| *Cloezia* | *aquarum* | 1 | New Caledonia |
| *Cloezia* | *artensis* | 2 | New Caledonia vars. *artensis, riparia* |
| *Cloezia* | *buxifolia* | 1 | New Caledonia |
| *Eucalyptus* | *deglupta* | 1 | Pacific |
| *Eugenia* | *uniflora* | 1 | South America |
| *Melaleuca* | *quinquenervia* | 1 | Australasia |
| *Pimenta* | *dioica* | 1 | Central America/Caribbean |
| *Psidium* | *cattleianum* | 2 | Brazil |
| *Psidium* | *guajava* | 1 | South/Central America |
| *Purpureostemon* | *ciliatus* | 1 | New Caledonia |
| *Rhodomyrtus* | *tomentosa* | 1 | SE Asia |
| *Syzygium* | *malaccense* | 2 | Asia/South Pacific |
| *Syzygium* | *sandwicense* | 1 | Hawaii |
| *Tristaniopsis* | *calobuxus* | 1 | New Caledonia |
| *Tristaniopsis* | *glauca* | 1 | New Caledonia |
| *Tristaniopsis* | *reticulata* | 1 | New Caledonia |
| *Xanthomyrtus* | cf. *compacta* | 1 | Indonesia |
| *Xanthostemon* | *chrysanthus* | 1 | Australia |

1Taxa named with single- or double-letter codes are morphotypes not formally described in Dawson and Stemmermann (1990).

2These two varieties of *M. polymorpha* occur on all islands (var. *glaberrima*) or all islands except Kauaʻi (var. *incana*) and are purported by Stacy and Sakishima (2019) to be the principal colonizers of new islands and the progenitors of island-endemic *Metrosideros* taxa within the Hawaiian Islands.

Table 2. Gene diversity (div. and standard error, SE) and sampling per Hawaiian Island; N is the number of adults genotyped, and “taxa” refer to described taxa and morphotypes. Island codes: HA: Hawaiʻi, MA: Maui, MO: Molokaʻi, OA: Oʻahu, and KA: Kauaʻi.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Island | Div. | SE | N | No. taxa |
| HA | 0.00245 | 0.00058 | 28 | 7 |
| MA | 0.00235 | 0.00117 | 7 | 2 |
| MO | 0.00262 | 0.00070 | 18 | 6 |
| OA | 0.00254 | 0.00044 | 44 | 11 |
| KA | 0.00125 | 0.00100 | 30 | 3 |

Fig. 1. Maximum likelihood consensus tree with summarized support values from all analyses. Shapes on nodes indicate “good node support,” numbered circles correspond to calibration points used in estimation of divergence times. Inset (“ASTRAL-III alternate basal topology”) shows basal portion of ASTRAL-III consensus tree, where the placement of *Xanthomyrtus* was the only well-supported difference between alternative heuristics. Taxon labels include species identification and collection locality/DNA identifier in square brackets; for Hawaiian *Metrosideros*, “G” or “P” before collection locality refers to glabrous or pubescent leaves. Tribal classifications are provided after collection locality/DNA identifiers in taxon names, except for Myrteae and Metrosidereae which are labeled at their crown nodes, and the genus *Cloezia* where tribal classification is uncertain. Tribe names with asterisks belong to the BKMMST clade.

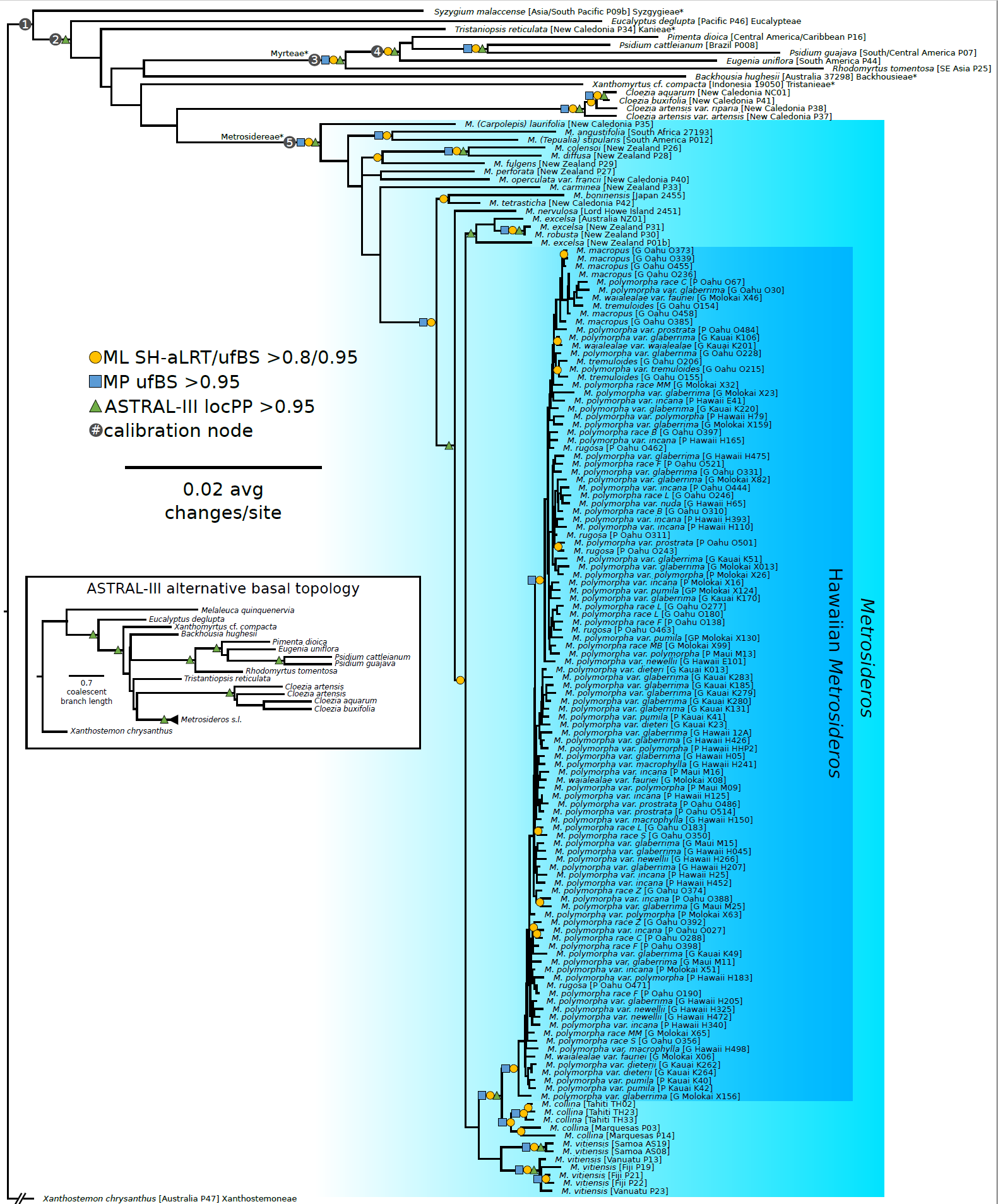


Fig. 2. Maximum clade credibility tree obtained from BEAST. Ages shown as median divergence time estimates with 95% credibility intervals (blue bars). Numbered circles correspond to calibration points used in estimation of divergence times.

