Selection on old variants drives adaptive radiation of *Metrosideros* across the Hawaiian Islands

Jae Young Choi[a,*], Xiaoguang Dai[b], Julie Z. Peng[c], Priyesh Rughani[b], Scott Hickey[d], Eoghan Harrington[b], Sissel Juul[b], Julien Ayroles[c], Michael Purugganan[a], and Elizabeth A. Stacy[e,*]

[a]Center for Genomics and Systems Biology, Department of Biology, New York University, New York, New York, USA
[b]Oxford Nanopore Technologies Inc., New York, New York, USA
[c]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey, USA
[d]Oxford Nanopore Technologies Inc., San Francisco, California, USA
[e]School of Life Sciences, University of Nevada Las Vegas, Las Vegas, Nevada, USA

*__Corresponding authors:__ JYC (jyc387@nyu.edu) and EAS (elizabeth.stacy@unlv.edu)
Competing Interest Statement: XD, PR, SH, EDH, and SJ are employees of Oxford Nanopore Technologies and are shareholders and/or share option holders. The other authors declare that they have no competing interests.
Classification: Biological Sciences/Evolution
Keywords: Metrosideros, adaptive radiation, ecological speciation, incipient speciation, speciation genomics.

Abstract

Some of the most spectacular adaptive radiations begin with founder populations on remote islands. How genetically limited founder populations give rise to the striking phenotypic and ecological diversity characteristic of adaptive radiations is a paradox of evolutionary biology. We conducted an evolutionary genomic analysis of genus *Metrosideros*, a landscape-dominant, incipient adaptive radiation of woody plants that spans a striking range of phenotypes and environments across the Hawaiian Islands. Using nanopore-sequencing, we created a chromosome-level genome assembly for *M. polymorpha* var. *incana* and analyzed whole-genome sequences of 131 individuals from 11 taxa sampled across the islands. We found evidence of population structure that grouped taxa by island. Demographic modeling showed concordance between the divergence times of island-specific lineages and the geological formation of individual islands. Gene flow was also detected within and between island taxa, suggesting a complex reticulated evolutionary history. We investigated genomic regions with increased differentiation as these regions may harbor variants involved in local adaptation or reproductive isolation, thus forming the genomic basis of adaptive radiation. We discovered differentiation outliers have arisen from balancing selection on ancient divergent haplotypes that formed before the initial colonization of the archipelago. These regions experienced recurrent divergent selection as lineages colonized and diversified on new islands, and hybridization likely facilitated the transfer of these ancient variants between taxa. Balancing selection on multiple ancient haplotypes–or time-tested variants–may help to explain how lineages with limited gene pools can rapidly diversify to fill myriad ecological niches on remote islands.

Significance statement

Some of the most spectacular adaptive radiations of plants and animals occur on remote oceanic islands, yet such radiations are preceded by founding events that severely limit genetic variation. How genetically depauperate founder populations give rise to the spectacular phenotypic and ecological diversity characteristic of island adaptive radiations is not known. We generated novel genomic resources for Hawaiian *Metrosideros*—a hyper-variable incipient adaptive radiation of woody taxa—for insights into the paradox of remote island radiations. We found that *Metrosideros* colonized each island shortly after formation and diversified within islands through recurrent selection on ancient variations that predate the radiation. Recurring use

of ancient variants may explain how genetically depauperate lineages can diversify to fill countless niches on remote islands.

Main Text

Introduction

Adaptive radiations exhibit extraordinary levels of morphological and ecological diversity (1). Although definitions of adaptive radiation vary (2–7), all center on ecological opportunity as a driver of adaptation and, ultimately, diversification (2, 8–10). Divergent selection, the primary mechanism underlying adaptive radiations, favors extreme phenotypes (11) and selects alleles that confer adaptation to unoccupied or under-utilized ecological niches. Differential adaptation results in divergence and, ultimately, reproductive isolation between populations (12). Adaptive radiations demonstrate the remarkable power of natural selection as a driver of biological diversity, and provide excellent systems for studying evolutionary processes involved in diversification and speciation (13).

Adaptive radiations on remote oceanic islands are especially interesting, as colonization of remote islands is expected to involve population bottlenecks that restrict genetic variation (14). Adaptive radiations in such settings are especially impressive and even paradoxical, given the generation of high species richness from an initially limited gene pool (15). Several classic examples of adaptive radiation occur on oceanic islands, such as Darwin's finches from the Galapagos islands (16), anole lizards from the Caribbean islands (9), Hawaiian Drosophilids (17), and Hawaiian silverswords (18), to name a few.

Recent advances in genome sequencing and analyses have greatly improved our ability to examine the genetics of speciation and adaptive radiations. By examining sequences of multiple individuals from their natural environment, it has become possible to "catch in the act" the speciation processes between incipient lineages (19). Genomic studies of early-stage speciation show that differentiation accumulates in genomic regions that restrict the homogenizing effects of gene flow between incipient species (20). The number, size, and distribution of these genomic regions can shed light on evolutionary factors involved in speciation (19). Regions of high genomic differentiation can also form from evolutionary factors unrelated to speciation, such as linkage associated with recurrent background selection or selective sweeps on shared genomic features (21, 22).

Genomic studies of lineages undergoing rapid ecological diversification have begun to reveal the evolutionary mechanisms underlying adaptive radiations. Importantly, these studies highlight the pivotal role of hybridization between populations and the consequent exchange of adaptive alleles that facilitates rapid speciation and the colonization of diverse niches (23–25). Most genomic studies of adaptive radiation involve animal systems, however, in particular, birds and fishes. In plants, genomic studies of adaptive radiations are sparse (26–28) and all examine continent-wide radiations, while there are no equivalent genomic studies of plant adaptive radiations in geographically restricted systems such as remote islands. Because the eco-evolutionary scenarios associated with adaptive radiations are diverse (5, 29), whether commonalities identified in adaptive radiations in animals (23, 30) are applicable to plants is an open question. For example, the genetic architecture of animal adaptive radiations typically involves differentiation at a small number of genomic regions (31–33). In contrast, the limited insights available for plants suggest a more complex genetic architecture (26).

We investigated the evolutionary genomics of adaptive radiation in *Metrosideros* Banks ex Gaertn. (Myrtaceae) across the Hawaiian archipelago. Hawaiian *Metrosideros* is a landscape-dominant, hypervariable and highly dispersible group of long-lived [possibly > 650 years] (34) woody taxa that are non-randomly distributed across Hawaii's heterogeneous landscape, including cooled lava flows, wet forests and bogs, subalpine zones, and riparian zones (35, 36). About 25 taxa or morphotypes are distinguished by vegetative characters ranging from prostate plants that flower a few centimeters above ground to 30-m tall trees, and leaves range dramatically in size, shape, pubescence, color, and rugosity (35, 37, 38). Variation in leaf mass per area within the four *Metrosideros* taxa on Hawaii Island alone matches that observed for woody species globally (39). Common garden experiments (38, 40–44) and parent-offspring analysis (45) demonstrate heritability of taxon-diagnostic vegetative traits, indicating that taxa are distinct genetic groups and not the result of phenotypic plasticity. *Metrosideros* taxa display evidence of local adaptation to contrasting environments (46, 47), suggesting ecological divergent selection is responsible for diversification within the group (48). This diversification, which spans the past ~3.1-3.9 million years (MY) (49, 50), has occurred in spite of the group's high capacity for gene flow by way of showy bird-pollinated flowers and tiny wind-dispersed seeds (36, 51). Lastly, the presence of partial reproductive isolating barriers between taxa is consistent with the early stages of speciation (52). Here, we generated several genomic resources

for Hawaiian *Metrosideros* and used these in population genomic analyses to gain deeper insights into the genomic architecture and evolutionary processes underlying this island adaptive radiation.

Results

An improved genome assembly for the Hawaiian *Metrosideros*. Using nanopore sequencing, an individual *Metrosideros polymorpha* var. *incana* was sequenced to 66× coverage (see Table S1 for genome sequencing statistics). The reads were assembled into a draft assembly that had high contiguity with a contig N50 of 1.85 Mbp (Table 1). We implemented Pore-C sequencing (53), which combines chromosome conformation capture with long-read nanopore sequencing, to assay the *Metrosideros*-specific chromosome contact map and anchor contigs to their chromosomal positions [see Table S2 for Pore-C sequencing statistics] (54). Using Pore-C contact maps, initial assembly contigs were scaffolded into 11 super-scaffolds (Fig. 1A) spanning 292.8 Mbps with an N50 of 25.9 Mbp. Compared to a previous genome assembly that was based only on Illumina sequencing (55), our assembly was similar in total genome size but had significantly higher contiguity. The number of super-scaffolds was consistent with the 11 chromosomes in *Metrosideros* (56). The assembly was evaluated with 2,326 Benchmarking Universal Single-Copy Ortholog (BUSCO) genes from eudicots, and 2,183 genes (93.8%) were present. These results suggest that our chromosome-level genome assembly was highly contiguous and complete. Gene annotation was conducted using nanopore sequencing of a cDNA library generated from leaf tissue (see Table S3 for cDNA sequencing statistics). A total of 28,270 genes were predicted with 94.2% of the transcripts showing an annotation edit distance (AED) of less than 0.5.

The population genomics of Hawaiian *Metrosideros* was investigated by whole-genome re-sequencing 89 individuals from the islands of Oahu and Kauai, and combining these data with previously published sequence data from Hawaii Island and Molokai (57). In addition, we sequenced 4 *Metrosideros* species outside the Hawaiian archipelago as outgroup genomes. In total, we analyzed 131 individuals belonging to 11 taxa (Taxa are abbreviated as *M. polymorpha* race B [B], *M. polymorpha* race C [C], *M. polymorpha* race F [F], *M. polymorpha* race L [L], *M. polymorpha* var. *glaberrima* [G], *M. polymorpha* var. *incana* [I], *M. macropus* [M], *M. polymorpha* var. *newellii* [N], *M. polymorpha* var. *waialealae* [WW], *M. rugosa* [R], and *M.*

tremuloides [T]) across the Hawaiian archipelago (Fig 1B). The median genome coverage was 14× per individual, and on average 93% of the sequencing reads were aligned to our reference genome (Table S4). The mapped population genomic data were used to call SNPs and after filtering there were 22,511,205 variable sites that were used for subsequent analysis.

Population structure and relationships across the Hawaiian archipelago. A neighbor-joining phylogenomic tree was reconstructed to examine evolutionary relationships among populations (Fig. 2C). The tree topology grouped individuals by island with little evidence of recent migration between islands. Within the phylogeny, individuals clustered according to taxonomic designations and were monophyletic with high confidence (>95% bootstrap support). Exceptions were paraphyletic relationships among taxa C, F, and I on Oahu and taxa G and N on Hawaii Island. Among the outgroup species, *M. vitiensis* from Fiji and American Samoa was sister to all Hawaiian *Metrosideros*.

Finer scale evolutionary relationships across Hawaiian *Metrosideros* were examined by estimating genomic ancestry proportions (K) for each individual (see Fig. S1 for K = 3 to K = 15 results). At K=3 most individuals from Kauai and Hawaii Island showed a single, island-specific ancestry component (Fig 2D). On Oahu, there were taxa with Oahu-specific ancestry (taxa C, F, and R), Kauai-specific ancestry (taxa M and T), and mixed Oahu and Kauai ancestry (taxa B, L, and I). With increasing K each taxon/population showed increasingly unique ancestry, and at K = 14, with few exceptions, individuals belonging to the same taxon/population shared a single, unique ancestry. One exception was taxon F, for which all individuals consistently showed admixed ancestry comprising components of both taxa C and R. Additionally, for taxa G and N on Hawaii Island, increasing K led to further subdivision of ancestry, consistent with the complex population substructure of Hawaii Island *Metrosideros* (57). On Hawaii Island, G belonged to two genetic groups designated G_{H1} and G_{H2} in our previous analysis, and G_{H2} represented a recently admixed population with taxon N (57). Taxon F on Oahu and population G_{H2} on Hawaii Island are likely to be hybrid populations formed from recent hybridization of genetically distinct populations (58). As such, they were removed from downstream analysis. We focused on 10 taxa/populations (B, C, G_{H1}, G from Kauai [G_K], I, L, M, N, R, and T), each of which had a largely single ancestry and sufficient sample sizes (minimum of 7 individuals in G_K and maximum of 14 individuals in taxon N) for population genomic analysis.

We investigated evidence for gene flow between taxa by calculating Patterson's D-statistics (ABBA-BABA D test) (59, 60) on all population trios following the species-wide topology (Fig. 2C). *M. vitiensis* was used as the outgroup, specifically the sample from Fiji due to its high genome coverage. Overall, 85 of the 159 (53.5%) trio topologies had significant D-statistics (Bonferroni corrected p-value < 0.05; see Fig. S2 for results of each trio). Of the 85 trios, 53 (62.4%) topologies involved admixture between taxa on different islands, indicating that reticulate evolution was pervasive within and between islands. Taxa M and T, however, were exceptions, as they had the fewest significant D-statistics and no evidence of admixture with lineages outside of Oahu (Fig. S2).

Divergence times of *Metrosideros* across the islands. We investigated the speciation history of Hawaiian *Metrosideros* through demographic modeling. The population divergence times were estimated using the generalized phylogenetic coalescent sampler (G-PhoCS) (61). From each island, a single population/taxon was chosen as a representative since we were interested in the colonization history of *Metrosideros* across the islands. Populations of the archipelago-wide taxon G were chosen from each island, except Oahu, for which G samples were not available. Thus, taxon M was chosen to represent Oahu, as it had no significant evidence of between-island admixture (Fig. S2), which reduced the number of admixture models to be tested with G-PhoCS.

Initially, we ran G-PhoCS models fitting migration bands (*i.e.* admixture) between terminal lineages. Results showed that parameterizing admixture had no effect on estimates of divergence times (Fig. S3). Given this result, we based our divergence time analysis on the simple no-migration model. The G-PhoCS-estimated divergence time between Hawaiian *Metrosideros* and the outgroup *M. vitiensis* was 4.36 million years ago [MYA] (95% Highest Posterior Density [HPD] 4.16 – 4.55 MYA), which is younger than the geological formation of Kauai (Fig. 2A). However, given that *M. vitiensis* is not the most closely related outgroup species to Hawaiian *Metrosideros* (50), colonization of the Hawaiian Islands likely occurred more recently than 4.36 million years ago. Additionally, the divergence time estimates within the Hawaiian archipelago suggested that colonization of each island occurred long after its formation. The exception was Hawaii Island for which the divergence time of *Metrosideros* pre-dated the geological formation of that island, consistent with the results of our previous study (57).

We used the program Multiple Sequentially Markovian Coalescent 2 [MSMC2] (63, 64) to estimate past changes in effective population size (N_e) for each taxon/population. Results showed that all taxa had identical trajectories in showing a decrease in N_e until ~3 MYA, followed by an increase and subsequent drop in N_e in a pattern unique to each taxon (Fig. 2B). This result suggests that all Hawaiian *Metrosideros* taxa share the same common ancestor that experienced a population bottleneck about 3 – 4 MYA when the ancestral population initially colonized the archipelago. Based on G-PhoCS and MSMC2 analysis the initial colonization of the Hawaiian archipelago was estimated to have occurred ~3 – 4.4 MYA (150,000–220,000 generations ago).

The genomic landscape of differentiation. To investigate the genetic architecture underlying the *Metrosideros* radiation we narrowed our analysis to pairs of taxa that were phylogenetic sisters, since for these sister pairs the pattern of genome-wide differentiation reflects the early stages of speciation/radiation. For these sisters we used δaδi (65) to fit 20 different demography models (see Table S5 for complete δaδi results) to find the best-fitting model. Results showed that pair G_{H1} and N and pair C and I were consistent with a speciation model in which lineage divergence occurred with continuous gene flow [*i.e.* primary gene flow] (66), while in pair B and L and pair M and T the populations have been largely isolated from each other with the exception of either a recent or ancient gene flow event (Fig 3A).

We investigated the genomic architecture of the *Metrosideros* adaptive radiation by quantifying genome-wide patterns of differentiation and signatures of divergent selection between sister taxa. We focused on differentiation (F_{ST}) outlier regions since these regions would harbor genetic variation associated with the divergence of sister pairs (67). Results showed that, in all four sister pairs, areas of high genomic differentiation were scattered across all 11 chromosomes (Fig. 3B). Pair M and T had the fewest outlier windows (52 zF_{ST} outlier windows), while taxon pairs G_{H1} and N, C and I, and B and L had over 200 zF_{ST} outlier windows each (257, 226, and 260, respectively). The median genome-wide F_{ST} between M and T ($F_{ST} = 0.16$) was more than twice the level of F_{ST} within other sister pairs (G_{H1} and N median $F_{ST} = 0.04$; C and I median $F_{ST} = 0.002$; B and L median $F_{ST} = 0.07$), suggesting that increased genome-wide differentiation between M and T due to their genetic isolation may have eroded outlier windows to undetectable levels (26, 68). The genomic positions of outlier windows generally did not

overlap across four sister-taxon pairs (*i.e.* >84% of outlier windows were found in only one taxon pair; see Fig. S4).

Evolutionary factors shaping genomic differentiation. Genomic outliers of differentiation do not always result from divergent selection (69–71), and F_{ST} can be a biased estimate of differentiation (72). Hence, as a complementary analysis we examined levels of absolute genetic divergence (D_{xy}) within differentiation outlier regions, as D_{xy} is expected to be elevated in regions under divergent selection or regions acting as genetic barriers between populations (69). Within each pair of sister taxa, D_{xy} levels were significantly elevated in differentiation outliers (MWU test p-value < 0.001) relative to genomic background (Fig. 4A left box). Further, genomic differentiation outlier regions had significantly elevated D_{xy} levels when compared with other non-sister taxa representing increasing phylogenetic distances [MWU test p-value < 0.05] (Fig. 4A right box). Interestingly, differentiation outliers had elevated D_{xy} even in comparisons with G_K on the oldest island, Kauai.

We examined evidence of selective sweeps in regions of elevated genomic differentiation using ω_{max} statistics. Results showed ω_{max} were significantly higher for differentiation outlier regions (MWU test p-value < 0.001) compared to genomic background (Fig. 4B left box). Further, the same regions had significantly elevated ω_{max} (MWU test p-value < 0.01) in other taxa as well (Fig. 4B right box), indicating that the differentiation outlier regions formed between sister pairs were also targets of positive selection in multiple other *Metrosideros* taxa.

These results suggest that an archipelago-wide selective pressure has acted recurrently on the same genomic regions across multiple *Metrosideros* taxa. Such recurrent selection can have profound effects on the genomic landscape of differentiation between populations (27, 73). In *Metrosideros* we found that π, D_{xy}, F_{ST}, and ω_{max} were all positively and significantly correlated among taxa (Fig. S5). Because archipelago-wide selection is unlikely to be involved in the early-stage divergence of lineages, we examined regions that had evidence of elevated genomic differentiation in specific lineages. Using the correlated differentiation landscape as a genomic control (21), we focused on regions that were significantly differentiated in individual sister pairs (see Table S6 for genomic positions). D_{xy} within these lineage-specific differentiation regions were significantly elevated compared to the genomic background (MWU test p-value < 0.001) [Fig. S6]. When compared to other non-sister taxa with increasing phylogenetic distances, the

lineage-specific differentiation regions still had elevated D_{xy} levels compared to genomic background (MWU test p-value < 0.01), except for the pair M and T.

A gene ontology- (GO) enrichment analysis was conducted to identify any over-represented functional categories associated with regions of genomic differentiation. Results showed that genes within the outlier windows were enriched for 19 GO terms in the 'biological process' category, and the functions were largely related to metabolic processes, cell cycle, and disease immunity responses (Fig. S7).

Discussion

The evolution of plant adaptive radiations on islands has been studied largely through phylogenetic approaches by comparing DNA sequences or genomes of a single representative from each species within the radiation (74, 75). In contrast, we investigated the evolution of *Metrosideros* by constructing a chromosome-level genome assembly and sampling genome-wide variation of 131 individuals across the Hawaiian Islands. Using a population genomics approach, we inferred the evolutionary history of Hawaiian *Metrosideros*, including population divergence times and demography associated with this island adaptive radiation. We discovered that selection on ancient variation played a major role in shaping regions of genome-wide differentiation during the recent and ancient radiation of this group across the Hawaiian Islands.

The timing of colonization of the Hawaiian Islands by *Metrosideros* may explain the contrast in current *Metrosideros* diversity and dominance on the two oldest main islands. Coalescent analysis yielded an estimate for the initial split between the common ancestor of all Hawaiian *Metrosideros* and the outgroup *M. vitiensis* at 4.4 MYA, when only Kauai, the oldest main island, was present for colonization, having completed the primary (shield-building) phase of island growth (62, 76). Phylogeography and population structure analysis indicated that the radiation was endogenous to the islands wherein colonization and subsequent lineage diversification occurred within newly formed islands in the order of their appearance. This progressive colonization from older to younger islands [*i.e.* progression rule] (77) is a hallmark phylogeographic pattern for many Hawaiian lineages (78). Our lower bound estimate of the initial colonization of the islands, however, was 3 MYA, consistent with the estimate (3.1 MYA) derived from a recent fossil-calibrated analysis of 40 nuclear genes (50). This estimate of a more recent colonization of the islands is consistent with the lower dominance of *Metrosideros* on

Kauai relative to the younger islands and the greater taxonomic richness of *Metrosideros* on islands of intermediate age, especially Oahu (35). Both of these patterns would be expected if *Metrosideros* colonized the islands more recently when Oahu was nascent and Kauai was fully formed and already colonized by lineages from the northwest Hawaiian Islands or more distant sources (*e.g.*, lobelioids (79), silverswords (80), *Psychotria* (81)). With the available data, it is not possible to determine the order or exact timing of island colonization for *Metrosideros*. Increased genomic sampling within the south Pacific island chains and along the hypothesized origins of the Hawaiian lineage (49, 50) could clarify the colonization history and ancestral diversity of this group.

The phylogenetic sister taxa that we examined inhabit contrasting environments [G_{H1} vs. N: wet forest vs. riparian; C vs. I: lower montane vs. low-elevation, dry, including new lava flows; B vs. L: cloud forest vs. upper montane; M vs. T: montane vs. steep windward slope] (35, 36), and differential local adaptation has been suggested as the primary driver of their phenotypic divergence (46, 47). The genome-wide distribution of differentiation outliers suggests that ecological diversification within *Metrosideros* either has a polygenic basis or involves multiple traits with simple genetic architectures. *Metrosideros* taxa display a rich diversity of vegetative traits that are expected to be polygenic (82, 83), consistent with the differentiation outliers that were well distributed throughout the eleven chromosomes. The number of differentiated genomic regions was not associated with migration rate between sister taxa, suggesting that differentiation outliers were not necessarily formed as barriers to gene flow (20). These patterns contrast with the genetic architecture observed for key traits in animal adaptive radiations. For instance, in butterflies (84), crows (85), and cichlids (86) where adaptive radiation involves different color morphs, one finds a small number of genomic regions that are strongly differentiated due to the simple genetic architecture of color-based traits, which also contribute to assortative mating and further build up differentiation (87). In animal adaptive radiations, in general, differentiation is seemingly localized to a few genomic regions with prominent broad peaks (31–33). In contrast, the outlier peaks in *Metrosideros* were narrow and their distribution was heterogeneous across the genome, a pattern that was also found in the continent-wide adaptive radiation of sunflowers (26). The narrow peaks suggest that fine-scale mapping of the genes underlying divergent phenotypes may be possible.

DNA divergence and selective sweep analyses of the differentiation outliers suggest a prominent role for ancient balanced polymorphisms in the radiation of *Metrosideros* in Hawaii. The differentiation outliers in *Metrosideros* showed evidence of selective sweeps in multiple taxa, suggesting that the outlier regions were targets of recurrent selection throughout the radiation. Under recurrent sweeps, however, D_{xy} levels of outliers are expected to be lower than those of the genomic background (69); yet D_{xy} levels of outlier regions in *Metrosideros* were significantly elevated. Elevated D_{xy} were even seen in comparisons with a population from the oldest island of Kauai, suggesting that these genomic regions were involved in ancient barriers to gene flow in the common ancestor, or that they harbor ancient divergent haplotypes that predate colonization of the islands (68). Long-term balancing selection or diversifying selection could have maintained increased D_{xy} within differentiation outlier regions as taxa radiated across the islands. Functionally, the enrichment of genes relating to immunity response within these differentiated regions is consistent with the hypothesis that archipelago-wide, shared balancing selection is driving the formation of these regions in *Metrosideros* (88).

Differentiated genomic regions that were lineage specific–and thus likely shaped by taxon-specific divergent selection–were also enriched for ancient variation. This finding, coupled with results from the shared differentiation outlier regions, indicates that both historical and contemporary diversification within *Metrosideros* involved selection on old divergent haplotypes. Several adaptive radiations have been shown to be facilitated by genetic variation that is older than the radiations themselves (23, 25). Selection on old variation may be a common evolutionary mechanism for both animals and plants during adaptive radiation, and may explain how, in remote islands with limited gene pools, lineages are able to repeatedly radiate into new species (15). Utilizing old variation may explain how *Metrosideros*, a long-lived tree, has been able to rapidly and repeatedly colonize novel environments, either through selection on standing ancestral variation (28, 89) or through admixture of ancient divergent haplotypes among hybridizing taxa (33).

The maintenance of ancestral polymorphisms in Hawaiian *Metrosideros* may be aided by the occurrence of a single widespread “generalist” taxon coupled with the rapid evolution of volcanic islands. *M. polymorpha* var. *glaberrima* (G), the only *Metrosideros* taxon found on all of the main islands, has high genetic variation within populations (48) and weak genetic differentiation across the islands, likely playing a role in the origin of the many island-endemic

species and varieties of *Metrosideros* (35). Observations on volcanically active Hawaii Island suggest that taxon G is the second taxon to colonize each new island, following rapidly behind early-successional *M. polymorpha* var. *incana* (I). Taxon G maintains the largest population, broadest elevation range, and most generalist niche of any Hawaii Island taxon, and as a result overlaps and hybridizes with all other (habitat-specialist) *Metrosideros* taxa on the island (48). Observations on the extinct volcanoes of the older islands reveal that where divergent selection wanes due to the cessation of lava flows, the differentiation between G and I that is observable on Hawaii Island disappears (35). More broadly, this results suggests that increased introgression would be expected between any co-occurring taxa that are isolated by incomplete intrinsic barriers (52) whenever divergent selection weakens as volcanoes subside and erode. Retention of ancient polymorphisms within taxon G through its massive population size and hybridization with habitat specialists is consistent with the transporter process model proposed by Schluter and Conte (90) and expanded to whole radiations by Martin and Richards (15). Alternating periods of divergent selection across heterogeneous environments and introgression where selection wanes on rapidly evolving volcanic islands would promote linkage among adaptive alleles. In the complex landscape of volcanic islands, this process could lead to the maintenance of linkage disequilibrium among alleles conferring adaptation to a range of environments (i.e., axes of ecological diversification), thus facilitating continued rapid radiation on each new island (15). Future research dissecting the genetic basis of the adaptive phenotype (*e.g.* GWAS (91) or QTL approaches (86)) could clarify the functionality of the ancient variations discovered in the current study and test the importance of the transporter process in the Hawaiian *Metrosideros* radiation.

Materials and Methods

Detailed description of materials and methods can be found in SI Appendix. Nanopore sequencing data are available from NCBI bioproject ID PRJNA670777. The population genomic sequencing data are available from NCBI bioproject ID PRJNA534153, specifically with the SRR identifiers SRR12673403 to SRR12673495. Data generated from this study, including the reference genome assembly, gene annotation, variant call file, and population genetics statistics, can be found at Zenodo data repository (https://doi.org/10.5281/zenodo.4264399).

Acknowledgements.

We thank the Hawaii Department of Forestry and Wildlife for permission to collect leaf samples from state forests. We also thank Jennifer Johansen, Yohan Pillon, Melissa Johnson, and Chrissen Gemmill for assistance with field collections, Tomoko Sakishima for assistance with greenhouse sample collection and DNA extractions, the College of Agriculture, Forestry, and Natural Resource Management at the University of Hawaii Hilo for greenhouse space, and Angalee Kirby for greenhouse management. We are also grateful to the Genomics Core Facility at Princeton University for sequencing support and the New York University IT High Performance Computing for supplying the computational resources, services, and staff expertise. This research was funded by National Science Foundation Plant Genome Research Program (IOS-1546218), the Zegar Family Foundation (A16-0051), and the NYUAD Research Institute (G1205) to MDP, and the National Science Foundation Faculty Early Career Development Program (DEB0954274) (PI) and Centers of Research Excellence in Science and Technology Program (HRD-0833211) (co-PI) to EAS.

References

1. R. G. Gillespie, *et al.*, Comparing Adaptive Radiations Across Space, Time, and Taxa. *J Hered* **111**, 1–20 (2020).

2. D. Schluter, *The Ecology of Adaptive Radiation* (Oxford University Press, 2000).

3. D. J. Futuyma, *Evolutionary Biology* (Sinauer Associates, 1998).

4. J. B. Losos, Adaptive radiation, ecological opportunity, and evolutionary determinism. American Society of Naturalists E. O. Wilson award address. *Am. Nat.* **175**, 623–639 (2010).

5. T. J. Givnish, Adaptive radiation versus “radiation” and “explosive diversification”: why conceptual distinctions are fundamental to understanding evolution. *New Phytol.* **207**, 297–303 (2015).

6. G. G. Simpson, *Tempo and Mode in Evolution* (Columbia University Press, 1944).

7. R. E. Glor, Phylogenetic Insights on Adaptive Radiation. *Annual Review of Ecology, Evolution, and Systematics* **41**, 251–270 (2010).

8. T. Givnish, “Adaptive radiations and molecular systematics: issues and approaches” in *Molecular Evolution and Adaptive Radiation*, T. Givnish, K. Sytsma, Eds. (Cambridge Univ. Press, 1997), pp. 1–54.

9. J. Losos, *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles* (University of California Press, 2009).

10. G. G. Simpson, *The Major Features of Evolution* (Columbia University Press, 1953).

11. H. D. Rundle, P. Nosil, Ecological speciation. *Ecology Letters* **8**, 336–352 (2005).

12. D. Schluter, Evidence for ecological speciation and its alternative. *Science* **323**, 737–41 (2009).

13. E. O. Wilson, *The Diversity of Life* (Harvard University Press, 1992).

14. R. Frankham, Do island populations have less genetic variation than mainland populations? *Heredity (Edinb)* **78 (Pt 3)**, 311–327 (1997).

15. C. H. Martin, E. J. Richards, The Paradox Behind the Pattern of Rapid Adaptive Radiation: How Can the Speciation Process Sustain Itself Through an Early Burst? *Annual Review of Ecology, Evolution, and Systematics* **50**, 569–593 (2019).

16. P. R. Grant, B. R. Grant, *How and why species multiply: The Radiations of Darwin's Finches* (Princeton University Press, Princeton, NJ, 2008).

17. H. L. Carson, K. Y. Kaneshiro, Drosophila of Hawaii: Systematics and Ecological Genetics. *Annu. Rev. Ecol. Syst.* **7**, 311–345 (1976).

18. R. H. Robichaux, G. D. Carr, M. Liebman, R. W. Pearcy, Adaptive Radiation of the Hawaiian Silversword Alliance (Compositae- Madiinae): Ecological, Morphological, and Physiological Diversity. *Annals of the Missouri Botanical Garden* **77**, 64–72 (1990).

19. J. B. W. Wolf, H. Ellegren, Making sense of genomic islands of differentiation in light of speciation. *Nature Reviews Genetics* **18**, 87–100 (2017).

20. J. L. Feder, S. P. Egan, P. Nosil, The genomics of speciation-with-gene-flow. *Trends in Genetics* **28**, 342–350 (2012).

21. R. Burri, Interpreting differentiation landscapes in the light of long-term linked selection. *Evolution Letters* **1**, 118–131 (2017).

22. M. Ravinet, *et al.*, Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology* **30**, 1450–1477 (2017).

23. D. Berner, W. Salzburger, The genomics of organismal diversification illuminated by adaptive radiations. *Trends Genet.* **31**, 491–499 (2015).

24. O. Seehausen, Conditions when hybridization might predispose populations for adaptive radiation. *Journal of Evolutionary Biology* **26**, 279–281 (2013).

25. D. A. Marques, J. I. Meier, O. Seehausen, A Combinatorial View on Speciation and Adaptive Radiation. *Trends Ecol. Evol. (Amst.)* **34**, 531–544 (2019).

26. S. Renaut, *et al.*, Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nature Communications* **4**, 1–8 (2013).

27. S. Stankowski, *et al.*, Widespread selection and gene flow shape the genomic landscape during a radiation of monkeyflowers. *PLOS Biology* **17**, e3000391 (2019).

28. J. B. Pease, D. C. Haak, M. W. Hahn, L. C. Moyle, Phylogenomics Reveals Three Sources of Adaptive Variation during a Rapid Radiation. *PLOS Biology* **14**, e1002379 (2016).

29. M. E. Olson, A. Arroyo-Santos, Thinking in continua: beyond the "adaptive radiation" metaphor. *Bioessays* **31**, 1337–1346 (2009).

30. C. R. Campbell, J. W. Poelstra, A. D. Yoder, What is Speciation Genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biol J Linn Soc* **124**, 561–583 (2018).

31. J. W. Poelstra, *et al.*, The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* **344**, 1410–1414 (2014).

32. M. Malinsky, *et al.*, Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science* **350**, 1493–1498 (2015).

33. S. Lamichhaney, *et al.*, Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**, 371–375 (2015).

34. P. J. Hart, Tree growth and age in an ancient Hawaiian wet forest: vegetation dynamics at two spatial scales. *Journal of Tropical Ecology* **26**, 1–11 (2010).

35. E. A. Stacy, T. Sakishima, Phylogeography of the highly dispersible landscape-dominant woody species complex, Metrosideros, in Hawaii. *Journal of Biogeography* **46**, 2215–2231 (2019).

36. J. Dawson, L. Stemmermann, "Metrosideros (Gaud)" in *Manual of the Flowering Plants of Hawai'i*, W. Wagner, D. Herbst, S. Sohmer, Eds. (Univ. Hawai'i Press, 1990), pp. 964–970.

37. G. L. Sur, R. Keating, N. Snow, E. A. Stacy, Leaf Micromorphology Aids Taxonomic Delineation within the Hypervariable Genus Metrosideros (Myrtaceae) on O'ahu. *pasc* **72**, 345–361 (2018).

38. E. A. Stacy, T. Sakishima, H. Tharp, N. Snow, Isolation of Metrosideros ('Ohi'a) Taxa on O'ahu Increases with Elevation and Extreme Environments. *J. Hered.* **111**, 103–118 (2020).

39. Y. Tsujii, Y. Onoda, A. Izuno, Y. Isagi, K. Kitayama, A quantitative analysis of phenotypic variations of Metrosideros polymorpha within and across populations along environmental gradients on Mauna Loa, Hawaii. *Oecologia* **180**, 1049–1059 (2016).

40. S. Cordell, G. Goldstein, D. Mueller-Dombois, D. Webb, P. M. Vitousek, Physiological and morphological variation in Metrosideros polymorpha, a dominant Hawaiian tree species, along an altitudinal gradient: the role of phenotypic plasticity. *Oecologia* **113**, 188–196 (1998).

41. S. Cordell, G. Goldstein, P. J. Melcher, F. C. Meinzer, Photosynthesis and Freezing Avoidance in Ohia (Metrosideros polymorpha) at Treeline in Hawaii. *Arctic, Antarctic, and Alpine Research* **32**, 381–387 (2000).

42. C. A. Corn, W. M. Hiesey, Altitudinal variation in hawaiian metrosideros. *American Journal of Botany* **60**, 991–1002 (1973).

43. K. Kitayama, R. Pattison, S. Cordell, D. Webb, D. Mueller-Dombois, Ecological and Genetic Implications of Foliar Polymorphism in Metrosideros polymorpha Gaud. (Myrtaceae) in a Habitat Matrix on Mauna Loa, Hawaii. *Ann Bot* **80**, 491–497 (1997).

44. L. Stemmermann, Ecological studies of Hawaiian Metrosideros in a successional context. *Pacific Science* **37**, 361–373 (1983).

45. E. A. Stacy, J. B. Johansen, T. Sakishima, D. K. Price, Genetic analysis of an ephemeral intraspecific hybrid zone in the hypervariable tree, Metrosideros polymorpha, on Hawai‘i Island. *Heredity* **117**, 173–183 (2016).

46. J. M. Ekar, D. K. Price, M. A. Johnson, E. A. Stacy, Varieties of the highly dispersible and hypervariable tree, Metrosideros polymorpha, differ in response to mechanical stress and light across a sharp ecotone. *Am. J. Bot.* **106**, 1106–1115 (2019).

47. K. R. Morrison, E. A. Stacy, Intraspecific divergence and evolution of a life-history trade-off along a successional gradient in Hawaii’s Metrosideros polymorpha. *Journal of Evolutionary Biology* **27**, 1192–1204 (2014).

48. E. A. Stacy, J. B. Johansen, T. Sakishima, D. K. Price, Y. Pillon, Incipient radiation within the dominant Hawaiian tree Metrosideros polymorpha. *Heredity* **113**, 334–342 (2014).

49. D. M. Percy, *et al.*, Progressive island colonization and ancient origin of Hawaiian Metrosideros (Myrtaceae). *Proceedings. Biological sciences* **275**, 1479–90 (2008).

50. J. R. Dupuis, *et al.*, Targeted amplicon sequencing of 40 nuclear genes supports a single introduction and rapid radiation of Hawaiian Metrosideros (Myrtaceae). *Plant Syst Evol* (2019) https:/doi.org/10.1007/s00606-019-01615-0 (October 7, 2019).

51. D. R. Drake, Seed Dispersal of Metrosideros polymorpha (Myrtaceae): A Pioneer Tree of Hawaiian Lava Flows. *American Journal of Botany* **79**, 1224–1228 (1992).

52. E. A. Stacy, B. Paritosh, M. A. Johnson, D. K. Price, Incipient ecological speciation between successional varieties of a dominant tree involves intrinsic postzygotic isolating barriers. *Ecol Evol* **7**, 2501–2512 (2017).

53. N. Ulahannan, *et al.*, Nanopore sequencing of DNA concatemers reveals higher-order features of chromatin structure. *bioRxiv*, 833590 (2019).

54. N. Kaplan, J. Dekker, High-throughput genome scaffolding from in vivo DNA interaction frequency. *Nature Biotechnology* **31**, 1143–1147 (2013).

55. A. Izuno, *et al.*, Genome sequencing of Metrosideros polymorpha (Myrtaceae), a dominant species in various habitats in the Hawaiian Islands with remarkable phenotypic variations. *Journal of Plant Research* **129**, 727–736 (2016).

56. G. D. Carr, Chromosome Numbers of Hawaiian Flowering Plants and the Significance of Cytology in Selected Taxa. *American Journal of Botany* **65**, 236–242 (1978).

57. J. Y. Choi, M. Purugganan, E. A. Stacy, Divergent Selection and Primary Gene Flow Shape Incipient Speciation of a Riparian Tree on Hawaii Island. *Mol Biol Evol* **37**, 695–710 (2020).

58. J. Mallet, Hybrid speciation. *Nature* **446**, 279–283 (2007).

59. R. E. Green, *et al.*, A draft sequence of the Neandertal genome. *Science* **328**, 710–22 (2010).

60. E. Y. Durand, N. Patterson, D. Reich, M. Slatkin, Testing for Ancient Admixture between Closely Related Populations. *Molecular Biology and Evolution* **28**, 2239–2252 (2011).

61. I. Gronau, M. J. Hubisz, B. Gulko, C. G. Danko, A. Siepel, Bayesian inference of ancient human demography from individual genome sequences. *Nat. Genet.* **43**, 1031–1034 (2011).

62. D. Clague, "The growth and subsidence of the Hawaiian-Emperor volcanic chain" in *The Origin and Evolution of Pacific Island Biotas, New Guinea to Eastern Polynesia: Patterns and Processes*, (SPB Academic Publishing, 1996), pp. 35–50.

63. S. Schiffels, R. Durbin, Inferring human population size and separation history from multiple genome sequences. *Nature Genetics* **46**, 919–925 (2014).

64. S. Schiffels, K. Wang, "MSMC and MSMC2: The Multiple Sequentially Markovian Coalescent" in *Statistical Population Genomics*, Methods in Molecular Biology., J. Y. Dutheil, Ed. (Springer US, 2020), pp. 147–166.

65. R. N. Gutenkunst, R. D. Hernandez, S. H. Williamson, C. D. Bustamante, Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* **5**, e1000695 (2009).

66. E. J. Richards, M. R. Servedio, C. H. Martin, Searching for Sympatric Speciation in the Genomic Era. *BioEssays* **41**, 1900047 (2019).

67. S. Via, Natural selection in action during speciation. *PNAS* **106**, 9939–9946 (2009).

68. F. Han, *et al.*, Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome research* **27**, 1004–1015 (2017).

69. T. E. Cruickshank, M. W. Hahn, Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology* **23**, 3133–3157 (2014).

70. M. W. Nachman, B. A. Payseur, Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philos Trans R Soc Lond B Biol Sci* **367**, 409–421 (2012).

71. M. a. F. Noor, S. M. Bennett, Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity* **103**, 439–444 (2009).

72. B. Charlesworth, Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution* **15**, 538–543 (1998).

73. R. Burri, *et al.*, Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Res.*, gr.196485.115 (2015).

74. B. Nevado, E. L. Y. Wong, O. G. Osborne, D. A. Filatov, Adaptive Evolution Is Common in Rapid Evolutionary Radiations. *Curr Biol* **29**, 3081-3086.e5 (2019).

75. M. Barrier, B. G. Baldwin, R. H. Robichaux, M. D. Purugganan, Interspecific hybrid ancestry of a plant adaptive radiation: allopolyploidy of the Hawaiian silversword alliance (Asteraceae) inferred from floral homeotic gene duplications. *Mol Biol Evol* **16**, 1105–1113 (1999).

76. J. P. Price, D. A. Clague, How old is the Hawaiian biota? Geology and phylogeny suggest recent divergence. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **269**, 2429–2435 (2002).

77. W. L. Wagner, V. A. Funk, *Hawaiian biogeography: evolution on a hotspot archipelago* (Smithsonian Institution Press, 1995).

78. K. L. Shaw, R. G. Gillespie, Comparative phylogeography of oceanic archipelagos: Hotspots for inferences of evolutionary process. *Proc Natl Acad Sci U S A* **113**, 7986–7993 (2016).

79. T. J. Givnish, *et al.*, Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *Proceedings of the Royal Society B: Biological Sciences* **276**, 407–416 (2009).

80. B. G. Baldwin, M. J. Sanderson, Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *PNAS* **95**, 9402–9406 (1998).

81. M. Nepokroeff, K. J. Sytsma, W. L. Wagner, E. A. Zimmer, 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 (2003).

82. M. C. Hall, C. J. Basten, J. H. Willis, Pleiotropic Quantitative Trait Loci Contribute to Population Divergence in Traits Associated With Life-History Variation in Mimulus guttatus. *Genetics* **172**, 1829–1844 (2006).

83. J. M. Pérez-Pérez, J. Serrano-Cartagena, J. L. Micol, Genetic Analysis of Natural Variations in the Architecture of Arabidopsis thaliana Vegetative Leaves. *Genetics* **162**, 893–915 (2002).

84. N. J. Nadeau, *et al.*, Population genomics of parallel hybrid zones in the mimetic butterflies, H. melpomene and H. erato. *Genome Res* **24**, 1316–1333 (2014).

85. N. Vijay, *et al.*, Evolution of heterogeneous genome differentiation across multiple contact zones in a crow species complex. *Nature Communications* **7**, 13195 (2016).

86. A. F. Kautt, *et al.*, Contrasting signatures of genomic divergence during sympatric speciation. *Nature*, 1–6 (2020).

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

88. D. Koenig, *et al.*, Long-term balancing selection drives evolution of immunity genes in Capsella. *eLife* **8**, e43606 (2019).

89. D. Brawand, *et al.*, The genomic substrate for adaptive radiation in African cichlid fish. *Nature* **513**, 375–381 (2014).

90. D. Schluter, G. L. Conte, Genetics and ecological speciation. *PNAS* **106**, 9955–9962 (2009).

91. Z. L. Fuller, *et al.*, Population genetics of the coral Acropora millepora: Toward genomic prediction of bleaching. *Science* **369** (2020).

Figures and Tables

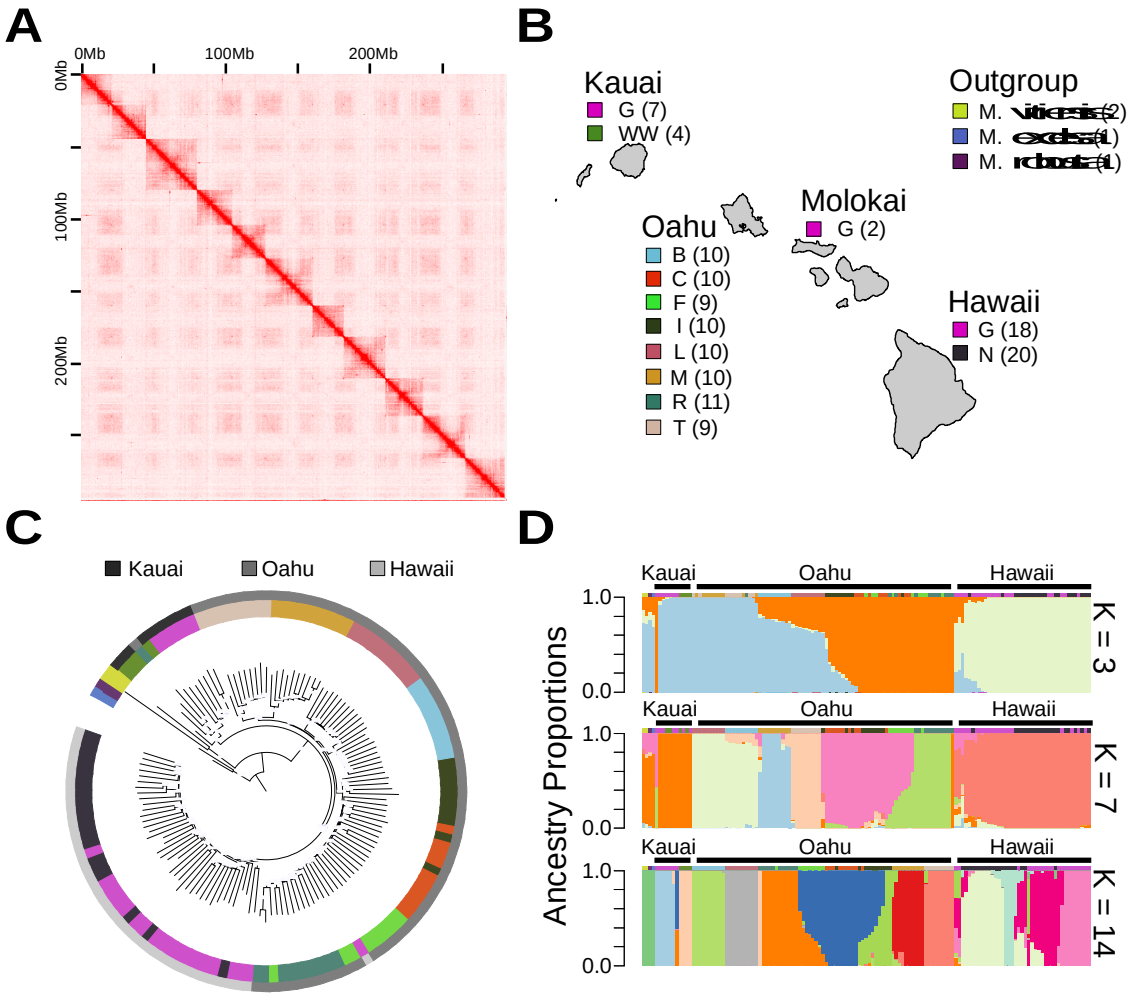

Figure 1. Genomics of Hawaiian *Metrosideros*. (A) Pore-C-based chromosome contact matrix for *M. polymorpha* var. *incana*. (B) Geographic distribution and taxon classification for the 135 samples that were analyzed in this study. Numbers in parentheses represent sample sizes. (C) Genome-wide neighbor-joining tree built using a distance matrix. Outer circle colors indicate island of origin for each sample, and inner circle colors indicate taxa as in panel (B). Nodes with greater than 95% bootstrap support are indicated with light blue circles. (D) Ancestry proportion estimates using the ADMIXTURE algorithm for K = 3, 7, and 14. Colors above admixture barplots represent taxa as in panel (B).

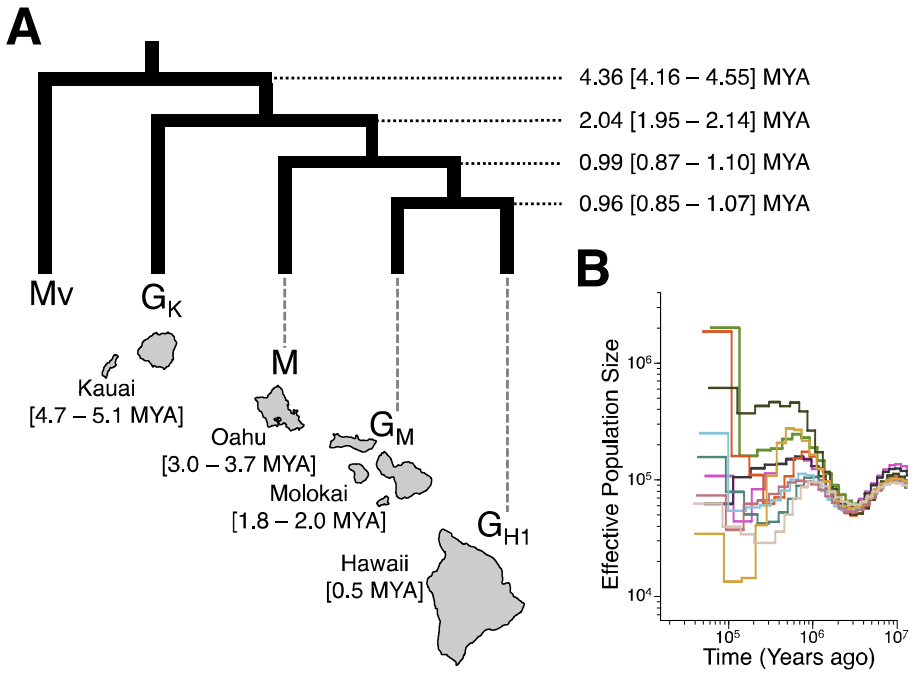

Figure 2. Divergence time and demographic history of Hawaiian *Metrosideros*. Relative times were converted to absolute times assuming a mutation rate of $7e^{-9}$ mutations per base pair per generation and a 20-year generation time. (A) G-PhoCS-estimated divergence times for representative taxa from each island (above), and time of geological formation of each island based on Clague (62) (below). (B) MSMC2-estimated effective population size for each Hawaiian taxon, color-coded as in Fig. 1B.

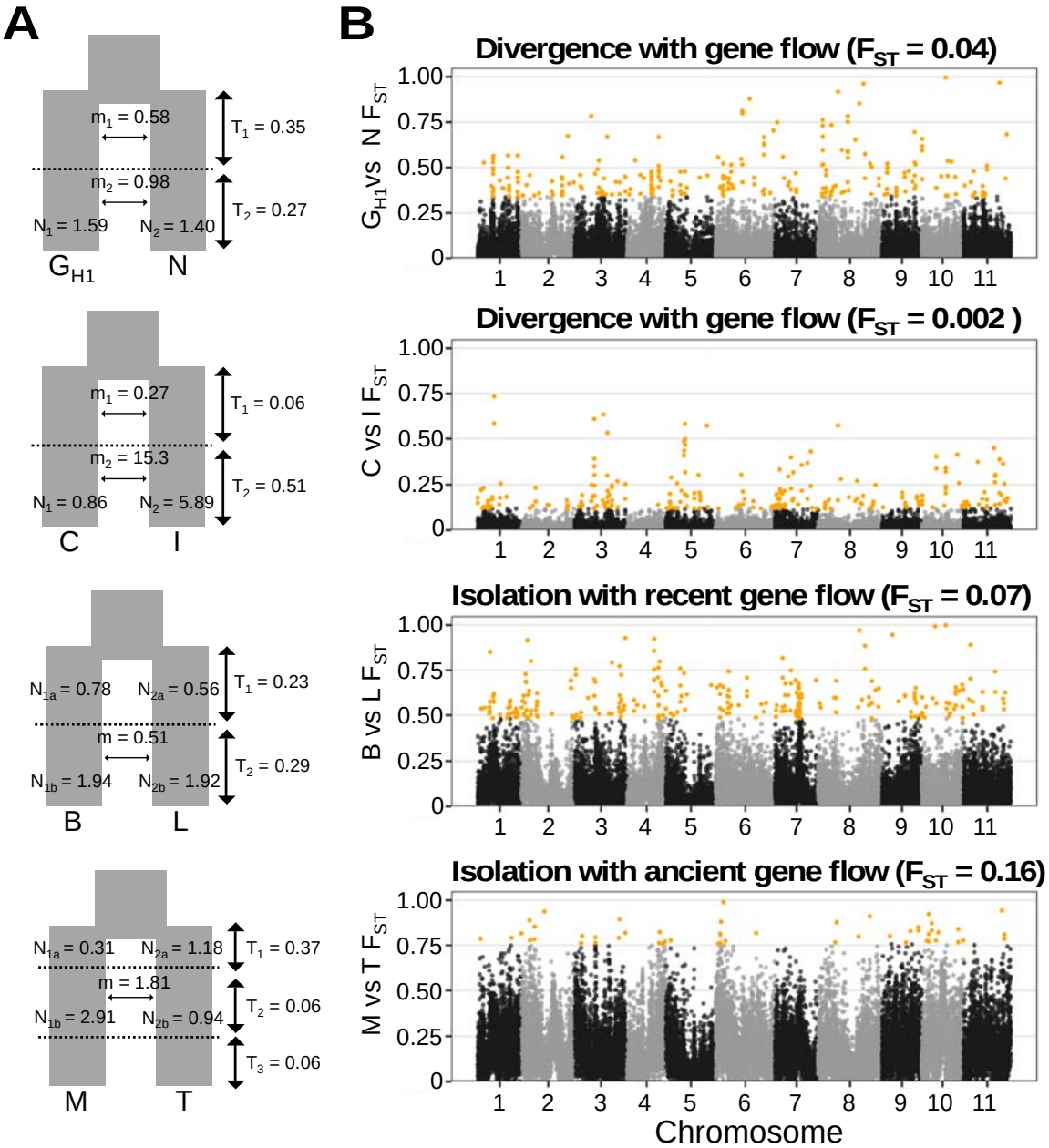

Figure 3. Genomic landscape of differentiation for the four phylogenetic sister pairs. (A) best-fitting demography model based on δaδi modeling. (B) Genome-wide F_{ST} in 10-kbp windows. Yellow dots are outliers identified with z-score-transformed F_{ST} values (zF_{ST}) > 4.

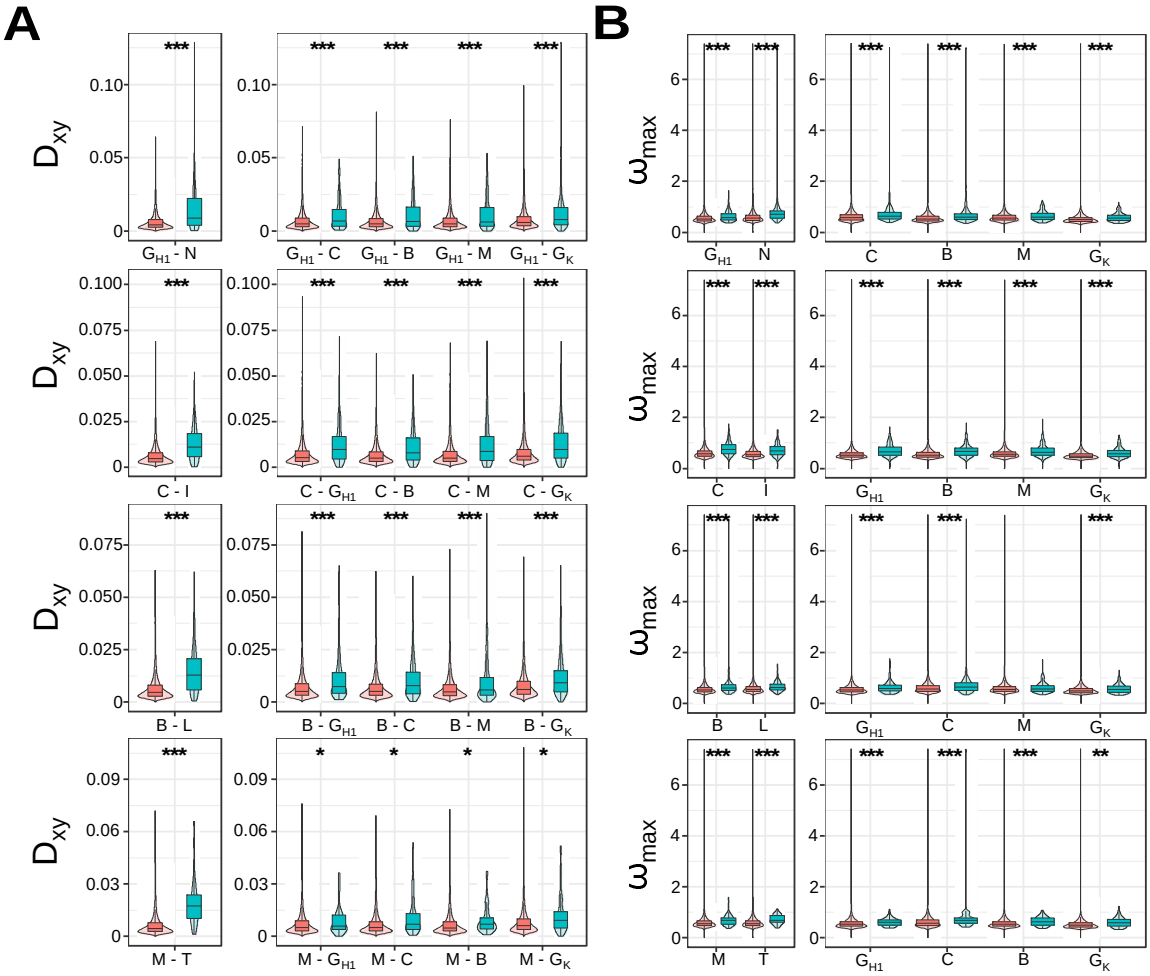

Figure 4. Sequence divergence and evidence of selective sweeps in differentiation outlier regions. Red boxes are statistics from the genomic background, and green boxes are statistics from the differentiation outlier regions. (A) Sequence divergence (D_{xy}) statistics identified in a sister pair (left window) and D_{xy} estimated from the same regions but between one member of the pair and four other taxa representing increasing phylogenetic distances (right window). (B) Selective sweep statistics (ω_{max}) identified in a sister pair (left window) and the same regions in other *Metrosideros* taxa (right window). * indicate $p < 0.05$, ** indicate $p < 0.01$, and *** indicate $p < 0.001$.

Table 1. Genome assembly statistics for *M. polymorpha* var. *incana*.

Assembly features	
# of contigs	1,035
# of scaffolded psuedomolecules	11
Total number of bases in contigs	307,350,354 bp
Total number of bases scaffolded	292,766,828 bp
Maximum contig length	8,272,703 bp
Contig N50 length	1.852 Mbp
Scaffold N50 length	25.93 Mbp
BUSCO score	93.8%
GC content	39.8%
Repeat content	36.3%
Number of annotated genes	28,270