# Evolution of *Melicope* J.R.Forst & G.Forst (Rutaceae), the largest adaptive radiation of woody plants on the Hawaiian Islands.

Dissertation

for the award of the degree

„*Doctor rerum naturalium*“ (Dr.rer.nat)

of the Georg-August-University Göttingen

within the doctoral program Biology

of the Georg-August University School of Science (GAUSS)

submitted by

Master of Science Biology

Claudia Pätzold

from Bad Dürrenberg

Göttingen, 2019

Thesis Committee

Prof. Dr. Elvira Hörandl

Systematics, Biodiversity, and Evolution of Plants (with Herbarium), Albrecht-von-Haller-Institute for Plant Sciences

Prof. Dr. Alexander Schmidt

Evolution of Landplants & Development of Terrestrial Ecosystems, Geoscience Centre, University of Göttingen (GZG)

Dr. Marc S. Appelhans

Systematics, Biodiversity, and Evolution of Plants (with Herbarium), Albrecht-von-Haller-Institute for Plant Sciences

Members of the Examination Board

Reviewer: Prof. Dr. Elvira Hörandl
Systematics, Biodiversity, and Evolution of Plants (with Herbarium), Albrecht-von-Haller-Institute for Plant Sciences

Second Reviewer: Prof. Dr. Alexander Schmidt
Evolution of Landplants & Development of Terrestrial Ecosystems, Geoscience Centre (GZG)

Further Members of the Examination Board:

Prof. Dr. Hermann Behling, Department of Palynology and Climate Dynamics, Albrecht-von-Haller-Institute for Plant Sciences

Prof. Dr. Christoph Bleidorn, Department of Animal Evolution and Biodiversity, Johann-Friedrich-Blumenbach Institute for Zoology & Anthropology

Prof. Dr. Oliver Gailing, Department of Forest Genetics and Forest Tree Breeding, Büsgen-Institute

Prof. Dr. Holger Kreft, Department of Biodiversity, Macroecology & Biogeography, Centre for Biodiversity and Sustainable Land Use

Date of oral examination: 18.02.2020

# Contents

## Abstract

Adaptive radiation describes the divergence of an ancestral taxon into multiple, phenotypically diverse species, adapted to a range of ecological niches by means of natural selection. The process is recognized as a fundamental reason for the origin of biodiversity. The main driver of adaptive radiation is ecological opportunity, though the specific agents are often poorly understood with the exception of some iconic lineages. Many well-studied adaptive radiations are island endemics, which makes island systems an ideal study system for adaptive radiation. Oceanic islands represent discrete replicates of the evolutionary process, as they are isolated, comparatively small, and often topographically complex. Species communities are formed by colonization and *in situ* diversification. The Hawaiian Islands are the most isolated archipelago on earth and home to a range of adaptively radiating lineages. The islands form as the Pacific plate passes over a magmatic hotspot with the eight current high islands originating within the last ca. 5-6 million years and the majority of the native biodiversity diverging within that time. The genus *Melicope* colonized numerous archipelagos throughout the Pacific including the Hawaiian Islands, where the lineage comprises currently 54 endemic species and represents the largest radiation of woody plants on the islands. Most species are single-island endemics and adapted to a variety of habitat types and elevational ranges. The lineage is monophyletic with an estimated crown age predating the rise of the current high islands, the oldest of which originated approximately 5 million years ago. As for many adaptively radiating lineages, phylogenetic inference based on Sanger sequencing has not been sufficient to resolve species or deeper level relationships in Hawaiian *Melicope.* Recent years have seen development of high throughput sequencing methods and their increasing application to solve recalcitrant relationships.

In this thesis, I examined the evolutionary trajectory of the Hawaiian *Melicope* adaptive radiation. I investigated the so-called 'island syndrome', which describes a set of traits commonly characterizing successful island colonizers, including recent polyploidy and shifts associated with subsequent establishment, in Hawaiian *Melicope*. I utilized restriction site-associated high throughput sequencing (RAD-seq) to reconstruct species relationships and historical biogeography in the lineage and estimate diversification rates and the impact of habitat adaption on species divergence.

RAD-seq datasets provided unprecedented resolution of species relationships in Hawaiian *Melicope*. However, the size and complexity of high throughput sequencing datasets require a high computational effort, which currently limits the applicability of algorithms for phylogenetic inference to concatenated analysis or

site-specific coalescence-based methods. I employed both methods and found them to result in incongruent relationships for the backbone of the Hawaiian *Melicope* topology. Concatenation violates the assumptions of the multispecies coalescent model, while site-based methods are statistically inconsistent but less accurate in simulated and empirical datasets. Considering the increased accuracy of the concatenated approaches as evaluated by quartet concordance methods and the synergistic effect of concatenation, I concluded that results of concatenated analysis reflect the relationships of Hawaiian *Melicope* best.

Results of flow cytometric screening of 32 Hawaiian species, representing 66% of the described diversity, and literature searches indicate that the ancestor of Hawaiian *Melicope* did not show traits associated with successful colonizers. The genus seemingly retained colonization success while exhibiting a combination of traits that typically characterize well-established island specialists. In particular, the ancestral *Melicope* colonist was not a recent polyploid. Neopolyploidy increases evolutionary flexibility and thus enhances chances for establishment and adaption. In Hawaiian *Melicope* flexibility is possibly facilitated by introgressive hybridization events.

Phylogenetic reconstruction based on RAD-seq datasets provides evidence for two ancient and several recent introgression events. Extant Hawaiian *Melicope* are divided into five fully supported main clades, two of which correspond to morphologically circumscribed infrageneric groups, whereas three morphologically defined taxonomic units are not monophyletic. All in all, 24 species were included with multiple samples, four of which were resolved as non-monophyletic. Finally, I confirmed that the *Melicope* radiation endemic to the Marquesas Islands originated from the Hawaiian radiation. These results highlight the necessity for a taxonomic revision in the lineage.

Estimated divergence times revealed that the Hawaiian archipelago was colonized prior to the origin of the current high islands. Inter-Island colonization patterns largely follow the progression rule from older to younger islands, but back colonizations to older islands occurred. Extant diversity results from recent divergence of a small number of taxa prevailing through the bottlenecks represented by the origin and colonization of the high islands. Long internal branches and estimated diversification rates indicate a high extinction rate, possibly related to the consequences of volcanic activity and the impact of glacial cycles. Consequently habitat types that are more vulnerable to climatic changes, i.e. dry ranges and bogs show high speciation and extinction rates. Increased rates of diversification are linked to habitat dissection and frequent ecological trait shifts.

# 1. | Introduction

Lava stream emitted by the Kilauea volcano, creating new landmass and flowing into the sea. Photograph: Marc Appelhans

## 1.1 Adaptive radiations

> *Seeing this gradation and diversity of structure in one small, intimately related group of birds, one might really fancy that, from an original paucity of birds in this archipelago, one species had been taken and modified for different ends.*
>
> C. Darwin (1842)

When Charles Darwin contemplated his observations of the avifauna on the Galápagos Islands (Darwin, 1842), he recognized a process that would prevail to this day and become to be regarded as one of the fundamental, if not only origin(s) of the biodiversity on earth. The concept was consolidated and christened in the modern synthesis (Dobzhansky, 1937; Mayr, 1942; Stebbins, 1951; Simpson, 1953): adaptive radiation.

Adaptive radiation describes the divergence of an ancestral taxon into multiple, phenotypically diverse species, adapted to a range of ecological niches by means of natural selection, (Simpson, 1953; Schluter, 2000). The selecting agent is the environment itself, as populations or closely related species compete for resources and develop traits to improve their exploitation while avoiding competition with each other (Simpson, 1953; Givnish, 1997; Schluter, 2000). Icons of adaptive radiation include such enigmatic lineages as Darwin's finches on the Galápagos Islands (Grant and Grant, 2002), the cichlid fishes in African rift lakes (Seehausen, 2006) or the marsupial fauna in Australia (Cássia-Silva and Sales, 2019).

Despite the prevalence of the concept spanning several decades and the ever-growing amount of research on adaptive radiation, or maybe because of it, its definition, diagnosis, and mechanisms have accumulated controversies (reviewed in Givnish, 1997, 2015; Glor, 2010; Losos and Mahler, 2010). According to the most recent major synthesis of the concept (Schluter, 2000) adaptive radiation is characterized by four features, illustrated here on Darwin's finches.

A) The component species share a common ancestor. All life on earth traces back to one common ancestor and might thus be regarded as one single adaptive radiation. However, given that the process is driven by natural selection operating on low taxonomic levels (Schluter, 2000), research of adaptive radiation in practice is often limited to clades of closely related, geographically cohesive species. Darwin's finches comprise 15 species distributed on the Galápagos Islands and derive from one common ancestor. However, they also include one derived species occurring in Cocos Island (Grant and Grant, 2002), illustrating that strict monophyly is not required for a group to represent an adaptive radiation (Schluter, 2000).

B) The species are morphologically divergent and different phenotypes correlate with features of their environments. These differences must be stable. Darwin's finches differ consistently in their body sizes and the size and shape of their beaks, even when transferred to other environments (Grant and Grant, 2008).

C) Phenotypic traits offer fitness advantages in their respective environments. This embodies the adaption component to adaptive radiation. Each species of Darwin's finches on the Galápagos Islands shows specific feeding habits. There are several insectivorous, granivorous and cactus-feeding species, each. Beak shape and sizes are adapted to the preferred food items. The three species of seed-eating finches feed on seeds, which differ in size and hardness. Beak size evolved to convey fitness advantages in specific niches defined by seed sizes (Grant and Grant, 2002, 2008).

D) Bursts of speciation associated with the process of phenotypic and ecological divergence. The ancestor to Darwin's finches arrived on the islands about 3 million years ago (mya). This is the fastest speciation rate recorded for any bird group on earth and changes in morphology can largely be traced to ecological conditions through time (Grant and Grant, 2008).

Since divergence is driven by the environment during adaptive radiation, one necessary, though not determinate, element to the process is the availability of open ecological niches. This is termed ecological opportunity and loosely defined as a cornucopia of available resources in the absence of competing taxa (Simpson, 1953). Ecological opportunity contains three distinct features, all of which have to be met by an ancestral species for adaptive radiation to occur: spatial, ecological and evolutionary opportunity (Simpson, 1953; Stroud and Losos, 2016). Spatial opportunity (also referred to as physical or geographical opportunity) specifies that an ancestral species must find itself in an area, where a range of underutilized resources exist. Ecological opportunity describes that resources cannot already be exploited by another species, or if so this species must be competitively inferior (Simpson, 1953). Evolutionary opportunity states that the ancestral species must be evolutionarily equipped to exploit the resources provided by ecological opportunity, i.e. to evolve the necessary features (Simpson, 1953).

The majority of adaptive radiations identified to date occur on islands or lakes, highlighting the importance of ecological opportunity. Examples include iconic lineages like Darwin's finches on the Galápagos Islands (Grant and Grant, 2002), the Hawaiian silverswords (Baldwin and Sanderson, 1998), or the African rift lake cichlids (Seehausen, 2006). Young islands and lakes are both characterized by a lack of inhabitants, which reduces competition for resources or pressure from predators

for newly arriving species (Carlquist, 1974). This allows radiating species to occupy ecological niches, that they might have been originally blocked from (Stroud and Losos, 2016). Mass extinctions are another possible cause of competitive release. Arguably one of the most prominent examples being the adaptive radiation of mammals following the extinction of the dinosaurs (Meredith et al., 2011). Further identified causes for adaptive radiations are key adaptions unlocking novel habitats, e.g. the evolution of CAM photosynthesis in bromeliads allowing the exploitation of drier areas (Silvestro et al., 2014).

In the Carribean *Anolis* lizards divergence follows identical trajectories on every single island following its initial colonization, producing convergent species adapted to similar niches across the islands, called ecomorphs (Losos, 1992). The emergence of a set of ectomorphs on each island seemingly follows the same progression; first divergence in body size, then in microhabitats and finally divergence along a climatic axis (Losos and Mahler, 2010). The evolution of many plant radiations on oceanic islands follows a trajectory of herbaceous colonizers to woody species (Carlquist, 1974; Dulin and Kirchoff, 2010). This indicates that the process of adaptive radiation may have overall patterns of diversification along identical axes. The identification of all elements to the selective process during adaptive radiation could then potentially explain the emergence of the majority of diversity of life.

## 1.2 Oceanic Island Systems

Oceanic Islands have long been emerging as an ideal study system for evolution, biogeography, and ecology, because they are comparatively small, have distinct boundaries and smaller species communities than continental areas. As such island species communities represent discrete replicates of the evolutionary process (McArthur and Wilson, 1967). Many well-studied adaptive radiations are island or archipelago endemics, which makes islands also an ideal study system for the process of adaptive radiation.

Oceanic islands may be broadly divided into three distinct categories: (1) continental shelf islands, situated on the same landmass as their neighboring continents, e.g. Borneo and Java. They are often connected to the continental landmasses during times of low sea levels, i.e. during glacial maxima. (2) Continental fragments; islands that were once connected to continental landmasses and drifted away due to plate tectonic effects, e.g. Madagascar. (3) Oceanic Islands, which emerge from the ocean and were never connected to a continental landmass. They are mostly volcanic in origin caused by a variety of tectonic processes, associated either with subduction or rift zones at plate margins or stationary mantel plumes, etc. (Whittaker et al., 2010).

With few exceptions volcanic islands are short-lived on geological time scales, and proceed through a typical life cycle of growth, submergence, erosion and final submergence leading to a steadily changing continuum of ecological opportunities.

Most studies about adaptive radiations on islands aim to correlate patterns of species richness and endemism to geologic and climatic island characteristics to elucidate evolutionary processes and their conditions. More than 50 years ago the seminal equilibrium theory of island biogeography provided a first conceptual framework (McArthur and Wilson, 1967). The equilibrium theory stresses the role of island area and isolation in shaping the biodiversity on islands. However, it does not adequately capture how the life cycle of volcanic islands affect biodiversity (Price and Clague, 2002; Stuessy et al., 2005). Consequently, a modification to the equilibrium theory was proposed: the general dynamic model of island biogeography (Whittaker et al., 2007, 2008).

The majority of remote volcanic islands, including the Hawaiian archipelago and the Galápagos Islands as well as the Canary Islands are shield volcanoes, a type of volcano built by effusive eruptions. With these type of eruptions highly fluid lava flows continuously in all directions gradually building up large cones with a very gentle slope (Pyle, 2015; Staudigel and Koppers, 2015). Shield volcanoes are typically a product of either volcanism related to continental rift zones or magmatic hot-spots. In both cases, the lava flow for a single caldera is transitory, lasting only while the volcano is directly above the magmatic chamber and ceases when tectonic plate movements carry the volcano away from its source.

A typical volcano ontogeny is divided into three main periods: pre-shield, shield, and post-shield stages. The pre-shield stage describes the initial, entirely submarine stages of volcanism. During the shield stage, the volcano grows to break the sea surface and towards its maximum area and height while acquiring the typical shield form (Staudigel and Koppers, 2015). Subaerial erosion and dissection due to rain and wind set in, dissecting the landscape and creating novel ecological niches. Initially, most of the inhabitant species are expected to result from immigration to the islands, with *in situ* speciation gradually increasing and reaching its maximum in the late shield stage (Figure 1.1) (Borregaard et al., 2017; Whittaker et al., 2007, 2008, 2010). Upon moving away from the magma chamber, eruptions will cease, the volcano will submerge rapidly and sometimes substantially (e.g. Moore and Clague, 1992), while continued erosion will further dissect the surface. During the post-shield stage erosion is the prominent process shaping the area resulting in further dissection and maximum topographic complexity. At this time the island reaches its peak species richness due to the increasingly dissected topography allowing adaptive radiation and allopatric speciation (Whittaker et al., 2010). (Figure 1.1). Further erosion will result in a gradual loss of height and area until finally the subaerial part of the

volcano either founders or remain as a reef (Morgan, 1996; Price and Clague, 2002; Stuessy et al., 2005).

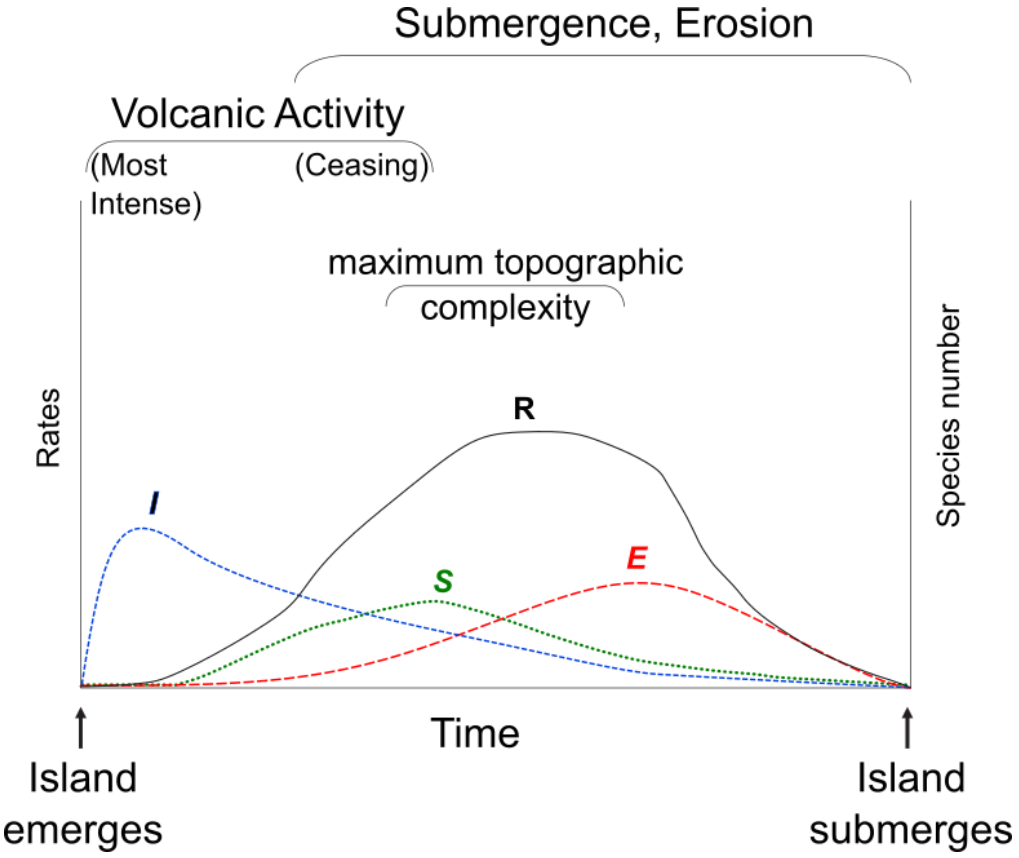



**Figure 1.1.** | Conceptual schematic of key aspects of island ontogeny shaping diversity through time. R: Absolute species numbers (solid black line; left axis). Island-level rates (dashed lines) are expressed as number of species per unit time (right axis) - *I*: immigration rate (blue); *S*: speciation rate (green); *E*: extinction rate (red). Modified from Borregaard et al. (2017).

Typically, real island ontogenies are additionally shaped by further processes, possibly involving catastrophic events like landslides and tsunamis (e.g. Moore et al., 1989; Krastel et al., 2001; Whelan and Kelletat, 2003; Le Friant et al., 2004), rejuvenated volcanism (Pyle, 2015) or fusion and fission of individual volcanos or entire islands related to sea-level fluctuations during glacial cycles (e.g. Fernández-Palacios et al., 2016; Price and Elliott-Fisk, 2004).

Oceanic islands are often clustered together into archipelagos, where each island originates from the same magmatic source but the timing of the formation varies depending on the movement of the tectonic plate traversing the volcanic hotspot (e.g. Carracedo, 1999; Price and Clague, 2002). Each island within an archipelago represents a different stage of the ontogeny and species composition trajectory. Examples include the Galápagos Islands, the Canary Islands, and the Hawaiian Archipelago. The geology of these systems has long been researched leading to a very detailed (if not yet complete) understanding of their histories (e.g. Borregaard et

al., 2017; Carracedo, 1999; Geist et al., 2014; Neall and Trewick, 2008; Price and Clague, 2002). Several theories were established aiming to explain how the biogeography of archipelagos relates to the speciation in endemic taxa, most prominently the progression rule, which states that taxa colonize the islands in the archipelago in an age-dependent pattern from older to younger with or without *in situ* speciation (Wagner and Funk, 1995).

Well resolved and time-calibrated phylogenies of island lineages are necessary to elucidate the relationship between island ecology, immigration, and adaptive radiation.

### Island Evolution

The application of molecular methods has increased our understanding of oceanic island evolution. Species communities of remote oceanic islands are not merely "downscaled" versions of those on the nearest continent, but are individual species assemblies compiled entirely from the descendants of successful colonizers (Carlquist, 1966a, 1974). In most island systems, only a scarce minority of colonizers give rise to radiations (Whittaker and Fernández-Palacios, 2010). For example, on the Hawaiian Islands, the ten largest lineages combined represent 57% of all endemic Angiosperm species, but only 4% of the successful colonizers (Price, 2004). As colonization and establishment are not easily observable processes, inferences are usually drawn from extant native island floras and their close relatives (e.g. Carlquist, 1969, 1974). Some traits have been identified as characterizing the colonizers giving rise to island radiations (e.g. Carlquist, 1974; Whittaker and Fernández-Palacios, 2010).

The most obvious trait is dispersal ability as it determines which taxa colonize at all and if so how frequently. Obviously, taxa, that fail to colonize in the first place, cannot form endemic island species. On the other hand, a high dispersal ability is not conducive to produce adaptive radiations, either. Ferns possess tiny spores that are effectively dispersed by wind over long distances. On the Galápagos Islands, less than 8 % of the native pteridophyte flora is endemic compared to almost 60 % of dicot Angiosperms (Porter, 1979, 1984). High dispersal ability and subsequently frequent immigrations to islands, maintain gene flow between source areas and island populations, thus preventing local speciation (Carlquist, 1974; Porter, 1979, 1984; Whittaker and Fernández-Palacios, 2010). Speciation on an island system seemingly requires the colonist to be in the "goldilocks zone" of dispersal ability, which must be high enough to reach island systems, but infrequent enough to prevent gene flow between remote populations. (Carlquist, 1974; Price and Wagner, 2004). This is the adaptive zone as defined by McArthur and Wilson (1967), which is a function of distance and thus highly taxon-specific. Many island lineages originate

from herbaceous colonizers as these generally have a high dispersibility (Carlquist, 1974).

Self-compatibility was argued to facilitate the establishment of colonizers on islands (Baker, 1955). Theoretically, one autogamous individual could establish a stable population, whereas an obligate outcrossing species requires the colonization of at least two specimens. Seemingly confirming this argument is the observation that the floras of New Zealand, the Hawaiian Islands, and the Galapágos Archipelago have lower proportions of self-incompatible species than the putative source continents (Carr et al., 1986; McMullen, 1987; Webb and Kelly, 1993).

Polyploidy was suggested as one of the most important traits for successful establishment and subsequent adaptive radiation on remote islands (Carr, 1998; Crawford et al., 2009). Estimates regarding the ploidy level of Angiosperm communities exist for several island systems. On the Canary Islands polyploidy was inferred to characterize 25.5% of endemic Angiosperms (Bramwell, 1976), compared to 63% in New Zealand (Hair, 1966), 66% in the Juan Fernandez Islands (Stuessy et al., 1992), and even 88% in the Hawaiian Archipelago (Carr, 1998). As island taxa often display chromosomal stasis (Stuessy and Crawford, 1998) these numbers reflect the successful establishment of polyploid colonizers. Polyploidy has been shown to offer advantages to colonizers including gene redundancy and heterosis leading to increased vitality (Comai, 2005). In addition, polyploid colonizers are expected to have a greater genetic diversity than diploids, which would increase the breadth of their response to the novel, insular environments. This effect is likely related to island characteristics like size, age or degree of isolation, as the different proportions of polyploidy on the archipelagos illustrate (Whittaker and Fernández-Palacios, 2010).

Characteristics promoting colonization of and establishment on oceanic islands do not necessarily also facilitate subsequent speciation. Comparative research of island biotas revealed common biological and niche shifts displayed by established, radiating island lineages, summarized under the term “island syndrome” (Carlquist, 1974). One component of the island syndrome is the loss of dispersal ability in both plants and animals (Carlquist, 1966b, 1966c). Members of the Asteraceae family are efficient dispersers, representing a significant fraction of the endemic species in many oceanic island systems, including Hawaii (e.g. *Bidens* L.; Crawford et al., 2009), Polynesia (e.g. *Fitchia* Meisn., Carlquist, 1974) and the Canaries (e.g. the *Sonchus* Alliance; Kim et al., 1996). Hawaiian *Bidens* likely arrived on the islands by external bird dispersal, but extant species are often characterized by the reduction or loss of achene awns - bristle or hook-like structures facilitating attachment of the fruit to feathers or hair (Carlquist, 1966b, 1974). One possible reason for the selection against dispersibility in island populations is that highly dispersive propagules or organisms

are more likely to be blown off the island, thereby getting lost from the gene pool (Carlquist, 1974).

Similarly, the initial advantage of self-compatibility is likely short-termed compared to the duration of lineage development on islands and selected against in order to avoid inbreeding depression (Barrett et al., 1996; Whittaker and Fernández-Palacios, 2010). Approximately 50% of the flora endemic to the Canary Islands shows floral features promoting outcrossing (Francisco-Ortega et al., 2000). A more striking example is the endemic flora of the Hawaiian Islands, which has the highest proportion of dioecious species world-wide (Sakai et al., 1995). The high incidence of dioecy results from both, dioecious colonists and at least 12 distinct shifts from hermaphroditic immigrants including in ancestors of several species-rich lineages (Sakai et al., 1995).

A third component of the island syndrome is insular woodiness (Carlquist, 1974). As Darwin (1859) already noted, herbaceous species make good island colonists because of their higher dispersal abilities compared to woody taxa. However, he also observed that many island taxa with exclusively herbaceous continental relatives grow as trees or shrubs (Darwin, 1859). Molecular studies have confirmed the evolution of woodiness from herbaceous ancestors in many lineages. Examples include the Hawaiian violets (Ballard and Sytsma, 2000), silverswords (Baldwin and Sanderson, 1998) and lobeliads (Carlquist, 1969; Givnish et al., 2009), *Senecio* on the Juan Fernandez Islands (Pelser et al., 2010), and the Macronesian *Sonchus* Alliance (Kim et al., 1996). On the Canary Islands, 38 distinct shifts to insular woodiness were reconstructed, representing 70% of the current woody flora (Lens et al., 2013). Several theories aim to explain this highly repetitive evolutionary shift, including that growing taller might convey a competitive advantage (Darwin, 1859), that the longevity enabled by woodiness increases the number of produces seeds or chances for reproduction when pollinators are scarce (Wallace, 1878), or that the usually less seasonal climate or the lack of large herbivores allows herbaceous colonists to grow longer and eventually become woody (Carlquist, 1974). The different hypotheses and subsequently suggested extensions (Böhle et al., 1996; Givnish, 1998) are not mutually exclusive, but experimental data supporting either one are scarce (Whittaker and Fernández-Palacios, 2010).

## 1.3 The Hawaiian Islands

The Hawaiian Islands are the most isolated archipelago on earth with a distance of >3200 km to the nearest continental landmass and still >2600 km to the nearest island system, the Marquesas Islands. The archipelago constitutes the windward islands of the larger Emperor mountain chain running in a north-western direction from the

Hawaiian Islands to a subduction zone at the border of the Eurasian plate. The islands form as the pacific plate moves over a stationary mantel plume. The hot spot has been active for ca. 85 million years (MA) leaving a string of 129 volcanos, most of which have been eroded to submarine mounts (Clague, 1996).

Twice the volcanic activity of the hot spot was reduced in combination with faster movement of the Pacific plate. As a result, there were no subaerial seamounts between 32-29 mya and in the two periods between 29-23 mya and again between 8-5 mya, only small, low-elevation and widely spaced islands existed (Clague, 1996; Price and Clague, 2002; Clague et al., 2010). The former represents the upper boundary for the age of the Hawaiian biota while the latter two represent significant bottlenecks. In accordance with the severity of the bottleneck, the majority of the Hawaiian biota is younger than 5million years (MA) (Price and Clague, 2002; Price and Wagner, 2004; Cowie and Holland, 2006).

The eight current main islands are formed by the 15 youngest islands in that chain. Their ages and sequence of formation have important implications for the biogeography and evolution in the archipelago: Kauaʻi/Niʻihau (5.8-4.3 mya), Oʻahu (3.9-1.8 mya), Molokaʻi (2.1-1.35 mya), Maui (2.0-1.1 mya), Lānaʻi (1.35 -1.3 mya), Kahoʻolawe (1.35-1.2 mya) and Hawaiʻi (since 1.1 mya) (Figure 1.2; Clague and Sherrod, 2014). Since the rise of Kauaʻi, there have always been summits over 1500 m in height, indicating a high amount of topographic complexity and ecological opportunity, which afforded opportunities for adaptive speciation processes. The islands of Molokaʻi, Lānaʻi, Kahoʻolawe, and Maui were connected for most of their existence, forming a singular island which was called Maui-Nui ("Big Maui") larger than the island of Hawaiʻi is currently. The four islands only became separated again less than 200,000 years ago. In addition, the island of Molokaʻi was connected to Oʻahu for a short period until ca. 1.9 mya (Price and Elliott-Fisk, 2004). These fusion and fission processes provided opportunities for allopatric speciation.

The Hawaiian native flora is characterized by the world-wide highest rate of endemicity, with over 90% of Angiosperms and over 70% of fern species being endemic to the archipelago. Due to the great morphological diversity displayed by extant taxa, the assumed number of colonization events per taxon to the archipelago was very high, e.g. up to five separate events for the Hawaiian lobeliads (6 genera, ca. 130 species; Campanulaceae; Givnish et al., 2009). Molecular phylogenetic research in recent years has revealed that many native lineages are monophyletic and estimations for the number of successful colonization events were reduced. As currently estimated, the approximately 1200 native flowering plant species have originated from 259 separate colonization events (Keeley and Funk, 2011; Price and Wagner, 2018).

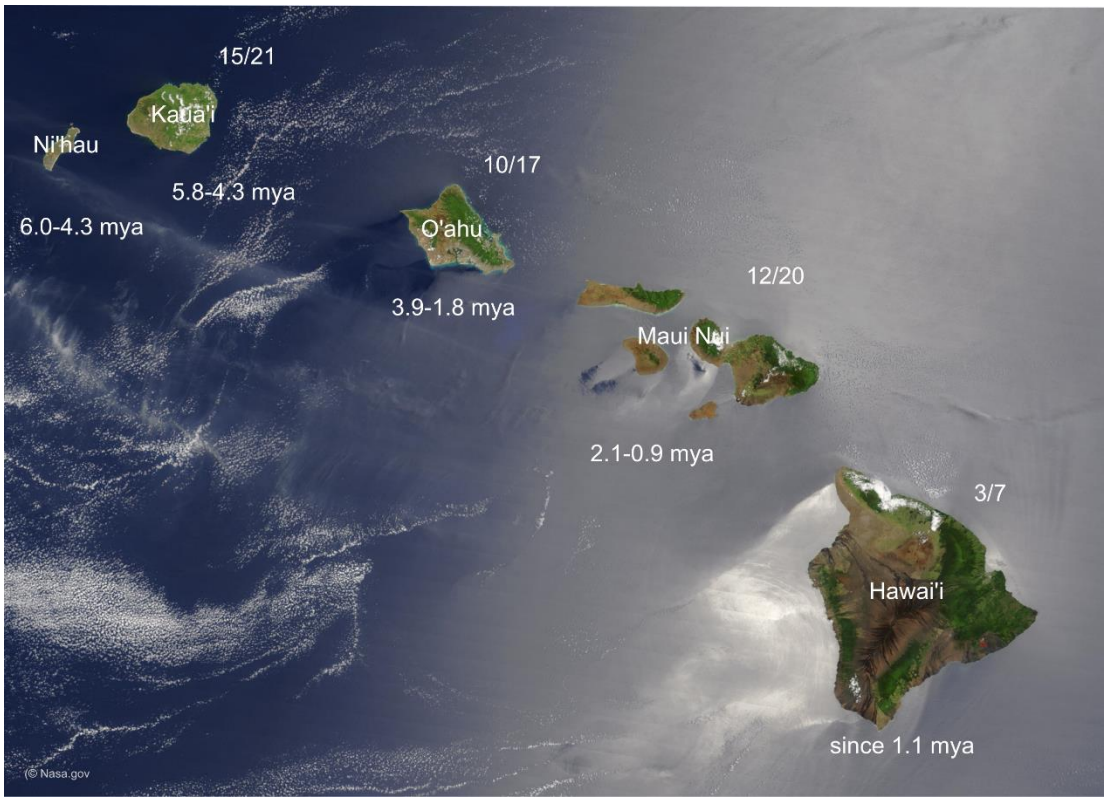



**Figure 1.2.** | Map of the Hawaiian Islands. Island names are given, with Maui Nui summarizing the Islands Molokaʻi, Maui, Kahoʻolawe and Lānaʻi. Times of subaerial growth is given below islands (Clague and Sherrod, 2014). Above islands, the number of *Melicope* species is shown (number of endemic species/total number of species per island) (Stone et al., 1999) (Photograph: nasa.gov).

Colonizers arrived exclusively by long-distance dispersal (LDD). Based on dispersal mechanisms displayed by extant native Angiosperm species, bird dispersal provided the majority of immigrants. A smaller amount of colonizers probably arrived by floating, while colonization by wind dispersal is extremely rare (Price and Wagner, 2018). Wind dispersal is, however, quite essential in ferns (Carlquist, 1966a, 1974). Successful colonizers have widespread origins, with the majority of immigrants originating in the Indo-Malayan and Pacific regions. A substantial amount of successful colonizations originated in North America and the Neotropics, respectively (Keeley and Funk, 2011; Price and Wagner, 2018). However, the origin of the majority of lineages (77% of extant species) is unknown due to limited resolution in phylogenetic analyses, insufficient taxon sampling or the fact that detailed research efforts are still lacking for them. Additionally, the origin of a substantial number of lineages could not be resolved precisely, due to widespread extant relatives or low resolution of phylogenetic relationships (Keeley and Funk, 2011; Price and Wagner, 2018). The majority of colonization events did not result in radiation on the islands, but in single endemic species (Price, 2004; Price and Wagner, 2004). With respect to the species-rich lineages, there is no apparent scale between the number of colonizers from each specific region and the subsequent radiations.

For example, there were only few successful colonizations from Eastern Asia, and only one of them radiated. This radiation, however was extensive, as it represents the Hawaiian lobeliads, the largest adaptively radiating plant lineage on any island system (Givnish et al., 2009; Price and Wagner, 2018). In contrast, three colonizers from North America radiated to a substantial degree on the Hawaiian Islands, including the ancestor to the iconic silversword alliance (Baldwin and Sanderson, 1998; Baldwin and Wagner, 2010; Price and Wagner, 2018).

Within the archipelago, the majority of radiations seems to adhere somewhat to the progression rule (Wagner and Funk, 1995; Nepokroeff et al., 2003; Dunbar-Co et al., 2008; Percy et al., 2008; Givnish et al., 2013; Landis, 2017; Johnson et al., 2019) with various proportions of inter- vs. intra-island speciation events (Price and Wagner, 2004). However, for several multi-species lineages, the biogeographic pattern is not consistent with the progression rule (Lindqvist et al., 2003; Havran et al., 2009; Morden and Harbin, 2013; Appelhans et al., 2014b; Roy et al., 2015).

While the application of molecular methods, especially Sanger-sequencing approaches have undoubtedly shed light on the evolution of many endemic Hawaiian lineages, especially with respect to establishing their monophyly (e.g. Baldwin and Sanderson, 1998; Givnish et al., 2009; Harbaugh et al., 2009; Lindqvist and Albert, 2002), resolution of relationships within each lineage is typically low or statistical support is lacking (e.g. Appelhans et al., 2014b; Baldwin and Sanderson, 1998; Cronk et al., 2005; Eggens et al., 2007; Knope et al., 2012; Nepokroeff et al., 2003; Percy et al., 2008), which prevents conclusive insights into evolution of island adaptive radiations. However, the recent advances in sequencing technologies provide the opportunity to generate genome-sized datasets to study the evolution of recalcitrant relationships (see 1.5).

## 1.4 *Melicope* J.R. Forst. & G. Forst.

The family Rutaceae in the order Sapindales contains ca. 161 genera and 2100 species (Stevens, 2001; Kubitzki et al., 2011). Rutaceae are characterized, amongst other features, by secretory civities on the leaves containing essential oils (Figure 1.3). The majority of species are woody; either trees or shrubs and very few herbaceous representatives (Kubitzki et al., 2011). Rutaceae are distributed worldwide in tropical, subtropical and (warm) temperate regions in a wide variety of habitats (Hartley, 2001; Kubitzki et al., 2011). The incidence of dioeciousness in the family is high, with 19 genera exclusively dioecious and several others containing both monoclinous and dioecious species (Kubitzki et al., 2011). One noted characteristic of the family is the large and diverse amount of secondary chemical compounds present (Price, 1963). Rutacean taxa produce a large number of structurally diverse quinolones, acridones,

coumarins, flavones, acetophenones, and limonoids as well as volatile oils, which confer the characteristic strong scent many species emit when crushing leaves.

**Figure 1.3.** | Secretory civities on leaves are the family characteristic of most Rutaceae, as here in the ornamental Calamondin (*Citrus x microcarpa* Bunge). Photograph: Claudia Paetzold

The most influental comprehensive treatment of the family recognized seven subfamilies and various tribes and subtribes (Engler, 1931) based on morphology, mainly fruit type. However, the classifications were increasingly disputed by results of phytochemical (Waterman, 1975; da Silva et al., 1988) and DNA sequence data (e.g. Appelhans et al., 2011; Bayly et al., 2013; Chase et al., 1999; Groppo et al., 2008; Poon et al., 2007) revealing non-monophyly of most subfamilies or tribes. In this thesis, I adopted the most recent classification proposed by Morton and Telmer (2014) due to the high statistical support of relationships recovered and phytochemical, karyological or morphological synapomorphies characterizing each taxon. The family is subdivided into four subfamilies: the basal Cneroideae Webb (8 genera), Amyridoideae Link (105-108 genera) sister to the sister-groups Rutoideae Arn. (6-7 genera), and Aurantoideae Horan. (24-26 genera).

*Melicope* J.R. Forst & G. Forst is the largest genus in Rutaceae containing some 230 species (Kubitzki et al., 2011) and was resolved in the Amyridioideae subfamily (Morton and Telmer, 2014). Kubitzki et al. (2011) placed the genus into the so-called

'*Euodia*-Alliance' together with 30 other genera, which are linked by morphological, chemotaxonomic and early molecular evidence, but refrained from proposing relationships between alliances or clades due to lack of evidence.

The genus contains mostly shrubs and trees with trifoliate, unifoliate or simple leaves (Figure 1.4a-c), bisexual or functionally unisexual flowers (Figure 1.4d-f), and exclusively dehiscent fruits (Hartley, 2001). The fruit is comprised of up to four basally connate follicles, to sub-syncarpous or syncarpous capsules. The seeds are persistent and prominently displayed upon dehiscence of the fruit (Hartley, 2001). The seed is covered by a shiny black pellicle, covering the spongy sarcotesta and the thick and robust sclerotesta. Fruit and seed anatomy have been interpreted as an adaption to bird dispersal with the sarcotesta representing the reward for the vector and the sclerotesta offering protection from the digestive system (Hartley, 2001; Kubitzki et al., 2011). *Melicope* species are distributed throughout the Pacific from New Zealand to Hawaii, from Japan throughout Southeast Asia to Madagascar and the Mascarene Islands. However, the majority of species are endemic to comparatively small regions, with only a few widespread representatives (Hartley, 2001).

The most recent taxonomic revision of the genus recognizes four sections diagnosed mainly by the type of seed attachment in the dehisced fruit (type A: partially detached pericarp strip or raphe or type B: no detached pericarp strip or raphe), the trichomes (simple or compound), the number of stamens (4 or 8), the adnation of the endocarp to the mesocarp, and carpel connation (Hartley, 2001).

*Melicope* sect. *Melicope* contains 38 species characterized by strictly follicular fruits, seed attachment type A and a separate endocarp, while stamen number and trichome structure are variable. The section is distributed (Hartley, 2001) mainly in Australasia with some species endemic to India and the Pacific to the outer Melanesian Islands and Tahiti (Hartley, 2001). *Melicope* sect. *Vitiflorae* T.G.Hartley is characterized by strictly four stamens and simple trichomes as well as basally connate follicles, seed attachment type A, and a separate endocarp. The section is the smallest containing only 8 species distributed in Australasia and the Pacific Islands from the Austral to the Society Islands (Hartley, 2001). *Melicope* section *Lepta* (Lour.) T.G.Hartley is the largest section containing 102 species distributed throughout the Pacific Islands up to the Society Islands, Australasia, Malesia, and Madagascar and the Mascarene Islands. The section is characterized by possessing exclusively 4 stamens, seed attachment type B, an endocarp adnate at the middle or towards the apex and fruits displaying the entire range of carpel connation. The section also contains both, monoclinous and dioecious species (Hartley, 2001; Appelhans et al., 2014a). Lastly, *Melicope* sect. *Pelea* (A.Gray) Hook comprises more than 80 species characterized by 8 stamens, the entire range of carpel connation, the endocarp being adnate ate least at base, middle or

apex and showing seed attachment type A or B. The section is mainly distributed on Island systems throughout the Pacific including Oceania, Melanesia, New Caledonia, the Hawaiian, and Marquesan Islands, thus representing the westwards maximum extension of the genus entire (Hartley, 2001). The section is named for the Hawaiian volcano goddess Pele. It comprises all Hawaiian species of the genus, which were originally regarded as a Hawaiian endemic genus *Pelea* and later reduced into *Melicope* (Hartley and Stone, 1989).

Molecular systematics has also reshaped our understanding of the relationships of *Melicope*. The '*Euodia*-Alliance' was shown to be not monophyletic (Poon et al., 2007; Groppo et al., 2008; Bayly et al., 2013). *Melicope* was resolved as closely related to some of the genera from the '*Euodia*-Alliance', notably *Euodia* J.R.Forst. & G.Forst. and *Acronychia* J.R.Forst. & G.Forst. as well as several other genera originally thought to belong to another alliance altogether (Bayly et al., 2013). In accordance with Appelhans et al. (2014a), I will refer to these putatively closely related genera as *Acronychia-Euodia-Melicope* group.

Several molecular phylogenies were generated to elucidate relationships between genera in the *Acronychia-Euodia-Melicope* group. The data sets contained up to seven nuclear and chloroplast coding and non-coding marker regions from up to 164 species from 26 genera and covering the entire geographic range of the group. The results show that the genera in the *Acronychia-Euodia-Melicope* group are not monophyletic in their current circumscription (Appelhans et al., 2014a, 2014b; Holzmeyer et al., 2015). The group is broadly divided into two main clades; the species-poor *Euodia* clade and the species-rich *Acronychia-Melicope* clade, which contains all but one described *Melicope* species included in the data sets. Several smaller genera are nested within *Melicope* and need to be reduced to achieve monophyly (Appelhans et al., 2014a, 2014b; Holzmeyer et al., 2015).

On the subgeneric level only section *Lepta* was inferred to be monophyletic; the remaining three sections were all revealed to be paraphyletic (Appelhans et al., 2014a, 2014b). The *Acronychia-Melicope* clade was divided into four clades with the earliest diverging clade comprising some species of section *Melicope*, the majority of section *Vitiflorae*, the New Caledonian species of section *Pelea* and four smaller genera (Appelhans et al., 2014a). The second clade comprises the genera *Acronychia* and *Maclurodendron* intermingled (Appelhans et al., 2014a; Holzmeyer et al., 2015). The third clade comprises all species of section *Lepta*, but the different reproductive systems are each not monophyletic (Appelhans et al., 2014a). Finally, the fourth clade comprises the remaining species of section *Melicope* in two subclades, the remaining species of section *Pelea*, as well as the genus *Platydesma* as sister to the Hawaiian *Pelea*. Consequently, the genus was merged into *Melicope* (Appelhans et al., 2017).

Members of the *Acronychia-Melicope* clade are efficient dispersers, colonizing almost all Pacific islands systems and other remote regions, sometimes repeatedly (Appelhans et al., 2018b). The majority of colonization events occurred comparatively recently in the Pleistocene (Appelhans et al., 2018b). Colonization is often followed by rapid adaptive radiation leading to high amounts of species diversity and endemism in the group. The most species-rich radiation within the lineage is the Hawaiian *Melicope* clade (Appelhans et al., 2018b).

### Hawaiian *Melicope*

Hawaiian *Melicope* currently comprises 54 accepted species (Stone et al., 1999; Harbaugh et al., 2009; Wood et al., 2016, 2017; Appelhans et al., 2017) endemic to the Hawaiian Islands, where they are traditionally called *Alani* or *Pilo kea*. The majority of the species are single island endemics, with only eight species occurring on multiple islands, *M. clusiifolia* (A.Gray) T.G.Hartley & B.C.Stone, *M. elliptica* (A.Gray) T.G.Hartley & B.C.Stone, *M. hawaiensis* (Wawra) T.G.Hartley & B.C.Stone, *M. ovata* (St.John & E.Hume) T.G.Hartley & B.C.Stone, *M. spathulata* A.Gray, *M. pallida* (Hillebr.) T.G.Hartley & B.C.Stone, *M. peduncularis* (H.Lév) T.G.Hartley & B.C.Stone and *M. pseudoanisata* (Rock) T.G.Hartley & B.C.Stone (Figure 1.2; Stone et al., 1999; Appelhans et al., 2017). They occur in a variety of habitats, from dry to wet forests, bogs, and even subalpine shrublands, and in elevations from 360-2073m (Stone et al., 1999; NTBG, 2019). As such they represent the third largest adaptive radiation on the Hawaiian Islands after the Hawaiian Lobelioids (Givnish et al., 2009) and *Cyrtandra* J.R.Forst & G.Forst (Gesneriaceae; Lorence and Perlman, 2007). It does, however, represent the largest radiation of exclusively woody plants (Figure 1.4a-c) on the islands.

Before the reduction into *Melicope*, the Hawaiian species of *Pelea* were divided into four sections, *Apocarpa* B.C.Stone, *Cubicarpa* B.C.Stone, *Megacarpa* B.C.Stone and *Pelea* (Stone et al., 1999), which are still used as informal taxonomic groups and I will refer to them as "Stone's sections" herein. Stone's sections are mainly characterized by the fruit morphology and leaf position (Figure 1.4g-j). The fruits of Stone's section *Apocarpa* present four distinct follicles (Figure 1.4g), whereas the remaining three sections show capsules with a varying degree of carpel connation. Stone's section *Megacarpa* is characterized by the carpels being connated up to 2/3 of their overall length (Figure 1.4h). Consequently, Stone's section *Cubicarpa* presents carpels connated (nearly) completely (Figure 1.4j). Finally, Stone's section *Pelea* is characterized by whorled leaves (Figure 1.4c), in contrast to the remaining sections, which show opposite leaves (Figure 1.4a, b; Stone et al., 1999). The latest taxonomic treatment, which included these sections was considered provisional by the authors due to the variability of diagnostic characters (Stone et al., 1999).

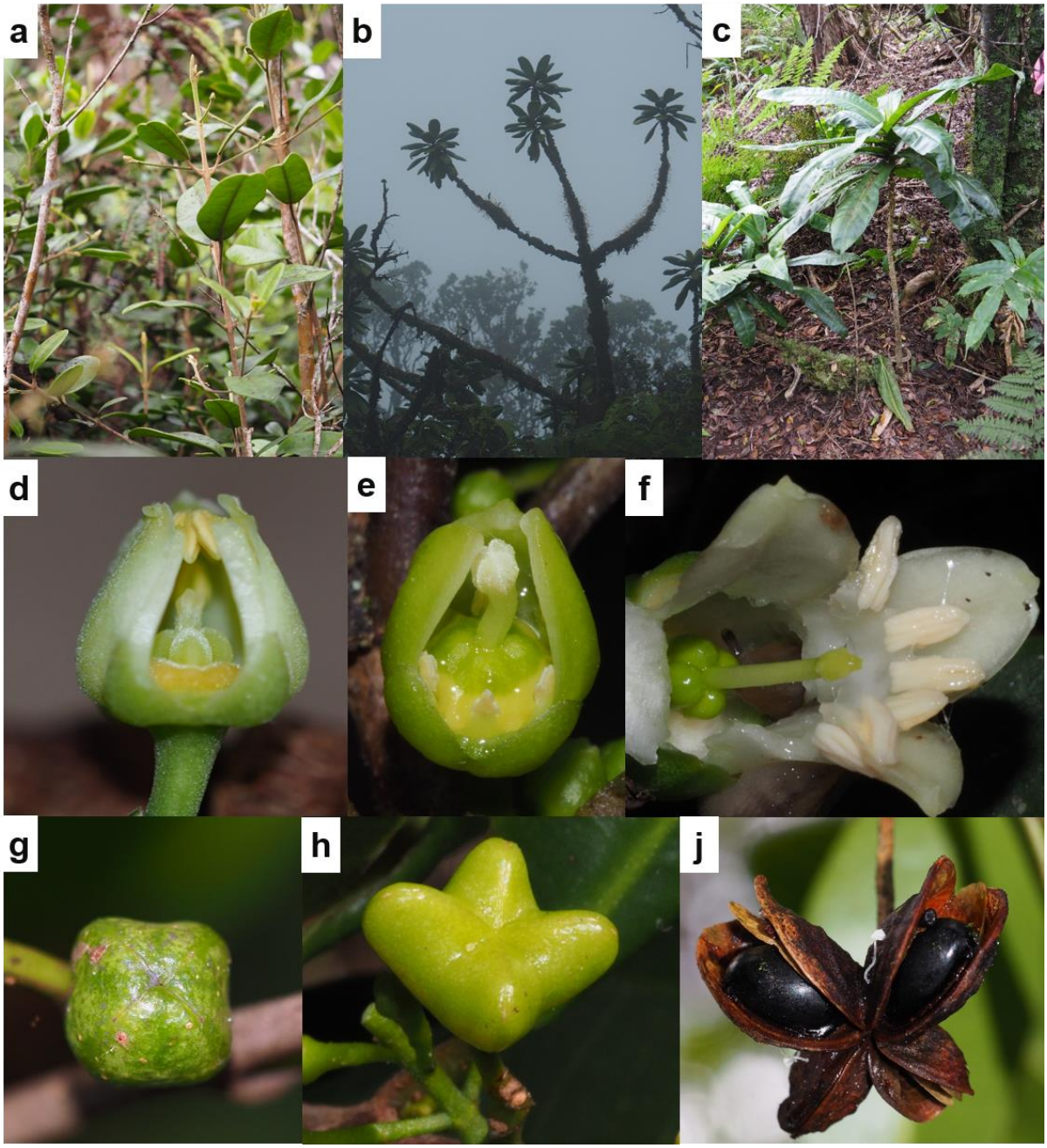



**Figure 1.4.** | Examples of habit (a-c), flowers with a petal removed (d-e), and fruits (g-f) in Hawaiian *Melicope*. (a) shrub-like habit in *M. kavaiensis*. (b) tree-like habit of *M. clusiifolia*. (c) unbranched shrub-like habit in *M. spathulata*. (d) functionally male flower in *M. hawaiiensis*. (e) functionally female flower in *M. clusiifolia*. (f) hermaphroditic flower in *M. spathulata*. (g) "cubicarpic" fruit in *M. anisata*. (h) "megacarpic" fruit in *M. feddei*. (f) apocarpous open fruits, showing the black seeds in *M. sessilis*.

A first molecular systematic effort in the group showed the congenerity of the genus *Platydesma* with *Melicope* but taxon sampling was insufficient to address subgeneric or interspecies relationships (Harbaugh et al., 2009). *Platydesma* had been described as a separate genus resulting from a different colonization event due to its perfect flowers (Mann, 1866) and a separate colonization event was proposed for its origin (Sakai et al., 1995; Kubitzki et al., 2011). A second molecular phylogeny with an

increased taxon sampling and based on seven marker regions provided first insights into the evolution of the lineage (Appelhans et al., 2014b). Hawaiian *Melicope* samples were resolved in six distinct, highly supported clades. The species of *Platydesma* were confirmed as the earliest diverging lineage within the radiation and thus representing a reversal from dioecy (Figure 1.4d, e) to bisexual flowers (Figure 1.3.f; Appelhans et al., 2014b, 2017). Species of Stone's section *Apocarpa* were resolved in three different clades, with one species, *M. elliptica* (A.Gray) T.G.Hartley & B.C.Stone forming a monotypic sister group to a clade comprising all species of Stone's sections *Cubicarpa* and *Megacarpa*. *Pelea* was the only Stone's section resolved to be monophyletic. Unfortunately, the informative content of the marker set was not sufficient to resolve relationships between clades or interspecies relationships (Appelhans et al., 2018b). Consequently, biogeographic patterns within the islands could not be tested in detail. However, the molecular evidence shows that the Marquesan *Melicope*, comprising seven species, are nested within the Hawaiian clade (Harbaugh et al., 2009; Appelhans et al., 2014b). These seven species likely represent two independent colonization events from the Hawaiian Archipelago, crossing a distance of over 3800km leading to an adaptive radiation on the Marquesas Islands (Appelhans et al., 2014b, 2018b).

Divergence time estimation revealed that the most recent common ancestor (MRCA) to the lineage most likely predates the origin of the current Hawaiian main islands. Hence, the group is one of a sparse selection of Hawaiian endemics to pass through the bottleneck in the period of 8 – 5 mya, when no high islands existed in the archipelago (Price and Clague, 2002; Appelhans et al., 2018b). The island of Kaua'i harbors the highest number of species with progressively fewer species occurring on younger islands (Figure 1.2). Kaua'i species were resolved in all major clades within Hawaiian *Melicope* indicating diversification might follow the progression rule (Appelhans et al., 2014b; Wagner and Funk, 1995; see 1.2) All in all Hawaiian *Melicope* represent an ideal model system to study adaptive radiation. However, reliably testing biogeographic and diversification patterns within the lineage requires a fully resolved phylogeny.

## 1.5 High-Throughput-Sequencing in Systematics

The development of the Polymerase Chain Reaction (Mullis et al., 1986) and Sanger-sequencing (Sanger et al., 1977) and its subsequent application to evolutionary research has drastically increased our knowledge regarding evolutionary patterns. For over 30 years these methods provided insights into diverse taxa across the tree of life including many adaptive radiations and island lineages. However, the comparatively small number of genomic marker regions generated by that approach often contains insufficient variation to resolve relationships on shallow taxonomic

levels or in rapid divergences (e.g. Appelhans et al., 2014; Baldwin and Sanderson, 1998; Cronk et al., 2005; Eggens et al., 2007; Knope et al., 2012; Nepokroeff et al., 2003; Percy et al., 2008).

The development of high-throughput sequencing (HTS; Next Generation Sequencing (NGS)) methods in the wake of the human genome project (Collins et al., 2003) might prove to have an even bigger impact on evolutionary research than Sanger biochemistry had three decades prior. HTS provides the opportunity to sequence millions of base pairs and thus obtain sufficient amounts of information to resolve even the most recalcitrant relationships. Sequencing costs per base have dropped drastically, amounting to around $1000 for a human genome (Wetterstrand, 2019). However, in systematic research, sequencing and assembly of entire genomes for an entire lineage of interest is generally too costly and labor-intensive, especially since the genomes of many organisms are bigger and more complex compared to humans (e.g. Bennett and Leitch, 2007; NCBI Resource Coordinators, 2018). Instead, a representative subset of the genome is targeted and sequenced with HTS techniques, providing hundreds or thousands of genomic loci at an improved cost-benefit ratio.

Consequently, a range of reduced representation strategies for sampling a representative fraction of the genome were developed. For hypothesis-driven application of HTS to systematics, RNA sequencing (RNAseq, transcriptomics) was initially the most popular approach (Zimmer and Wen, 2015). Transcriptomic approaches reduce genomic complexity by sequencing only the set of transcripts in a sample, most commonly messenger RNA (Wang et al., 2009). The approach requires living tissues from the same organs and identical developmental stages to ensure a wide overlap of expressed genes (Zimmer and Wen, 2015). In addition, due to the conserved nature of coding regions, the method is best used at deep and possibly medium phylogenetic scales (Zimmer and Wen, 2015). The plasticity of eukaryotic, especially plant genomes with the common occurrence of duplication events, makes orthology inference the most critical step (Yang et al., 2018).

Among reduced representation methods, target enrichment (TE; target capture, anchored hybrid enrichment) and restriction site-associated DNA sequencing (RAD-seq) are currently most commonly used for systematic research. The phylogenetic range of applicability for either method is still under exploration (e.g. Cariou et al., 2013; Leaché et al., 2015b).

Target enrichment uses biotinylated ‘baits’ or ‘probes’ (70-120 bp RNA oligonucleotides) to capture genomic fragments via hybridization to complementary sections (Gnirke et al., 2009). Captured fragments are then isolated and sequenced by

HTS (Gnirke et al., 2009). The approach is theoretically applicable at all taxonomic levels, provided orthologous genomic regions are identified in the taxon of interest showing a rate of variability matching the research hypothesis. Consequently, TE methods require the presence of more or less closely related reference genomes and/or transcriptomes and intensive bioinformatics efforts to identify suitable genomic regions (Mayer et al., 2016). In general, TE methods are mostly applied at deep to medium taxonomic levels (e.g. Leaché et al., 2015b).

RAD-seq targets genomic DNA adjacent to a restriction enzyme recognition site. Genomic DNA is digested and fragments of a targeted size subjected to high throughput sequencing (Miller et al., 2007; Baird et al., 2008). The method allows sampling regions over the entire genome and no reference is needed. Mutations may affect enzyme recognition sites and cause locus dropout with increasing phylogenetic distance, and limit the applicability of the method to shallower phylogenetic levels (Ree and Hipp, 2015). To date, phylogenetic RAD-seq studies have been primarily employed in population-level research or in small sets of closely related species (Ree and Hipp, 2015; Díaz-Arce et al., 2016; Hodel et al., 2017). However, simulated RAD investigation using *Drosophila,* mammals and yeast genomes revealed the potential applicability of the method in groups aged up to 60 MA (Rubin et al., 2012). Since then, the application at deeper taxonomic levels has increased (Hipp et al., 2014; Leaché et al., 2015b; Eaton et al., 2017). A range of strategies modifying the original RAD-seq approach (Baird et al., 2008) have been proposed, including the so-called 'genotyping-by-sequencing' (GBS) mathods. The individual approaches differ mainly in the number of digestion enzymes (two in ddRAD), the frequency of their targeted cut sites (ezRAD), the approach to shearing (original RAD), and the approach for size selection (Davey et al., 2011; Puritz et al., 2014). In addition, a strategy aiming combine the advanteges of RAD-seq and TE approaches by employing RAD loci as baits for hybridization has been proposed: hyRAD (Suchan et al., 2016).

One of the key issues of using RAD-seq is the assembly of reads into a matrix of orthologous loci. In rare cases, a closely related reference genome exist, which can be employed to guide the assembly (Ree and Hipp, 2015). In a majority of taxa the assembly is achieved *de novo* purely by bioinformatic means (e.g. Catchen et al., 2011; Eaton, 2014; reviewed in Ree and Hipp, 2015). The basic outline for *de novo* RAD-seq assembly can be summarized in three steps:

1) Identification of loci within samples. Reads of individual samples must be collected into clusters (also 'stacks') representing genomic loci. Clustering must account for sequencing errors and allelic variation. Finally, the consensus sequence of each locus is called (genotyping).

2) Orthology assessment between samples. Consensus sequences from individual samples are compared across samples and grouped into clusters representing putative orthologous loci. If multiple sequences from the same sample are clustered into the same locus, this can indicate duplication events or repetitive regions and these loci should be discarded to avoid comparing paralogs (Eaton, 2014). The algorithms and parameters governing this step have a deep impact on the final matrix and any subsequent analysis. Setting the allowed genetic distance between sequences too high leads to overmerging of loci, i.e. grouping of paralogous regions into one cluster. Setting the allowed distance too low results in undermerging, i.e. loci actually representing one ortholog being split into different clusters, each with a potentially reduced sample coverage. A conclusive strategy for identifying the best clustering threshold is still lacking but under active research (Mastretta-Yanes et al., 2015; Paris et al., 2017; Suchan et al., 2017). Finally, putatively orthologous sequences are aligned.

3) Assembling the data matrix. By definition, RAD matrices are incomplete, as most of the identified loci will not be present in all samples. Assembling the final matrix requires considering that including more loci means also including more missing information. The criterion for including a locus is sample coverage; i.e. a locus is included in the final matrix if it is recovered from at least a specified number of samples. Several studies have evaluated the trade-off between the number of loci and amount of missing data and its impact on the resolution of trees (Wagner et al., 2013; Hipp et al., 2014; Eaton et al., 2017).

Several bioinformatics pipelines have been developed to implement these general steps. Each pipeline is usually focused on a specific research approach either regarding taxon set, RAD-method, implementation of clustering steps or available computational resources. However, most of the software is targeted at population level genomics. The most commonly used pipeline (Ree and Hipp, 2015; Andrews et al., 2016) is *stacks*, explicitly designed for "genetic analysis of crosses or populations" (Catchen et al., 2011, 2013), i.e. of highly similar individuals. In the presence of indel variation, *stacks* tends to undermerge loci when clustering between samples and thus fails to accurately detect orthologs (Eaton, 2014). In general, *stacks* is likely best suited for shallow taxonomic scales. The program *pyRAD*, on the other hand, is designed for phylogenetic research. The identification of genomic loci is implemented as a global alignment clustering algorithm, which allows for indel variation while identifying homology and for more efficient computation time compared to *stacks* (Eaton, 2014). Detection of paralogs is performed during both within and between samples by application of a number of explicit filters regarding a number of allowed

heterozygous sites in a locus or number of haplotypes. In general, *pyRAD* has a better ability to detect homologs in phylogenetic RAD-seq sampling (Eaton, 2014).

RAD-seq datasets have some unique features compared to other HTS or traditional approaches. As RAD-seq is typically combined with short-read HTS, individual loci are only about 100-250 bp in length. Due to the shortness RAD loci are expected to carry little individual phylogenetic signal. This characteristic is expected to decrease with further advances in sequencing technology or the application of long-read sequencing platforms (Ree and Hipp, 2015). Compared to TE and transcriptomic approaches, the number of loci is increased by at least an order of magnitude, as RAD-seq approaches may yield hundreds of thousands of loci. These loci originate from the entire with of the genome and are thus expected to provide a genome-wide view of sequence divergence. Finally, RAD matrices contain substantial amounts of missing data, as sampling errors and mutation of restriction sites cause locus dropout in samples. In a nutshell, RAD-seq datasets are very large, incomplete by nature and contain short individual loci (e.g. Eaton et al., 2017; Harvey et al., 2016; Mastretta-Yanes et al., 2015). These characteristics pose some limitations with respect to which methods of species tree inference are applicable. The shortness of individual loci and the amount of missing data are not well tolerated by algorithms using gene resolved trees to infer the species tree, (e.g. Larget et al., 2010; Liu et al., 2010; Liu and Yu, 2011; Mirarab et al., 2014a; Mirarab and Warnow, 2015). In addition, most algorithms scale badly to the sizes of RAD-seq datasets. The most frequently used and arguably most practical approach is concatenating all loci into a single matrix, which is then used for tree inference using Maximum Likelihood or Bayesian approaches (Ree and Hipp, 2015).

The analysis of large-scale, phylogenomic datasets in a concatenated matrix has been criticized as it may produce erroneous results with high statistical support (Gadagkar et al., 2005; Kubatko and Degnan, 2007; Seo, 2008). Since concatenation assumes all loci in the dataset share the same evolutionary history, it is inconsistent in the presence of incomplete lineage sorting (ILS) (Kubatko and Degnan, 2007). Accordingly, short branches are likely especially susceptible to the resolution of erroneous results with concatenation (Kumar et al., 2012). Interestingly, the effect may be driven by only a handful of loci (Shen et al., 2017). In contrast, simulation studies have shown, that concatenated RAD-seq matrices are very robust to gene tree/species tree discord and produce the correct species tree even when a substantial amount of discordant loci are present. Inclusion of these loci is unlikely to result in highly supported, false species tree topologies, but rather in unresolved ones (Huang and Knowles, 2009).

Coalescent-based species tree inference methods explicitly address gene-tree conflict due to ILS and hence are designed to be more robust than concatenated methods (Kubatko and Degnan, 2007). The majority of algorithms implementing the coalescence theory aim to infer the species tree by reconciling individual gene trees. These algorithms have become increasingly popular due to their comparative speed and accuracy, provided model assumptions are not violated. Examples include ASTRAL (Mirarab et al., 2014b), ASTRAL-II (Mirarab and Warnow, 2015), MP-EST (Liu et al., 2010) or NJst (Liu and Yu, 2011). The shortness and low informative content of individual RAD-loci often lead to poorly resolved locus trees, which may have a negative impact on species tree estimation via gene tree methods (Salichos and Rokas, 2013; Mirarab et al., 2016). In addition, their demands for computational resources are high (Liu et al., 2015) and hardly feasible for datasets containing tens to hundreds of thousands of loci.

Finally, coalescence based methods are consistent only when the assumptions of the coalescence model are met. Consequently, they are proven to be inconsistent when reticulate evolution is the cause for gene tree discord, and not ILS only (Solís-Lemus et al., 2016; Fernández-Mazuecos et al., 2018).

Despite methodological differences and challenges, HTS methods provide researchers with large, datasets containing hundreds to tens of thousands of genomic loci and are expected to stimulate phylogenetic research in all branches of the tree of life, including Hawaiian lineages. To date, only a few studies have applied HTS methods to Hawaiian plant radiations. The majority of studies focused on either on smaller lineages or in population-level studies. Whole Genome sequencing was used to study population genomics in *Metrosideros polymorpha* Gaudich. (‘Ōhi‘a, Myrtaceae), the most iconic and most common Hawaiian plant (Izuno et al., 2016, 2017; Choi et al., 2019). A population-level study of two closely related species in the Hawaiian lobelioids utilized RAD-seq (Jennings et al., 2016). A Transcriptomic approach was applied to identify informative genes in *Cyrtandra* and *Clermontia* Gaudich. (Campanulaceae, Lobelioidae). A target enrichment approach was employed to research species-level systematics on a limited sampling in the diverse Hawaiian *Cyrtandra* focusing on hybridization (Johnson et al., 2019; Kleinkopf et al., 2019). Whole plastomes were sequenced for species-level relationships in Hawaiian mints but yielded little resolution in the young lineage (Welch et al., 2016).

## 1.6 Aims and Scope

In this thesis, I investigated the Hawaiian lineage of *Melicope* comprising currently 54 described species (Wood et al., 2016, 2017; Appelhans et al., 2017). I used integrated multidisciplinary approaches from karyology, phylogenomics, and evolutionary modeling to infer the trajectory from colonization and establishment on the island system to the evolution of the current species diversity. The results were integrated into the existing framework of adaptive radiation on oceanic islands as predicted by well-researched lineages, e.g the Hawaiian Lobeliods (Givnish et al., 2009) or *Cyrtandra* (Johnson et al., 2017, 2019).

Chapter 2 – *Melicope* lack traits characteristic for island colonization.
In this chapter, I reviewed traits characterizing island colonizers that subsequently give rise to adaptive radiations. These traits include herbaceousness, self-compatibility, and neo- or mesopolyploidy (Carlquist, 1974; Carr et al., 1986; Baldwin and Sanderson, 1998; Carr, 1998). Establishment of colonizers on the islands is accompanied by specific trait shifts, e.g. from herbaceousness to secondary woodiness (Lens et al., 2013) and a reduction in dispersal ability (Carlquist, 1966b), collectively referred to as 'island syndrome' (Carlquist, 1974). I investigated the island syndrome in *Melicope* throughput the Pacific with a special emphasis on the Hawaiian Islands. I used Flow Cytometry to infer the karyotype of Hawaiian and other *Melicope* to detect ancestral polyploidization events. The remaining traits were compiled from the literature to test if Hawaiian *Melicope* complies with the island syndrome as currently described.

Chapter 3 – RAD-seq phylogeny of Hawaiian *Melicope.*
Previous attempts to resolve species relationships in Hawaiian *Melicope* were based on a comparatively small number of nuclear and plastid marker regions generated by Sanger sequencing. The resulting datasets comprised several thousand base pairs and established the monophyly of the lineage including the formerly endemic genus *Platydesma* and resolved six major clades within the lineage (Harbaugh et al., 2009; Appelhans et al., 2014a, 2014b). However, the number of informative sites in the datasets was neither sufficient to resolve relationships between these clades nor at the species-level (Harbaugh et al., 2009; Appelhans et al., 2014a, 2014b). In this chapter, I employed an HTS approach of restriction-site associated genomic loci (RAD-seq) to assemble datasets comprising several thousand to tens of thousands of genomic loci and millions of base pairs. Both, coalescence-based and concatenated Maximum Likelihood or Bayesian inference methods were applied. Results were evaluated in a statistical framework assessing concordance, informative capacity, and quality of discord on a per branch basis. In addition, partitioned *D*-statistics were used to investigate signals of introgressive hybridization.

Chapter 4 – Biogeography and Diversification of Hawaiian *Melicope.*

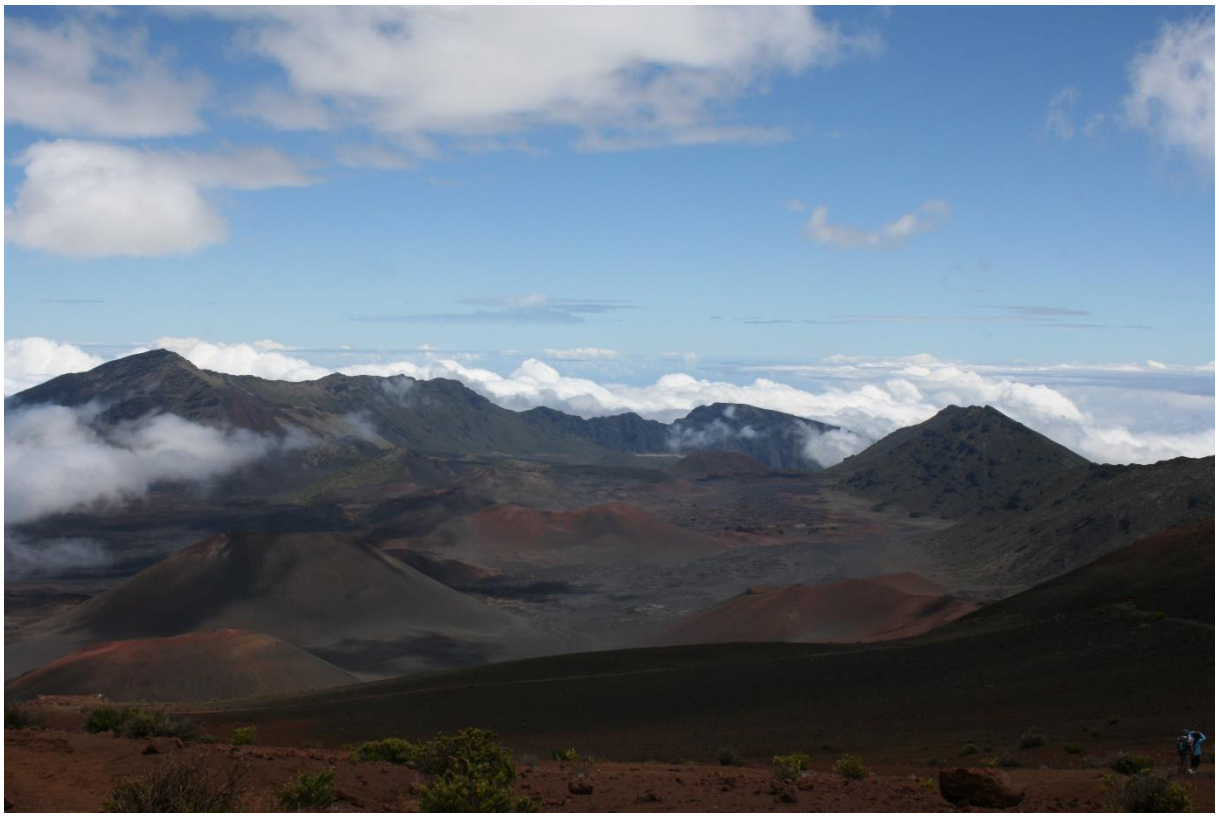
The RAD-seq datasets generated in chapter 3 served as a foundation to analyze patterns of diversification in Hawaiian *Melicope*. I estimated the divergence times in a Bayesian framework to infer when the ancestor colonized the archipelago and discuss the origin of the lineage considering that the majority of Hawaiian radiations are younger than the current high islands (Price and Clague, 2002). In addition, I modeled the historical biogeography of Hawaiian *Melicope* to investigate dispersal patterns within the archipelago. Results are were discussed with respect to common patterns of island biogeography, notably the progression rule (Wagner and Funk, 1995). Estimated diversification rates were used to explore the rapidity of speciation within the clade. Differences in diversification rate are discussed with respect to the geology of the islands. Finally, I tested how species habitats influence diversification in a probabilistic framework.

## 2. | The odd one out or a hidden generalist: Hawaiian *Melicope* (Rutaceae) do not share traits associated with successful island colonization

Claudia Paetzold, Michael Kiehn, Kenneth R. Wood, Warren L. Wagner & Marc S. Appelhans

Published in *Journal of Systematics and Evolution* 2018: 56(6), 621-636

doi:10.1111/jse.12454



The caldera of the dormant Haleakala volcano (Maui). Photograph: Marc Appelhans

**Abstract**

Oceanic islands are unique in their species composition, which is defined by arrival of colonizers via long distance dispersal followed by establishment of species followed in some cases by adaptive radiation. Evolutionary biologists identified traits facilitating successful colonization of islands as including polyploidy, self-compatibility, herbaceousness and ability for long-distance dispersal. Successful establishment and evolutionary diversification of lineages on islands often involves shifts to woodiness and shifts in methods of outcrossing as well as changes in dispersal ability. The genus *Melicope* colonized numerous archipelagos throughout the Pacific including the Hawaiian Islands, where the lineage comprises currently 54 endemic species and represents the largest radiation of woody plants on the islands. The wide distributional range of the genus illustrates its high dispersibility, most likely due to adaption to bird dispersal. Here we investigate ploidy in the genus using flow cytometry and chromosome counting. We find the genus to be paleopolyploid with $2n = 4x = 36$, a ploidy level characterizing the entire subfamily Amyridoideae and dating back to at least the Palaeocene. Therefore Hawaiian *Melicope* have not undergone recent polyploidization prior to colonization of the islands. Thus *Melicope* retained colonization success while exhibiting a combination of traits that typically characterize well established island specialists while lacking some traits associated to successful colonizers.

## 2.1 Introduction

Ever since Charles Darwin wrote about his observations upon visiting the Galápagos Islands (Darwin, 1859), oceanic islands have been a focal point for biologists in their quest to unravel the process of evolution. The study of islands that have never been connected to a continental land mass, especially those that are greatly isolated and of volcanic origin, offer several unique advantages (Emerson, 2002). Islands are discrete systems with oceanic boundaries restricting gene flow between land masses. In spite of their small size (compared to continents), many oceanic islands offer a wealth of habitats and ecological niches, which are often in a constant flux due to influences of outside forces, e.g., plate tectonics, volcanic activity, erosion, flooding and tropical storms.

Yet, island floras are not merely 'downscaled' versions of the neighbouring continental ones. In contrast, islands possess unique species compositions differing remarkably from those of the continental land mass and typically with a high degree of endemism. For example, in the Canary Islands about 40% of all Angiosperm taxa are endemic (Francisco-Ortega et al., 2000) and about 90% in the Hawaiian Islands (Wagner et al., 1999b; Keeley and Funk, 2011). The species composition of an island is dependent on three main factors: distance, geology (incl. altitudinal variation) and age. Distance refers to the distance between an island and other landmasses serving as a possible origin of colonizers. Increasing distance decreases the frequency of successful colonization events and restricts the diversity of possible colonizers to those with propagules 'equipped' to travel the distance. The geology and size of the island determines the quality and the quantity of ecological niches it provides. The age of an island represents the time frame available for colonization, establishment, adaptive radiation and even extinction of species (Carlquist, 1966a).

Successful colonizations of oceanic islands are rare, so that arrivals, especially to young islands, probably experience less selective pressure from other species than in their continental environment (Baldwin, 1998). When a viable seed reaches a given island and meets conditions allowing its establishment, the colonizer may undergo extensive adaptive radiation giving rise to a lineage of diverse species (Carlquist, 1966a; Givnish et al., 2009).

The synergy of colonization by few founders, along with establishment in available ecological niches and adaptive radiation result in unique island floras that are vastly different from their source areas – both morphologically, ecologically and in terms of species richness (Carr, 1998). Yet, despite the individuality of each island system, after close to two centuries of island evolution research, several evolutionary trends have become apparent. In the past island biodiversity has been associated with

multiple colonization events per lineage based on the presence of divergent morphological characters. However, more recent molecular phylogenetic and biogeographic studies revealed that this not the case and that, e.g., the 1192 species of vascular plants native to the Hawaiian Islands are the result of only 263-270 colonization events (Keeley and Funk, 2011). Most island lineages are monophyletic descending from one successful colonization event, e.g., the Hawaiian lobeliads (Campanulaceae; Givnish et al., 2009), *Dendroseris* D.Don (Asteraceae) on the Juan Fernandez Islands (Crawford et al., 1998) or the woody *Sonchus* L. (Asteraceae) alliance in Macaronesia (Kim et al., 1996). In many cases, island colonizers seem to be single, broadly adapted, often herbaceous, generalist species that radiated into several highly specialized, locally adapted and restricted species (Grant, 1998). Common traits of successful colonizers and the subsequent evolutionary shifts during establishment and radiation on oceanic islands include:

(1) Polyploidization. The advent of modern sequencing techniques has revealed that a whole genome duplication (WGD) event predated the diversification of all Angiosperms, rendering all flowering plants 'polyploid' (Amborella Genome Project, 2013). For simplicity in this paper these most ancient events will be ignored and polyploidy will concern only chromosome number changes post-dating them. That being said, many oceanic island floras are characterized by a high number of polyploid plant taxa. Conventional estimations of polyploidy are often based on identifying the most likely base number of Angiosperms by widespread comparison of numerous lineages combined with chromosome pairing analysis and postulating a threshold. Using this method Grant (1963) postulated that plants with a basic chromosome number of $n = 14$ or higher are most likely polyploid. While detailed comparisons and genomic and cytological estimations are required to identify ploidy levels on a lineage-by-lineage basis, this approach serves as an adequate approximation.

Employing this approach, more than 80% of Hawaiian endemics (including *Melicope* J.R.Forst. & G.Forst.) are polyploid (Carr, 1998), as are 66% of all endemics on the Juan Fernandez Islands (Stuessy et al., 1992), while on the Canary Islands the fraction is only 24.5% (Bramwell, 1976). These numbers indicate that polyploidy has a different impact or prevalence on islands depending on island age and distance to continental land masses (Whittaker, 1998). High levels of polyploidy on many oceanic islands do not reflect high instances of in situ polyploidization, as island lineages often display chromosomal stasis during speciation (Stuessy and Crawford, 1998; Kiehn, 2005). Instead, the high percentage of polyploid endemics indicates the success of polyploid immigrants (Stuessy and Crawford, 1998) in the competition for

colonization and adaptive radiation. In the grass subfamily Danthonioideae, polyploidization events were shown to facilitate Long-Distance Dispersal (LDD; Linder and Barker, 2014). Polyploidy offers advantages that may be particularly potent for establishment on oceanic islands, including increased vigour through heterosis and gene redundancy (Comai, 2005). Although detailed molecular studies for many lineages are still lacking, the few radiations that have been investigated indicate an allo- or autopolyploidization event directly predates the colonization of oceanic islands. While the phenomenon is fairly well researched in Asteraceae (Crawford et al., 2009), it is perhaps most striking in the sandalwoods (*Santalum* L., Santalaceae). Members of *Santalum* colonized islands throughout the Pacific in a stepwise fashion, following at least six polyploidization events leading to three additional ploidy levels (Harbaugh and Baldwin, 2007; Harbaugh, 2008). Hawaiian examples include the silversword alliance originating from an allopolyploidization event in California ca. 15 million years ago (mya) (Baldwin et al., 1991; Baldwin and Sanderson, 1998), or the Hawaiian violets arriving as recently as ca. 1.2-2 mya (Havran et al., 2009). Following the classification of Ehrendorfer (1980) on those few investigated lineages, colonizers classify as neo-or mesopolyploids.

(2) Dispersibility. Immigrants to remote oceanic islands arrive by definition via LDD. While there is an element of chance to that, the likelihood of successful LDD event(s) increases with diaspores adapted to efficient dispersal, as evidenced by families or genera that colonized multiple islands. Adaptions of highly dispersible diaspores include smallness of spores or seeds for wind dispersed taxa (e.g., ferns, orchids), hooks, barbs and adhesive layers for exozoochory (e.g., *Bidens* L. (Asteraceae); *Peperomia* Ruiz & Pav. (Piperaceae)), or pulpous parts (often containing many tiny seeds) attracting feeders for endozoochory (e.g. *Rubus* L. (Rosaceae)). Regardless of vector, a small seed size is a common factor among efficient dispersers, both because this makes them easy to carry or swallow and because most immigrants are herbs (see 4). Weedy or herbaceous open habitat species tend to have small seeds as seedlings are exposed to sunlight shortly after germination. In contrast woody species tend to have larger seed sizes, as the seed contains stored nutrients, from which the seedling will grow until it reaches higher forest strata and exposure to sunlight (Carlquist, 1966a). Though detailed studies are scarce, trends for island species to drastically reduce their dispersal ability as pertaining to LDD and water barriers have been observed (e.g. Carlquist, 1966b, 1966c; Fresnillo and Ehlers, 2008; Price and Wagner, 2004). In several fern genera an increase in spore size has been observed as well as reduction or loss of pappus awns in *Bidens* (Carlquist, 1966b) or an increase in fruit size in, e.g., *Polyscias* J.R.Forst. & G.Forst. (Araliaceae; as

*Tetraplasandra* A.Gray) or *Zanthoxylum* L. (Rutaceae; as *Fagara* L.) (Carlquist, 1966c). Reducing dispersibility is an advantageous adaptation in an island setting as it decreases the likelihood of seeds becoming 'lost at sea' and reflects the condition where the habitable area of most species is often much smaller than the total island size (Carlquist, 1966a; Price and Wagner, 2004).

(3) Self-Compatibility to Outbreeding. In 1955 Herbert Baker proposed the hypothesis (later widely known as Baker's law) that self-compatibility is an advantageous trait for an island colonizer to possess. Since colonization events are rare and typically involve only one or a small number of individual(s), being self-compatible allows establishment on an island in the absence of potential mates and/or pollinators or when potential mates are present but incompatible (Pannell, 2015). However, high instances of outbreeding mechanisms observed on oceanic islands (Carlquist, 1966a) seem to point towards the development of said mechanisms following establishment to counter possible negative effects of small population sizes and gene pools. In New Zealand 12-13% of species are dioecious (Webb and Kelly, 1993) as are 14% of species on the Hawaiian Islands (Sakai et al., 1995), where the worldwide ratio is at 4% (Yampolski and Yampolski, 1922). On the Hawaiian archipelago approximately one third of all dimorphic species evolved from a monomorphic colonizer (Sakai et al., 1995).

(4) Herbaceousness to insular woodiness. Stuessy and Crawford (1998) argued that in many cases successful island colonizers are predominantly herbs. Decreased generation times of herbs, as compared to woody species, should enable them to adapt to a new environment more quickly. Upon establishment, however, a shift to a woody growth form can often be observed, which Carlquist (1974) termed 'insular woodiness'. It has been observed in several Angiosperm families and islands, and evolved in numerous lineages independently. In Asteraceae this pattern is highly prominent with the woody *Sonchus* alliance on the Macaronesian islands (Kim et al., 1996), *Dendroseris* and *Robinsonia* DC. (Asteraceae) on the Juan Fernandez Islands (Crawford et al., 1998), the Hawaiian silversword alliance (Baldwin, 1998) or Hawaiian *Schiedea* Cham & Schltdl. (Caryophyllaceae, Wagner et al., 2005).

Of course, not all successful radiations exhibit all of these traits, and research is incomplete for a majority of lineages. While some species of Hawaiian mints are shrubby or herbs with a "somewhat a woody base", others are herbaceous (Wagner et al., 1999b), and as such the lineage as a whole does not exhibit insular woodiness (Lens et al., 2013). Since no detailed study exists regarding the woodiness in Hawaiian mints, and as the boundary between herbaceous and woody is considered fuzzy (Lens et al., 2013), evaluation of this trait is not final. On the other hand

Hawaiian mints are of allopolyploid origin and share the same chromosome number ($2n = 64$) as their closest relatives in the genus *Stachys* L. (Lamiaceae). However, with chromosome numbers ranging from $2n = 10$ to 102 in the genus *Stachys* (Wagner et al., 1999b; Lindqvist and Albert, 2002; Roy et al., 2015), the ancestor of Hawaiian mints may be classified as a mesopolyploid. And while we do know that the largest oceanic radiation in the world, Hawaiian lobeliads (Campanulaceae), is polyploid (Lammers, 1988; Carr, 1998), we do not know whether polyploidization occurred prior to colonization. We can surmise, however, that successful adaptive radiations on oceanic islands seem to show at least one or several, if not necessarily all of these traits.

*Melicope* J.R.Forst. & G.Forst. (Rutaceae) in its traditional circumscription comprises ca. 230 species of shrubs and trees distributed in east-west-extension from Madagascar to the Hawaiian Islands and in north-south-extension from Japan to New Zealand. Currently (Hartley, 2001) the genus is subdivided into four sections: *Lepta* (Lour.) T.G.Hartley, *Melicope, Pelea* (A.Gray) Hook. and *Vitiflorae* (F.Muell.) T.G.Hartley. Recent molecular work has revealed that several genera are nested within *Melicope* and that the enlarged genus now contains about 300 species (Appelhans et al., 2014a). The Hawaiian genus *Platydesma* H.Mann was one of these genera and has recently been included in *Melicope* (Appelhans et al., 2017). *Melicope* has its origin in the Australasian region but has colonized numerous archipelagos throughout the Pacific and even Madagascar and the Mascarene Islands (Appelhans et al., 2018b). The Hawaiian Island lineage of *Melicope* is monophyletic and nested deeply within the genus. The clade belongs to section *Pelea* and comprises 54 currently accepted species (Hartley, 2001; Appelhans et al., 2017; Wood et al., 2017). It represents the largest radiation of woody plants on the Hawaiian Islands (Wagner et al., 1999b) and colonization predates the age of the current high islands (Appelhans et al., 2018b). At first glance the lineage seems to match the pattern for insular specialist very nicely; the species are woody, mostly distributed in forests and about 80% of the species are endemic to a single island (when Maui Nui is treated as a single island) with only small distributional ranges on the islands. Also they are mostly dioecious and their capsular/follicular fruits display shiny black seeds in varying sizes with a spongy and nutritious sarcotesta and a thick sclerotesta, which have been interpreted as an adaption to bird dispersal (Hartley, 2001). However, all species of not only the genus *Melicope* but also all related genera (Appelhans et al., 2014a) are woody and bird dispersed and all species of *Melicope* section *Pelea* are dioecious (Hartley, 2001). Therefore these traits are ancestral and not acquired following colonization of the archipelago. Whether the same is true regarding the

ploidy is not yet clear. Up until now chromosome counts exist for 20 *Melicope* species, two species of *Acronychia* J.R.Forst. & G.Forst., which is nested in *Melicope* as well as one recorded count for *Comptonella* Baker f., which was revealed to be nested within *Melicope* sect. *Vitiflorae* (Appelhans et al., 2014a) (Table 2.1). Altogether these records span the entire distributional range of *Melicope* (except Madagascar and the Mascarene Islands) and all four sections of the genus. The 14 species representing non-Hawaiian lineages of *Melicope*, the two specimens of *Acronychia* as well as the record for *Comptonella* revealed a base chromosome number of $2n = 36$; with the exception of one count for *M. semecarpifolia* (Merr.) T.G.Hartley ($n = 12$; Hsu, 1968) and the result for *M. brassii* T.G.Hartley ($2n = 32$; Borgmann, 1964). Though an ancestral state of $n = 18$ has also been suggested (Stace et al., 1993), the ancestral haploid chromosome number in Rutaceae is most likely nine (Kubitzki et al., 2011), as the most closely related sister clades (Meliaceae, Simaroubaceae) also show a base chromosome number of $n = 9$ (Figure 2.1). Within Rutaceae only the species-poor subfamilies Aurantioideae and Rutoideae (Morton and Telmer, 2014; ~300 species in 33 genera) possess $n = 9$ (or more rarely $n = 10$). The vast majority of Rutaceae (including *Melicope*) is represented by subfamily Amyridoideae (Morton and Telmer, 2014), a clade of 1800 species in 113 genera with $n = 18$ as base chromosomal number (Kubitzki et al., 2011). The shift from $n = 9$ to $n = 18$ likely happened in the Paleocene or even the Late Cretaceous (Appelhans et al., 2012; Figure 2.1). Therefore the Amyridoideae genera including *Melicope* can be considered paleopolyploids.

**Table 2.1.** | Chromosome counts for 12 Hawaiian and 13 non-Hawaiian Melicope species, two species of Acronychia and one species of Comptonella, both of which are nested within Melicope. Details on origin of specimens, collection numbers including deposition of Herbarium vouchers for new records and references are given. Herbarium acronyms are according to Index Herbariorum (http://sweetgum.nybg.org/science/ih/).

| Species | n | 2n | section | origin | Coll. No. (Herbarium of voucher deposition for new counts) | Reference |
|---|---|---|---|---|---|---|
| **Hawaiian taxa** | | | | | | |
| *Melicope adscendens* (St.John & Hume) T.G Hartley & B.C.Stone | | 36 | *Pelea* | Maui | *Oppenheimer* H20907 & *Perlman* (BISH, WU) | Kiehn, this paper |
| *M. anisata* (H.Mann) T.G.Hartley & B.C.Stone | | 34-36 | *Pelea* | Kauaʻi | *Perlman & Kiehn* SP 21325 (PTBG, WU) | Kiehn, this paper |
| *M. barbigera* A.Gray | | 36 | *Pelea* | Kauaʻi | | Kiehn, 2005 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *M. clusiifolia* (A.Gray) T.G.Hartley & B.C.Stone | | 36 | *Pelea* | Kauaʻi | *Perlman & Kiehn* SP 21328 (PTBG, WU) | Kiehn, this paper |
| *M. cornuta* (Hillebr.) Appelhans, K.R.Wood & W.L.Wagner | 18 | | *Pelea* | Oʻahu | | Carr, 1978 (as *Platydesma c.*) |
| *M. elliptica* A.Gray | 18 | | *Pelea* | Oʻahu | | Carr, 1978(as *Pelea e.*) |
| *M. ovata* (St.John & Hume) T.G.Hartley & B.C.Stone | | 34-36 | *Pelea* | Kauaʻi | *Perlman & Kiehn* SP 21333 (PTBG, WU) | Kiehn, this paper |
| *M. ovata* (St.John & Hume) T.G.Hartley & B.C.Stone | | 18, 36 | *Pelea* | Kauaʻi | | Kiehn, 2005 (as *M.* sp.) |
| *M. puberula* (H. St. John) T.G. Hartley & B.C. Stone | 18 | 36 | *Pelea* | Kauaʻi | *Perlman & Kiehn* SP 21327 (PTBG, WU) | Kiehn, this paper |
| *M. rostrata (*Hillebr.) Appelhans, K.R.Wood & W.L.Wagner | | 36 | *Pelea* | Kauaʻi | | Guerra, 1984 (as *Platydesma rostratum* Hillebr.) |
| *M.* sp. indet | | 36 | *Pelea* | Oʻahu | | Kiehn, 2005 |
| *M. wawraeana* Rock | | 72 | *Pelea* | Kauaʻi | | Guerra, 1984 (as *Pelea w.*) |
| *M. zahlbruckneri* (Rock) T.G.Hartley & B.C.Stone | | 36 | *Pelea* | Hawaiʻi | *Kiehn & Pratt* MK-090211-4/4 (BISH, WU) | Kiehn, this paper |
| **Non-Hawaiian taxa** | | | | | | |
| | | | | | | |
| *M. brassii* T.G.Hartley | | 32 | *Pelea* | Papua New Guinea | | Borgmann, 1964 |
| *M. bonwickii* (F.Muell.) T.G.Hartley | | 36 | *Lepta* | Philippine: Luzon | | Pancho, 1971 (as *Euodia villamillii* Merr.) |
| *M. frutescens* (Blanco) Appelhans & J.Wen | | 36 (38?) | *Lepta* | Philippine: Luzon | | Pancho, 1971 (as *Euodia confusa* (Blco.) Merr.) |

| | | | | | | |
|---|---|---|---|---|---|---|
| *M. grisea* (Planch.) T.G.Hartley | | 36 | *Lepta* | Japan: Bonin Islands | | Ono and Masuda, 1981 (as *Boninia grisea* Planch.) |
| *M. lunu-ankenda* (Gaertn.) T.G.Hartley | | 36 | *Lepta* | Sri Lanka | | Morawetz, 1986 (as *Evodia roxburghiana* Benth.) |
| *M. mantellii* Buchanan | 18 | | *Melicope* | New Zealand: cult. Auckland University College | | Rattenbury, 1957 |
| *M. micrococca* (F.Muell.) T.G.Hartley | 18 | | *Lepta* | Australia | | Smith-White, 1954 (as *Euodia micrococca* F.Muell.) |
| *M. quadrilocularis* (Hook. & Arn.) T.G.Hartley | | 36 | *Lepta* | Japan: Bonin Islands | | Ono & Masuda, 1981 (as *Boninia glabra* Planch.) |
| *M. retusa*(A.Gray) T.G.Hartley | | 36 | *Pelea* | Philippine: Luzon | | Pancho, 1971 |
| *M. rubra* (Lauterb. & K.Schum.) T.G.Hartley | | 36 | *Lepta* | Papua New Guinea | | Borgmann, 1964 |
| *M. semecarpifolia* (Merr.) T.G.Hartley | 12 | | *Lepta* | Taiwan | | Hsu, 1968 (as *Euodia confusa* (Blco.) Merr.) |
| *M. semecarpifolia* (Merr.) T.G.Hartley | | 36 | *Lepta* | Philippine: Luzon | | Pancho, 1971 |
| *M. simplex* A.Cunn. | | 36 | *Melicope* | New Zealand | | Rattenbury, 1957 |
| *M. ternata* J.R.Forst. & G.Forst. | 18 | | *Melicope* | New Zealand: cult. Auckland University College | | Rattenbury, 1957 |
| *M. ternata* J.R.Forst. & G.Forst. | | 36 | *Melicope* | cult. Botanical Garden University of Vienna (WU) | | Guerra, 1984 |
| *Acronychia suberosa* C.T.White | | 36 | | Australia: Queensland | | Guerra, 1984 |

| | | | | | | |
|---|---|---|---|---|---|---|
| *A. pubescens* | | 34 | | Australia: Queensland | | Guerra, 1984 |
| *Comptonella Baker f.* | 18 | | | France: New Caledonia | | Kubitzki, 2011 |

The observation of a depauperate sisterclade to a highly diverse, species-rich, polyploid one with a delay between the polyploidization event and the onset of diversification fits the WGD radiation lag-time model (Schranz et al., 2012). One hypothesis for this lag phase is, that this time is required for diploidization to take place (Dodsworth et al., 2016). Diploidization is a post-genome-duplication process that includes operations between duplicated genes, e.g. neofunctionalization, subfunctionalization and non-functionalization as well as operations between duplicated genomes, e.g. genome downsizing (Ma and Gustafson, 2005; Dodsworth et al., 2016).

Comparing chromosome numbers and DNA content in Rutaceae (C-value database Kew, http://data.kew.org/cvalues; assessed on 01. 16. 2017) shows no linear relationship. *Ruta graveolens* L. for example shows a chromosome number of $n = 78$ at a DNA content of 0.75 pg and illustrates the effects of genome downsizing. So far the only *Melicope* species for which both a chromosome count as well as genome size have been measured is *Melicope ternata* J.R.Forst. & G.Forst. This species has a chromosome number of $n = 18$ and shows a genome size of 0.93 pg (Guerra, 1984) comparable to that of many diploid members of Rutaceae.

The seven *Melicope* species from Hawaii investigated previously do not show a consistent picture. Three different chromosome numbers were reported ($2n = 18$, 36 or 72; Table 2.1) indicating possible polyploidization or hybridization events within the lineage (Kiehn, 2005).

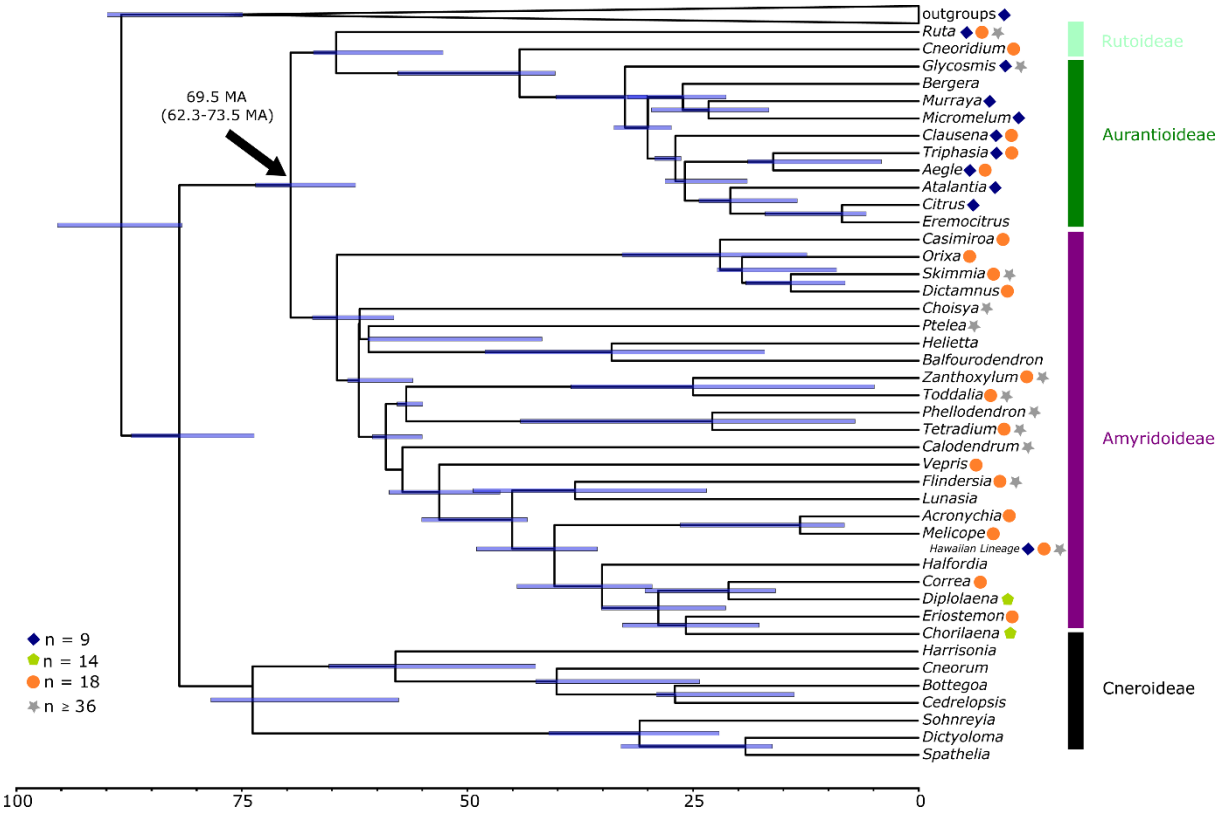



**Figure 2.1.** | Phylogenetic relationships of Rutaceae genera (modified from Appelhans et al., 2012) with known ploidy levels as inferred from known chromosome numbers (Kew C-value database (http://data.kew.org/cvalues; assessed on 01. 16. 2017), Kubitzki, 2011) plotted to each genus. Rutaceae subfamilies are indicated by black (Cneorideae), violet (Amyridoideae), mint (Rutoideae) and green (Aurantioideae) bars. Outgroups refer to the most closely related families Meliaceae and Simaroubaceae (Appelhans et al., 2012). A black arrow marks the split of the Aurantioideae and Amyridoideae subfamilies in the Palaeocene and the coinciding polyploidization event.

The main aim of this study is to investigate evolutionary trends characteristic for oceanic islands in the Hawaiian lineage of the genus *Melicope*. To that end we also investigate if specimens show traits specific for colonization of and/or establishment on islands, and if these traits are unique to the Hawaiian radiation or characteristic for the genus as a whole. We have conducted a literature search regarding traits of insular woodiness, dispersibility, and reproductive systems. We further investigate ploidy levels in the Hawaiian radiation of the genus as well as representatives of the non-Hawaiian species to infer whether the colonizer of the archipelago was a neo- or mesopolyploid.

## 2.2 Material and Methods

### Flow cytometry

DNA content was assessed for 61 samples representing 66% of the Hawaiian radiation of *Melicope* as well as nine samples of non-Hawaiian species via flow cytometry. Table 2.2 details geographic origins and collection details for the samples. Due to scarcity of material, only one measurement was taken per sample.

Leaf material was ground with a TissueLyzerII (Quiagen, Hilden, Germany) at 15 Hz for 45 s using a steal bead (Ø 3 mm) in a 2 mL Eppendorf cap. Nuclei were isolated by 8 min incubation in 300 µL Otto I buffer (Otto, 1990). After filtering the mix (30 µm mesh, CellTrics® Partec GmbH, Münster, Germany), 800 µL staining solution (Otto II buffer, Doležel and Göhde, 1995) was added and the solution again incubated for 8 min on ice. The solution was then measured on the flow cytometer (CyFlow® Ploidy Analyser, Sysmex Deutschland GmbH, Norderstedt, Germany) using the blue UV LED channel. Fluorescence intensity was measured and peaks medians were calculated using the program CyFlow Cube v 1.5.7.3 (Sysmex Deutschland GmbH, Norderstedt, Germany).

Samples were measured at gain 450 with *Pisum sativum* L. (Fabaceae) as internal standard. Several samples failed to produce a peak due to inference of debris particles, and the measurement was repeated for those at gain 480 with *Paspalum notatum* Flugge (Poaceae) as external standard. The mean peak value of all reference measurements (> 15) was used to calculate the DNA content of samples using the formula (sample mean peak * reference mean peak) / reference DNA content. Reference mean C1 values were obtained from the Kew C-value database (http://data.kew.org/cvalues; assessed on 16. 01. 2017) as 1C = 4.88pg for *Pisum sativum* and 1C = 0.89pg for *Paspalum notatum*. The software Past v 3.17 (Hammer et al., 2001) was used to test for normal distribution of measurements.

### Chromosome counts

Chromosome counts are based on field fixations or fixations from plants cultivated at the Botanical Garden of the University of Vienna (Austria). Fixations of meristematic tissues (actively growing root tips, young flowers or apices for counts of mitotic numbers, young flower buds for meiotic investigations) were made in a freshly mixed 3:1 solution of ethanol (96%):glacial acetic acid or in a 4:3:1 mixture of chloroforme:100% ethanol:glacial acetic acid. Some germinating seeds were pretreated with 0.002 M 8-hydroxyquinoline solution for 6 h at 8-10 °C in the dark before fixations were made (Table 2.1). Each fixation represents one individual in the case of field fixations, or individually distinguishable seedlings in the case of fixations of germinating seeds. Chromosome staining was performed with Feulgen

reagent, Giemsa, or aceto-carmine (for details on staining procedures see Kiehn, 2005). Exact counts could not be achieved in some cases because of limited material. A range of chromosome numbers is given in such cases. Permanent slides for the counts are deposited in the personal collection of MK. Reference voucher specimens for each investigated collection have been deposited in at least one of the following herbaria: Bishop Museum, Honolulu, Hawaii (BISH), National Tropical Botanical Garden, Kalāheo (Kauaʻi), Hawaii, (PTBG), University of Hawaiʻi (HAW), or University of Vienna (WU).

## 2.3 Results

Table 2.2 summarizes the genome sizes for 61 samples of *Melicope* as estimated by flow cytometry. With the exception of *M. ternata* (Guerra, 1984) none of these species have been assessed regarding their genome sizes before. The results are normally distributed ($p = 0.71$; Shapiro-Wilk = 0.988). The mean 1C value of all samples is 0.76 pg with a standard deviation of 0.05. The lowest and highest genome sizes were estimated for the Hawaiian *M. haupuensis* (St.John) T.G.Hartley & B.C.Stone and *M. peduncularis* (H.Lév.) T.G.Hartley & B.C.Stone with 1C = 0.65 pg and 1C = 0.87 pg respectively. In samples using *Pisum sativum* as reference, estimated genome sizes were slightly higher (mean 1C = 0.8 pg). Four samples (*M. anisata* (H.Mann) T.G.Hartley & B.C.Stone [Appelhans MA665], *M. barbigera* A.Gray [Appelhans MA664], *M. barbigera* [Wood KW 16718] and *M. peduncularis* [Appelhans MA652]), that could be measured successfully with both available references, show a slightly higher 1C value when measured with *P. sativum* as a reference, indicating that there seems to be a slight bias introduced due to the different genome sizes of the references (Figure 2.2).

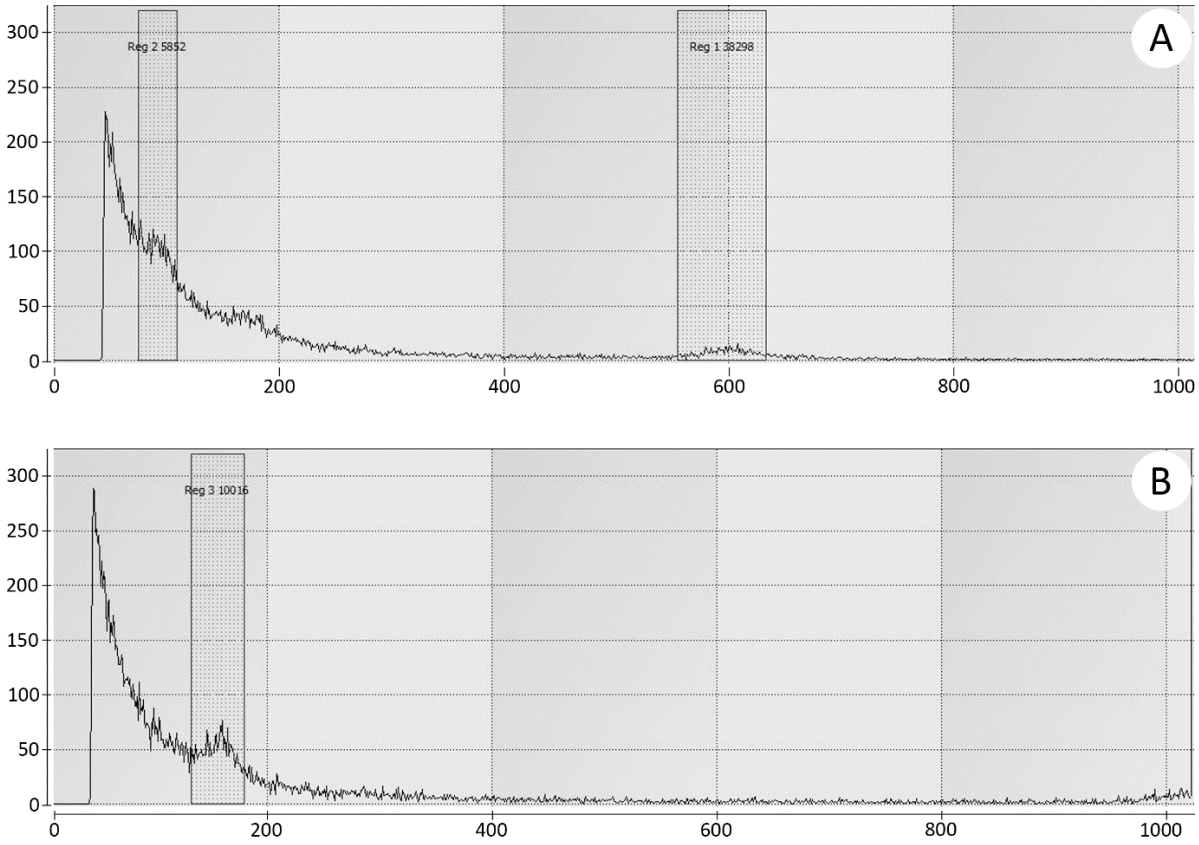



**Figure 2.2.** | Flow Cytometry measurements of *Melicope barbigera* A.Gray [Appelhans MA664] at gains 450 (A) and 480 (B). X-axis shows amount of particles at a given fluorescence intensity. Intensity peaks are marked (Reg 2 & 3, *M. barbigera*; Reg 1, reference *Pisum sativum*).

Chromosome numbers for six *Melicope* species were newly determined, increasing the total number of assessed species to 25 (including *Acronychia* and *Comptonella*), of which 12 represent the Hawaiian lineage (Table 2.1). All new reports reveal chromosome numbers of $n = 18$ or $2n = 36$, as did the majority of the previous counts for the genus. Altogether 12 species with known chromosome numbers are represented in the flow cytometry taxon sampling, including two of the four species showing varying chromosome numbers (*M. ovata* (St.John & Hume) T.G.Hartley & B.C.Stone and *M. wawraeana* (Rock) T.G.Hartley & B.C.Stone). DNA content measured in these species does not deviate (compare Table 2.1 and Table 2.2).

**Table 2.2.** | DNA content of 62 Hawaiian and 11 non-Hawaiian *Melicope* specimens as measured by flow cytometry using Pisum sativum (†) or Paspalum notatum (‡) as reference. Details for placement of herbarium vouchers and origin of samples are given. Herbarium acronyms are according to Index Herbariorum (http://sweetgum.nybg.org/science/ih/).

| Species | Herbarium voucher | Origin | C(pg) † | C(pg) ‡ |
|---|---|---|---|---|
| **Hawaiian species** | | | | |
| *Melicope adscendens* (H. St. John & E.P. Hume) T.G. Hartley & B. C. Stone | Appelhans MA628 (silica sample only, cultivated at Olinda Rare Plant Facility) | Maui | | 0.71 |
| *Melicope anisata* (H. Mann) T. G. Hartley & B. C. Stone | Appelhans MA665 (GOET [GOET019849], PTBG [PTBG 1000057433]) | Kauaʻi | 0.75 | 0.79 |
| *Melicope anisata* (H. Mann) T. G. Hartley & B. C. Stone | Appelhans MA668 (GOET [GOET019850], PTBG [PTBG 1000057439], US) | Kauaʻi | | 0.78 |
| *Melicope barbigera* A. Gray | Appelhans MA664 (GOET [GOET019851], PTBG [PTBG 1000057432], US) | Kauaʻi | 0.72 | 0.69 |
| *Melicope barbigera* A. Gray | Appelhans MA666 (BISH, GOET [GOET019852], PTBG [PTBG 1000057437], US) | Kauaʻi | | 0.79 |
| *Melicope barbigera* A. Gray | Wood 16718 (PTBG) | Kauaʻ | 0.78 | 0.73 |
| *Melicope christophersenii* (H. St. John) T. G. Hartley & B. C. Stone | Appelhans MA617 (BISH, GOET [GOET019853], PTBG [PTBG 1000057596], US) | Oʻahu | | 0.75 |
| *Melicope christophersenii* (H. St. John) T. G. Hartley & B. C. Stone | Appelhans MA621 (silica sample only, cultivated at Puʻu Kaʻala) | Oʻahu | | 0.73 |
| *Melicope christophersenii* (H. St. John) T. G. Hartley & B. C. Stone | Takahama s.n. (silica sample only) | Oʻahu | 0.86 | |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | Appelhans MA615 (GOET [GOET019855], PTBG [PTBG 1000057517]) | Oʻahu | | 0.82 |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | Appelhans MA634 (PTBG [PTBG 1000057507]) | Maui | | 0.78 |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | Appelhans MA650 (GOET [GOET019857], PTBG [PTBG 1000057504], US) | Maui | 0.82 | |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | Appelhans MA651 (BISH, GOET [GOET019856], PTBG [PTBG 1000057511], US) | Maui | 0.85 | |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | Appelhans MA655 (silica sample only) | Maui | | 0.76 |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | Appelhans MA657 (GOET [GOET019858], PTBG [PTBG 1000057572], US) | Maui | 0.80 | |

| | | | | |
|---|---|---|---|---|
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | Oppenheimer H91641 (US) | Lānaʻi | | 0.67 |
| *Melicope cruciata* (A. Heller) T.G. Hartley & B.C. Stone | Wood 16251 (PTBG) | Kauaʻi | | 0.76 |
| *Melicope feddei* (H. Lév.) T. G. Hartley & B. C. Stone | Appelhans MA688 (BISH, GOET [GOET019864], PTBG [PTBG 1000057431], US) | Kauaʻi | | 0.74 |
| *Melicope haleakalae* (B. C. Stone) T. G. Hartley & B. C. Stone | Appelhans MA637 (BISH, GOET [GOET019866], PTBG [PTBG 1000057497], US) | Maui | | 0.79 |
| *Melicope haleakalae* (B. C. Stone) T. G. Hartley & B. C. Stone | Appelhans MA641 (BISH, GOET [GOET019865], PTBG [PTBG 1000057502]) | Maui | | 0.74 |
| *Melicope haleakalae* (B. C. Stone) T. G. Hartley & B. C. Stone | Appelhans MA645 (BISH, GOET [GOET019867], PTBG [PTBG 1000057495]) | Maui | | 0.75 |
| *Melicope haleakalae* (B. C. Stone) T. G. Hartley & B. C. Stone | Appelhans MA646 (BISH, GOET [GOET019868], PTBG [PTBG 1000057496], US) | Maui | | 0.76 |
| *Melicope haupuensis* (H. St. John) T. G. Hartley & B. C. Stone | Appelhans MA687 (BISH) | Kauaʻi | | 0.73 |
| *Melicope haupuensis* (H. St. John) T. G. Hartley & B. C. Stone | Wood 16794 (PTBG) | Kauaʻi | | 0.65 |
| *Melicope hawaiensis* (Wawra) T.G.Hartley & B.C.Stone | Appelhans MA633 (BISH, GOET [GOET019869], PTBG [PTBG 1000057494], US) | Maui | | 0.78 |
| *Melicope kavaiensis* (H. Mann) T. G. Hartley & B. C. Stone | Appelhans MA679 (BISH, GOET [GOET019871], PTBG [PTBG 1000057501], US) | Kauaʻi | 0.77 | |
| *Melicope knudsenii* (Hillebr.) T.G. Hartley & B.C. Stone | Appelhans MA629 (silica sample only, cultivated at Olinda Rare Plant Facility) | Maui | | 0.71 |
| *Melicope knudsenii* (Hillebr.) T.G. Hartley & B.C. Stone | Oppenheimer H41610 (BISH) | Maui | | 0.66 |
| *Melicope knudsenii* (Hillebr.) T.G. Hartley & B.C. Stone | Wood 17119 (PTBG) | Kauaʻi | | 0.65 |
| *Melicope lydgatei* (Hillebr.) T.G. Hartley & B.C. Stone | Ching s.n. (silica sample only) | Oʻahu | | 0.71 |
| *Melicope makahae* (B. C. Stone) T. G. Hartley & B. C. Stone | Takahama s.n. (silica sample only) | Oʻahu | | 0.71 |
| *Melicope makahae* (B. C. Stone) T. G. Hartley & B. C. Stone (cf.) | Appelhans MA609 (GOET [GOET019872], PTBG [PTBG 1000057509]) | Oʻahu | | 0.76 |

| | | | | |
|---|---|---|---|---|
| *Melicope molokaiensis* (Hillebr.) T. G. Hartley & B. C. Stone | Appelhans MA635 (BISH, GOET [GOET019875], PTBG [PTBG 1000057498]) | Maui | | 0.74 |
| *Melicope molokaiensis* (Hillebr.) T. G. Hartley & B. C. Stone | Appelhans MA643 (BISH, GOET [GOET019874], PTBG [PTBG 1000057560], US) | Maui | | 0.72 |
| *Melicope mucronulata* (H. St. John) T.G. Hartley & B.C. Stone | Appelhans MA630 (silica sample only, cultivated at Olinda Rare Plant Facility) | Maui | | 0.71 |
| *Melicope oahuensis* (H. Lév.) T. G. Hartley & B. C. Stone | Appelhans MA610 (BISH, GOET [GOET019876], PTBG [PTBG 1000057508], US) | Oʻahu | | 0.82 |
| *Melicope orbicularis* (Hillebr.) T. G. Hartley & B. C. Stone | Appelhans MA656 (BISH, GOET [GOET019877], PTBG [PTBG 1000057584], US) | Maui | 0.80 | |
| *Melicope orbicularis* (Hillebr.) T. G. Hartley & B. C. Stone | Appelhans MA659 (GOET [GOET019878], PTBG [PTBG 1000057578]) | Maui | 0.79 | |
| *Melicope ovata* (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | Appelhans MA662 (GOET [GOET019880], PTBG [PTBG 1000057460], US) | Kauaʻi | 0.75 | |
| *Melicope ovata* (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | Appelhans MA663 (BISH, GOET [GOET019879], PTBG [PTBG 1000057427], US) | Kauaʻi | 0.78 | |
| *Melicope ovata* (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | Appelhans MA684 (BISH, GOET [GOET019881]) | Kauaʻi | | 0.73 |
| *Melicope ovata* (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | Wood 17082 (PTBG) | Kauaʻi | | 0.77 |
| *Melicope pallida* (Hillebr.) T. G. Hartley & B. C. Stone | Appelhans MA689 (silica sample only) | Kauaʻi | | 0.77 |
| *Melicope pallida* (Hillebr.) T. G. Hartley & B. C. Stone | Wood 16789 (PTBG) | Kauaʻi | | 0.75 |
| *Melicope paniculata* (H. St. John) T. G. Hartley & B. C. Stone | Perlman 19387 (PTBG) = Appelhans MA660 (silica sample) | Kauaʻi | 0.85 | |
| *Melicope peduncularis* (H. Lév.) T. G. Hartley & B. C. Stone | Appelhans MA613 (BISH, GOET [GOET019882], PTBG [PTBG 1000057524], US) | Oʻahu | | 0.79 |
| *Melicope peduncularis* (H. Lév.) T. G. Hartley & B. C. Stone | Appelhans MA652 (BISH, GOET [GOET019883], PTBG [PTBG 1000057547], US) | Maui | 0.87 | 0.80 |
| *Melicope peduncularis* (H. Lév.) T. G. Hartley & B. C. Stone | Appelhans MA653 (BISH, GOET [GOET019884], PTBG [PTBG 1000057513], US) | Maui | | 0.79 |
| *Melicope pseudoanisata* (Rock) T.G. Hartley & B.C. Stone | Appelhans MA632 (silica sample only, cultivated at Olinda Rare Plant Facility) | Maui | | 0.70 |

| | | | | |
|---|---|---|---|---|
| *Melicope pseudoanisata* (Rock) T.G. Hartley & B.C. Stone | Appelhans MA636 (silica sample only) | Maui | | 0.71 |
| *Melicope pseudoanisata* (Rock) T.G. Hartley & B.C. Stone | Appelhans MA642 (GOET [GOET019885], PTBG [PTBG 1000057554], US) | Maui | | 0.79 |
| *Melicope puberula* (H. St. John) T. G. Hartley & B. C. Stone | Appelhans MA680 ( GOET [GOET019886], PTBG [PTBG 1000057484], US) | Kauaʻi | | 0.73 |
| *Melicope radiata* (H. St. John) T. G. Hartley & B. C. Stone | Appelhans MA698 (BISH, GOET [GOET019888], PTBG [PTBG 1000057523], US) | Hawaiʻi | | 0.71 |
| *Melicope rostrata* (Hillebr.) Appelhans, K.R. Wood & W.L. Wagner | Appelhans MA683 (BISH, GOET [GOET019889]) | Kauaʻi | | 0.85 |
| *Melicope rotundifolia* (A. Gray) T.G. Hartley & B.C. Stone | Ching s.n. (silica sample only) | Oʻahu | 0.72 | |
| *Melicope sandwicensis* (Hook. & Arn.) T.G. Hartley & B.C. Stone | Ching s.n. (silica sample only) | Oʻahu | | 0.69 |
| *Melicope sessilis* (H. Lév.) T. G. Hartley & B. C. Stone | Appelhans MA644 (BISH, GOET [GOET019890], PTBG [PTBG 1000057483], US) | Maui | 0.79 | |
| *Melicope sessilis* (H. Lév.) T. G. Hartley & B. C. Stone | Appelhans MA654 (BISH, GOET [GOET019891], PTBG [PTBG 1000057519], US) | Maui | 0.77 | |
| *Melicope spathulata* A. Gray | Wood 16836 (PTBG [PTBG 1000059483]) | Kauaʻi | | 0.77 |
| *Melicope stonei* K.R. Wood, Appelhans & W.L. Wagner | Wood 17505 (PTBG) | Kauaʻi | 0.81 | |
| *Melicope volcanica* (A. Gray) T.G. Hartley & B.C. Stone (cf.) | Oppenheimer s.n. (silica sample only) | Lānaʻi | | 0.69 |
| *Melicope wawreana* (Rock) T.G. Hartley & B.C. Stone | Wood 17478 (PTBG) | Kauaʻi | 0.86 | |
| **Non-Hawaiian species** | | | | |
| *Melicope elleryana* (F. Muell.) T.G. Hartley | Lorence 6602 (PTBG) | cultivated: NTBG, Kalaheo, Kauaʻi, Hawaii | | 0.70 |
| *Melicope elleryana* (F. Muell.) T.G. Hartley | Appelhans MA404 (LAE, US) | New Guinea | | 0.71 |
| *Melicope elleryana* (F. Muell.) T.G. Hartley | Appelhans MA413 (LAE, US) | New Guinea | | 0.74 |
| *Melicope frutescens* (Blanco) Appelhans & J.Wen | Brambach 464 (GOET) | Indonesia: Sulawesi | | 0.74 |

| | | | | |
|---|---|---|---|---|
| *Melicope latifolia* (DC.) T.G. Hartley | Lorence 10298 (PTBG [PTBG 1000027858]) | cultivated: NTBG, Kalaheo, Kaua'i, Hawaii | | 0.77 |
| *Melicope mantellii* Buchanan | Pelser 3122 (GOET) | New Zealand | | 0.81 |
| *Melicope maxii* T.G. Hartley | Brambach 1916 (GOET) | Indonesia: Sulawesi | | 0.77 |
| *Melicope ternata* J.R. Forst. & G. Forst. | Appelhans MA487 (GOET) | cultivated: Botanical Garden Göttingen | | 0.81 |
| *Melicope triphylla* (Lam.) Merr. | Appelhans MA394 (GOET) | cultivated: Hortus Botanicus Leiden | | 0.87 |

## 2.4 Discussion

All newly reported chromosome numbers of Hawaiian *Melicope* exhibit $n = 18$ or $2n = 36$. Most Amyridoideae (Morton and Telmer, 2014) show identical or similar chromosome numbers (Kubitzki et al., 2011), so that we confirm *Melicope* to be a Palaeocene paleopolyploid. The DNA content of the genus *Melicope* as measured by flow cytometry is also reasonably uniform. None of the estimated DNA amounts represents one and a half times (3n) or twice (4n) that of any other. That includes the assessed specimens of *M. wawraeana* and *M. ovata*, of which earlier studies had indicated a shift in chromosome numbers (Guerra, 1984; Kiehn, 2005). Therefore we conclude that *Melicope* is characterized by a mean DNA amount of 2C = 0.76 pg, which corresponds to the chromosome number $2n = 36$ (Figure 2.3).

Guerra (1984) reported $2n = 72$ for *M. wawraeana*, which might indicate a polyploidization event on the Hawaiian Islands. Since our measurements did not support this result, we conclude that the species as a whole likely did not experience a shift in ploidy level. Instead, our result could indicate that there is an individual or a population of *M. wawraeana* originating from a recent polyploidization event resulting in $2n = 72$ chromosomes. At least 11 genera in Rutaceae are facultative apomicts (Carman, 1997), a reproductive strategy highly associated with polyploidy (Asker and Jerling, 1992). As of yet reproduction in *Melicope* has not been researched, but *Zanthoxylum*, a distantly related genus within the same subfamily (Bayly et al., 2013) reproduces both by facultative apomixis and adventitious embryony, a strategy strongly associated with paleopolyploidy (Carman, 1997).

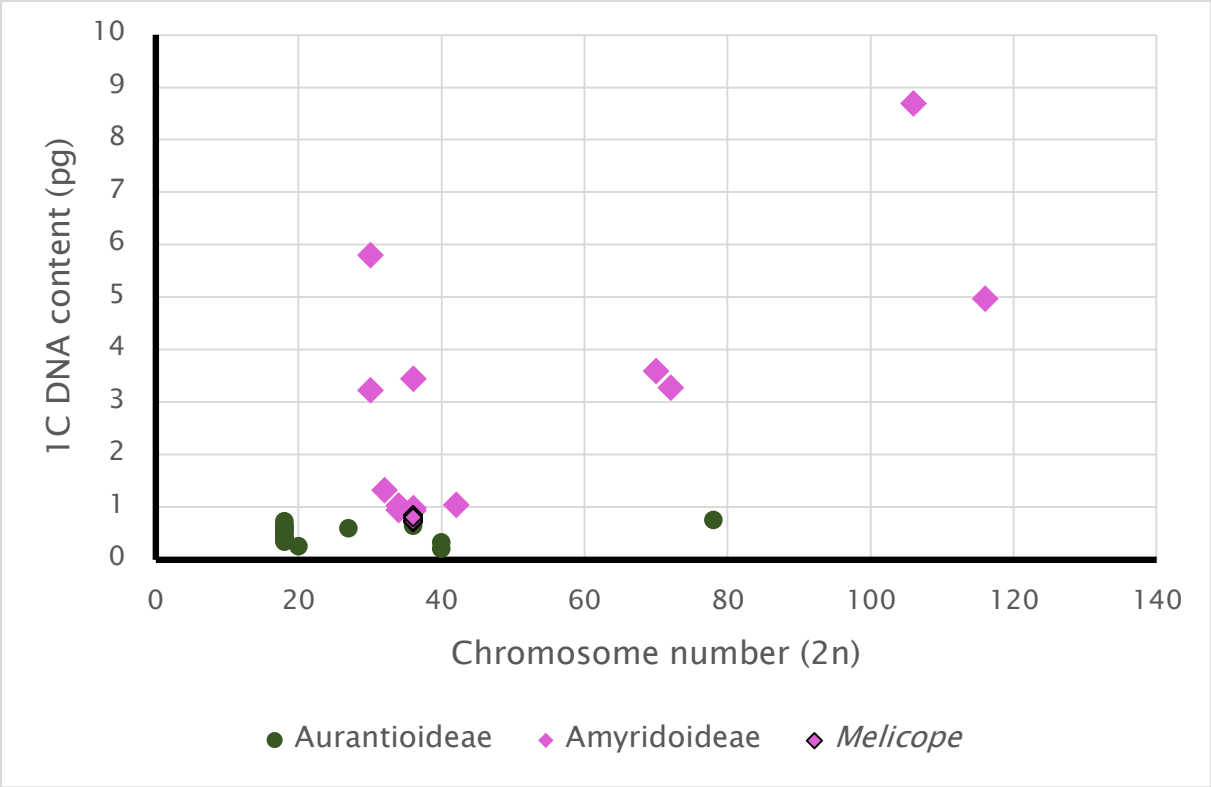



**Figure 2.3.** | Comparison of chromosome numbers and DNA content in 49 species of Rutaceae including newly assessed Hawaiian *Melicope* specimens. Values were extracted from the Kew C-value database (http://data.kew.org/cvalues; assessed on 01. 16. 2017). Green circles represent species in Aurantioideae and Rutoidae (base chromosome number $n = 9$). Violet diamonds represent species in Amyridoideae (base chromosome number $n = 18$). *Melicope* species are indicated by a black frame around the violet diamonds. There is no linear increase of DNA content with increasing ploidy levels. Instead the effects of diploidization can be observed in paleopolyploids with C-values comparable to diploids.

With $2n = 136$-$144$ several species of *Zanthoxylum* have the highest chromosome number known in the family (Kiehn and Lorence, 1996). The observed chromosome number of $2n = 72$ in an individual of *M. wawraeana* (Guerra, 1984) might therefore indicate the influence of apomixis or a recent hybridization event. However, since the species is a member of the youngest clade within Hawaiian *Melicope* (Appelhans et al., 2014b), this putative polyploidization event is not basal in the lineage but would have occurred on the Islands.

The only report of a lower ploidy level in a seedling of *M. ovata* (Kiehn, 2005; as *M. spec.*: $2n = 18$ for one seedling with three other seedlings from the same fruit exhibiting $2n = 36$) cannot be explained with certainty, but might be an effect of irregularities in embryogenesis.

There are only two other reports for *Melicope* of chromosome numbers deviating from $n = 2x = 18$. One is for *M. semecarpifolia*, which was assessed by Pancho (1971) with $n = 18$, but with $n = 12$ by Hsu (1968; as *Euodia confusa* Merr.). Figure 37 of this latter publication shows a drawing of an anaphase I stadium of pollen mother cell meiosis. While it cannot be excluded that the count is correct, the drawing could also be interpreted to show a higher number of chromosomes (personal observation M. Kiehn). All other accounts within section *Lepta* (Table 2.1) revealed $n = 18$ and $2n = 36$ respectively, so this seems to be an isolated deviation, as does the second deviating count of $2n = 32$ in *M. brassii* (Borgman, 1964).

In summary it can be stated that Hawaiian *Melicope* are uniform in terms of chromosome numbers and 1C values (Figure 2.3). Aberrations likely represent local events, e.g., disruptions in embryogenesis, possible hybridization events, chromosome loss or putative effects of apomixis. Also, there is no indication for a difference between Hawaiian representatives of the genus and the remainder of the genus indicating there was no polyploidization event prior to the colonization of the islands.

In terms of the traits for successful island colonization and adaptive radiation, it seems that at least Hawaiian *Melicope* do not exhibit features characteristic for many examples of lineages that colonized distant islands. While sampling herein is not sufficient to exclude polyploidy in all island radiations of the genus, we have shown that the Hawaiian colonizer was not a recently formed neo- or mesopolyploid. Woodiness is a pervading character of the whole genus (Hartley, 2001). Dioecy is present in two subsections of *Melicope, Pelea* and *Lepta* with the latter also containing monoecious species (Hartley, 2001). While the genus as a whole seems to show several shifts in breeding system (compare Appelhans et al., 2014a; Hartley, 2001), the dioecy of the Hawaiian lineage seems to be a trait acquired before the colonization. Up to now, a detailed study on the dispersibility of *Melicope* species has not been undertaken. However, the whole genus displays dehiscent fruits (follicles or capsules), with shiny black seeds, which remain attached upon dehiscence (Hartley, 2001). This, together with the spongy sarcotesta and the thick sclerotesta, likely represents an adaption to bird dispersal (Carlquist, 1966c; Hartley, 2001). This hypothesis is supported by field observations (Frith et al., 1976; Floyd, 1989; Innis, 1989; Hartley, 2001; Medeiros, 2004). Seed size varies in the genus – and indeed within the Hawaiian lineage ranging from relatively small (Ø 2.5 mm) to several times that size (Stone et al., 1999) showing no clear trend for reduction of spatial dispersibility by seed size on the island (Carlquist, 1966a).

There are three possible explanations for the apparent deviation of the genus from the generalist-colonizer-to-specialist-island-endemic pattern.

1.) The-odd-one-out. LDD events are very rare and therefore not governed by regular migration patterns (e.g. Carlquist, 1966a; Appelhans et al., 2018b). Unusual behavior of vectors, catastrophic events or uncommon vectors are suspected causes (Higgins et al., 2003; Nathan et al., 2008). Thus there is a significant element of chance to migration and establishment of a lineage on an island. While certain prerequisites increasing the likelihood of an establishment followed by adaptive radiation exist and researchers seem to have made strides in identifying them, chance might also be an influencing factor here. On the Juan Fernandez Islands 35.6% of the endemic flora is represented by species directly derived from their continental relatives (Stuessy and Crawford, 1998) without any apparent radiation, despite some of them being a member of families renowned for successful island adaptive radiations. Chance may prevent an adaptive radiation in a lineage despite it meeting all identified predispositions or it may allow an 'unexpected' radiation in a lineage not exhibiting any of the facilitating factors. However, that is unlikely the case in *Melicope*, as the Hawaiian radiation is not an exception in an otherwise poorly distributed group. The genus has colonized numerous islands throughout the Pacific, and even colonized Madagascar and the Mascarene Islands radiating into ca. 20 spp. There (Appelhans et al., 2018b). That many successful colonization events followed by adaptive radiation seem unlikely without the genus exhibiting predisposing traits. Due to the rarity of LDD events, exceptional occurrences (Higgins et al., 2003; Nathan et al., 2008), or vectors (Wenny et al., 2016) cannot be ruled out as causes for colonization of an island. However, the adaptations of *Melicope* to bird dispersal (Hartley, 2001) seem to be the key feature facilitating high dispersibility as evidenced by the high number of successful island colonizations (Appelhans et al., 2014a, 2018b).

2.) The hidden generalist. The vast majority of Hawaiian *Melicope* species are highly endemic (about 80% single-island endemics), with only a small number of species being more widespread (Hartley, 2001; Appelhans et al., 2014b). The relatively small distributional niches most of these species occupy certainly fit the picture of the island specialist with a very narrow distributional range. Carlquist (1966a) also observed a loss of dispersibility manifested as an increase in seed size in some species of *Melicope*. On the other hand these specialist Hawaiian lineages spawned two successful independent colonizations of the remote Marquesas Islands (distance > 3500 km) resulting in a local radiation of seven species (Appelhans et al., 2014b, 2018b). Successful colonizations of oceanic islands with subsequent adaptive radiations originating from an insular lineage is a repeated occurrence in the genus

(Appelhans et al., 2014b, 2018b). This indicates the possibility of some species having a broader ecological capacity than suggested by the niches they are observed to occupy. The comparatively small distributional ranges of these species would then likely be due to competition. If this pressure is removed by transmission to another island system with a different species composition, the colonizer may occupy any fraction in a comparatively wide range of ecological conditions. This is corroborated by the fact that both colonizers of the Marquesas Islands are from clades comprising narrowly distributed species (Appelhans et al., 2014b).

3.) The incomplete picture. Although evolutionary patterns on oceanic islands have been a research focus of biologists for more than 200 years, the application of molecular methods has been comparatively recent. Applying these methods to insular radiations and their continental relatives has helped confirm some and rescind other long standing theories. The high morphological diversity in island lineages has often lead to overestimation of the frequency of colonization events, e.g. in the Hawaiian lobeliads (Givnish et al., 2009) or Hawaiian *Cyrtandra* J.R.Forst. & G.Forst. (Cronk et al., 2005; Johnson et al., 2017), or of phylogenetic affiliations as in *Melicope* (Appelhans et al., 2014a), which are rectified by results of molecular investigations. However, most studies focus on resolving phylogenetic relationships (e.g., Givnish et al., 2009) or one specific trait of the island pattern, e.g., dispersal routes (e.g., (Appelhans et al., 2018b) or ploidy levels in lineages (e.g., Harbaugh, 2008) or archipelagos (e.g., Carr, 1998). Attempts of identifying underlying patterns are then made by synergy of these studies. Continued research into adaptive island radiations, especially comparison of displayed traits between species rich lineages and colonizers not undergoing radiation, could help to ultimately identify traits facilitating island adaptive radiations. As of now, the picture is most likely incomplete. For instance, the high proportion of polyploid lineages on islands (e.g., Carr, 1998; Stuessy et al., 1992) indicates polyploidy to be a positive trait. However, we do not have a clear picture here, yet. *Melicope* are paleopolyploid having likely undergone extensive diploidization already as indicated by comparing chromosome counts and genome sizes in Rutaceae (Figure 2.3). All investigated species of *Melicope* including all Hawaiian representatives show genome sizes highly similar to diploid Rutaceae. Therefore the genus has most likely undergone profound post-ploidization diploidization and may be regarded as genetically and cytologically diploidized. However as of yet there are no studies on the formation of bivalents during meiosis; so whether the species' are functionally diploid remains unclear. Research of the Hawaiian silversword alliance (Sakai et al., 1995), the Canarian *Argyranthemum* Webb (Asteraceae; Francisco-Ortega et al., 1997) or Pacific sandalwoods (Harbaugh, 2008)

suggest a recent polyploidization prior to colonization. However, it is entirely unknown whether the colonizer spawning the polyploid Hawaiian lobeliads (Lammers, 1988; Kiehn, 2005) should be considered a neo-, meso- or paleopolyploid. Long term effects of polyploidization and the cytological mechanisms responsible for them are poorly understood (Wendel, 2015). While neopolyploids may exploit the effects of heterosis and gene redundancy (Comai, 2005), meso- and paleopolyploids may exploit ongoing diploidization to maintain genetic diversity over long periods of time (Hohmann et al., 2015). In fact there seems to be a correlation between increased genome downsizing, even beyond the size of the diploid ancestor, and increased diversification rates (Hohmann et al., 2015; Dodsworth et al., 2016). In *Arabidospis thaliana* (L.) Heybh. ($n$ = 5) and several other Angiosperm species' genome reduction during post-polyploidization diploidization has led to a small number of chromosomes and obscured several WGD events (Leitch and Bennett, 2004; Hohmann et al., 2015). The same might be the case in several Hawaiian lineages, possibly even including *Melicope*. Applying genomic methods to Hawaiian plant lineages is required to reliably identify polyploids, their origin and diversity. In addition, even identifying the trait as 'polyploidy' might be misleading. It is entirely plausible, that polyploidy is merely a 'casualty' of the actual trait: hybrid origin. All of the aforementioned neo- and mesopolyploid lineages are allopolyploid and hybridization is suspected to facilitate adaptive radiations (Seehausen, 2013). Seemingly non-polyploid colonizers spawning successful lineages may still be the result of a homoploid hybridization. It has been shown that homoploid hybrid speciation can rapidly reach stability, especially when spatially separated from the parents (Seehausen, 2004). While there are no investigations yet regarding hybridization within Hawaiian *Melicope*, *M. mantellii* Buchanan on New Zealand was suggested to be a hybrid of the closely related *M. simplex* A.Cunn. and *M. ternata* (Cockayne and Allan, 1934). If this is indeed true, it would constitute a case of homoploid hybrid speciation within the genus. Further investigations are needed to reach definitive conclusions regarding not just the trait polyploidy, but the entire pattern. Once we have clearly identified the pattern, we might find Hawaiian *Melicope* to meet it very well.

## Conclusion

With successful colonizations of nearly all Pacific archipelagos, including the remote Hawaiian Islands in the East and Madagascar and the Mascarene Islands in the West, as well as the only known instance of two independent colonizations of the Marquesas Islands within a single genus, *Melicope* shows a very high dispersal ability. Characteristics of successful colonizers were identified as the genomic

flexibility a polyploidization event facilitates, herbaceousness, self-compatibility and high dispersal ability. Successful establishments are characterized by shifts to reduced dispersibility, outcrossing and secondary woodiness. In the case of *Melicope* the main driving factor for successful colonizations seems to be the adaption to bird dispersal. We have shown that the Hawaiian radiation of *Melicope* did not experience a recent polyploidization event prior to colonization of the islands. As the genus is woody and several lines show adaptions to outcrossing (i.e., dioecy), including the clade spawning the Hawaiian lineage, evolutionary shifts characteristic to establishment are observed in the entire genus, not merely in oceanic island lineages. In terms of reduced dispersibility on islands, the picture is not yet clear. Both an increase and a decrease in seed size have been observed, the latter being attributed to an adaption to bog habitats by Carlquist (1966c), but as to how this might affect dispersibility on a case by case basis is unclear. Future research of oceanic lineages will reveal, whether *Melicope* represent a lineage thriving on islands despite not expressing most traits associated with successful colonizations or if we have not yet identified important parts of the island evolution picture.

# 3. | Phylogeny of Hawaiian *Melicope* (Rutaceae): RAD-seq resolves species relationships and reveals ancient introgression

Claudia Paetzold, Kenneth R. Wood, Deren A. R. Eaton, Warren L. Wagner & Marc S. Appelhans

Published in *Frontiers in Plant Science* 2019: 10:1074

doi:10.3389/fpls.2019.01074

Pu‘u Ka‘ala Bog, O‘ahu. Photograph: Marc Appelhans

**Abstract**

Hawaiian *Melicope* are one of the major adaptive radiations of the Hawaiian Islands comprising 54 endemic species. The lineage is monophyletic with an estimated crown age predating the rise of the current high islands. Phylogenetic inference based on Sanger sequencing has not been sufficient to resolve species or deeper level relationships. Here we apply Restriction-site Associated DNA sequencing (RAD-seq) to the lineage to infer phylogenetic relationships. We employ Quartet Sampling to assess information content, statistical support, and to quantify discordance as well as partitioned ABBA-BABA tests to uncover evidence of introgression. Our new results drastically improved resolution of relationships within Hawaiian *Melicope*. The lineage is divided into five fully supported main clades, two of which correspond to morphologically circumscribed infrageneric groups. We provide evidence for both ancestral and current hybridization events. We confirm the necessity for a taxonomic revision of the *Melicope* section *Pelea*, as well as a re-evaluation of several species complexes by combining genomic and morphological data.

## 3.1 Introduction

Oceanic islands have long been a focal point of evolutionary studies as they represent a microcosm for examining the process of speciation. This microcosm is shaped by a combination of factors: (1) islands are geographically small and discrete units, sometimes far removed from continental landmasses; (2) colonizations or secondary arrivals are relatively infrequent, thus gene flow between the source areas and island systems is restricted; and (3) islands often have dynamic geological histories that give rise to extensively varying landscapes with numerous ecological niches (Emerson, 2002; Price and Wagner, 2018). These factors can often lead to high levels of endemism, which is often the result of adaptive radiation of a limited number of colonizers (Price and Wagner, 2004; Losos and Ricklefs, 2009; Keeley and Funk, 2011). Synthesizing the unique aspects of island evolution and extrapolating results to larger scales may allow us to better uncover general patterns and processes in evolution. Such phenomena include identifying factors affecting successful colonization and adaptive radiation (Carlquist, 1967, 1974; Paetzold et al., 2018), morphological or ecological shifts (e.g. "insular woodiness", (Carlquist, 1974; Lens et al., 2013)), the spatiotemporal origins of lineages (Appelhans et al., 2018b), reconstructing colonization events (Harbaugh et al., 2009), and studying co-evolution (Roderick, 1997). These insights may result in further questions regarding taxonomy, species richness, medicinal or technical applications and conservation (e.g. Francisco-Ortega et al., 2000).

Adaptive radiations on islands are of special interest for connecting changes in morphology and ecology through time (Givnish, 1998), but require well-resolved phylogenies to do so. In the Hawaiian Islands, phylogenetic studies based on morphology and taxonomy have sometimes overestimated the number of colonization events, because high levels of morphological diversity led researchers to overestimate lineage diversity and the number of colonization events (Price and Wagner, 2018). In contrast, molecular phylogenetic studies have revealed that many enigmatic Hawaiian plant radiations colonized the islands only once followed by adaptive radiation: the Hawaiian lobeliads (Campanulaceae; Givnish et al., 2009), *Psychotria* (Rubiaceae; Nepokroeff et al., 2003), *Silene* (Caryophyllaceae; Eggens et al., 2007), *Touchardia*/*Urera* (Urticaceae; Wu et al., 2013), and *Melicope* (Harbaugh et al., 2009; Appelhans et al., 2014b). Polyploidization and hybridization events were also discovered to predate colonization and radiation in several island lineages, including the Hawaiian silverswords (Asteraceae; Baldwin and Sanderson, 1998; Barrier et al., 1999) and mints (Lamiaceae; Roy et al., 2015) along with the pan-Pacific sandalwoods

(Santalaceae; Harbaugh, 2008) suggesting evolutionary success in young hybrid or polyploid colonists (Carr, 1998; Paetzold et al., 2018).

Time-scaled phylogenies have revealed that most Hawaiian radiations are ≤5 Myr old, which corresponds to the age of the oldest current main islands, Kauaʻi and Niʻihau. This suggests a bottleneck for dispersal from older (and now largely submerged) leeward islands to the current main islands. However, there are several known exceptions of lineages older than 5 Myr, including *Drosophila*, damselflies, lobeliads, *Zanthoxylum* (Rutaceae), as well as *Melicope* (Price and Clague, 2002; Keeley and Funk, 2011; Appelhans et al., 2018b, 2018a). Most phylogenetic studies of Hawaiian flora, however, have relied on few sequenced loci, and have thus lacked sufficient power to resolve recent rapid radiations where hybridization, incomplete lineage sorting, and polyploidy may be common. Newer genomic tools are likely to provide more accurate estimates that may transform our understanding of island radiations.

The genus *Melicope* comprises about 235 species of shrubs and trees distributed throughout SE Asia and Australasia, extending to the Mascarene Islands and Madagascar in the West and most of the Pacific Archipelagos in the East (Hartley, 2001). There are 54 species of *Melicope* endemic to the Hawaiian Islands (Appelhans et al., 2017; Wood et al., 2017), 41 of which are single island endemics (Stone et al., 1999). Hawaiian *Melicope* were initially placed in the genus *Pelea* together with species from the Marquesas Islands (Stone, 1969; Stone et al., 1999) but later incorporated into *Melicope* forming the majority of the section *Pelea* (Hartley, 2001). Hawaiian *Pelea* was divided into four sections based mainly on the grade of carpel connation: *Apocarpa, Cubicarpa, Megacarpa,* and *Pelea.* Since the incorporation of the genus *Pelea* into *Melicope,* these sections have not been formally recognized within the larger infrageneric taxonomy for *Melicope* as recognized by Hartley (2001), but are still being used informally as species groups (Appelhans et al., 2014b) and we refer to them as Stone's sectional species groups (Stone's sections) from here on. The most current recent and comprehensive taxonomic treatment of Hawaiian *Melicope* was considered 'provisional' by the authors (Stone et al., 1999), as species boundaries are difficult to define in some cases. Examples include three described species complexes, where the incorporated species vary from each other primarily in the degree of fruit pubescence; the *M. elliptica* complex based mainly in Oʻahu (6 species), the Hawaiian based *M. volcanica* complex (4 species) and the Kauaʻi based *M. kavaiensis* complex (5 species) (Stone et al., 1999).

In contrast to other successful island radiations, the colonization of the Hawaiian Archipelago in *Melicope* was not preceded by a recent polyploidization event. In

general, the genus *Melicope* shows a uniform chromosome number (Paetzold et al., 2018). To date, phylogenetic relationships in Hawaiian *Melicope* have been investigated in four molecular studies (Harbaugh et al., 2009; Appelhans et al., 2014a, 2014b, 2018b), with a combination of up to six nuclear and plastid genomic regions amplified using PCR. Hawaiian *Melicope* was shown to be derived from a single colonization event (Harbaugh et al., 2009). The origin of the lineage was dated to the Mid or Late Miocene (Appelhans et al., 2018b), predating the age of Kauaʻi and Niʻihau (Price and Clague, 2002). In addition, the Hawaiian endemic genus *Platydesma* is nested within *Melicope* as a monophyletic sister group to the Hawaiian species and has since been reduced (Appelhans et al., 2017). Statistically supported incongruences between individual genomic regions were not observed, yet the resolution of relationships within and among the clades was in general medium to poor (Harbaugh et al., 2009; Appelhans et al., 2014a, 2014b, 2018b). However, two independent Hawaiian origins of the Marquesan *Melicope* radiation, which encompasses 7 species, were inferred (Appelhans et al., 2014a, 2014b, 2018b).

Restriction-site associated sequencing (RAD-seq; Baird et al., 2008; Miller et al., 2007) is among the most frequently used reduced representation methods employed in plant systematics. To date, most phylogenetic RAD-seq studies have focused mostly on populations or closely related species (Ree and Hipp, 2015; Díaz-Arce et al., 2016; Hodel et al., 2017). However, a simulated RAD investigation in *Drosophila* revealed the method to be potentially applicable in groups aged up to 60 million years (Rubin et al., 2012). Since then, application to deeper species-level relationships has increased (Eaton and Ree, 2013; Hipp et al., 2014; Eaton et al., 2017) facilitated by the development of RAD-seq assembly pipelines targeted at phylogenetic research (Eaton, 2014).

Incongruence between datasets has been a long-standing occurrence in molecular phylogenetic inference, traditionally manifesting as incongruences between different gene trees. The advance of NGS technology has shown that the issue is not solved by merely incorporating more data (Jeffroy et al., 2006). There are three possible categories of confounding information in a phylogenetic study: noise, systematic error and an underlying biological signal. Noise is an effect of the inherently stochastic nature of sequence evolution and leads to a deterioration of phylogenetic signal over time. As such, noise most heavily impacts very small datasets and deep nodes (Misof et al., 2014). Incongruence may also reflect a true biological signal, e.g. the presence of incomplete lineage sorting (ILS) or non-tree-like evolution, i.e. introgression, hybridization, or recombination (Misof et al., 2014; Salichos et al., 2014) (Misof et al., 2014; Salichos et al., 2014). Effects of hybridization range from

introgression of individual alleles, organelle capture, to hybrid speciation (Currat et al., 2008; Stegemann et al., 2012; Twyford and Ennos, 2012). Either of these processes will result in discordant gene trees and several approaches have been proposed to unravel them. Based on the distributions of conflicting phylogenetic patterns in the genome, it is possible to distinguish the more stochastic signal of ILS from the directional and asymmetric signal of hybridization (Durand et al., 2011).

Here we apply RAD-seq to Hawaiian *Melicope*, a lineage with a crown age of ca. 10 Myr (Appelhans et al., 2018b). We use RAD-seq to infer species-level relationships in the lineage; in a phylogenetic context of several colonization events of individual islands, multiple possible bottlenecks and adaptive radiations within a lineage. The taxonomic implications of our phylogenetic results are discussed within the framework of evidence for both ancient and current introgression.

## 3.2 Material & Methods

### Taxon Sampling

Table 3.1 details the identity and origin of the 101 samples of this study; 6 outgroup and 95 ingroup specimens representing 41 Hawaiian species (81% of the lineage). Two samples represent the two independent colonization events to the Marquesas Islands (28% of Marquesan species). Taxonomic treatment follows species recognized in Wood et al. (2016) plus a recently described species (Wood et al., 2017) and including *Platydesma* (Appelhans et al., 2017). Additionally, morphologically divergent specimens of *M. barbigera* (KW16722, KW16718) and *M. ovata* (KW16762, KW17082, MA663) were included (Table 3.1, asterisk) to elucidate whether these might represent separate taxa. We also included two specimens, KW17111 and KW15733, which correspond closely, though not entirely to the description of *M. wawraeana* as delimited by Stone et al. (1999). Even the O‘ahu populations that were considered the core of *M. wawraeana* are variable, suggesting it is a potentially artificial taxon (Stone et al., 1999). Since the morphology of the two specimens did not correspond entirely to the O‘ahu populations considered to be *M. wawraeana*, we included them here as *M.* sp. (Table 3.1).

**Table 3.1.** | Samples within this study including origin, voucher placement and assignment to Stone's sections. Asterisks mark morphologically deviating specimens. Samples in bold were used in parameter optimization. ORPF: cultivated at Olinda Rare Plant Facility.

| Species | Stone's section | Collection number, Herbarium voucher | Origin |
|---|---|---|---|
| ***Melicope adscendens* (H. St. John & E.P. Hume) T.G. Hartley & B. C. Stone** | ***Apocarpa*** | **Appelhans MA628 (silica sample only, ORPF)** | **Maui** |
| *Melicope anisata* (H. Mann) T. G. Hartley & B. C. Stone | *Cubicarpa* | Appelhans MA665 (GOET, PTBG) | Kauaʻi |
| *Melicope anisata* (H. Mann) T. G. Hartley & B. C. Stone | *Cubicarpa* | Appelhans MA668 (GOET, PTBG, US) | Kauaʻi |
| *Melicope balloui* (Rock) T.G.Hartley & B.C.Stone | *Megacarpa* | Wood KW7685 (PTBG) | Maui |
| ***Melicope barbigera* A. Gray** | ***Apocarpa*** | **Appelhans MA666 (BISH, GOET, PTBG, US)** | **Kauaʻi** |
| *Melicope barbigera* A. Gray | *Apocarpa* | Wood KW15333 (PTBG) | Kauaʻi |
| *Melicope barbigera* A. Gray | *Apocarpa* | Wood KW15449 (PTBG) | Kauaʻi |
| *Melicope barbigera* A. Gray | *Apocarpa* | Wood KW15961 (PTBG) | Kauaʻi |
| *Melicope barbigera** A. Gray | *Apocarpa* | Wood KW16722 (PTBG) | Kauaʻi |
| *Melicope barbigera** A. Gray | *Apocarpa* | Wood KW16718 (PTBG) | Kauaʻi |
| *Melicope christophersenii* (H. St. John) T. G. Hartley & B. C. Stone | *Megacarpa* | Appelhans MA618 (BISH, GOET, PTBG, US) | Oʻahu |
| *Melicope christophersenii* (H. St. John) T. G. Hartley & B. C. Stone | *Megacarpa* | Appelhans MA621 (silica sample only, cultivated at Puʻu Kaʻala) | Oʻahu |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA615 (GOET, PTBG) | Oʻahu |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA617 | Oʻahu |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA634 (PTBG) | Maui |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA650 (GOET, PTBG, US) | Maui |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA651 (BISH, GOET, PTBG, US) | Maui |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA655 (silica sample only) | Maui |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA657 (GOET, PTBG, US) | Maui |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA670 | Kauaʻi |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA672 | Kauaʻi |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA693 | Hawaiʻi |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA695 | Hawaiʻi |

| | | | |
|---|---|---|---|
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Oppenheimer s.n. (silica sample only) | Maui |
| ***Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone** | ***Pelea*** | **Oppenheimer H91641 (US)** | **Lānaʻi** |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | *Pelea* | Wood KW16146 (PTBG) | Kauaʻi |
| ***Melicope clusiifolia* (Gray) T.G.Hartley & B.C.Stone** | ***Pelea*** | **Appelhans MA675** | **Kauaʻi** |
| *Melicope cornuta* (Hillebr.) Appelhans, K.R.Wood & W.L.Wagner | *Platydesma* | Ching s.n. (silica sample only) | Oʻahu |
| ***Melicope cornuta* var. *decurrens* (B.C. Stone) Appelhans, K.R. Wood & W.L. Wagner** | ***Platydesma*** | **Takahama s.n. (silica sample only)** | Oʻahu |
| *Melicope cruciata* (A. Heller) T.G. Hartley & B.C. Stone | *Megacarpa* | Wood KW16251 (PTBG) | Kauaʻi |
| *Melicope degeneri* (B.C.Stone) T.G.Hartley & B.C.Stone | *Cubicarpa* | Wood KW15903 (PTBG) | Kauaʻi |
| *Melicope degeneri* (B.C.Stone) T.G.Hartley & B.C.Stone | *Cubicarpa* | Wood KW15984 (PTBG) | Kauaʻi |
| *Melicope feddei* (H. Lév.) T. G. Hartley & B. C. Stone | *Megacarpa* | Appelhans MA688 (BISH, GOET, PTBG, US) | Kauaʻi |
| *Melicope feddei* (H. Lév.) T. G. Hartley & B. C. Stone | *Megacarpa* | Wood KW15844 (PTBG) | Kauaʻi |
| ***Melicope haleakalae* (B. C. Stone) T. G. Hartley & B. C. Stone** | ***Pelea*** | **Appelhans MA645 (BISH, GOET, PTBG)** | **Maui** |
| *Melicope haleakalae* (B. C. Stone) T. G. Hartley & B. C. Stone | *Pelea* | Appelhans MA646 (BISH, GOET, PTBG, US) | Maui |
| ***Melicope haupuensis* (H. St. John) T. G. Hartley & B. C. Stone** | ***Apocarpa*** | **Appelhans MA687 (BISH)** | Kauaʻi |
| *Melicope haupuensis* (H. St. John) T. G. Hartley & B. C. Stone | *Apocarpa* | Wood KW16791 (PTBG) | Kauaʻi |
| *Melicope haupuensis* (H. St. John) T. G. Hartley & B. C. Stone | *Apocarpa* | Wood KW16794 (PTBG) | Kauaʻi |
| *Melicope hawaiensis* (Wawra) T.G.Hartley & B.C.Stone | *Apocarpa* | Appelhans MA633 (BISH, GOET, PTBG, US) | Maui |
| *Melicope hawaiensis* (Wawra) T.G.Hartley & B.C.Stone | *Apocarpa* | Appelhans MA700 | Hawaiʻi |
| *Melicope hawaiensis* (Wawra) T.G.Hartley & B.C.Stone | *Apocarpa* | Oppenheimer s.n. (silica sample only) | Maui |
| *Melicope hiiakae* (B.C.Stone) T.G.Hartley & B.C.Stone | *Megacarpa* | Ching s.n. (silica sample only) | Oʻahu |
| *Melicope hivaoaensis* J.Florence | | Meyer 826 | Hivaoa, Marquesas Islands |
| *Melicope inopinata* J.Florence | | Meyer 887 | Hivaoa, Marquesas Islands |

| | | | |
|---|---|---|---|
| ***Melicope kavaiensis* (H. Mann) T. G. Hartley & B. C. Stone** | ***Megacarpa*** | **Appelhans MA679 (BISH, GOET, PTBG, US)** | **Kauaʻi** |
| *Melicope knudsenii* (Hillebr.) T.G. Hartley & B.C. Stone | *Apocarpa* | Appelhans MA629 (silica sample only, ORPF) | Maui |
| ***Melicope knudsenii* (Hillebr.) T.G. Hartley & B.C. Stone** | ***Apocarpa*** | **Oppenheimer H41610 (BISH)** | **Maui** |
| *Melicope knudsenii* (Hillebr.) T.G. Hartley & B.C. Stone | *Apocarpa* | Wood KW17119 (PTBG) | Kauaʻi |
| *Melicope lydgatei* (Hillebr.) T.G. Hartley & B.C. Stone | *Megacarpa* | Ching s.n. (silica sample only) | Oʻahu |
| ***Melicope makahae* (B. C. Stone) T. G. Hartley & B. C. Stone** | ***Apocarpa*** | **Takahama s.n. (silica sample only)** | **Oʻahu** |
| *Melicope makahae* (B. C. Stone) T. G. Hartley & B. C. Stone (cf.) | *Apocarpa* | Appelhans MA609 (GOET, PTBG) | Oʻahu |
| *Melicope molokaiensis* (Hillebr.) T. G. Hartley & B. C. Stone | *Megacarpa* | Appelhans MA635 (BISH, GOET, PTBG) | Maui |
| *Melicope molokaiensis* (Hillebr.) T. G. Hartley & B. C. Stone | *Megacarpa* | Appelhans MA643 (BISH, GOET, PTBG, US) | Maui |
| ***Melicope molokaiensis* (Hillebr.) T. G. Hartley & B. C. Stone** | ***Megacarpa*** | **Oppenheimer s.n. (silica sample only)** | **Maui** |
| *Melicope mucronulata* (H. St. John) T.G. Hartley & B.C. Stone | *Apocarpa* | Appelhans MA630 (silica sample only, ORPF) | Maui |
| *Melicope munroi* (St.John) T.G.Hartley & B.C.Stone | *Megacarpa* | Oppenheimer s.n. (silica sample only) | Lānaʻi |
| *Melicope oahuensis* (H. Lév.) T. G. Hartley & B. C. Stone | *Cubicarpa* | Appelhans MA610 (BISH, GOET, PTBG, US) | Oʻahu |
| *Melicope oahuensis* (H. Lév.) T. G. Hartley & B. C. Stone | *Cubicarpa* | Ching s.n. (silica sample only) | Oʻahu |
| ***Melicope oppenheimeri* K.R.Wood, Appelhans & W.L.Wagner** | ***Megacarpa*** | **Wood KW7419 (PTBG)** | **Maui** |
| *Melicope oppenheimeri* K.R.Wood, Appelhans & W.L.Wagner | *Megacarpa* | Wood KW7408 (PTBG) | Maui |
| *Melicope orbicularis* (Hillebr.) T. G. Hartley & B. C. Stone | *Megacarpa* | Appelhans MA656 (BISH, GOET, PTBG, US) | Maui |
| *Melicope orbicularis* (Hillebr.) T. G. Hartley & B. C. Stone | *Megacarpa* | Appelhans MA659 (GOET, PTBG) | Maui |
| *Melicope ovalis* (St.John) T.G.Hartley & B.C.Stone | *Cubicarpa* | Wood KW13724 (PTBG) | Maui |
| ***Melicope ovata* (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone** | ***Apocarpa*** | **Appelhans MA662 (GOET, PTBG, US)** | **Kauaʻi** |
| *Melicope ovata* (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | *Apocarpa* | Appelhans MA684 (BISH, GOET) | Kauaʻi |

| | | | |
|---|---|---|---|
| *Melicope ovata** (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | *Apocarpa* | Appelhans MA663 (BISH, GOET, PTBG, US) | Kauaʻi |
| *Melicope ovata** (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | *Apocarpa* | Wood KW17082 (PTBG) | Kauaʻi |
| *Melicope ovata** (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | *Apocarpa* | Wood KW16762 (PTBG) | Kauaʻi |
| *Melicope pallida* (Hillebr.) T. G. Hartley & B. C. Stone | *Apocarpa* | Appelhans MA689 (silica sample only) | Kauaʻi |
| *Melicope pallida* (Hillebr.) T. G. Hartley & B. C. Stone | *Apocarpa* | Wood KW16789 (PTBG) | Kauaʻi |
| *Melicope pallida* (Hillebr.) T. G. Hartley & B. C. Stone | *Apocarpa* | Wood KW15571 (PTBG) | Kauaʻi |
| ***Melicope paniculata* (H. St. John) T. G. Hartley & B. C. Stone** | ***Cubicarpa*** | **Perlman 19387 (PTBG) = Appelhans MA660 (silica sample)** | **Kauaʻi** |
| *Melicope paniculata* (H. St. John) T. G. Hartley & B. C. Stone | *Cubicarpa* | Wood KW16155 (PTBG) | Kauaʻi |
| *Melicope peduncularis* (H. Lév.) T. G. Hartley & B. C. Stone | *Cubicarpa* | Appelhans MA652 (BISH, GOET, PTBG, US) | Maui |
| *Melicope peduncularis* (H. Lév.) T. G. Hartley & B. C. Stone | *Cubicarpa* | Appelhans MA653 (BISH, GOET, PTBG, US) | Maui |
| ***Melicope pseudoanisata* (Rock) T.G. Hartley & B.C. Stone** | ***Megacarpa*** | **Appelhans MA632 (silica sample only, ORPF)** | **Maui** |
| *Melicope pseudoanisata* (Rock) T.G. Hartley & B.C. Stone | *Megacarpa* | Appelhans MA636 (silica sample only) | Maui |
| *Melicope pseudoanisata* (Rock) T.G. Hartley & B.C. Stone | *Megacarpa* | Appelhans MA642 (GOET, PTBG, US) | Maui |
| *Melicope puberula* (H. St. John) T. G. Hartley & B. C. Stone | *Megacarpa* | Appelhans MA680 ( GOET, PTBG, US) | Kauaʻi |
| *Melicope puberula* (H. St. John) T. G. Hartley & B. C. Stone | *Megacarpa* | Wood KW16058 (PTBG) | Kauaʻi |
| *Melicope radiata* (H. St. John) T. G. Hartley & B. C. Stone | *Megacarpa* | Appelhans MA696 | Hawaiʻi |
| ***Melicope rostrata* (Hillebr.) Appelhans, K.R. Wood & W.L. Wagner** | ***Platydesma*** | **Appelhans MA683 (BISH, GOET)** | **Kauaʻi** |
| ***Melicope rotundifolia* (A. Gray) T.G. Hartley & B.C. Stone** | ***Megacarpa*** | **Ching s.n. (silica sample only)** | **Oʻahu** |
| *Melicope sandwicensis* (Hook. & Arn.) T.G. Hartley & B.C. Stone | *Apocarpa* | Ching s.n. (silica sample only) | Oʻahu |
| *Melicope sessilis* (H. Lév.) T. G. Hartley & B. C. Stone | *Megacarpa* | Appelhans MA644 (BISH, GOET, PTBG, US) | Maui |
| *Melicope* sp. (Rock) T.G.Hartley & B.C.Stone | *Megacarpa* | Wood KW17111 (PTBG) | Kauaʻi |
| *Melicope* sp. (Rock) T.G.Hartley & B.C.Stone | *Megacarpa* | Wood KW15733 (PTBG) | Kauaʻi |

| | | | |
|---|---|---|---|
| *Melicope spathulata* A. Gray | *Platydesma* | Appelhans MA697 | Hawai‘i |
| *Melicope spathulata* A. Gray | *Platydesma* | Wood KW16743 (PTBG) | Kaua‘i |
| *Melicope spathulata* A. Gray | *Platydesma* | Wood KW16836 (PTBG) | Kaua‘i |
| *Melicope stonei* K.R.Wood, Appelhans & W.L.Wagner | *Apocarpa* | Appelhans MA691 | Kaua‘i |
| *Melicope stonei* K.R.Wood, Appelhans & W.L.Wagner | *Apocarpa* | Wood KW16727 (PTBG) | Kaua‘i |
| *Melicope volcanica* (A. Gray) T.G. Hartley & B.C. Stone (cf.) | *Megacarpa* | Oppenheimer s.n. (silica sample only) | Lāna‘i |
| *Melicope waialealae* (Wawra) T.G.Hartley & B.C.Stone | *Pelea* | Wood KW16015 (PTBG) | Kaua‘i |
| **outgroup** | | | |
| *Melicope aneura* (Lauterb.) T.G.Hartley | | Appelhans MA418 (LAE, US) | Papua New Guinea |
| *Melicope brassii* T.G.Hartley | | Appelhans MA436 (LAE, US) | Papua New Guinea |
| *Melicope durifolia* (K.Schum.) T.G.Hartley | | Appelhans MA455 (LAE, US) | Papua New Guinea |
| *Melicope durifolia* (K.Schum.) T.G.Hartley | | Appelhans MA465 (LAE, US) | Papua New Guinea |
| *Melicope polyadenia* Merr. & L.M.Perry | | Appelhans MA438 (LAE, US) | Papua New Guinea |
| *Melicope triphylla* Merr. | | Appelhans MA394 (GOET) | cultivated Hortus Botanicus Leiden |

### RAD library preparation

DNA was extracted from silica-dried material using the Qiagen DNeasy Plant Mini Kit® (Qiagen, Hilden, Germany) as per the manufacturer's instructions with incubation in lysis buffer elongated to two hours. DNA concentration was measured using the Qubit® fluorometer and the Qubit® dsDNA BR Assay Kit (ThermoFisher Scientific, Darmstadt, Germany) and adjusted to 30 ng/µl. Floragenex Inc. (Portland, Oregon, USA) generated RAD libraries using the restriction enzyme *Sfb*I. Employing a method following Baird et al. (2008), including the use of sample-specific barcodes, the samples were sequenced on an Illumina® GAIIx platform to produce 100 bp single-end reads.

### RAD locus assembly

Quality of raw reads was checked using FastQC (Andrews, 2010). The program *ipyrad* v.0.7.21 (Eaton, 2014) was used to demultiplex raw reads allowing a mismatch of 1 bp. Raw reads were trimmed using cutadapt v.1.9.1 (Martin, 2011) as implemented in *ipyrad* by removing adapter sequences, trimming bases with Phred-Scores < 30 and removing reads shorter than 35 bp after trimming. Trimmed reads were assembled *de novo* using the *ipyrad* pipeline. The software attempts to evaluate

orthology by scoring alignments of reads or sequences, as opposed to assessing purely sequence identity (Eaton, 2014). The alignment score is the user-determined clustering threshold to be met. To reduce the risk of introducing assembly error to our dataset, we performed a modified clustering optimization approach (Paris et al., 2017). We iterated over core clustering parameters and plotted assembly matrices (cluster depth, heterozygosity, number of putatively paralogous loci, number of SNPs) to identify parameters introducing excessive assembly errors (Paetzold et al., unpublished; Paris et al., 2017). In addition, we optimized the clustering of reads within each individual sample and the clustering of consensus sequences across loci separately, reasoning that the divergence found within each individual genome might be significantly different from the ca. 10 Myr of divergence (Appelhans et al., 2018b) within the lineage as a whole. Thus, the assembly was generated using a clustering threshold of 95 for in-sample-clustering and 90 for between-sample-clustering. Final filtering of loci was performed for values 10, 32, 50, 67 and 85 as the minimum number of samples per locus.

**Phylogenetic inference & Quartet sampling**

Phylogenetic inference was performed on all resulting alignments using Maximum Likelihood (ML) and Bayesian inference (BI). As individual loci are very short and may comprise a high fraction of missing data, a partitioned analysis is neither computationally feasible nor expected to produce reliable results. Thus, all datasets were analyzed solely concatenated. ML was performed using ExaML v3.0.2 (Kozlov et al., 2015) using the new rapid hill-climbing algorithm, a random number seed, the GAMMA model of rate heterogeneity and using the median for discrete approximation of rate heterogeneity. For datasets containing a minimum number of 10, 32 and 50 samples, the memory saving option for gappy alignments was activated (-S). Parsimony starting trees were generated using RAxML v8.2.4 (Stamatakis, 2014). RAxML was also used to generate 100 bootstrap replicate alignments and their corresponding Parsimony Starting Trees. ExaML searches were run on every replicate alignment with the above-mentioned settings.

BI was performed using ExaBayes v 1.5 (Aberer et al., 2014). Four independent runs were carried out with a convergence stopping criterion (split frequencies average <5% in three subsequent generations) and for a minimum of 100,000 generations sampling every 100$^{th}$ generation under the GTR + I + G model. Majority rule consensus trees were drawn on topologies of all four runs combined after the first 25% was discarded as burn-in.

Analysis of large-scale, concatenated datasets can result in erroneous relationships with high bootstrap support because of a failure to model the effects of ILS

(Gadagkar et al., 2005; Kubatko and Degnan, 2007; Seo, 2008). These effects can be driven by only a few loci (Shen et al., 2017) and especially pertain to short branches (Kumar et al., 2012). On the other hand, a simulation study has shown, that concatenated analysis of datasets containing loci with anomalous gene trees will more likely result in unresolved species tree topologies, rather than highly supported false ones (Huang and Knowles, 2009).

Methods implementing the Multispecies Coalescent Model (MSC) explicitly incorporate gene tree conflict into species tree inference and are thus more robust to ILS than concatenation approaches (Kubatko and Degnan, 2007), but are often intractable for large datasets (Liu et al., 2015). Summary methods of species tree inference under the MSC e.g., ASTRAL (Mirarab et al., 2014b) or NJst (Liu and Yu, 2011), are based on the analysis of individual gene trees and have become popular due to their comparative speed and accuracy. However, the limited information content of individual RAD loci often limits their application for gene tree inference, which may negatively impact species tree estimation (Salichos and Rokas, 2013; Mirarab et al., 2016). Alternatively, site-based methods avoid estimation of gene trees, instead using SNP data directly, and so are expected to be well suited to short, low-variability loci (Molloy and Warnow, 2018). We employed the SVDQuartets method, which infers quartet trees from SNPs using phylogenetic invariant patterns under the coalescent model and then infers the species tree by quartet joining of the subtrees using algebraic statistics (Chifman and Kubatko, 2014). We converted the SNP datasets into nexus format using the Ruby script *convert_vcf_to_nexus.rb* (Matschiner, 2019). The SVDQuartets analysis was computed as implemented in the software PAUP*4.0a (Swofford, 2002, 2018). We analysed 250,000 randomly selected quartets and assessed statistical support using 100 non-parametric bootstrap support replicates. For ambiguous positions in the SNP matrix, we chose the "Distribute" option, as these positions represent heterozygous sites.

To estimate the robustness of resolved relationships, we employed the Quartet Sampling method, which aims to measure branch support in large sparse alignments (Pease et al., 2018). As each internal branch divides all samples within a phylogeny into four non-overlapping subsets, the method randomly samples one taxon per subset to produce a quartet phylogeny. The topology of each quartet is either concordant with the tree topology or discordant. Discord is measured and quantified to produce four metrics - Quartet Concordance (QC), Quartet Differential (QD), Quartet Informativeness (QI), and Quartet Fidelity (QF) – allowing effective assessment of branch-related (QC, QD, QI) and taxon-related (QF) discordance in the dataset (Pease et al., 2018). The method is implemented in the python script

*quartet_sampling.py* (https://www.github.com/fephyfofum/quartetsampling).We performed quartet sampling on all datasets and the respectively resolved topologies using 500 replicates per branch with a minimum required overlap of 300,000 bp in the min10, min32, min50 and min67 concatenated datasets. The minimum overlap was lowered to 140,000 bp in the min85 concatenated dataset, as otherwise 5 samples would have been excluded from the analysis.

### Test for introgression

The *D*-statistics (Durand et al., 2011) is a site-based test for introgression. In a four-taxon topology (((P1, P2), P3), O), a derived allele in the P3 lineage is expected to occur also in either P1 or P2 with equal frequency, giving rise to either an ABBA or BABA discordant site pattern (Durand et al., 2011).

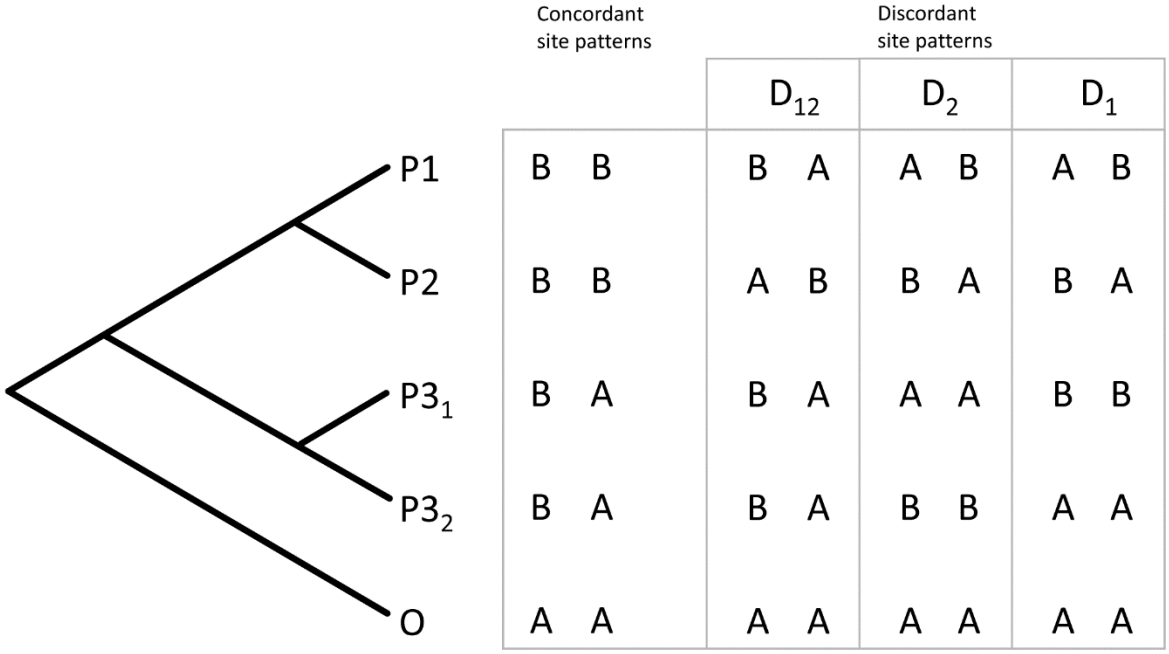



**Figure 3.1.** | The principle of five-taxon D-statistics test. Biallelic site patterns are quantified, which support or contradict the underlying symmetric phylogeny. Asymmetry of discordant site patterns is quantified to calculate three separate D-statistics characterizing introgression from the $P3_1$ taxon ($D_1$), the $P3_2$ taxon ($D_2$) or their common ancestor ($D_{12}$) into the taxa designated $P_1$ and $P_2$ (Eaton and Ree, 2013).

A statistically significant imbalance in these site pattern frequencies provides evidence of introgression, while equal frequencies are associated with neutral processes like ILS. Unfortunately, this test is not well suited for deeper evolutionary time scales, where the P3 lineage has diverged into multiple sub-lineages, and it also does not allow inference of direction of introgression. Partitioned *D*-statistics are a system of multiple four-taxon *D*-statistics in a symmetric, five-taxon phylogeny with

the ingroup taxa forming two pairs (P1, P2) and ($P3_1$, $P3_2$) and an outgroup taxon (O) (Figure 3.1) (Eaton and Ree, 2013). The partitioned *D*-statistics identifies sites, in which either or both of the P3 lineages share a derived allele with either P1 or P2, but not both (Figure 3.1) (Eaton and Ree, 2013).

We used partitioned *D*-statistics to infer whether discordant relationships inferred between major clades (see below) are caused by ILS or introgression. We defined entire clades as lineages and tested all combinations obeying the symmetric topology.

## 3.3 Results

### Raw data and assembly

Illumina Sequencing yielded an average of 10,439,082 reads per sample (342,914 - 34,663,109). After quality trimming an average of 10,327,562 reads per sample (271,257 - 34,542,777) were left. The assembled dataset contained a total of 786,169 clusters prior to filtering by sample coverage. Filtering reduced the number of loci by over 90 % (Table 2). The final datasets contained between 7,266 (min85) and 59,041 (min10) loci. The number of variable sites (SNPs) ranged from 529,045 (min10) to 82,760 (min85) (Table 3.2).

**Table 3.2.** | Differences between the number of loci, their concatenated length and the number of SNPs resulting from filtering by minimum samples per locus (10, 32, 50, 67 and 85).

| | **total** | **min10** | **min32** | **min50** | **min67** | **min85** |
|---|---|---|---|---|---|---|
| **number of loci** | 786,169 | 59,041 | 36,622 | 30,801 | 23,401 | 7,266 |
| **concatenated length (bp)** | NA | 4,800,367 | 2,986,760 | 2,506,242 | 1,892,473 | 584,086 |
| **number of SNPs** | NA | 529,045 | 385,871 | 332,935 | 256,276 | 82,760 |

**Phylogenetic Inference**

All five final datasets were used for phylogenetic inference in concatenated BI, ML, and SVDquartets analyses. Statistical support for inferred relationships was assessed using Posterior Probabilities (PP), Nonparametric Bootstrap (ML-NBS, SVD-NBS) and Quartet Sampling. Analyses of the five datasets resulted in mostly congruent relationships, with few exceptions (see below). NBS and PP values are very high across the trees. QI values are high for all nodes (>0.9), and QF scores average between 0.83 and 0.88 across datasets. Figure 2 shows the result of phylogenetic inference in the concatenated min32 dataset.

Hawaiian *Melicope* are divided into five main clades corresponding to those previously resolved by Appelhans et al. (2014b). These five clades are fully supported by all statistical methods. The former genus *Platydesma* represents the earliest diverging lineage (clade V; Figure 3.2). Clade IV corresponds to Stone's section *Pelea*, characterized by whorled leaves. The remaining Stone's sections appear to be non-monophyletic. Species ascribed to Stone's section *Apocarpa* are resolved as two independent lineages (Clades II and III). Clade I comprises all species of Stone's sections *Cubicarpa* and *Megacarpa* intermingled (Figure 3.2). Relationships of clade III were resolved incongruently between datasets and analyses. BI and ML analyses resolved clade III as sister to clade IV, and the resulting monophyletic lineage again in a sister-group relationship to clades I + II with maximum PP and high ML-NBS support in four of the datasets (min10, min32, min50, min85), yet with some discord detected by Quartet Sampling. (Figure 3.2, Figure 3.3, Supplemental Figures 3.1, 3.2, 3.4). The concatenated min67 dataset resolves clade III as sister to clades I + II, and clade IV as sister to clades I + II + III (Supplemental Figure 3.3) with medium statistical support. Coalescent based SVDQuartets analysis of SNP datasets resolved a third alternative topology. Here, clade II is resolved as sister to clade III and the resulting lineage is sister to clades I + IV. This topology receives medium to low SVD-NBS support across all SNP datasets, as well as medium to high negative QC values, indicating substantial counter-support for this relationship (Supplemental Figures 3.1-3.9).

The relationship of clade III is highly discordant over quartet replicates (Supplemental Figure 3.3). Across all datasets, the discord detected by Quartet Sampling for the ancestral branch is skewed favoring one of the tested alternative quartet topologies (QD; Figure 3.3, Supplemental Figures 3.1-3.9). The remaining relationships within individual clades are fully resolved, improving resolution to the species and intraspecies level (Figure 3.2). The majority of all Hawaiian *Melicope* are resolved in clade I and relationships among species show many nodes with notable discord and very short branches (Figure 3.2). Most of the nodes show low QC and medium to low QD values (Figure 3.3).

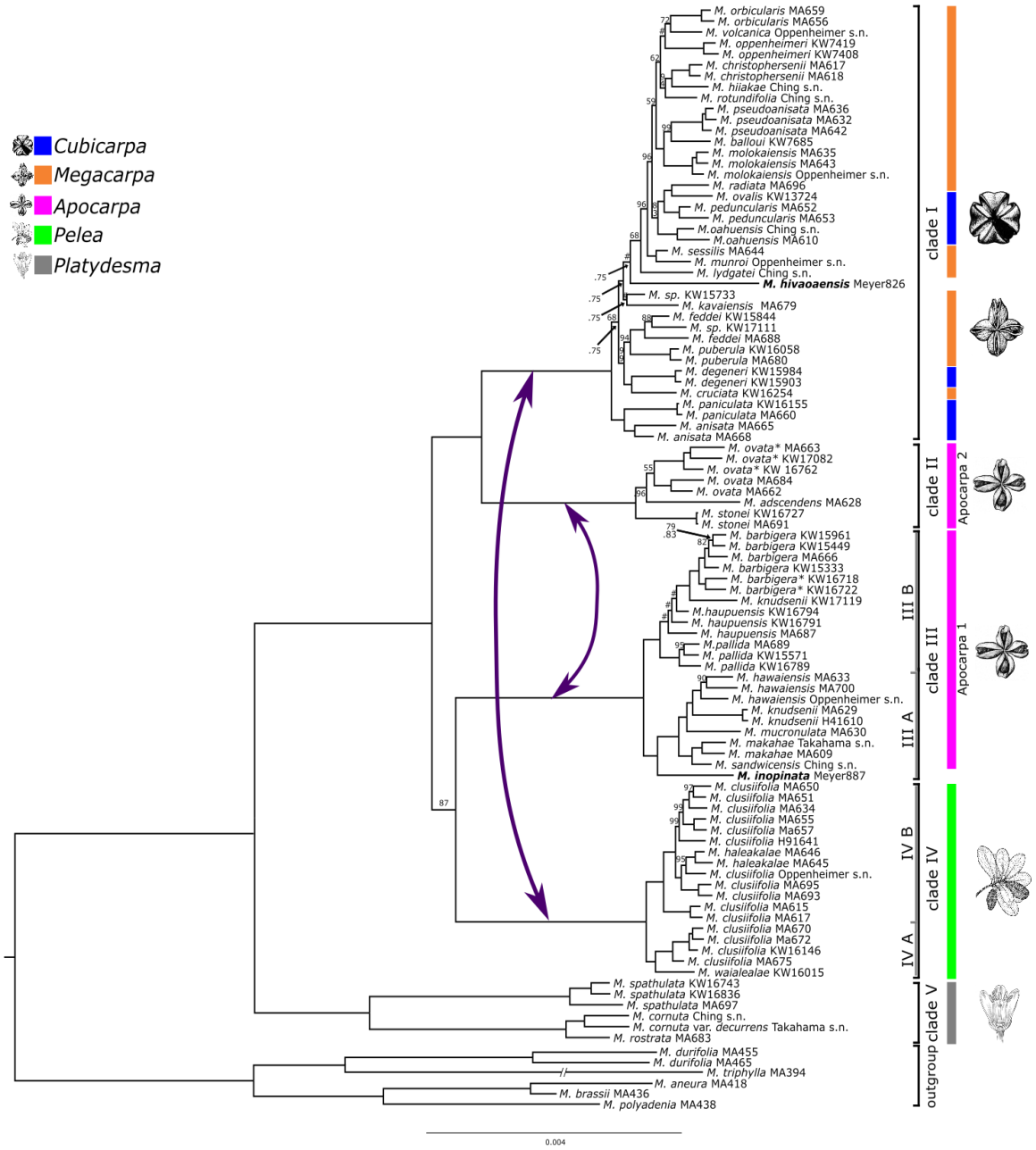



**Figure 3.2.** | Phylogeny of Hawaiian *Melicope* based on the concatenated min32 dataset. Bayesian posterior probability (pp) values are indicated above, ML nonparametric bootstrap support (ML-NBS) below branches. Support values are not shown for maximally supported clades (1.00pp/100BS). A Hashtag (#) represents incongruent species relationships between Bayesian and ML analysis. Clade colours and line drawings correspond to morphologically limited Stone's sections. Bold samples represent Marquesan species. Asterisks mark specimens differing morphologically from the typical representatives of these species. Purple arrows mark putative introgression events.

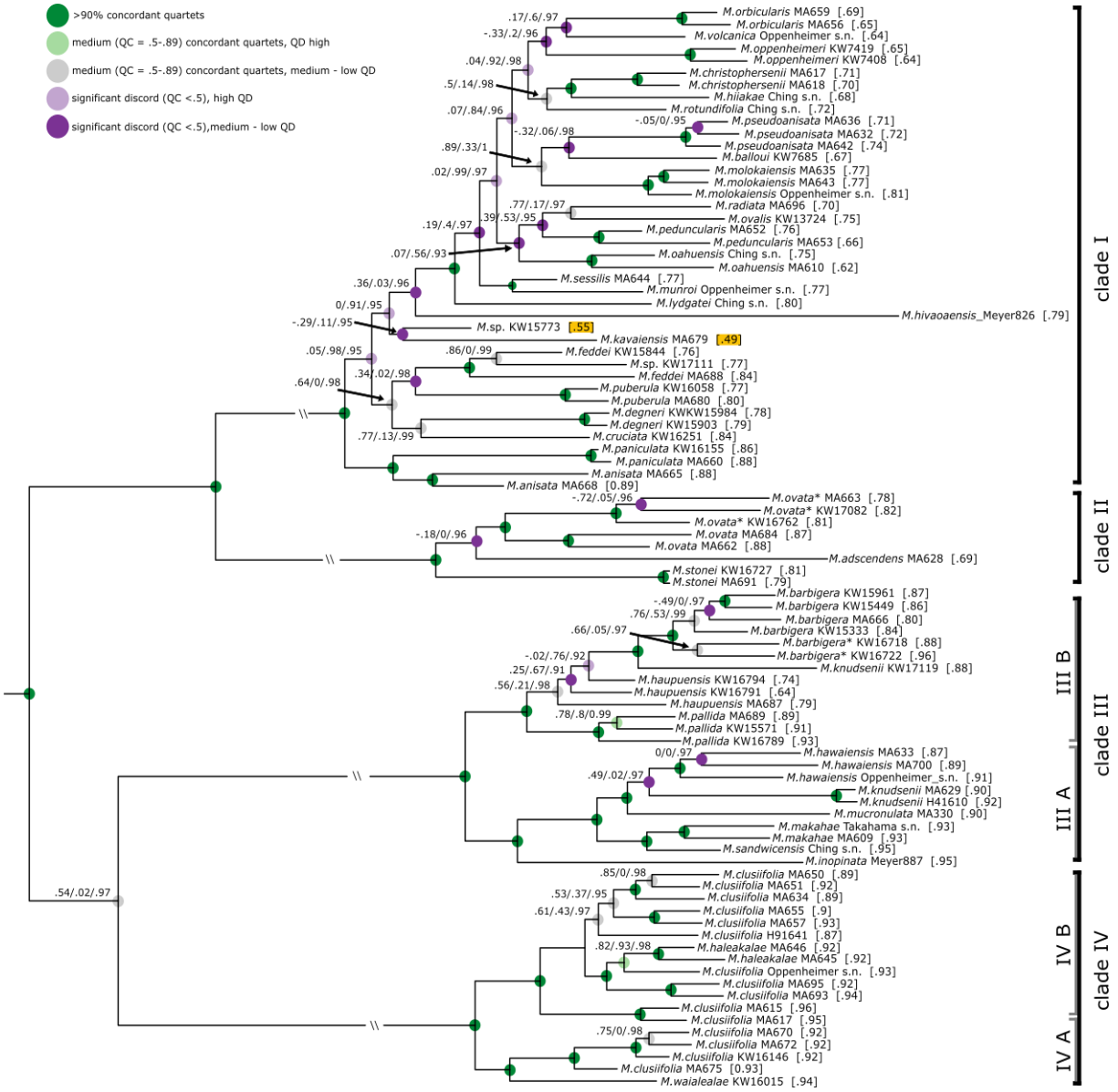



**Figure 3.3.** | Phylogeny of Hawaiian *Melicope* based on the min32 dataset. Quartet sampling results (Quartet concordance (QC)/Quartet Differential (QD)/Quartet Informativeness (QI)) are indicated on branches, Quartet Fidelity (QF) values behind samples. Nodes are colored according to QC & QD values. Results are not shown for branches with QC > 0.9. Lowest QF values are highlighted. Outgroup specimens are removed for graphical purposes. All outgroup relationships receive maximum QC values (1/-/1).

Three samples show incongruent relationships between datasets. This pertains to the Marquesan *M. hivaoaensis*, which is resolved in clade I as either sister to the remaining species (Supplemental Figures 3.3-3.9) within the clade or diverging prior to *M. lydgatei* (Figures 3.2, 3.3, Supplemental Figures 3.1, 3.2) as well as to *M. kavaiensis* and *M.* sp. KW15773 (Figure 3.3, Supplemental Figures 3.1-3.9). In all datasets, QC values show high discord or even counter-support for the placement of these three specimens. However, while QD and QF values are high for *M. hivaoaensis*,

for both *M. kavaiensis* and *M.* sp. KW15773, QD values are low and QF scores are below average (0.47-0.6 for *M. kavaiensis*) (Figure 3.3, Supplemental Figures 3.1-3.9). The remaining relationships in clade I are congruent among all concatenation based analyses. Site-specific coalescence analysis however resolved largely incongruent relationships for taxa in this clade, especially pertaining to the most recent divergences. The inferred relationships receive medium to very low SVD-NBS values and show a high amount of discord in quartet sampling (Supplemental Figures 3.5-3.9).

Clades III and IV are subdivided into two subclades each. Most species sampled with multiple accessions are resolved as monophyletic with high support and no discord detected in quartet sampling. Exceptions are *M. clusiifolia*, *M. haupuensis, M. knudsenii,* and *M. feddei. Melicope clusiifolia* is resolved paraphyletic with respect to *M. haleakalae,* which is nested within clade IVB with high to maximum support. Specimens of *M. haupuensis* are resolved as polyphyletic within clade IIIB. The relationships among the three sampled taxa are not resolved consistently across datasets and poorly supported. Quartet sampling reveals a high level of discord and below-average QF scores (Figure 3.3, Supplemental Figures 3.1-3.9). *Melicope knudsenii* is also resolved as polyphyletic with two Maui specimens (MA629, H41610) monophyletic in clade IIIA, while the third sample from Kauaʻi (KW17119) is resolved as sister to *M. barbigera* in clade IIIB (Figure 3.2). Either relationship is virtually uncontested (Figures 3.2, 3.3, Supplemental Figures 3.1-3.9). *Melicope feddei* is paraphyletic with respect to one of the Kauaʻi *M. wawraeana*-like specimens (KW17111). The three individuals form a fully supported, monophyletic unit (Figures 3.2, 3.3, Supplemental Figures 3.1-3.9).

None of the three species complexes (*M. elliptica*, *M. kavaiensis,* and *M. volcanica* complexes) are resolved as monophyletic. Species of both, the *M. kavaiensis* and *M. volcanica* complexes, are resolved in clade I (Figure 3.2) in proximity to each other, but not sister to each other. Species of the *M. elliptica* complex are resolved in different subclades of clade III (Figure 3.2). Both, *M. barbigera* and *M. ovata* were resolved as monophyletic and the morphologically divergent specimens (Table 3.1, asterisk) are resolved as sister clades to the samples with the typical morphology of the respective species with high support (Figure 3.2).

The species from the Marquesas Islands are deeply nested within the Hawaiian clade. *Melicope inopinata* is resolved in clade III as sister to the rest of subclade IIIA. *Melicope hivaoaensis* represents a group of six morphologically similar species that form a highly supported monophyletic clade (Appelhans et al., 2014b, 2018b) and is nested within clade I here (Figure 3.2).

## Test for Introgression

The min32 dataset was used for the ABBA-BABA test, since it produced the highest number of fully supported nodes. The tree topology in Figure 2 was chosen to represent the species tree topology, as it was recovered by the majority of analyses. The *D*-statistics were only used to test the incongruent position of clade III as for incongruent species within clade I, the sampling of the respective populations is not sufficient to draw reliable conclusions. Samples within clades were pooled and single nucleotide polymorphism (SNP) frequencies were used for *D*-statistic calculations (Durand et al. 2011). All possible relationships complying with the *D*-statistic assumptions were tested. A total of 24,673 loci covered at least one-third of all samples per clade and thus, contributed to the test results. Table 3 summarizes the tested topologies and inferred partitioned *D*-statistics. When clades III and IV are tested as donors for introgressed loci, values for $D_{12}$ are small and not significant ($Z_{12} < 2.55$), while values for $D_1$ and $D_2$ are significant, respectively. For tests with either of the clades I and II designated as P3 lineages, $D_{12}$, as well as $D_1$ and $D_2$, are all significant (Table 3.3). For all tested configurations, the dataset exhibits more than 3,000 discordant site patterns (Table 3.3).

**Table 3.3.** | Partitioned D-statistics for introgression involving clades I-IV. Z scores ≥ 2.55 represent a significant value for $D_x$. The respective numbers of concordant and discordant site patterns are listed. Clade numbers refer to those in Figure 2 and the group they are assigned to in the partitioned D-statistics test is indicated (compare Figure 1). (O = outgroup).

| **((P1, P2),(P3$_1$, P3$_2$), O)** | **$D_{12}$** | **$Z_{12}$** | ***n* ABAAA** | ***n* BABBA** |
|---|---|---|---|---|
| **((I, II), (III, IV), V&O)** | 0.020 | 0.95 | 809.84 | 778.1 |
| **((I, II), (IV, III), V&O)** | -0.020 | 0.96 | 778.1 | 809.84 |
| **((IV,III), (I, II), V&O)** | 0.066 | 3.28 | 1273.58 | 1115.49 |
| **((IV, III), (II, I), V&O)** | -0.066 | 3.29 | 1273.58 | 1115.5 |
| | | | | |
| **((P1, P2),(P3$_1$, P3$_2$), O)** | **$D_1$** | **$Z_1$** | ***n* ABBAA** | ***n* BABAA** |
| **((I, II), (III, IV), V&O)** | 0.276 | 8.48 | 261.01 | 437.67 |
| **((I, II), (IV, III), V&O)** | -0.276 | 8.17 | 437.67 | 261.01 |
| **((IV,III), (I, II), V&O)** | -0.242 | 7.07 | 403.07 | 222.05 |
| **((IV, III), (II, I), V&O)** | 0.290 | 7.61 | 271.16 | 444.45 |
| | | | | |
| **((P1, P2),(P3$_1$, P3$_2$), O)** | **$D_2$** | **$Z_2$** | ***n* ABABA** | ***n* BAABA** |
| **((I, II), (III, IV), V&O)** | -0.253 | 7.08 | 505.48 | 286.73 |
| **((I, II), (IV, III), V&O)** | 0.253 | 7.05 | 286.73 | 505.48 |
| **((IV,III), (I, II), V&O)** | 0.290 | 7.57 | 271.16 | 444.45 |
| **((IV, III), (II, I), V&O)** | -0.242 | 7.24 | 403.06 | 222.05 |

## 3.4 Discussion

### Phylogeny and Introgression

Analysis of *ipyrad* assemblies consistently resolved five major clades within Hawaiian *Melicope* (Figure 3.2). However, the relationships of clade III were incongruent among the five datasets and analysis methods (Figure, 3.2, Supplemental Figures 3.1-3.9). Incongruence between datasets may be caused by one of three factors: noise, ILS, or non-tree-like evolution. As noise is expected to impact small datasets and deep nodes most severely (Misof et al., 2014), it is unlikely a sufficient cause of the incongruence observed here, since our RAD-seq alignments are substantial in size (Table 3.2) and the remaining deep nodes are not affected.

The QD values of the branch illustrate that one of the discordant topologies is inferred significantly more often (0.0-0.4; Figure 3.3, Supplemental Figures 3.1-3.9), which indicates non-tree-like evolution as the cause for the discord. Thus, we used the partitioned *D*-statistics to test for signals of ancient introgression between clades I through IV with all clades tested as putative donor (P3) lineages. In all cases, values for $D_1$ and $D_2$ were each significant, yet values for $D_{12}$ were only significant when clades I and II were defined as P3 (Table 3.3). Positive values of $D_1$ represent introgression between P2 and $P3_1$, while negative values indicate introgression between P1 and $P3_1$, and values for $D_2$ represent events analogous for $P3_2$ and P2 (Eaton and Ree, 2013; Pease and Hahn, 2015). The significant values for $D_1$ and $D_2$ indicate introgression between the respective ancestors of clades I and IV as well as between respective ancestors of clades II and III. Significant values for $D_{12}$ represent shared ancestral alleles from the clade I+II progenitor introduced into the respective ancestor of clades III and IV (Figure 3.2, Table 3.3). All taxa in clades II and III have apocarpous fruits, while all taxa in clades I and IV have syncarpous fruits (Stone et al., 1999), providing a morphological connection between either of the two pairs, which might be linked to introgressed information. However, we interpret these result cautiously, as *D*-statistic results are sensitive to confounding signals from multiple introgressive events due to phylogenetic non-independence of tests (Eaton et al., 2015).

The origin of the Hawaiian *Melicope* lineage predates the rise of the current high islands (Appelhans et al., 2018b). Thus, the inferred introgressive events are associated with a time when the ancestral species were either still relegated to refugial areas on small, low islands or shortly after they colonized the young island of Kaua‘i. The time frame under consideration presents a ‘bottleneck’ scenario, where the ancestral lineages were likely in close spatial proximity. Additionally, increased volcanic activity of the Hawaiian hotspot coincided with the rise of Kaua‘i (Price and

Clague, 2002). This volcanic activity could have produced lava flows, earthquakes, tsunamis, and other catastrophic events, which may have additionally promoted hybridization (Stuessy et al., 2014). The ancestral hybridization events may even have promoted subsequent adaptive radiation on the islands (Kagawa and Takimoto, 2018). Estimation of divergence times in Hawaiian *Melicope* will be needed to infer the time frame for hybridization events in ancestral lineages. While there is strong evidence for ancient hybridization events within Hawaiian *Melicope*, the nature of *de novo* RAD-seq data currently limits our analytic methods. Further information may be obtained through gene tree based approaches applied to target capture or Whole-Genome-Sequencing data (Meng and Kubatko, 2009), or by examining SNP based patterns, as they vary spatially along a reference genome (Martin et al., 2013).

Bootstrap and Posterior Probability support values were generally high across trees inferred from different datasets but genrally increased with dataset size. Lenient filtering in RAD-seq data is often practiced, as there is a correlation between the size of a data matrix and resolution and support of relationships (Wagner et al., 2013; Hodel et al., 2017). RAD locus dropout is expected to increase with increasing divergence times, as enzyme cut sites will be lost or gained through mutation (Cariou et al., 2013). Loci with a small amount of missing data are therefore expected to represent the conserved spectrum of genomic sites and thus, provide a limited capacity of resolution. On the other hand, sparse loci are expected to increase resolution of relationships despite also introducing noise, as they are assumed to represent the more rapidly evolving genomic fractions (Cariou et al., 2013; Wagner et al., 2013; Eaton et al., 2015). However, including all loci is not advisable either, as there seems to be a point at which inclusion of increasingly more sparse loci might start to decrease support. At this point, noise, due to missing data introduced by the inclusion of more sparse loci, will overpower the informative value these loci provide. However, the Quartet Sampling method seems an adequate approach to evaluating the reliability of the dataset as the QC value showed the same trend in all datasets regardless of size and offer the QI score to assess the amount and impact of missing data.

We detected some discord between relationships resolved by concatenation and site-specific coalescence based methods (Figure 3.2, Supplemental Figures 3.1-3.9). The evaluation of the performance of different species-tree inference methods is a matter of ongoing research, especially with regards to genomic datasets. Concatenation based ML inference can be statistically inconsistent under some conditions in the MSC, i.e. ILS causing gene trees to differ from the true species tree (Kubatko and Degnan, 2007). However, the limits of the concatenated approach are poorly

understood (Molloy and Warnow, 2018) and the performance of concatenated Bayesian analysis has yet to be formally assessed. Some simulation studies show that concatenated RAD-seq data are robust to gene tree/species tree discord when inferring relationships among taxa (Rivers et al., 2016). In addition, concatenated approaches potentially offer hidden support as a feature overriding gene tree/species conflict (Gatesy and Springer, 2014; Rivers et al., 2016), although hidden support has not been addressed in plant phylogenomic research yet. Coalescence based methods are statistically consistent under the MSC. Bayesian co-estimation of gene trees and the species tree under the MSC is currently considered the most effective approach, yet computationally very demanding and thus less applicable to large datasets. Hence, summary and site-specific MSC methods have become popular and several algorithms implementing the concepts do exist (Liu et al., 2015). However, the assessment of the performance of these methods under empirical and simulated conditions is still a matter of active research. For example, gene tree methods have proven to be statistically inconsistent if the cause of gene tree discord is horizontal gene transfer, instead of ILS (Solís-Lemus et al., 2016; Fernández-Mazuecos et al., 2018). Several recent simulation studies compared the accuracy of multiple summary and site-based coalescent methods, including SVDQuartets, as well as concatenated ML under varying levels of ILS and Gene Tree Estimation Error (GTEE). Concatenated ML was at least competitive with MSC methods under most conditions and outperformed SVDQuartets under all tested conditions, including high GTEE (Chou et al., 2015; Mirarab et al., 2016; Molloy and Warnow, 2018). The latter would be expected in RAD-seq datasets and should also be present herein.

With respect to species relationships inferred for Hawaiian *Melicope* and considering the observed lower accuracy of SVDQuartets compared to concatenation based approaches under conditions typically characterizing RAD datasets, we suggest that the results from concatenated BI and ML are probably more accurate than those based on SVDQuartets and will be discussed below. However, we do stress that none of the approaches have proven to be statistically consistent under conditions observed herein, i.e. ILS, GTEE and horizontal gene transfer (Figure 3.2).

### Taxonomic Implications

The former small genus *Platydesma* and Stone's section *Pelea,* are each monophyletic (Figure 3.2), while the three remaining sections of Stone, comprising the majority of all Hawaiian *Melicope* species, are not. *Apocarpa* is divided into two lineages with the majority of species resolved in *Apocarpa* 1 (Figure 3.2). The three species of the *Apocarpa* 2 clade share a number of morphological traits, though neither of them is either exclusive or inclusive. All species of *Apocarpa* 2 occur in mesic forests only and,

with the exception of *M. stonei,* share a sprawling, shrubby habit (Stone et al., 1999; Wood et al., 2017). Finally, in all *Apocarpa* 2 species both, endo- and exocarp are glabrous and inflorescences are few-flowered, though both of these traits also appear outside of this group (Stone et al., 1999; Wood et al., 2017). In a previous analysis apocarpous species were resolved in three different clades (Appelhans et al., 2014b), one of which, consisting of *M. elliptica* only, could not be sampled in this study. Further research will be necessary to identify morphological character combinations distinguishing these lineages. Stone's sections *Cubicarpa* and *Megacarpa* are paraphyletic with respect to each other (Figure 3.2) with species of each resolved intermingled throughout the clade. The two groups differ by the degree of carpel connation, with carpels "connate from base up to 2/3 of their length" (Stone et al., 1999) characterizing *Megacarpa* and carpels "nearly to completely" connate (Stone et al., 1999) characterizing *Cubicarpa*. Carpel connation clearly represents a continuum and not two discrete units. As there is no pattern to the degree of carpel fusion apparent in clade I, the separation of these two of Stone's sections seems artificial.

Inter-species relationships within clade I are less well supported than in the remaining clades and Quartet Sampling reveals measurable discord at nearly every branch in the backbone of this clade. For many of the nodes with low QC values, QD values are high (Figure 3.3, Supplemental Figures 3.1-3.9), which characterizes ILS and corresponds to the shortness of these branches. On the other hand, many branches show low QD values indicating widespread introgression between these lineages. Unfortunately, sampling herein is not sufficient to test individual relationships.

Of the 24 species represented by multiple accessions, 20 were resolved as monophyletic, while four species were either para- or polyphyletic. *Melicope clusiifolia* is the most widespread and morphologically diverse of all Hawaiian *Melicope* (Stone et al., 1999) and it is paraphyletic with both of the other species of Stone's section *Pelea, M. haleakalae* and *M. waialealae* (clade IV, Figure 3.2). Several attempts have been made to subdivide *M. clusiifolia* into varying constellations of subspecies, varieties, and forms (St. John, 1944; Stone, 1969). In the most recent taxonomic treatment, Stone et al. (1999) synonymized all subdivisions of the species, arguing that the variable characters seem to represent a continuum rather than distinguishable, discrete units. However, the authors also issued the recommendation that the overall pattern of variability in *M. clusiifolia* should be studied in detail (Stone et al., 1999). *Melicope haleakalae* is characterized as differing from *M. clusiifolia*, mainly in its persistent sepals (Stone et al., 1999). Considering that *M. haleakalae* is nested deeply within *M. clusiifolia* (clade IV, Figure 3.2). The two

might be regarded as conspecific and included in an overall evaluation of the complex. *Melicope waialealae* differs from *M. clusiifolia* mainly in leaf shape (Stone et al., 1999). However, since the leaf shape of *M. clusiifolia* is highly variable, *M. waialealae* might represent one end of a continuum across both taxa rather than one of two distinct states. On the other hand, these three species might represent a case of speciation in progress. In this case, the deep nesting, especially of *M. haleakalae*, within *M. clusiifolia*, would represent speciation following a progenitor-derivative scenario (Crawford, 2010). The widespread, morphologically variable *M. clusiifolia* would meet all criteria of the progenitor (p) species. The persistent petals in *M. haleakalae* and the leaf shape in *M. waialealae* would represent a variable, morphological feature in the parent being fixed in the respective derivative (d) species. Identification of a true p-d relationship is difficult and rare However, several candidate species pairs do exist (Crawford, 2010). The p-d species pair *Layia glandulosa* (Hook.) Hook. & Arn. and *L. discoidea* D.D.Keck (Asteraceae) show not only a shift in morphology between progenitor and derivative species, but also geographic isolation due to a shift in habitat (Baldwin, 2005). This could be the same for *M. waialealae*, which is restricted to bogs, whereas the putative progenitor *M. clusiifolia* occurs in mesic to wet forests (Stone et al., 1999). Unfortunately, there is no data available regarding breeding system or pollinator communities in these species creating potential barriers to gene flow. Detailed studies of morphological characters, gene flow, and abiotic habitat factors are necessary to determine whether these taxa are separate p-d species pairs or conspecific, as already indicated in previous studies (Appelhans et al., 2014b).

*Melicope knudsenii,* delimited by Stone et al. (1999) as the only species occurring on non-adjacent islands, was resolved as polyphyletic with three samples resolved as two distinct lineages within clade III. Appelhans et al. (2014b) already showed that this taxon is polyphyletic, consisting of three taxa. One of these was recently described as *M. stonei* (Wood et al., 2017). Our results confirm the previously resolved pattern with the two specimens of *M. knudsenii* from Maui resolved as sister to *M. hawaiiensis* and the specimen from Kauaʻi as sister to *M. barbigera* (clade III, Figure 2). We confirm that these specimens clearly represent different species. The Maui species will be resurrected under one of the names used in an earlier treatment by Stone (1969), wherein he adopted a narrower species concept than in the later classification (Stone et al., 1999), leaving *M. knudsenii* restricted to only populations on Kauaʻi.

The three specimens of *M. haupuensis* included in this study are resolved as paraphyletic. Moreover, they are the only species resolved with incongruent

topologies of the individual samples associated with the different datasets (compare Figure 3.2, Supplemental Figures 3.1-3.9). Quartet Sampling shows strong discord for either of the inferred relationships with medium QD values (Figure 3.3, Supplemental Figures 3.1-3.9), indicating the possibility of introgressed sites. Moreover, QF scores for the three specimens are considerably lower than the average, indicating a rogue behavior (Aberer et al., 2013; Pease et al., 2018) of the three taxa. Additionally, several specimens in the field were observed presenting morphologically intermediate forms between *M. haupuensis* and *M. barbigera* (personal observation K.R. Wood). QD values for the latter are also low (Figure 3.3, Supplemental Figures 3.1-3.9). Both, the morphological intermediates as well as the incongruence associated with different datasets, indicate potential hybridization between these species. However, conclusively identifying putative hybridization events would require sampling at the population level, including any morphological intermediates.

Multiple samples of *M. ovata* and *M. barbigera,* were included in our analyses, representing both, the typical morphology and a deviating morphotype. For either species, the morphologically deviating samples were resolved as the sister group to the samples with the typical habit. Variant morphotypes of *M. ovata* displayed a pubescent lower leaf surface, whereas leaves are typically glabrous in this species. *Melicope barbigera* usually has few-flowered inflorescences (Stone et al., 1999). In contrast, the variant morphotype has inflorescences with a considerably larger number of flowers. Genomic divergence is comparable to other species pairs within the lineage. Both groups might be another case of speciation in progress within Hawaiian *Melicope.* In both cases, detailed morphological studies will be necessary to investigate if the morphologically divergent populations of the two species should be recognized as separate taxa.

The two *M. wawraeana*-like specimens are resolved in clade I, but not closely related to each other. One specimen (KW17111) is nested within the two samples of *M. feddei* with high support (Figures 3.2, 3.3). *M. wawraeana* is very similar to *M. feddei* and differs mainly in pedicel length (Stone et al., 1999). The present results suggest that some populations might be conspecific with *M. feddei*, while others (e.g. from the herein unsampled type location) are not. The relationships of the second *M. wawraeana*-like specimen (KW15733) are resolved incongruently among datasets, as are the relationships of the sampled specimen of *M. kavaiensis.* The two samples are resolved either as sister groups (Figure 3.3, Supplemental Figures 3.1, 3.4, 3.5, 3.7-3.9) or as consecutive sister clades within clade I (Supplemental Figures 3.2, 3.3, 3.6). There is a substantial amount of discord in the dataset for either of the resolved

relationships. QD values are low, indicating the possibility of introgression between these morphologically distinct species. Additionally, QF scores for either of the specimens are low corresponding to the rogue behavior of the samples.

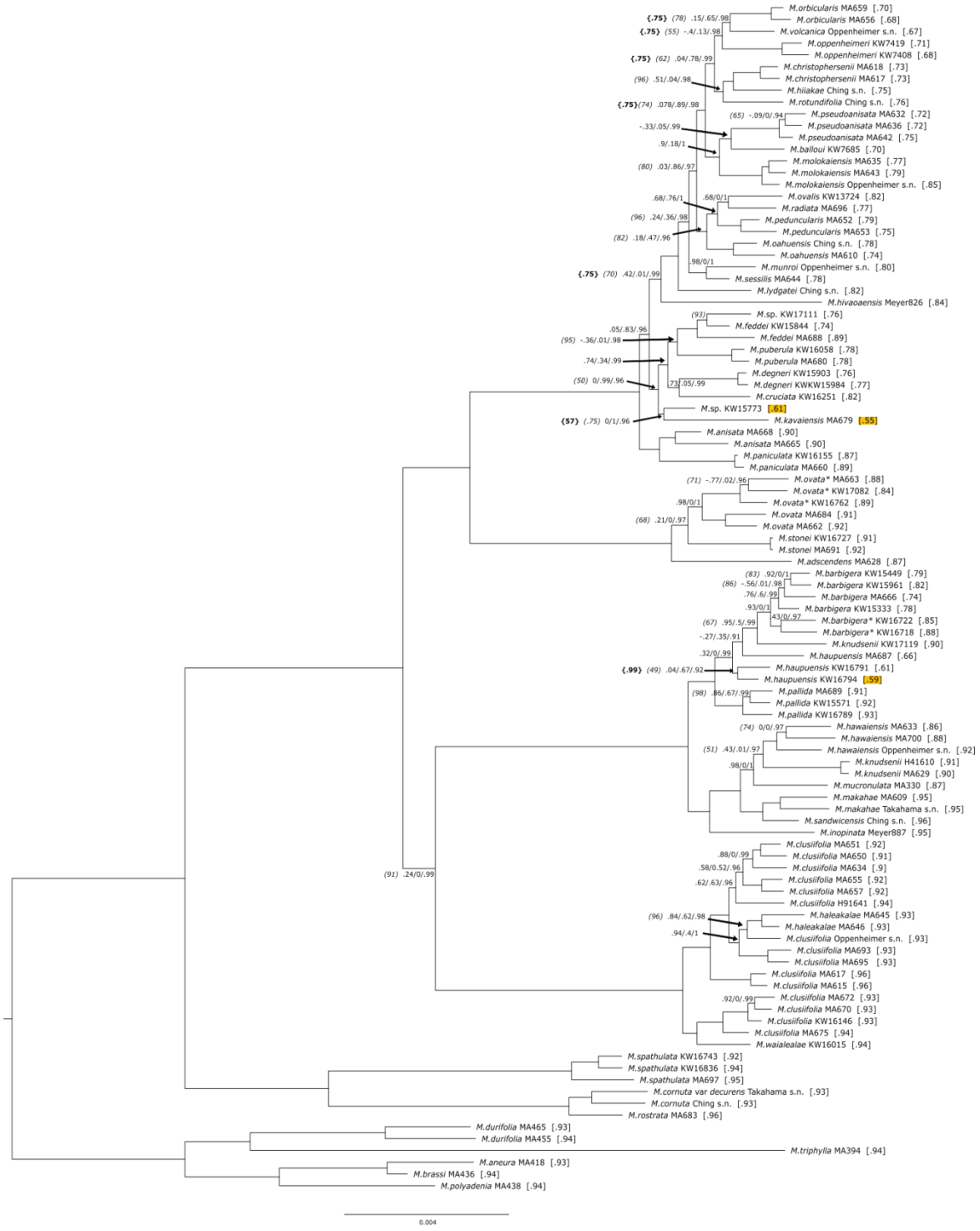
The rogue behavior of the aforementioned samples (*M. kavaiensis*, *M.* sp KW15733) might also be related to the incongruent placement of *M. hivaoaensis*, as the three taxa are inferred as closely related, regardless of the relation to the remainder of clade I. For this specimen QC and QD values are low, however, QF is high (Figure 3.3, Supplemental Figures 3.1-3.9). *Melicope hivaoaensis* represents an adaptive radiation of five species endemic to the Marquesas Islands, whose predecessor colonized from the Hawaiian Islands (Appelhans et al., 2014b, 2018b). Successful island colonizations have been associated with recent hybridization or polyploidization events (Paetzold et al., 2018). There was no polyploidization event immediately prior to the colonization of the Hawaiian Islands itself (Paetzold et al., 2018), making a polyploidization event prior to the colonization of the Marquesas Islands unlikely. Chromosome counts for Marquesan species are not available for a conclusive answer. However, results herein indicate the presence of several hybridization events within the lineage. Thus, a hybridization event might have predated the colonization of the Marquesas Islands as well. As the incongruent position of *M. hivaoaensis* seems to correspond to the rogue behavior of *M.* sp. KW15733 and *M. kavaiensis*, the latter two might represent the parental lineages of the Marquesan *Melicope* radiation. A conclusive answer to the question is contingent on a thorough sampling of all concerned lineages as well as a prior revision of the *M. wawraeana* species concept.

We confirm previous results showing that Hawaiian *Melicope* colonized the Marquesas Islands twice independently, negating the hypothesis that the remote Hawaiian Islands constitute a dispersal sink (Harbaugh et al., 2009; Appelhans et al., 2014b, 2018b). The nesting of Marquesan species in different Hawaiian clades is corroborated by fruit morphology (Hartley, 2001), since *M. hivaoaensis* and its close relatives from the Marquesas Islands have syncarpous fruits as do the species in clade I, while *M. inopinata* has apocarpous fruits like the species in clade III.

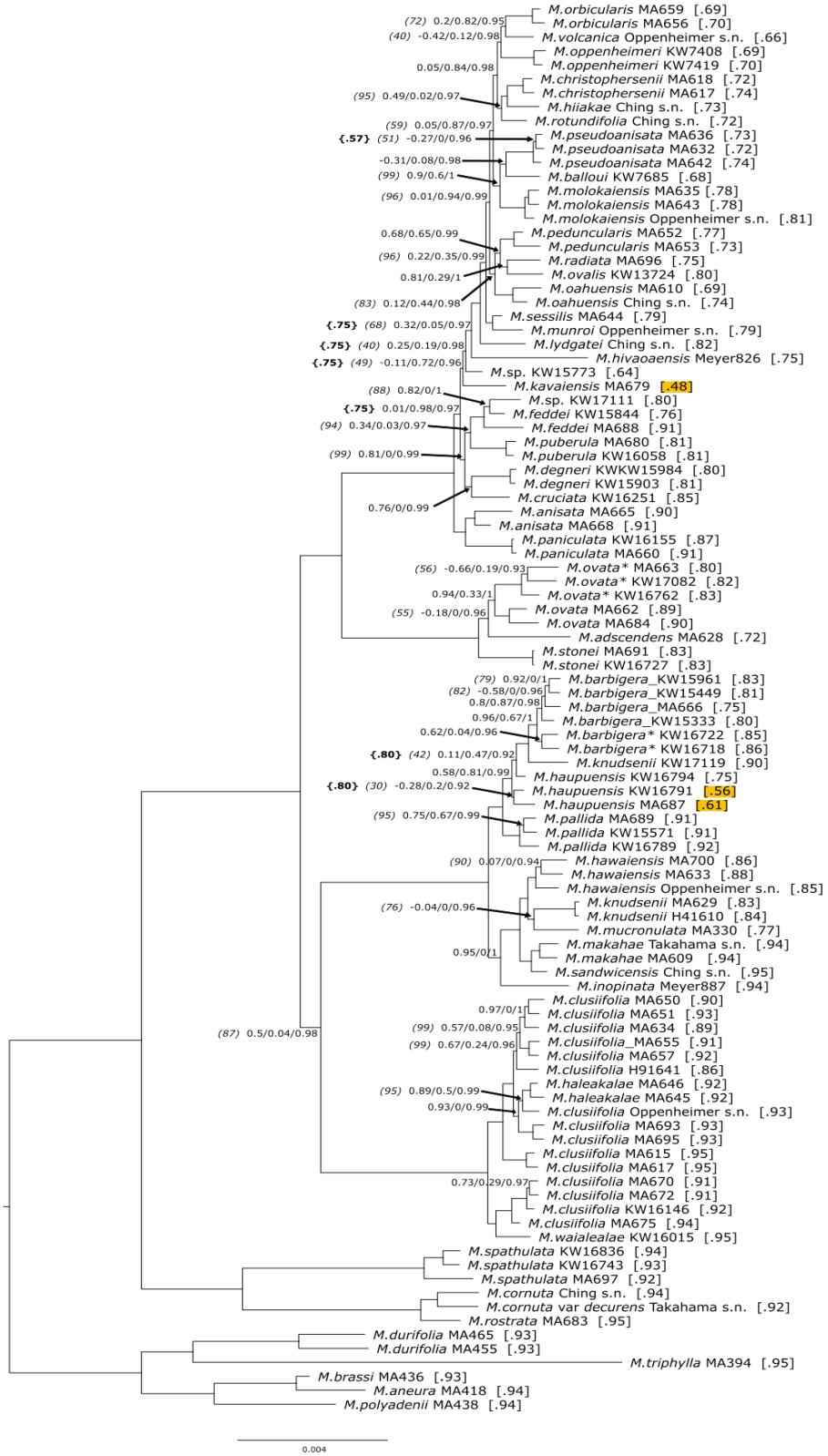
The present study provides unprecedented insight into the relationships of Hawaiian *Melicope.* Several previous findings could be corroborated and firmly supported by genome-wide data, including the non-monophyly of most of Stone's sections, which cannot be held up as delimited (Stone, 1969; Stone et al., 1999). The lineage is in need of a taxonomic revision. Understanding the relationships of Hawaiian *Melicope* would be enhanced by some formal recognition of the subclades with corresponding morphological features. However, the creation of novel formal subgroups within

*Melicope* section *Pelea* must also include the extra-Hawaiian members of the section. The former genus *Platydesma* is the most distinctive group within *Melicope* sect. *Pelea* and should receive some level of formal recognition. *Apocarpa* species need to be split into two groups, one of which would include the Marquesan species *M. inopinata*. However, conclusive treatment of *Apocarpa* should be adjourned until an improved understanding of the separation within the *M. elliptica* complex is attained. Delimitations of species within the *Pelea* group, *M. barbigera, M. ovata* as well as *M. haupuensis* may need revision, but levels of hybridization should also be investigated as part of that process. *Melicope wawraeana* requires revision as well as a prerequisite to test the putative hybrid character of the Marquesan radiation. Furthermore, the other six *Melicope* species endemic to the Marquesas Islands would need to be included in a novel taxonomic recognition of Stone's former sections *Megacarpa* and *Cubicarpa*.

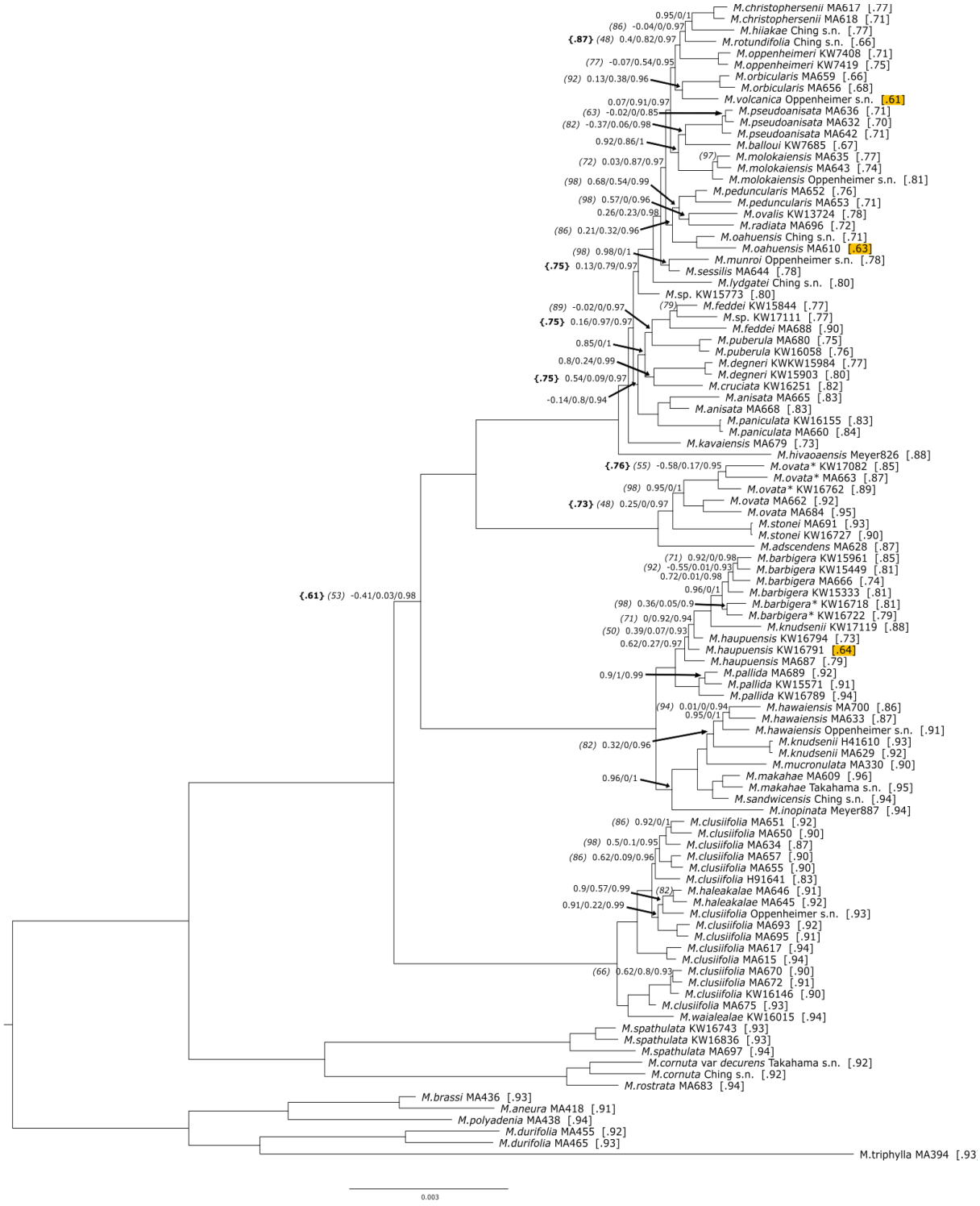
## 3. | Supplemental Information



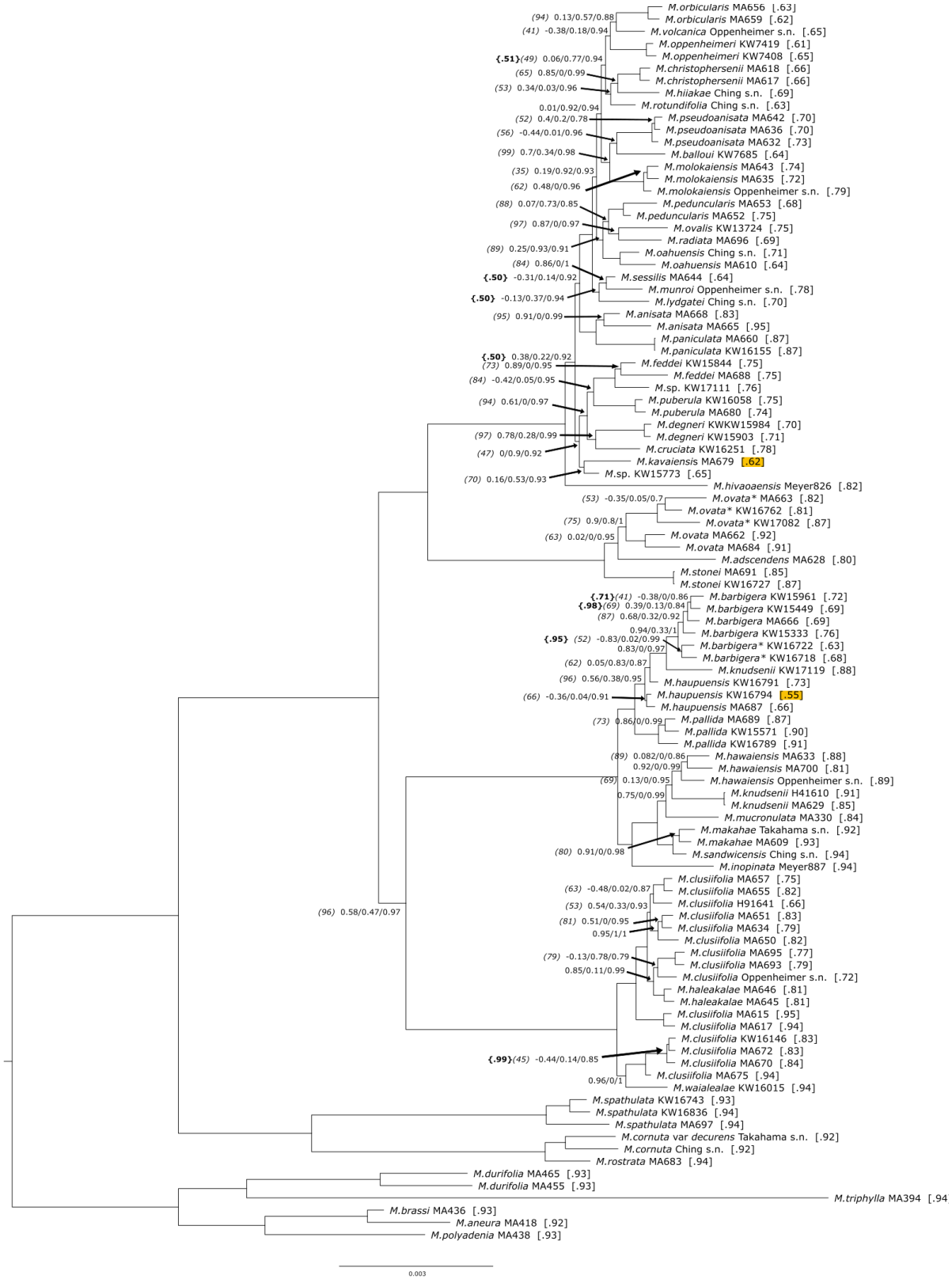

**Supplemental Figure 3.1** | Phylogeny of Hawaiian *Melicope* based on the concatenated min10 dataset. Statistical support values are indicated on branches; Bayesian posterior probability (pp) values in curly brackets and bold, Maximum Likelihood bootstrap support (ML-NBS) values in round brackets and in italics. Quartet Fidelity (QF) values are indicated in square brackets behind samples. For graphical purposes, values for maximally supported clades (100 ML-NBS, 1.00pp, 1/-/1) are not shown.



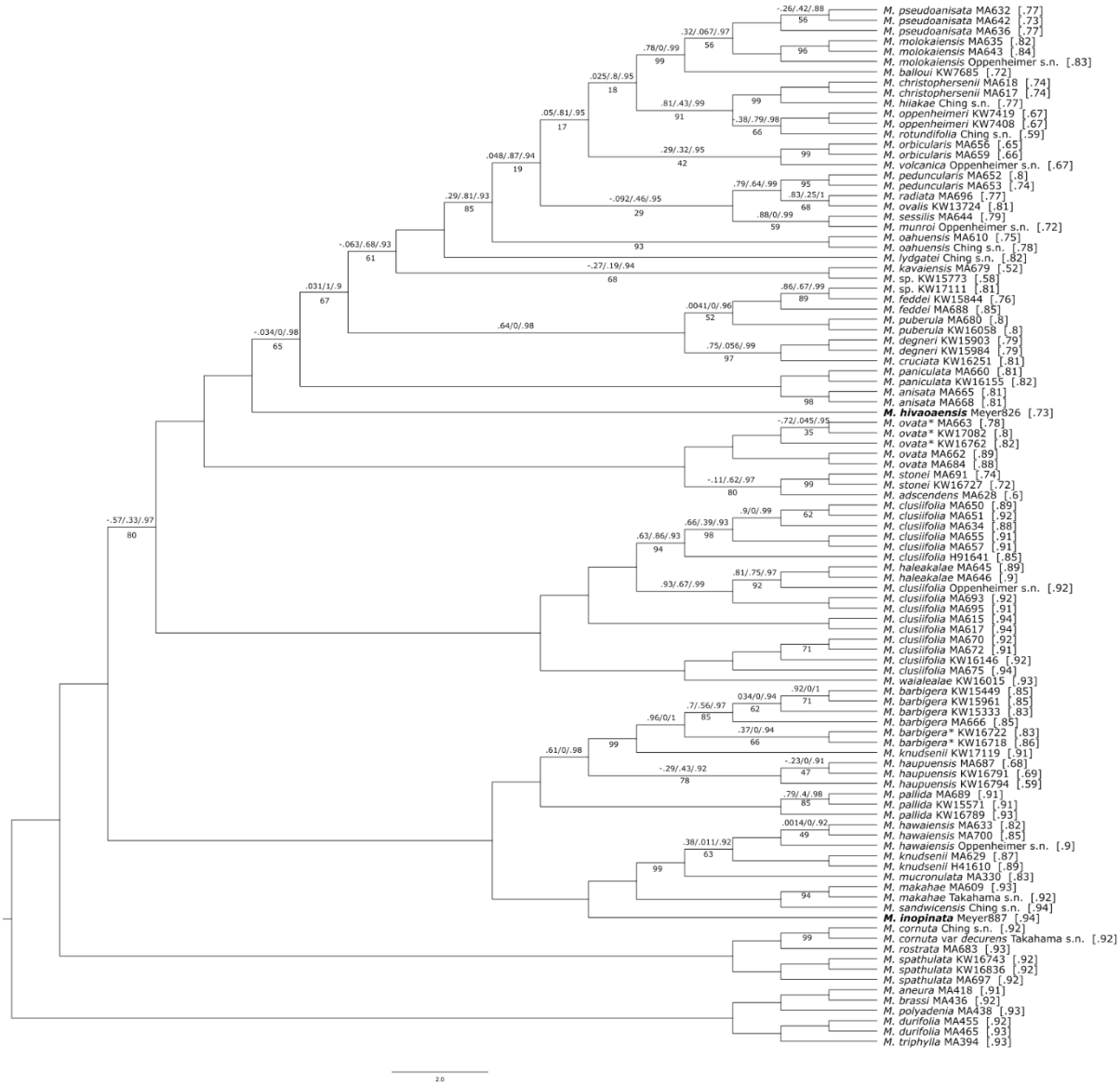

**Supplemental Figure 3.2** | Phylogeny of Hawaiian *Melicope* based on the concatenated min50 dataset. Statistical support values are indicated on branches; Bayesian posterior probability (pp) values in curly brackets and bold, Maximum Likelihood bootstrap support (ML-NBS) values in round brackets and in italics. Quartet Fidelity (QF) values are indicated in square brackets behind samples. For graphical purposes, values for maximally supported clades (100 ML-NBS, 1.00pp, 1/-/1) are not shown.



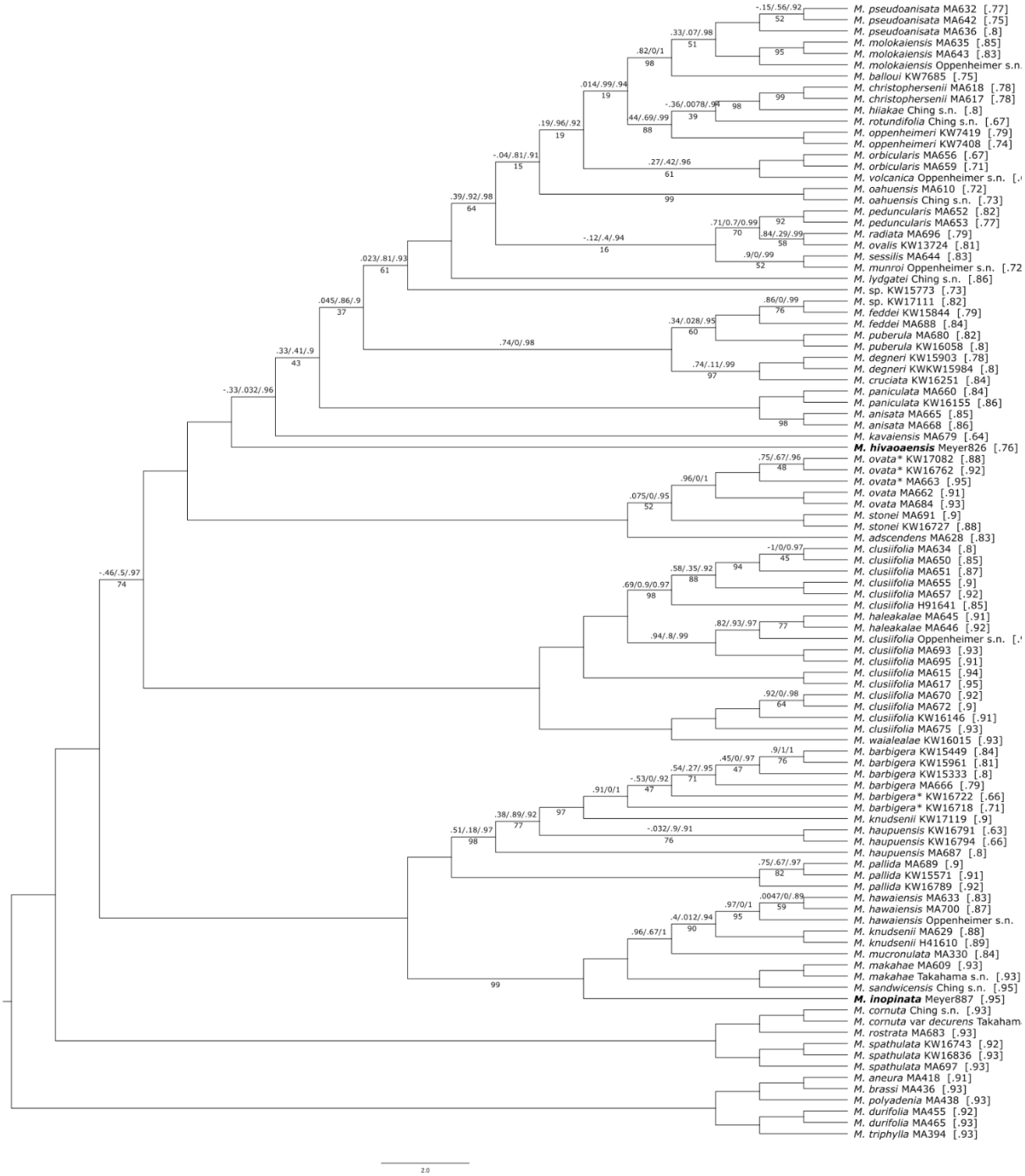

**Supplemental Figure 3.3** | Phylogeny of Hawaiian *Melicope* based on the concatenated min67 dataset. Statistical support values are indicated on branches; Bayesian posterior probability (pp) values in curly brackets and bold, Maximum Likelihood bootstrap support (ML-NBS) values in round brackets and in italics. Quartet Fidelity (QF) values are indicated in square brackets behind samples. For graphical purposes, values for maximally supported clades (100 ML-NBS, 1.00pp, 1/-/1) are not shown.



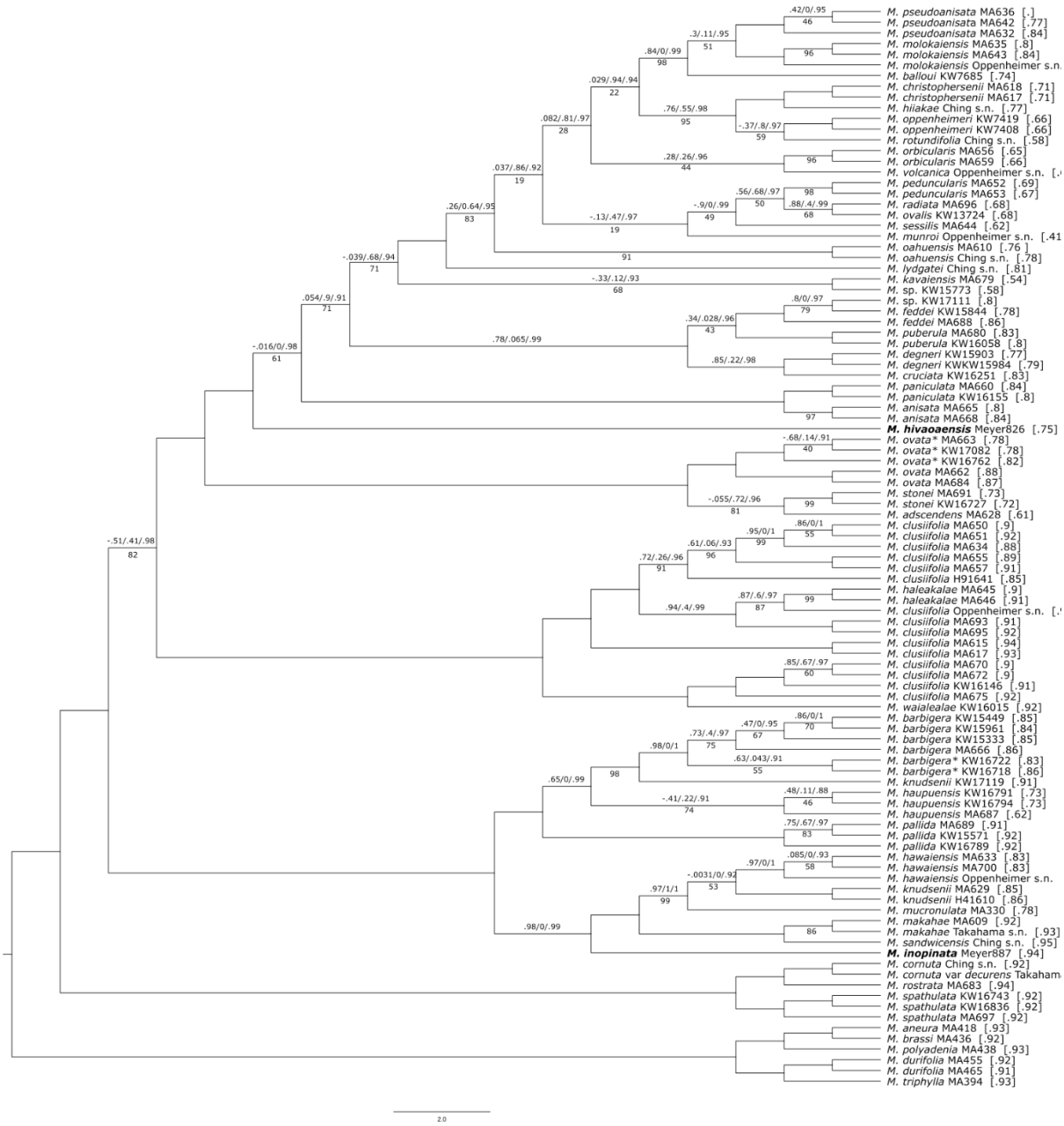

**Supplemental Figure 3.4** | Phylogeny of Hawaiian *Melicope* based on the concatenated min85 dataset. Statistical support values are indicated on branches, Bayesian posterior probability (pp) values in curly brackets and bold, Maximum Likelihood L bootstrap support (ML-NBS) in round brackets and in italics. Quartet Fidelity (QF) values are indicated in square brackets behind samples. For graphical purposes, values for maximally supported clades (100 ML-NBS, 1.00pp, 1/-/1) are not shown.



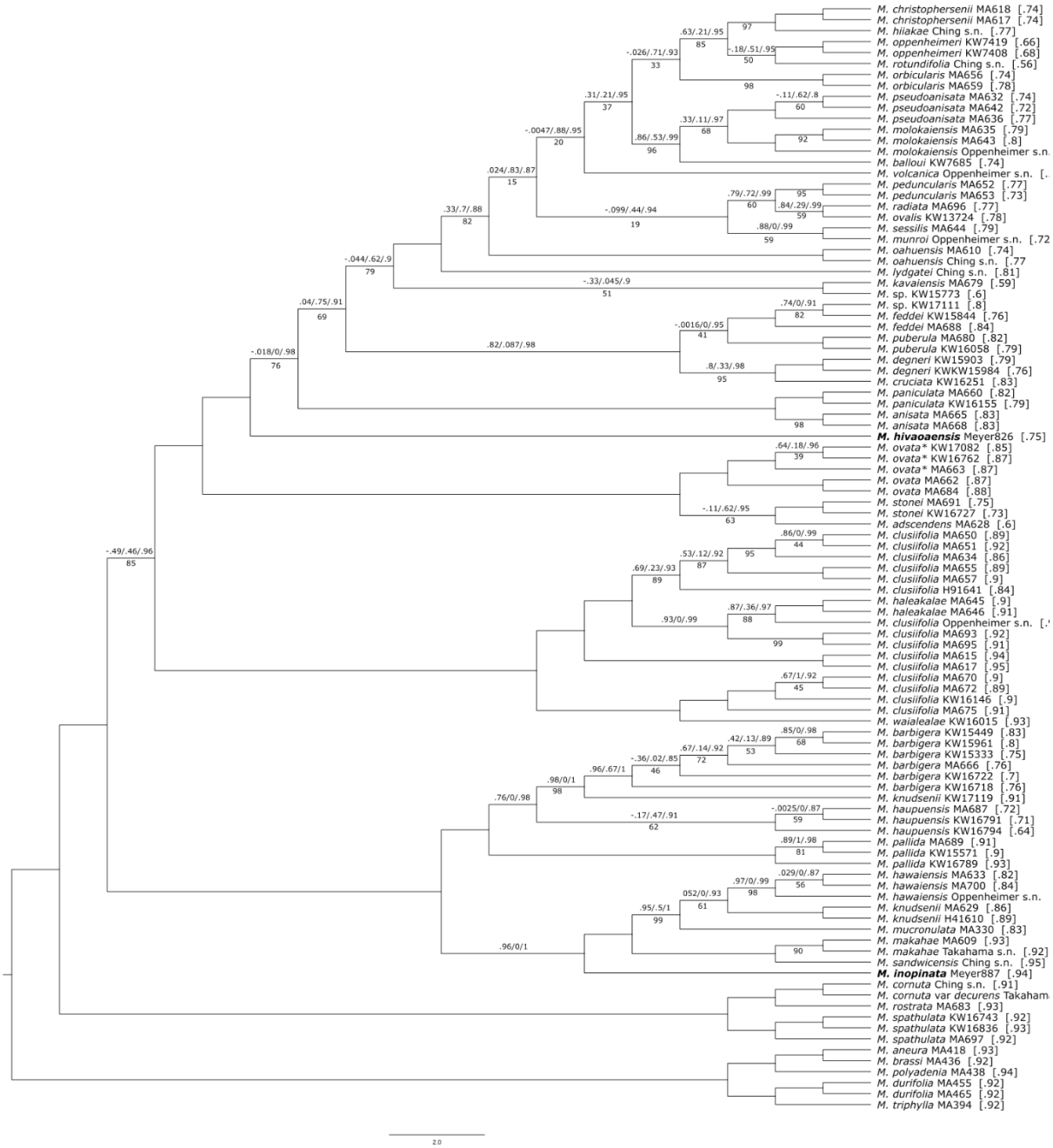

**Supplemental Figure 3.5** | Phylogeny of Hawaiian *Melicope* based SVDQuartet analysis of min10 SNP matrix. Statistical support values are indicated on branches; bootstrap support (SVD-NBS) values below and Quartet Sampling scores above branches. Quartet Fidelity (QF) values are indicated in square brackets behind samples. For graphical purposes, values for maximally supported clades (100 SVD-NBS, 1/-/1) are not shown.



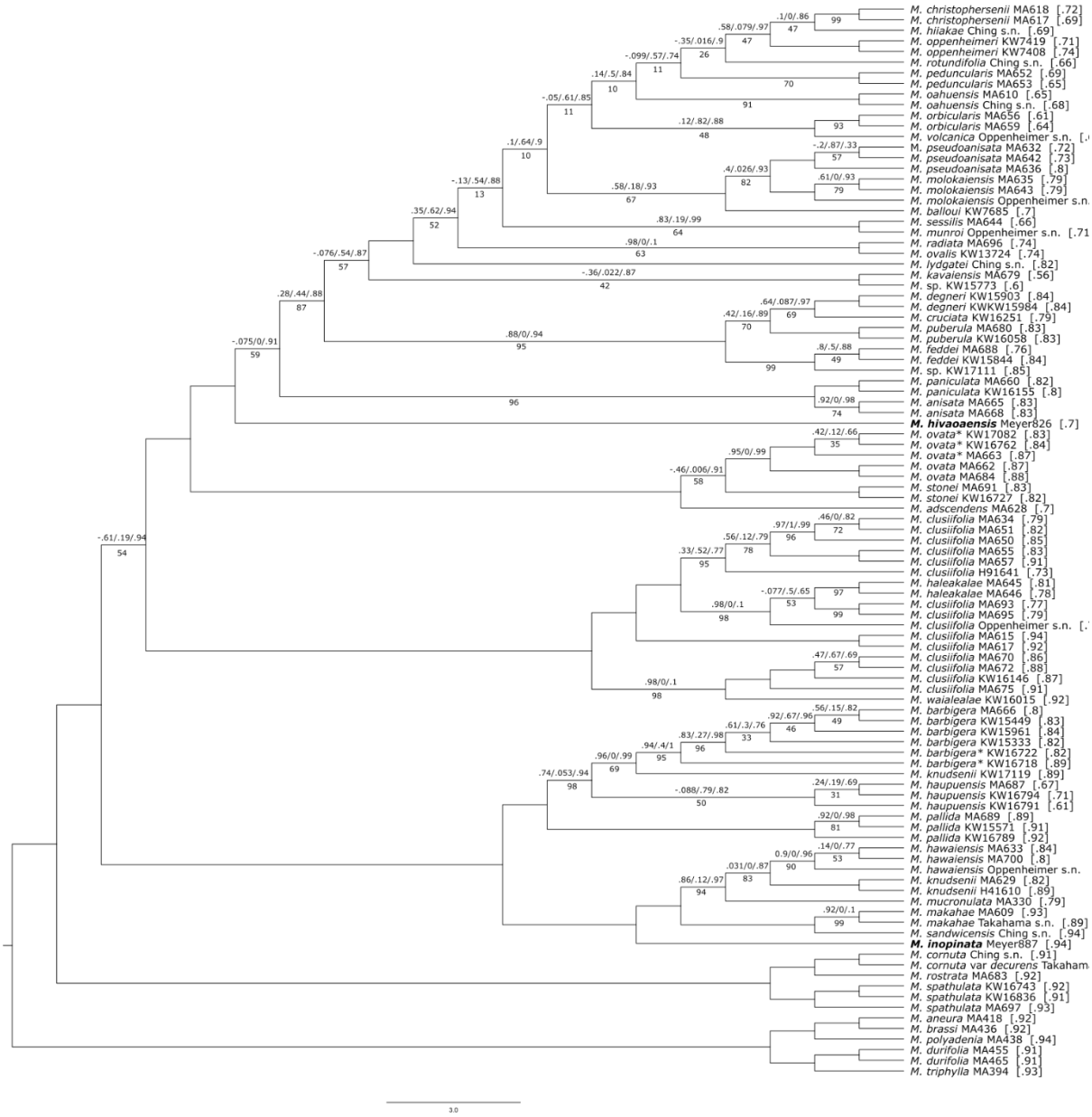

**Supplemental Figure 3.6** | Phylogeny of Hawaiian *Melicope* based SVDQuartet analysis of min32 SNP matrix. Statistical support values are indicated on branches; bootstrap support (SVD-NBS) values below and Quartet Sampling scores above branches. Quartet Fidelity (QF) values are indicated in square brackets behind samples. For graphical purposes, values for maximally supported clades (100 SVD-NBS, 1/-/1) are not shown.




**Supplemental Figure 3.7** | Phylogeny of Hawaiian *Melicope* based SVDQuartet analysis of min50 SNP matrix. Statistical support values are indicated on branches; bootstrap support (SVD-NBS) values below and Quartet Sampling scores above branches. Quartet Fidelity (QF) values are indicated in square brackets behind samples. For graphical purposes, values for maximally supported clades (100 SVD-NBS, 1/-/1) are not shown.




**Supplemental Figure 3.8** | Phylogeny of Hawaiian *Melicope* based SVDQuartet analysis of min67 SNP matrix. Statistical support values are indicated on branches; bootstrap support (SVD-NBS) values below and Quartet Sampling scores above branches. Quartet Fidelity (QF) values are indicated in square brackets behind samples. For graphical purposes, values for maximally supported clades (100 SVD-NBS, 1/-/1) are not shown.




**Supplemental Figure 3.9** | Phylogeny of Hawaiian *Melicope* based SVDQuartet analysis of min85 SNP matrix. Statistical support values are indicated on branches; bootstrap support (SVD-NBS) values below and Quartet Sampling scores above branches. Quartet Fidelity (QF) values are indicated in square brackets behind samples. For graphical purposes, values for maximally supported clades (100 SVD-NBS, 1/-/1) are not shown.

# 4. | Historical biogeography and diversification of Hawaiian *Melicope* (Rutaceae): flexibility is key.

Claudia Paetzold, Kenneth R. Wood, Warren L. Wagner & Marc S. Appelhans

Submitted to *New Phytologist*

Na Pali Coast, Kaua'i. Photograph: Claudia Paetzold

## Abstract

- Island radiations represent ideal case studies for the investigation of diversification of lineages through time, yet many inferences suffer from poorly resolved species relationships. Application of high-throughput sequencing datasets to investigations of island lineage evolutionary histories and diversification is still scarce.
- We investigate the spatio-temporal diversification patterns of the monophyletic *Melicope* clade endemic to the Hawaiian Islands using RAD-seq datasets sampling >80% of the currently accepted species of Hawaiian *Melicope* including two species from the Marquesas Islands. We compute divergence times and infer historical biogeography. We analyse relative diversification rates in the lineage and state-specific speciation and extinction rates linked to specific habitats.
- Hawaiian *Melicope* colonized the archipelago prior to the rise of the current main islands. Inter-Island colonization patterns largely follow the progression rule from older to younger islands, but back colonizations to older islands have occurred. Diversification is characterized by lag and burst times. Dry and bog habitats show high speciation and extinction rates.
- Extant diversification results from recent divergence of a small number of taxa prevailing through bottlenecks and high amounts of species turnover. Increased rates of diversification are linked to habitat dissection, and frequent ecological trait shifts.

## 4.1 Introduction

Adaptive radiation is the divergence of an ancestral species into multiple descendants via natural selection in response to "ecological opportunity". The latter encompasses three distinct spatio-temporal factors: physical, ecological and evolutionary opportunity (Glor, 2010). Consequently, the timing of evolutionary events, e.g. colonizations or trait shifts, is important.

Oceanic Island systems have played a key role in the development of the theory of evolution from its inception (Darwin, 1859) and have been a focal point for evolutionary biologists ever since. De-novo island systems are of particular interest as extant biodiversity largely originated from a relatively small number of founder species (Price and Wagner, 2018) followed by subsequent diversification (Whittaker et al., 2008; Warren et al., 2015). Traditional icons of adaptive radiation include many island lineages, e.g. Hawaiian lobeliads (Givnish et al., 2009) or Darwin's finches (Galapagos) (Grant and Grant, 2002). However, identification of the patterns of diversification is far from conclusive.

The fundamental challenge for the integrative research of adaptive radiation is a fully resolved and dated phylogenetic tree. For many adaptive radiations resolution of phylogenetic relationships remains problematic for at least some nodes, e.g. in Hawaiian lobeliads (Givnish et al., 2009) or Hawaiian silverswords (Landis et al, 2018). Lack of resolution often concerns nodes related to "bursts of speciation", where lineages have diverged rapidly and traditional genomic marker regions did not accumulate a sufficient amount of diagnostic variation. The application of high-throughput sequencing (HTS) approaches has improved phylogenetic resolution drastically, e.g. in African cichlid fishes (Wagner et al., 2013), Hawaiian mints (Welch et al., 2016) or Hawaiian *Melicope* (Paetzold et al., 2019).

Another challenge is the estimation of divergence times. While methods for dating phylogenies have been employed in island radiations for more than 20 years (Baldwin and Sanderson, 1998), algorithms for diversification rate estimation currently scale badly to HTS datasets and improvements are under active development. Nevertheless, dating phylogenies will remain a non-trivial matter, as the fossil record is very sparse for many lineages (Allison and Bottjer, 2011) and methods for incorporation of other lines of evidence, e.g. palaeogeographical features, (Landis, 2017) may introduce bias when lineages are older than the islands they inhabit (Heads, 2011).

The Hawaiian Islands are the most remote archipelago in the world and formed by the gradual northwest movement of the Pacific plate over a stationary mantel plume (Price and Clague, 2002). Each island progresses through a distinctive cycle of growth, subsidence, erosion and eventual submergence (Price and Clague, 2002)

resulting in a continuum of available ecological niches, which in turn determine species richness. There were no high islands (> 1000m) in the period between 8-5 million years ago (mya), hence most of the Hawaiian endemic biodiversity originates from adaptive radiation following colonizations within the last 5 million years (MA; Price and Clague, 2002). The current high islands age from the oldest islands Ni'ihau (6.0-4.3 mya) and Kaua'i (5.8-4.0 mya) to O'ahu (3.9-1.8 mya) followed by the Maui Nui complex, which was one single island during most of its existence and only became separated in the Pleistocene (Maui (2.0-0.9 mya), Moloka'i (2.1-1.5 mya), Lāna'i (1.32 mya) and Kaho'olawe (1.35 mya)) and finally Hawai'i (1.1 mya and ongoing) (Clague and Sherrod, 2014). Considering these decreasing island ages and the number of species per island, Wagner & Funk (1995) proposed the "progression rule" to explain diversification in the archipelago, stating that taxa colonize older islands first, and then disperse to younger islands as they emerge. As intra-island diversification occurs subsequent to colonization, a correlation between island age and species number or diversification patterns might be expected (Wagner & Funk, 1995). For many Hawaiian lineages the progression rule is a tested and true concept, including *Cyanea* (Givnish et al., 2009), *Cyrtandra* (Johnson et al., 2019; Kleinkopf et al., 2019), or damselflies (Jordan et al., 2005).

The genus *Melicope* (Rutaceae) comprises ca. 235 species (Hartley, 2001) distributed throughout Southeast Asia, Malesia, Australasia, the Pacific region as well as Madagascar and the Mascarene Islands. Hawaiian *Melicope* represents the fourth largest Hawaiian plant radiation (Wagner et al., 1999b) comprising currently 54 Hawaiian species (Stone et al., 1999; Appelhans et al., 2017; Wood et al., 2017) and seven species endemic to the Marquesas Islands, originating from two independent colonization events from the Hawaiian Islands (Appelhans et al., 2014b; Paetzold et al., 2019). The species are distributed across the islands in dry, mesic and wet habitats, including bogs and in elevational ranges from 300-1400 (-2100) m. The majority of the species are single-island endemics and their distributional range is often highly restricted (Stone et al., 1999). Approximately 40% of the species are distributed on Kaua'i with decreasing species numbers on the younger islands (Stone et al., 1999), indicating that diversification might follow the progression rule. The species are divided into five morphologically defined, informal groups, which we refer to as 'Stone's sections' (Paetzold et al., 2019). Stone's sections are defined largely by the degree of carpel connation (apocarpous to fully syncarpous), and the sexuality of flowers, (perfect vs. functionally unisexual and plants dioecious) (Stone et al., 1999).

Species relationships within the lineage have recently been resolved using RAD-seq (Paetzold et al., 2019). The analysis resolved five main clades in Hawaiian *Melicope*, only two of which corresponded to Stone's sections. The fully resolved relationships in the group allow biogeographic and diversification hypothesis testing. The main

aim of this investigation is to elucidate the spatio-temporal diversification in Hawaiian *Melicope* in the context of island geology. We will also test whether specific habitat types impact diversification.

## 4.2 Material and Methods

We used the RAD-seq datasets generated in Paetzold et al. (2019) containing 101 specimens, including two Marquesan species representing the two independent colonization events from the Hawaiian to the Marquesas Islands (Appelhans et al., 2014a, b; Paetzold et al., 2019) and six non-Hawaiian members of *Melicope* sect. *Pelea* as outgroup. The five datasets resulted from filtering the assembled loci by species coverage to a minimum of 10, 32, 50, 67 and 85, respectively (Paetzold et al., 2019). All computational analyses were run on the high performance computing cluster of the GWDG, Goettingen.

### Divergence Time Estimation

We used BEAST v. 1.10.4 (Drummond and Rambaut, 2007) for divergence time estimation in Hawaiian *Melicope*. As there are no suitable fossils, we used a secondary calibration to constrain the root age (Appelhans et al., 2018b). Secondary calibrations are sensitive to a potential primary error introduced by the previous analysis (Forest, 2009). To address this we used a normal distribution as prior of the root age with an upper truncation at 15 mya and a lower truncation at 5 mya, corresponding to the 95% credibility interval (CI) inferred by Appelhans et al. (2018b).

Two independent BEAST runs were performed with 5,000,000 generations each and every 500$^{th}$ generation stored under a constant size Coalescent Tree Prior, the GTR+G+I model of sequence evolution and 1/x as prior for the population size. We tested several clock rates and tree prior combinations, but only runs with a strict clock reached convergence. To evaluate the impact of dataset size on estimated node ages, we selected 11 nodes, representing the origin of major lineages and one colonization event to the Marquesas Islands (Figure 4.1), and compared the inferred ages. We used the Shapiro-Wilk-Test (Shapiro and Wilk, 1965) and multiple linear regression models (LM) to test whether the estimated mean node ages were biased to dataset size using the functions shapiro.test() and lm() in *R* v 3.6.1 (R Core Team, 2019).

Currently, the size of genome-sized datasets often poses a struggle for Bayesian methods, which promoted the development of less complex algorithms (To et al., 2016; Volz and Frost, 2017). To evaluate the results from divergence time estimation

using BEAST, considering that only a strict clock reached convergence (see Discussion), the least-squares dating method (LSD; To et al., 2016) implemented in the software IQTree v. 2.0.6 (Nguyen et al., 2015; Minh et al., 2020) was employed as well. The method combines the effective algorithm for resolving Maximum Likelihood (ML) trees provided by IQ-Tree (Nguyen et al., 2015; Minh et al., 2020) with the relaxed lognormal clock, Gaussian-noise, least-squares approach provided by LSD2 (To et al., 2016). In simulations that the LSD algorithm proved robust to uncorrelated violations of the molecular clock and performed similar to BEAST in most scenarios (To et al., 2016). Least-squares dating was performed with the min32 ML phylogeny from (Paetzold et al., 2019) as starting tree and 100 bootstrap replicates to attain confidence intervals. The algorithm may fail to produce a unique solution with an age interval and a single calibration point (Hien, personal communication). Consequently separate analyses were computed on each input alignment, one for the upper and lower bounds of the calibrated node, respectively and results were merged. This approach will be available in future versions of the LSD software (Hien, personal communication).

### Ancestral Area Reconstruction

For ancestral area reconstruction (AAR), we defined entire islands as ranges of endemism. The four islands of the Maui Nui complex were treated as a single unit (Price and Elliott-Fisk, 2004) as were the Marquesas Islands. For species occurring on multiple islands, we specified the entire distributional range as area with the exception of the widespread and morphologically very variable *M. clusiifolia*. A geographically stratified pattern within this taxon might provide insights towards its taxonomic treatment. Hence, we chose the sample origin as its respective area.

We used the *R* package BioGeoBEARS (Matzke, 2013), which implements three different models of range evolution, differing in assumptions of range evolution and inheritance. We tested three models, diversification-extinction-cladogenesis (DEC) (Ree et al., 2005; Ree and Smith, 2008), DIVA-like (Ronquist, 1997) and BayArea-like (Landis et al., 2013) with maximum range size limited to five areas. We conducted another analysis under each model including the free parameter j with a starting value of 0.5 (Matzke, 2014) to account for possible jump-dispersal speciation events. The reconstruction was time-stratified to incorporate island ages as estimated by Price and Elliott-Fisk (2004).We compared the results of the six models (DEC, DEC+J, DIVALIKE, DIVALIKE+J, BAYAREALIKE, BAYAREALIKE+J) using the Akaike Information Criterion.

### Diversification analysis

We estimated evolutionary rate regimes across Hawaiian *Melicope* using BAMM (Rabosky, 2014; Rabosky et al., 2014a). We pruned the min32 dataset to include only one sample per species with the exception of taxa previously resolved as non-monophyletic (Supplemental Table 4.1; Paetzold et al., 2019) and only the three most closely related outgroup samples. For multi-sampled species, we selected the specimen with the higher number of loci to increase informative content. Paetzold et al. (2019) showed that *M. clusiifolia* is paraphyletic with respect to both, *M. haleakalae* and *M. waialealae*, and this complex likely requires taxonomic revision. Currently, it is unclear whether *M. clusiifolia* represents several distinct species or one morphologically highly variable taxon. We conducted the analysis twice to account for both possibilities by either including one sample per described species (DA1) or including one sample per monophyletic lineage within the clade (DA2) (Supplemental Table 4.1). We inferred a dated phylogenetic tree for each of the reduced datasets using BEAST v. 1.10.4 (Drummond and Rambaut, 2007) as described above. To account for incomplete taxon sampling, we assigned missing species to one of the five main clades based on molecular or morphological evidence (Supplemental Table 4.2) and used the percentage of sampled species per clade to inform BAMM. We ran 10,000,000 generations in 4 Markov Chain Monte Carlo (MCMC) and summarized results using the *R* package BAMMtools (Rabosky et al., 2014b).

### State-specific speciation and extinction

To test whether habitat specificity might be linked to diversification rates within Hawaiian *Melicope,* we used state-dependent speciation and extinction (SSE) inference. SSE models a birth-death process, where diversification rates are dependent on the state of an evolving character in a Bayesian framework. Originally developed for binary characters (Maddison, 2006; Maddison et al., 2007), several variants exist including one for multistate characters, MuSSE (FitzJohn, 2014). We implemented MuSSE for four habitat states: dry, mesic, wet, and bog in the *RevBayes* language (Höhna et al., 2016). The distribution of character states among species was collected from the groups latest taxonomic revision (Stone et al., 1999; Hartley, 2001), augmented with information from herbarium specimens from PTBG and US (http://sweetgum.nybg.org/science/ih/) and personal observations (K. Wood)( Supplemental Table 4.3). We specified a log-uniform distribution as priors to both, speciation and extinction rates. An exponential distribution with a mean of 10 was chosen as the prior for the transition rates. The root state priors for each state were drawn from a Dirichlet distribution. We specified the proportion of sampled species as an approximation of the probability of taxa being sampled. We ran MCMC

analyses on each DA1 and DA2 pruning the outgroup specimens (Supplemental Table 4.1), for two runs of 1,000,000 generations sampling every 100$^{th}$ generation. We summarized results using the *R* package *RavGadgets* (Höhna and Freyman, 2016) and a burnin of 20%.

## 4.3 Results

### Phylogeny and Divergence times

The min10 dataset did not finish computing in available CPU time and will not be further considered here. Both, BEAST an IQ-Tree analyses (Figure 4.1, Supplemental Figures 4.1-4.7) of the remaining four datasets resolve the same five main clades as Paetzold et al. (2019) with clade V (*Platydesma*) as sister to the remaining Hawaiian *Melicope*. Clades I (Stone's sections *Megacarpa* and *Cubicarpa*) and II (*Apocarpa* I) are sister to each other and again sister to the lineage comprising clades III (*Apocarpa* II) + IV (*Pelea*) (Fig. 2; Paetzold et al., 2019). One difference between topologies herein is the relationship of the Marquesan species *M. hivaoaensis.* While BEAST resolves the species consistently as an early divergent lineage in clade I (Figure 4.1, Supplemental Figures 4.1-4.3), IQ-Tree- ML inference places the species as sister to *M. kavaiensis* and *M*. sp. KW15733 within clade I (Supplemental Figures 4.4-4.7).

Divergence times estimated for the four RAD-seq datasets (min32, 50, 67 and 85) using BEAST differed slightly, though 95% HPD intervals were always overlapping. The Shapiro-Wilk-Test revealed that estimated ages of the eleven focal nodes (Figure 4.1) were normally distributed. While some more recent nodes tended to receive younger age estimates from smaller datasets (Supplemental Figures 4.1-4.3), LM showed the relationship to be not or only weakly significant (Supplemental Table 4.4). Divergence times as estimated using LSD2 follow a similar pattern of recent nodes being estimated younger based on smaller datasets (Supplemental Figures 4.4-4.7). However, 95% CI's are widely overlapping as well. Mean divergence times based on LSD2 are consistently older for all nodes and datasets, and fall into the upper end of the 95% HPD-intervals resulting from BEAST (Figure 4.1, Supplemental Figures 4.1-4.7).

Divergence times estimated with the BEAST algorithm from the reduced dataset for diversification analyses were largely comparable to those estimated from the full dataset. Recent nodes received nearly identical divergence times, while ancestral node ages were generally slightly older in the reduced dataset, though always by less than 500,000 years and with 95% CI's widely overlapping.

Thus, only the divergence time estimates for the full min32 dataset are discussed here (Figure 4.1, Supplemental Figure 4.4). The root age was estimated close to the calibration point at 9.7 mya (CI: 6.7-12.9 mya) using BEAST and 12.0 mya (CI: 9.5-14.7

mya) using LSD2, respectively. Our estimates suggest a Late Miocene origin of Hawaiian *Melicope* with a crown age of 8.0 mya (5.5-10.6 mya)/10.0 mya (8.5-11.5 mya).

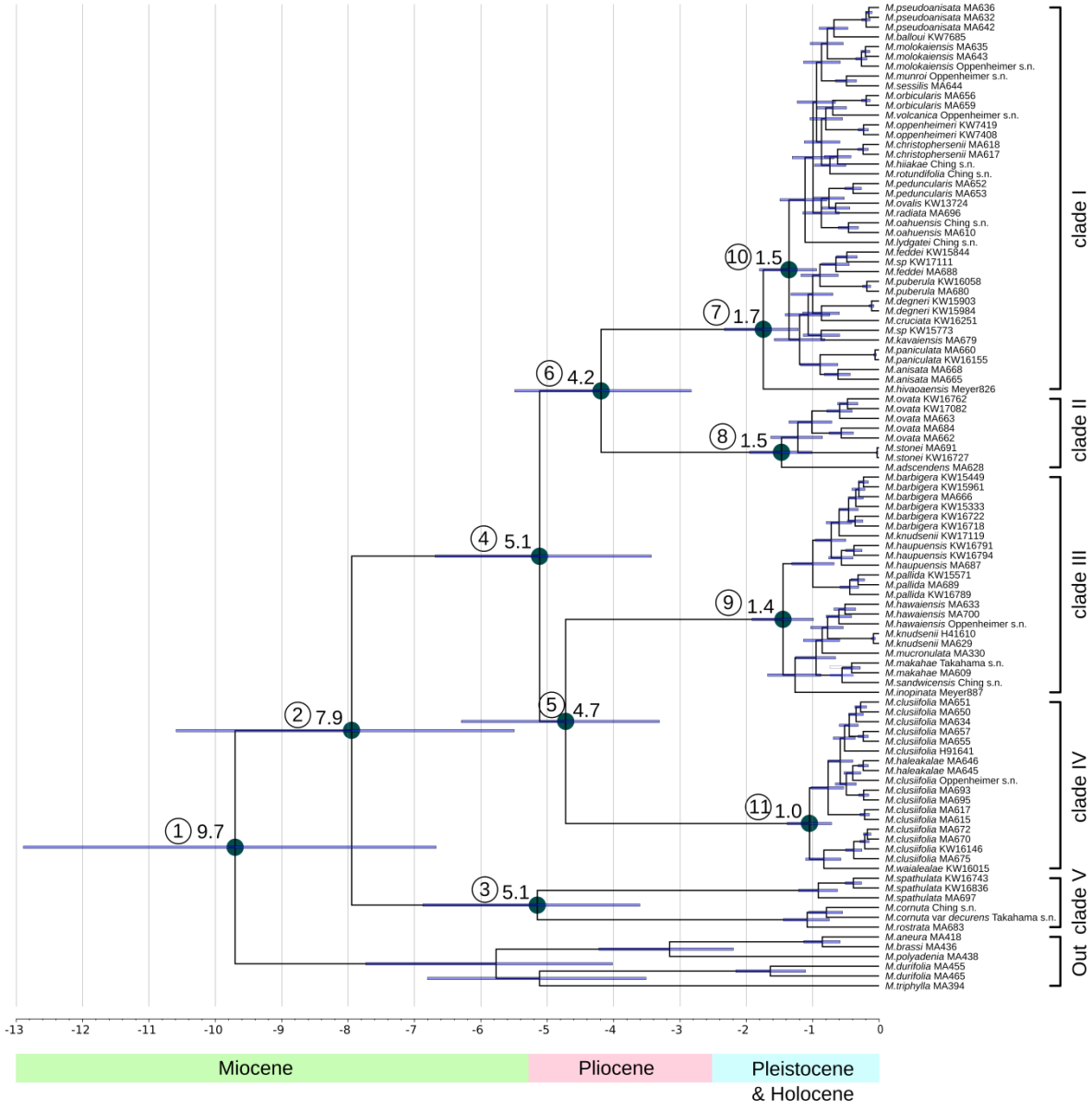



**Figure 4.1.** | Divergence times in Hawaiian *Melicope* as inferred by BEAST in the min32 dataset. The maximum clade credibility consensus tree is shown with the credibility intervals of estimated ages displayed as light blue bars. Nodes corresponding to the origin of major lineages are marked and their mean ages are shown. Node numbers refer to respective linear model tests for impact of dataset size on estimated node ages. Clades are indicated on the right. Outgroup including calibration point not shown fro graphical purposes.

The five main clades diverged within a period of less than 1 million years (MA) in the early Pliocene. We estimated the crown age of clade V at 5.2 mya (3.6-6.9 mya)/6.6 mya (5.0-8.9 mya) and of the MRCA to the four remaining groups (node 4; Figure 4.1)

at 5.1 mya (3.4-6.7 mya)/6.2 mya (4.6-8.0 mya). The estimated crown ages for the clades I+II (node 6) and clades III+IV (node 5) date back to 4.2 mya (2.8-5.5 mya)/5.1 mya (3.7-6.7 mya) and 4.7 mya (3.3-6.3 mya)/ 5.8 mya (4.3-7.5 mya), respectively.

Our results indicate that extant diversification commenced in the Mid Pleistocene. The crown ages of the main clades I-IV were estimated as: clade I: 1.7 mya (1.2-2.3 mya)/1.6 mya (1.1-2.1 mya), clade II: 1.5 mya (1.0-1.9 mya)/1.7 mya (1.1-2.4 mya), clade III: 1.5 mya (1.0-1.9)/1.5 mya (1.1-2.2 mya), and clade IV: 1.0 mya (0.7-1.4 mya)/1.2 mya (0.8-1.6 mya. Most of the extant species originated within the last 1 MA. The origin of both Marquesan Island lineages was estimated to the Mid Pleistocene as well with a stem age of1.7 mya (1.2-2.3 mya)/1.2 mya (0.8-1.6 mya) for *M. hivaoaensis* and 1.3 mya (0.9-1.7 mya)/1.5 mya (1.0-2.1 mya) for *M. inopinata*.

### Ancestral Area Reconstruction

Results of AAR revealed that adding the jump-dispersal parameter improved model fit in general, though only for BAYAREALIKE model was the likelihood significantly better. The BAYAREALIKE + J model fit our data best (Table 4.1), though inferred ancestral areas are similar between all models considering jump-dispersal (Supplemental Figures 4.8-4.12).

**Table 4.1.** | Comparison of the six models for Ancestral area reconstruction implemented in BioGeoBEARS.

| model | LnL | No. params | d | e | j | AIC | AIC_wt |
|---|---|---|---|---|---|---|---|
| **DEC** | -151.2 | 2 | 0.05 | 1.00E-12 | 0 | 306.5 | 2.80E-15 |
| **DEC+J** | -150 | 3 | 0.046 | 1.00E-12 | 0.0033 | 306.1 | 3.40E-15 |
| **DIVALIKE** | -152.9 | 2 | 0.058 | 1.00E-12 | 0 | 309.9 | 5.00E-16 |
| **DIVALIKE+J** | -151.7 | 3 | 0.053 | 1.00E-12 | 0.0041 | 309.3 | 6.60E-16 |
| **BAYAREALIKE** | -157.1 | 2 | 0.03 | 0.22 | 0 | 318.3 | 7.50E-18 |
| **BAYAREALIKE+J** | -116.7 | 3 | 0.014 | 1.00E-07 | 0.019 | 239.4 | 1 |

Figure 4.2 shows the inferred historical biogeography for Hawaiian *Melicope*. The MRCA to Hawaiian *Melicope* was inferred to originate outside of the current distributional range. The ancestor of clade V may have originated outside of the current distributional range as well, while the MRCA to all remaining species originated on Kaua'i (Figure 4.2). Within each of the five main clades biogeographic patterns are characterized by – sometimes repeated – progressive colonizations to and diversification within younger islands. Some instances of dispersal back to older islands or shifts from single-island endemism to widespread occupation of multiple islands were inferred (Figure 4.2).

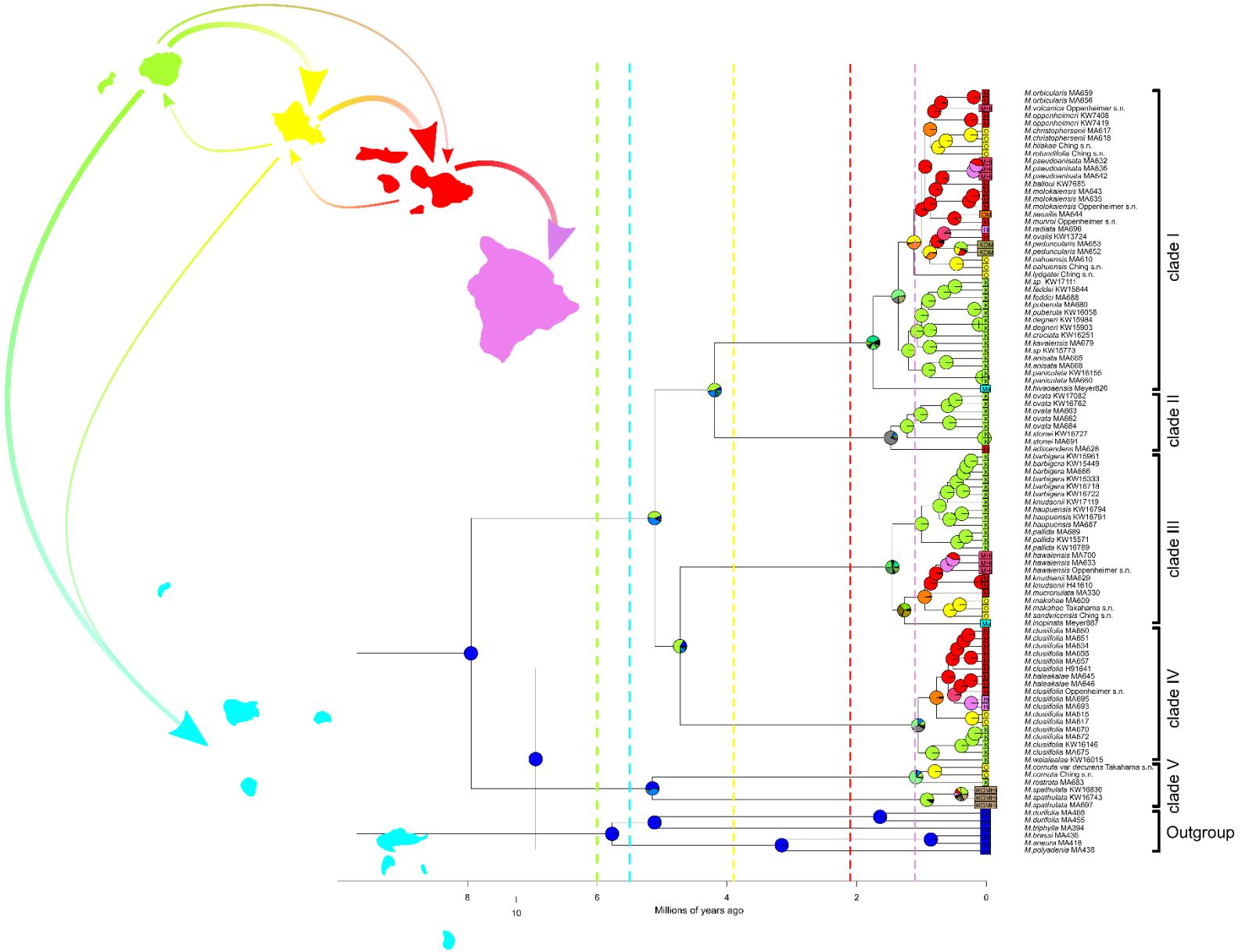



**Figure 4.2.** | Ancestral area reconstruction according to the BAYAREALIKE + J model implemented in BioGeoBEARS. A schematic view of the islands and dispersal routes is shown on the left. The origins of individual islands are indicated in the phylogeny as vertical, dashed lines. The distance between the Hawaiian

In clade V (Figure 4.2), the MRCA spawned two lineages: the widespread *M. spathulata,* which originated on Kaua'i and another taxon (*M. rostrata*) originating and remaining on Kaua'i while one daughter lineage dispersed to O'ahu. Clade IV follows the progression rule, as does clade III, except that one taxon (*M. inopinata*) dispersed to the Marquesas Islands, but its origin is not clearly inferred. In clade II (Figure 4.2)one lineage representing a single species (*M. adscendens*) dispersed from Kaua'i to Maui, while the remaining taxa diversified on Kaua'i. Clade I (Figure 4.2), containing the majority of species, shows a slightly more complex biogeographical pattern. The MRCA to the clade dispersed to the Marquesas Islands and spawned a diversification on Kaua'i and one dispersal event to O'ahu. This was followed by two independent colonizations from O'ahu to Maui followed by diversification on the island and eventual colonization of Hawai'i. The first lineage to disperse to Maui spawned *M. peduncularis*, a multi-island species occurring on Kaua'i, O'ahu, Moloka'i and Maui. Populations of *M. peduncularis* occurring in Kaua'i might represent a successful backwards dispersal. As the individuals herein were collected on Maui (Supplemental Table 4.1), the dispersal routes within the species could not be tested. The second lineage to disperse to Maui comprises the majority of all species endemic to Maui including three separate colonizations of the island of Hawaii with one resulting in a speciation event (*M. radiata*). The lineage also comprises one backwards colonization from Maui to O'ahu involving a speciation event and resulting in the diversification of three extant species (Figure 4.2).

**Diversification analysis**

The results of the BAMM analyses are plotted as lineage through time plots (LTT) and as a heat map on the branches of the phylogeny, warmer colours represent higher rates of diversification (Figure 4.3). For the DA1 dataset, seven distinct shift configurations wre identififed in the 95% credible (Supplemental Figure 4.9). The majority of inferred shift configurations correspond to the one represented by the mean phylorate plot (Figure 4.3a), with the time points for the shifts varying along branches and each clade within Hawaiian *Melicope* showing a different relative diversification rate (Figure 4.3a), the lowest inferred for clade V and the highest in clades I and III. The LTT plot (Figure 4.3a) shows a relatively constant diversification rate, with a slight increase ca. 5.3 mya, and a steep increase between ca. 1.4 and 0.3 mya. For the DA2 dataset, two distinct shift configurations were found in the 95% credible shift set (Supplemental Figure 4.14). The mean phylorate plot (Figure 4.3b) shows three different rates of diversification for the five clades in Hawaiian *Melicope.* Clade V has the lowest relative rate of diversification, and rates increased in the MRCA to the remaining four clades. Clade I then shows a second increase in relative diversification rate. The LTT plot (Figure 4.3b) shows a similar pattern as described for DA1.

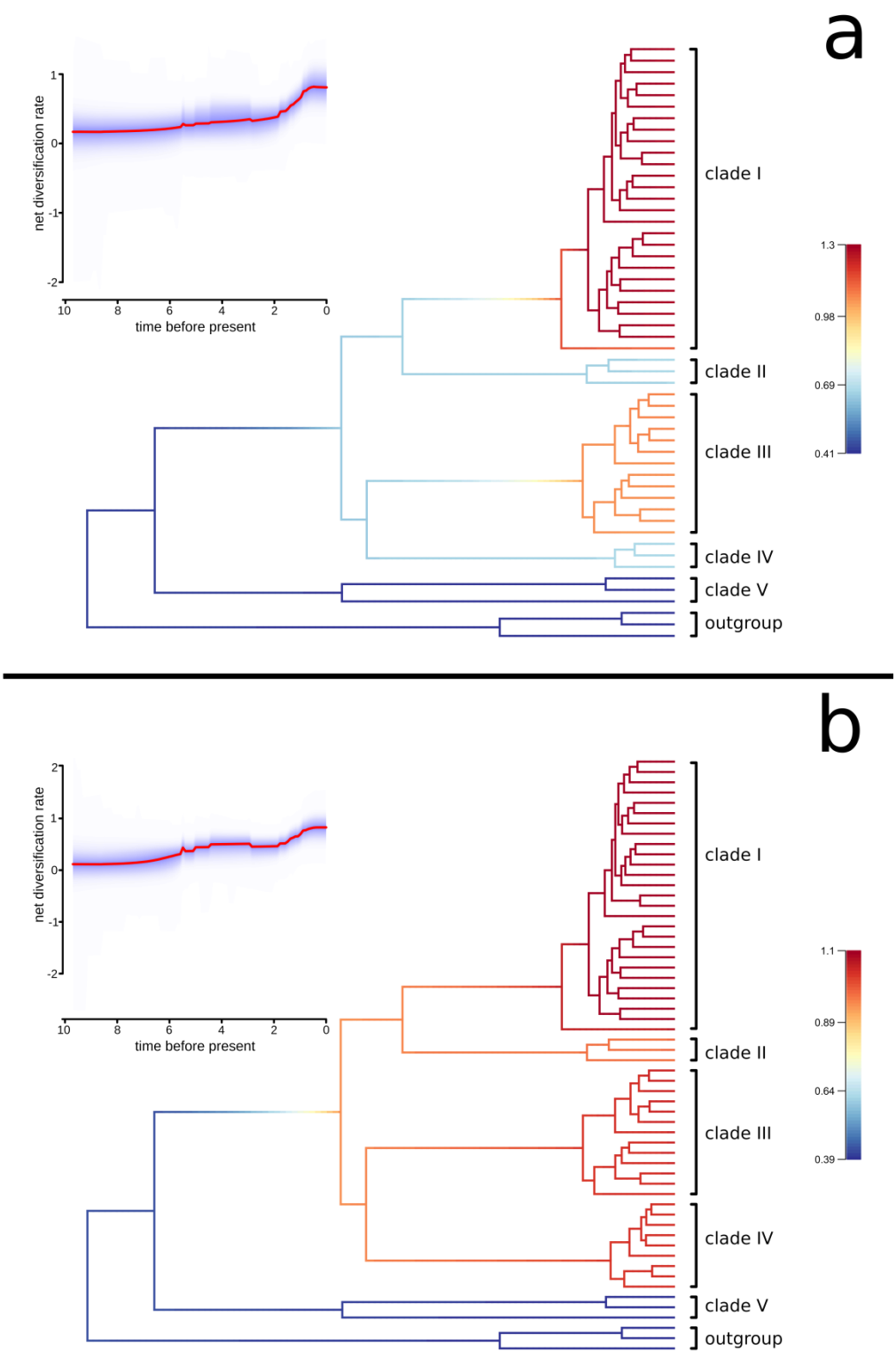



**Figure 4.3.** | Results of diversification analysis of a: DA1 (current taxonomic description of *M. clusiifolia* and close relatives) and b: DA2 (*M. clusiifolia* as several distinct lineages) using BAMM. The branches in the phylogenetic tree are color-coded by their relative diversification rates. Warm colors represent higher relative diversification rates, cold colors lower, relative diversification rates. The lineage through time plot shows the time in millions of years on the x-axis and relative

diversification rates on the y-axis. The mean diversification rate is red and the credibility interval shaded in blue.

### Multiple state-dependent Speciation and Extinction (MuSSE)

Analyses of DA1 and DA2 yielded highly comparable results with values for rates showing the same trend across datasets (Table 4.2). A mesic habitat was revealed as the ancestral state in Hawaiian *Melicope* with medium support (Figure 4.4). Posterior probability (PP) values for ancestral states are generally high in clades II and III and generally low in clades I, IV and V (< 0.5 PP).

**Table 4.2.** | Results of Multiple character dependent Speciation and Extinction analysis in DA1 and DA2 for different habitats. Subscript numbers 0, 1, 2, and 3 refer to dry, mesic, wet and bog habitat, respectively. Posterior estimates for rates of speciation (λ), extinction (μ), diversification (r) and state transition ($q_{xy}$) are given as mean with their 95% credibility interval (CI).

| | DA1 | | DA2 | |
|---|---|---|---|---|
| | **mean** | **CI** | **mean** | **CI** |
| $\lambda_0$ | 20.27 | $10^{-6}$-76.2 | 2.52 | $10^{-6}$-14.60 |
| $\mu_0$ | 29.36 | $10^{-6}$-99.94 | 15.15 | $10^{-5}$-73.81 |
| $r_0$ | -9.09 | -68.7-4.45 | -12.63 | -79.25-4.06 |
| $\lambda_1$ | 0.55 | 0.31-0.81 | 0.521 | 0.29-0.81 |
| $\mu_1$ | 0.022 | $10^{-6}$-0.93 | 0.021 | $10^{-5}$-1.05 |
| $r_1$ | 0.532 | 0.29-0.79 | 0.449 | 0.253-0.77 |
| $\lambda_2$ | 0.66 | $10^{-6}$-1.78 | 1.43 | 0.89-2.11 |
| $\mu_2$ | 0.06 | $10^{-6}$-0.38 | 0.028 | $2*10^{-5}$-0.16 |
| $r_2$ | 0.6 | 0.45-1.94 | 1.401 | 0.86-2.06 |
| $\lambda_3$ | 4.03 | $10^{-9}$-35.21 | 0.128 | $10^{-6}$-0.36 |
| $\mu_3$ | 7.18 | $10^{-6}$-42.97 | 0.399 | $10^{-5}$-1.78 |
| $r_3$ | -3.15 | -99.4-75.4 | -0.27 | -2.08-0.79 |
| $q_{01}$ | 0.131 | $6.3*10^{-5}$-0.39 | 0.152 | $10^{-5}$-0.44 |
| $q_{02}$ | 0.11 | $10^{-6}$-0.36 | 0.13 | $3*10^{-5}$-0.39 |
| $q_{03}$ | 0.043 | $6.5*10^{-6}$-0.13 | 0.055 | $4.7*10^{-5}$-0.13 |
| $q_{10}$ | 0.278 | $5.6*10^{-5}$-0.72 | 0.149 | $7.9*10^{-6}$-0.45 |
| $q_{12}$ | 0.089 | $4.4*10^{-6}$-0.96 | 0.06 | $9.6*10^{-6}$-0.19 |
| $q_{13}$ | 0.11 | $1.9*10^{-6}$-0.31 | 0.112 | $3.9*10^{-5}$-0.25 |
| $q_{20}$ | 0.088 | $9.7*10^{-6}$-0.26 | 0.063 | $1.2*10^{-6}$-0.2 |
| $q_{21}$ | 0.239 | $4.58*10^{-5}$-0.51 | 0.293 | 0.05-0.57 |
| $q_{23}$ | 0.171 | $3.8*10^{-5}$-0.5 | 0.177 | $3.5*10^{-6}$-0.49 |
| $q_{30}$ | 0.065 | $1.3*10^{-5}$-0.22 | 0.053 | $10^{-6}$-0.016 |

| | | | | |
|---|---|---|---|---|
| $q_{31}$ | 0.131 | $2.8*10^{-5}$-0.39 | 0.188 | $6.6*10^{-6}$-0.36 |
| $q_{32}$ | 0.186 | $1.9*10^{-5}$-0.61 | 0.15 | $7.9*10^{-6}$-0.42 |

For species occurring in habitat types dry and bog, the analysis revealed a high negative mean diversification rate resulting from high speciation rates and very high extinction rates (Table 4.2). For habitat types mesic and wet, mean diversification rates were positive, but lower by at least an order of magnitude and result from comparatively low speciation and even lower extinction rates (Table 4.2). Transition rates vary by an order of magnitude as well but are generally lower than speciation rates. The highest transition rates are inferred for shifts from mesic to dry and from bog to wet (Table 4.2).

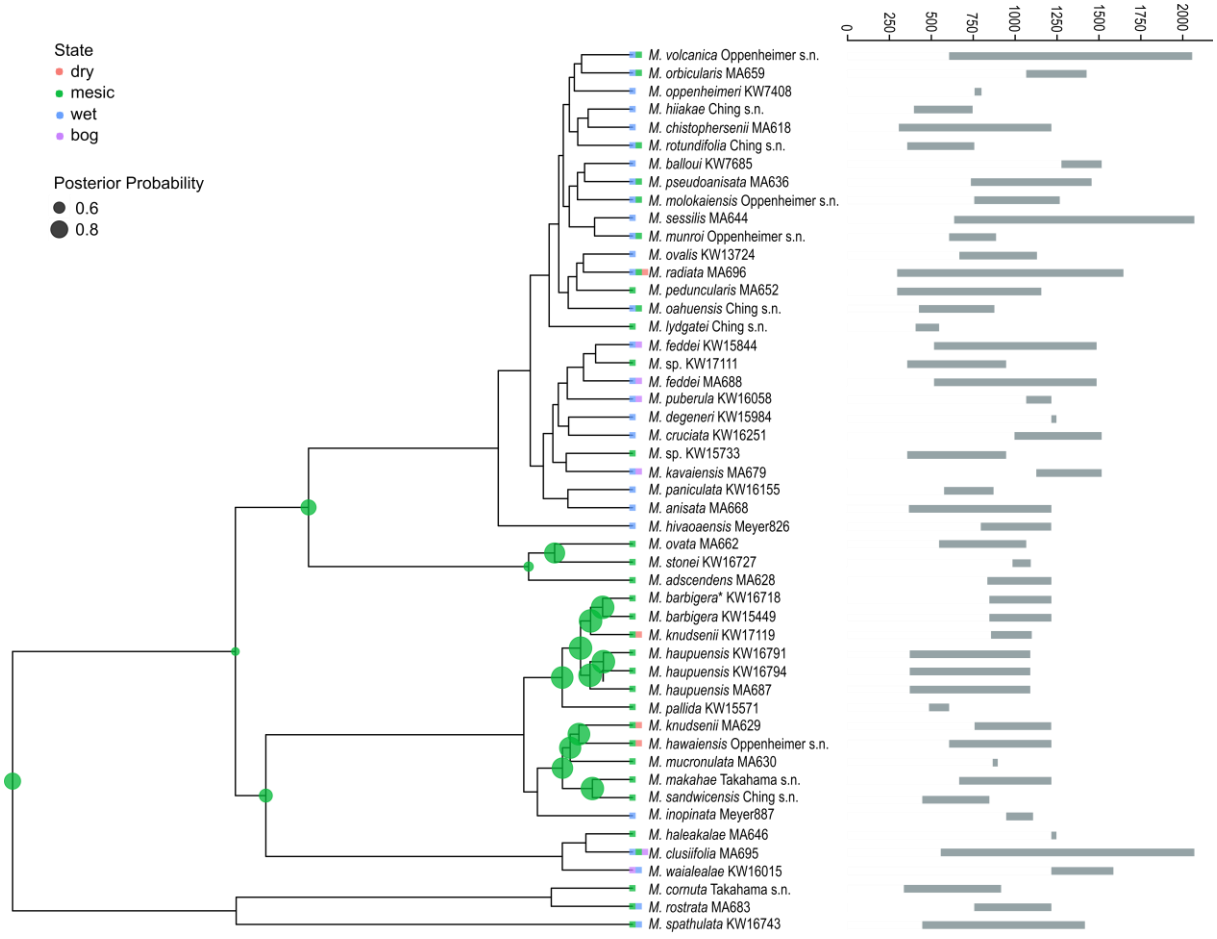



**Figure 4.4.** | Results of Multiple character dependent Speciation and Extinction (MuSSE) analyses for different habitats in the DA1 datasets. The elevational range for species included in the study is indicated.

## 4.4 Discussion

The tree topologies inferred herein are largely congruent with previous results based on the same datasets (Paetzold et al., 2019). One difference between topologies inferred by BEAST and the IQ-Tree-ML algorithm is the placement of the Marquesan species M. hivaoaensis within clade I, which was also recovered by Paetzold et al. (2019) and interpreted as possibly indicating reticulate evolution prior to the colonization of the Marquesas Islands. Sampling of Marquesan species is insufficient herein to conclusively address this question, which therefore must remain unresolved and results, including divergence time estimation as well as the historical biogeography remain preliminary. However, the relationships of the five main clades as resolved herein are congruent with those resulting from concatenated ML and Bayesian approaches of the same datasets (Paetzold et al., 2019).

For BEAST analyses we only achieved convergence of MCMC runs under the assumption of a strict clock. Runs assuming an uncorrelated lognormal relaxed clock did not converge within available computational time. The strict clock assumes that the evolutionary rate of the underlying data is the same across all branches and the entire length of the tree. This assumption has been proven inadequate in many empirical datasets (Ho and Duchêne, 2014). The uncorrelated lognormal relaxed clock model allows every branch to evolve at a unique, independent rate (Ho and Duchêne, 2014), resulting in a large parameter space. Compared to simpler models adequately exploring the entire parameter space drastically increases the number of required MCMC moves. Considering the size and complexity of the datasets herein (Paetzold et al., 2019), it is entirely reasonable, that 5,000,000 MCMC iterations are not sufficient to explore the parameter space of a relaxed model and reach convergence of rate estimates.

On the other hand, the dataset herein might be best (or at least well enough) characterized by a model assuming a strict clock. RAD datasets contain thousands of short loci, of which the vast majority originate from non-coding regions of the nuclear genome (Hipp et al., 2014). Non-coding regions are assumed to approach a neutral evolutionary rate better than coding regions. In this case, the assumption of a strict clock-like evolution might increase model fit and thus lead to convergence of MCMC runs. Ideally, model testing using a Stepping Stone algorithm and Bayes factors could estimate which clock model fits best. However, this requires that MCMC runs converge to sample from an effective power posterior (Baele et al., 2013; Barido-Sottani et al., 2018) and was thus not feasible here.

We compared the strict-clock BEAST results to those estimated by the LSD2 algorithm, which is less complex but uses a lognormal relaxed clock (To et al., 2016; Minh et al., 2020). Node ages were generally higher than those estimated by BEAST, falling into the upper third of the 95% HPD intervals estimated by the latter.

However, the general pattern of divergence times across the tree is comparable, with no node age estimated younger by LSD2 than BEAST or beyond the 95% HPD interval, which might be expected if a specific lineage violates the strict-clock assumption.

A recent study based on Sanger-sequencing of five marker regions and extensive sampling in the entire genus *Melicope* employed a similar calibration scheme and a relaxed lognormal clock (Appelhans et al., 2018b). The analysis resolved the monophyly of Hawaiian *Melicope* and the five main clades with very high statistical support, while inter-species and inter-clade-relationships were not well resolved. Age estimations for the supported Hawaiian clades are highly congruent to our BEAST estimates and in the case of clades I and III even identical, with 95% HPD intervals largely overlapping. The exception is the MRCA to clade V (*Platydesma*), estimated to be 2.4 MA old in Appelhans et al. (2018b), but nearly twice (Figure 4.1) or three times (Supplemental Figure 4.4) that in here. This could be caused by an outgroup effect, as the outgroup comprised the entirety of the genus plus several related genera in Appelhans et al. (2018b) but only six samples herein.

To date studies employing RAD-seq datasets for divergence time estimation are scarce or limited to a small number of samples (Lecaudey et al., 2018; Zhou et al., 2018). We expect an improvement of the computational efficiency of the multispecies coalescent model triggered by the increasing application of HTS and the resulting availability of phylogenomic datasets.

We investigated the relationship between the number of RAD loci in the alignment and the inferred node ages (Supplemental Table 4.4). The number of loci in the dataset had a weakly significant impact on the estimated node ages for more recent divergences, with inferred node ages an average 100,000 years older in the smallest dataset compared to the largest. This is probably related to the nature of RAD-seq assembly filtering by the number of samples per locus. Increasing evolutionary distance between taxa causes RAD locus dropout due to enzyme recognition site mutations. Thus, loci found in a comparatively large number of samples are conservative, while loci recovered from a limited number of samples have a higher rate of sequence evolution. Herein, the inferred divergence times for the most ancient nodes (1-6; Figure 4.1) are highly similar across all datasets (Supplemental Table 4.4), which likely reflects the fact that the oldest divergences are informed by the conservative loci in the dataset. The estimated differences in node ages are not altering the general picture of diversification patterns in *Melicope*, as they do not skirt time spans of geological or climatological changes in the region and 95% HPD intervals overlap widely.

A larger taxon sampling spanning the entire genus *Melicope*, comprehensive sampling of outgroups allowing the primary calibration of several nodes and improvements to molecular dating algorithms will be required to decisively evaluate our estimation. Nevertheless, our inferences provide the best resolved results to date.

### Divergence and biogeography

As node ages estimated using BEAST are lower compared to the results of LSD2, they represent the more conservative estimates for divergence times. Thus, the discussion will focus on these conservative estimates as they represent the 'minimum' estimated age.

Hawaiian *Melicope* originated at least ca. 7.9 mya (6.7-12.9 mya) (Figure 4.1). This is the second analysis inferring the origin of the lineage to pre-date the rise of the current high islands by more than 2 MA (Appelhans et al., 2018b). The colonization of the Hawaiian archipelago was seemingly not followed by immediate diversification (Figure 4.1). This is opposite to observations for the majority of Hawaiian adaptive radiations, where colonization was often followed by rapid diversification into distinct lineages, e.g. silverswords (Baldwin and Sanderson, 1998; Landis et al., 2013), honeycreepers (Lerner et al., 2011), or *Schiedea* (Willyard et al., 2011).

The period between 8-5 mya was characterized by a reduced Hawaiian archipelago with small islands lower than 1000 m (Price and Clague, 2002; Garcia et al., 2015) and represented a bottleneck for diversity on the islands. Hawaiian *Melicope* seems to represent one of the few lineages prevailing on older, low islands as did e.g. Hawaiian lobeliads (Givnish et al., 2009). As the *Melicope* colonist arrived between 4.7-11.7 (Appelhans et al., 2018b) or 6.7-12.9 mya (herein), the most likely islands to have been colonized were either Necker (ca. 10.3 mya), Twin Banks (ca. 9.6 mya), Nihoa (ca. 7.5 mya); even the French Frigate Shoals (ca. 12.0 mya) or the Gardner pinnacles (ca. 12.3 mya) (Garcia et al., 2015), although the latter two possibilities seem less likely due to their ages falling into upper end of the 95% HPD interval (Figure 4.1). These islands were all comparatively small and low in elevation, which suggests a paucity of ecological niches compared to higher islands (Price and Clague, 2002). In addition, the average times of existence for these islands are shorter compared to the present, higher elevation islands (Clague, 1996). By the time the *Melicope* colonist arrived on the archipelago, these islands might already have been in decline with increasingly high ecological niche turnover due to erosion in addition to niche saturation by earlier colonizers. Under these circumstances, the initial population of the *Melicope* colonizer might have encountered only few niches to exploit via adaptive diversification. Depending on where the colonization of the island chain occurred, the ancestor to the lineage might have arrived on the current high islands

by using *several* low islands as stepping stones. This process would represent a prolonged bottleneck.

The ancestors to the extant five main clades diverged within a short period of less than 1 MA, around the emergence of Kaua'i. The following 3 MA are characterized by seemingly low diversification before the onset of the current burst of diversification 1.5 mya (Figure 4.1). The recent period of divergence was accompanied by the colonization of the remaining islands of the Hawaiian Islands (and even the Marquesas Islands), so that the colonization of the younger islands seems to have started long after their emergence with the partial exception of Hawai'i (Figure 4.2). However, the estimated diversification rates for the last 5 MA (Figure 4.3) and the handle-and-broom pattern in the topology imply high rates of extinction in the previous intervals (Crisp and Crone, 2009). In this case, earlier lineages might have colonized O'ahu and Maui Nui, but became extinct and the extant species are descendants from subsequently colonizing lineages. The Hawaiian Archipelago is a geologically highly active region, which causes a range of high-impact events, e.g. substantial landslides and (Mega-) Tsunamis (Moore and Clague, 1992). These events might have caused substantial extinction in taxa with high degrees of endemism like *Melicope* (Stone et al., 1999). Additional strain might have been put on taxa by changing sea levels during the glacial-interglacial cycles, which reshaped or even discarded entire habitat regimes on the islands (Price and Elliott-Fisk, 2004).

Biogeographical patterns in Hawaiian *Melicope* are overall characterized by a per-clade progression rule (Figure 4.2). Within each main clade, the major route of colonization is from older to younger islands, which in this case might not necessarily reflect the emergence of the islands but rather their relative proximity. Hawaiian *Melicope* show adaptions to bird dispersal (Stone et al., 1999; Hartley, 2001).While the native Hawaiian avifauna was rich at least prior to the arrival of humans, many species adapted to island life by evolving a low-cost reduced fly apparatus, limiting their dispersal ranges (McNab, 1994).

Colonization events to younger islands resulted in speciation in most cases (Figure 4.3). However, the majority of species richness originated from intra-island diversification of established colonists. There is no general pattern of divergence; rather each clade seems to represent a unique pattern. Clade IV mainly comprises the morphologically variable *M. clusiifolia* and might represent a "generalist" strategy. The species is morphologically variable and occurs in a wide range of habitats and elevations on all islands (Figure 4.4). The diversification rate within this widespread taxon is comparatively low (Figure 4.3a). However, the taxonomic treatment of *M.*

*clusiifolia* requires revision and might represent several distinct species with speciation along the progenitor-derivative concept (Paetzold et al., 2019).

Clade III mostly comprises species adapted to mesic habitats, with the majority of taxa endemic to Kaua'i. The island has long been subjected to erosion processes, which might act as a driver of diversification by creating a plethora of dissected microhabitats exploitable by vicariant speciation.

Clade I is the most diverse clade of all, showing the highest rates of diversification (Figure 4.3) in the entire lineage. One subclade, comprising approximately half of all taxa, is endemic to Kaua'i, while the other comprises all species occurring on the younger islands. The taxa in clade I show a wide range of habitat preferences including a comparatively high amount of habitat shifts and range expansions (Figure 4.4, Table 4.2). In clade I, there is also some evidence for species-to-species matching, where different species occupy identical habitat types on different islands (Wagner and Funk, 1995), e.g. *M. molokaiensis* and *M. oahuensis* or *M. hiiakae* and *M. ovalis* (Figure 4.4). The high diversity in this clade might signify a pattern of repeated adaptive shifts to different height ranges and precipitations.

Clades II and V show low rates of diversification and species numbers. *Platydesma* (clade V) is characterized by hermaphroditic flowers with extensive nectar production as an adaption to bird pollination (Stone et al., 1999), while the remaining species of *Melicope* section *Pelea*, including the extra-Hawaiian representatives, are functionally dioecious (Hartley, 2001). Dioecy is frequent in island floras and has been associated with avoidance of inbreeding depression in small island populations and thus presenting a selective advantage (Sakai et al., 1995). In Hawaiian *Melicope* the shift towards hermaphroditism seems to have occurred only once. Divergence time estimation suggests a crown age of 5.1 - 6.5 mya for *Platydesma* (Figure 4.1), placing the shift during the time of colonization of a young, active Kaua'i. The shift to hermaphroditism might have presented a short-term advantage when population sizes were small. On the other hand, species-rich *Melicope* section *Lepta* comprises both monoclinous and dioecious species, with several shifts between the flower types and some plasticity in a number of species (Hartley, 2001; Appelhans et al., 2014a). Plasticity in flower type continues to be documented in the field for some Hawaiian *Melicope* species (Stone et al., 1999; Hartley, 2001; K. Wood, personal observation), perhaps implicating selective suppression of flower parts. A formal assessment of the trait and a genus-wide taxon sampling would be required to assess the effect of the trait on speciation.

There is an ongoing debate regarding estimation of diversification rates from dated phylogenetic trees using model-based methods like BAMM (e.g. Meyer and Wiens, 2016; Moore et al., 2016; Meyer et al., 2018). While the majority of the criticism was shown to be founded in shortcomings of either mathematical or statistical

framework, or experimental design (Rabosky, 2018, 2019; Mitchell et al., 2019), one aspect has been addressed, but remains unsolved: unobserved rate shifts. Likelihood estimation for model based rate-shift algorithms all follow the logic introduced in the BiSSE method (Maddison et al., 2007), and condition the likelihood on the finite set of observed branches on the tree (i.e. the methods assume no rate shifts in unsampled or extinct lineages). The problem has been recognized (Moore et al., 2016; Rabosky et al., 2017; Mitchell et al., 2019) but as of yet it is unclear how that information might be gained (Mitchell et al., 2019). However, recent simulations indicated, that unobserved rate shifts do not bias rate estimations significantly (Mitchell et al., 2019) and that BAMM provides more accurate rates compared to other methods andacross a range of diversification scenarios (Rabosky, 2018; Title and Rabosky, 2019). Nevertheless, our understanding of diversification progresses and is not complete and, like others before, we encourage improvements of methods and models (Rabosky et al., 2017; Louca et al., 2018; Mitchell et al., 2019).

The four different habitat types Hawaiian *Melicope* are adapted to, have a significant impact on species diversification. In mesic and wet habitats, which represent the majority of available ecological niches on the island group, speciation rates are an order of magnitude higher than extinction rates (Table 4.2), leading to a steady, but slow increase in species adapted to these habitats. In more extreme habitats, dry ranges and bogs, the inferred diversification rate is negative, resulting in a loss of species. However, both speciation and extinction rates associated with these more extreme habitats are up to 40 times higher compared to mesic and wet habitats, illustrating a high rate of species turnover and possibly a shorter time of individual species prevalence. Bog habitats do not initially exist on young, volcanic islands, as their formation requires a long-term primary succession of rainforests as well as high amounts of rainfall. In addition, bog habitats are vulnerable to decreases in moisture (Mueller-Dombois and Boehmer, 2013). Drier conditions have been common on the islands repeatedly during glacial maxima (Price and Elliott-Fisk, 2004). As such these habitats are transient, colonized late and highly competitive.

The highest rates of speciation and extinction were inferred for species occurring in dry habitats, which are generally associated to leeward, lowland areas on the Hawaiian Islands (Wagner et al., 1999a). These areas are transient in the initial states of island formation due to rapid subsidence when an island moves away from the hot-spot (Clague, 1996). Dry habitats, especially low elevation ones, are most extensive during glacial maxima, as wide ranges of land emerge from the ocean and conditions are generally cooler and much drier (Price and Elliott-Fisk, 2004). Repeated glacial cycles combined with changing habitat availability on young islands might contribute to high species turnover associated with dry habitats (Table 4.2). Dry and lowland habitats have also been exceptionally impacted by anthropogenic

changes since human arrival with many native species adapted to these conditions extinct or at high risk of extinction (Sakai et al., 2002), which might be reflected by the extremely high inferred rates of extinction.

However, considering that the highest diversification rates are observed in clade I (Figure 4.3), one key to diversification in Hawaiian *Melicope* might be trait flexibility. The ability to shift habitat types, elevational ranges and possibly other characters frequently might create an abundance of small-range specialists, with the capacity to spawn new taxa that successfully compete for different ecological niches. This evolutionary flexibility would result in a lineage that is adapted to exploit continuously emerging and changing habitats typical for a geologically active environment such as the Hawaiian Islands.

# 4. | Supplemental Information

**Supplemental Table 4.1** | Samples within this study including assignment to Stone's sections. Samples in bold are included in reduced datasets for diversification analyses: bold: DA1, bold + shaded: DA2. Asterisks denote multi-sample species included because of non-monophyly in Paetzold et al. (2019).

| Species | Coll. No. | Stone's section |
|---|---|---|
| ***Melicope adscendens* (H. St. John & E.P. Hume) T.G. Hartley & B. C. Stone** | MA628 | *Apocarpa* |
| *Melicope anisata* (H. Mann) T. G. Hartley & B. C. Stone | MA665 | *Cubicarpa* |
| ***Melicope anisata* (H. Mann) T. G. Hartley & B. C. Stone** | MA668 | *Cubicarpa* |
| ***Melicope balloui* (Rock) T.G.Hartley & B.C.Stone** | KW7685 | *Megacarpa* |
| *Melicope barbigera* A. Gray | MA666 | *Apocarpa* |
| *Melicope barbigera* A. Gray | KW15333 | *Apocarpa* |
| ***Melicope barbigera* A. Gray** | KW15449 | *Apocarpa* |
| *Melicope barbigera* A. Gray | KW15961 | *Apocarpa* |
| *Melicope barbigera* A. Gray | KW16722 | *Apocarpa* |
| *Melicope barbigera* A. Gray | KW16718 | *Apocarpa* |
| ***Melicope christophersenii* (H. St. John) T. G. Hartley & B. C. Stone** | MA618 | *Megacarpa* |
| *Melicope christophersenii* (H. St. John) T. G. Hartley & B. C. Stone | MA621 | *Megacarpa* |
| ***Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone** | MA615 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | MA617 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | MA634 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | MA650 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | MA651 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | MA655 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | MA657 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | MA670 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | MA672 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | MA 693 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | MA695 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | Oppenheimer s.n. | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | Oppenheimer H91641 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | KW16146 | *Pelea* |
| *Melicope clusiifolia* (A. Gray) T. G. Hartley & B. C. Stone | MA675 | *Pelea* |
| *Melicope cornuta* (Hillebr.) Appelhans, K.R.Wood & W.L.Wagner | Ching s.n. | *Platydesma* |
| ***Melicope cornuta* var. *decurrens* (B.C. Stone) Appelhans, K.R. Wood & W.L. Wagner** | Takahama s.n. | *Platydesma* |

| | | |
|---|---|---|
| ***Melicope cruciata* (A. Heller) T.G. Hartley & B.C. Stone** | KW16251 | *Megacarpa* |
| *Melicope degeneri* (B.C.Stone) T.G.Hartley & B.C.Stone | KW15903 | *Cubicarpa* |
| ***Melicope degeneri* (B.C.Stone) T.G.Hartley & B.C.Stone** | KW15984 | *Cubicarpa* |
| ****Melicope feddei* (H. Lév.) T. G. Hartley & B. C. Stone** | MA688 | *Megacarpa* |
| ****Melicope feddei* (H. Lév.) T. G. Hartley & B. C. Stone** | KW15844 | *Megacarpa* |
| *Melicope haleakalae* (B. C. Stone) T. G. Hartley & B. C. Stone | MA637 | *Pelea* |
| *Melicope haleakalae* (B. C. Stone) T. G. Hartley & B. C. Stone | MA645 | *Pelea* |
| ***Melicope haleakalae* (B. C. Stone) T. G. Hartley & B. C. Stone** | MA646 | *Pelea* |
| ****Melicope haupuensis* (H. St. John) T. G. Hartley & B. C. Stone** | MA687 | *Apocarpa* |
| ****Melicope haupuensis* (H. St. John) T. G. Hartley & B. C. Stone** | KW16791 | *Apocarpa* |
| ****Melicope haupuensis* (H. St. John) T. G. Hartley & B. C. Stone** | KW16794 | *Apocarpa* |
| *Melicope hawaiensis* (Wawra) T.G.Hartley & B.C.Stone | MA633 | *Apocarpa* |
| *Melicope hawaiensis* (Wawra) T.G.Hartley & B.C.Stone | MA700 | *Apocarpa* |
| ***Melicope hawaiensis* (Wawra) T.G.Hartley & B.C.Stone** | Oppenheimer s.n. | *Apocarpa* |
| ***Melicope hiiakae* (B.C.Stone) T.G.Hartley & B.C.Stone** | Ching s.n. | *Megacarpa* |
| ***Melicope hivaoaensis* J.Florence** | Meyer826 | |
| ***Melicope inopinata* J.Florence** | Meyer887 | |
| ***Melicope kavaiensis* (H. Mann) T. G. Hartley & B. C. Stone** | MA679 | *Megacarpa* |
| ****Melicope knudsenii* (Hillebr.) T.G. Hartley & B.C. Stone** | MA629 | *Apocarpa* |
| *Melicope knudsenii* (Hillebr.) T.G. Hartley & B.C. Stone | Oppenheimer H41610 | *Apocarpa* |
| ****Melicope knudsenii* (Hillebr.) T.G. Hartley & B.C. Stone** | KW17119 | *Apocarpa* |
| ***Melicope lydgatei* (Hillebr.) T.G. Hartley & B.C. Stone** | Ching s.n. | *Megacarpa* |
| ***Melicope makahae* (B. C. Stone) T. G. Hartley & B. C. Stone** | Takahama s.n. | *Apocarpa* |
| *Melicope makahae* (B. C. Stone) T. G. Hartley & B. C. Stone (cf.) | MA609 | *Apocarpa* |
| *Melicope molokaiensis* (Hillebr.) T. G. Hartley & B. C. Stone | MA635 | *Megacarpa* |
| *Melicope molokaiensis* (Hillebr.) T. G. Hartley & B. C. Stone | MA643 | *Megacarpa* |
| ***Melicope molokaiensis* (Hillebr.) T. G. Hartley & B. C. Stone** | Oppenheimer s.n. | *Megacarpa* |
| ***Melicope mucronulata* (H. St. John) T.G. Hartley & B.C. Stone** | MA630 | *Apocarpa* |
| ***Melicope munroi* (St.John) T.G.Hartley & B.C.Stone** | Oppenheimer s.n. | *Megacarpa* |
| *Melicope oahuensis* (H. Lév.) T. G. Hartley & B. C. Stone | MA610 | *Cubicarpa* |
| ***Melicope oahuensis* (H. Lév.) T. G. Hartley & B. C. Stone** | Ching s.n. | *Cubicarpa* |
| *Melicope oppenheimeri* K.R.Wood, Appelhans & W.L.Wagner | KW7419 | *Megacarpa* |
| ***Melicope oppenheimeri* K.R.Wood, Appelhans & W.L.Wagner** | KW7408 | *Megacarpa* |
| *Melicope orbicularis* (Hillebr.) T. G. Hartley & B. C. Stone | MA656 | *Megacarpa* |
| ***Melicope orbicularis* (Hillebr.) T. G. Hartley & B. C. Stone** | MA659 | *Megacarpa* |
| ***Melicope ovalis* (St.John) T.G.Hartley & B.C.Stone** | KW13724 | *Cubicarpa* |

| | | |
|---|---|---|
| ***Melicope ovata* (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone** | MA662 | *Apocarpa* |
| *Melicope ovata* (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | MA684 | *Apocarpa* |
| *Melicope ovata* (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | MA663 | *Apocarpa* |
| *Melicope ovata* (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | KW17082 | *Apocarpa* |
| *Melicope ovata* (H. St. John & E. P. Hume) T. G. Hartley & B. C. Stone | KW16762 | *Apocarpa* |
| *Melicope pallida* (Hillebr.) T. G. Hartley & B. C. Stone | MA689 | *Apocarpa* |
| *Melicope pallida* (Hillebr.) T. G. Hartley & B. C. Stone | KW16789 | *Apocarpa* |
| ***Melicope pallida* (Hillebr.) T. G. Hartley & B. C. Stone** | KW15571 | *Apocarpa* |
| *Melicope paniculata* (H. St. John) T. G. Hartley & B. C. Stone | MA660 | *Cubicarpa* |
| ***Melicope paniculata* (H. St. John) T. G. Hartley & B. C. Stone** | KW16155 | *Cubicarpa* |
| ***Melicope peduncularis* (H. Lév.) T. G. Hartley & B. C. Stone** | MA652 | *Cubicarpa* |
| *Melicope peduncularis* (H. Lév.) T. G. Hartley & B. C. Stone | MA653 | *Cubicarpa* |
| *Melicope pseudoanisata* (Rock) T.G. Hartley & B.C. Stone | MA632 | *Megacarpa* |
| ***Melicope pseudoanisata* (Rock) T.G. Hartley & B.C. Stone** | MA636 | *Megacarpa* |
| *Melicope pseudoanisata* (Rock) T.G. Hartley & B.C. Stone | MA642 | *Megacarpa* |
| *Melicope puberula* (H. St. John) T. G. Hartley & B. C. Stone | MA680 | *Megacarpa* |
| ***Melicope puberula* (H. St. John) T. G. Hartley & B. C. Stone** | KW16058 | *Megacarpa* |
| ***Melicope radiata* (H. St. John) T. G. Hartley & B. C. Stone** | MA696 | *Megacarpa* |
| ***Melicope rostrata* (Hillebr.) Appelhans, K.R. Wood & W.L. Wagner** | MA683 | *Platydesma* |
| ***Melicope rotundifolia* (A. Gray) T.G. Hartley & B.C. Stone** | Ching s.n. | *Megacarpa* |
| ***Melicope sandwicensis* (Hook. & Arn.) T.G. Hartley & B.C. Stone** | Ching s.n. | *Apocarpa* |
| ***Melicope sessilis* (H. Lév.) T. G. Hartley & B. C. Stone** | MA644 | *Megacarpa* |
| ****Melicope* sp. (*wawraeana*-like)** | KW17111 | *Megacarpa* |
| ****Melicope* sp. (*wawraeana*-like)** | KW15733 | *Megacarpa* |
| *Melicope spathulata* A. Gray | MA697 | *Platydesma* |
| ***Melicope spathulata* A. Gray** | KW16743 | *Platydesma* |
| *Melicope spathulata* A. Gray | KW16836 | *Platydesma* |
| *Melicope stonei* K.R.Wood, Appelhans & W.L.Wagner | MA691 | *Apocarpa* |
| ***Melicope stonei* K.R.Wood, Appelhans & W.L.Wagner** | KW16727 | *Apocarpa* |
| ***Melicope volcanica* (A. Gray) T.G. Hartley & B.C. Stone (cf.)** | Oppenheimer s.n. | *Megacarpa* |
| ***Melicope waialealae* (Wawra) T.G.Hartley & B.C.Stone** | KW16015 | *Pelea* |
| *outgroup* | | |
| ***Melicope aneura* (Lauterb.) T.G.Hartley** | MA418 | |
| ***Melicope brassii* T.G.Hartley** | MA436 | |
| *Melicope durifolia* (K.Schum.) T.G.Hartley | MA455 | |

| | | |
|---|---|---|
| *Melicope durifolia* (K.Schum.) T.G.Hartley | MA465 | |
| ***Melicope polyadenia* Merr. & L.M.Perry** | MA438 | |
| *Melicope triphylla* Merr. | MA394 | |

**Supplemental Table 4.2** | Assignment of unsampled species to clades herein for BAMM analysis.

| **Taxon** | **assorted to clade** | **Reference** |
|---|---|---|
| *M. cinerea* A.Gray | III | Appelhans et al., 2014b |
| *M. elliptica* (A.Gray) T.G.Hartley & B.C.Stone | I | Appelhans et al., 2014b |
| *M. hosakae* (H.St.John) W.L.Wagner & R.K.Shannon | I | Stone et al., 1999 |
| *M. kaalaensis* (H.St.John) T.G.Hartley & B.C.Stone | I | Stone et al., 1999 |
| *M. macropus* (Hillebr.) T.G.Hartley & B.C.Stone | I | Stone et al., 1999 |
| *M. nealae* (B.C.Stone) T.G.Hartley & B.C.Stone | I | Stone et al., 1999 |
| *M. obovata* (H.St.John) T.G.Hartley & B.C.Stone | III | Appelhans et al., 2014b |
| *M. quadrangularis* (H.St.John & E.P.Plume) T.G.Hartley & B.C.Stone | I | Stone et al., 1999 |
| *M. reflexa* (H.St.John) T.G.Hartley & B.C.Stone | I | Stone et al., 1999 |
| *M. remyi* (Sherff) Appelhans, K.R.Wood & W.L.Wagner | V | Appelhans et al., 2017 |
| *M. saint-johnii* (E.P.Plume) T.G.Hartley & B.C.Stone | III | Appelhans et al., 2014b |
| *M. wailauensis* (H.St.John) T.G.Hartley & B.C.Stone | I | Stone et al., 1999 |
| *M. zahlbruckneri* (Rock) T.G.Hartley & B.C.Stone | I | Stone et al., 1999 |

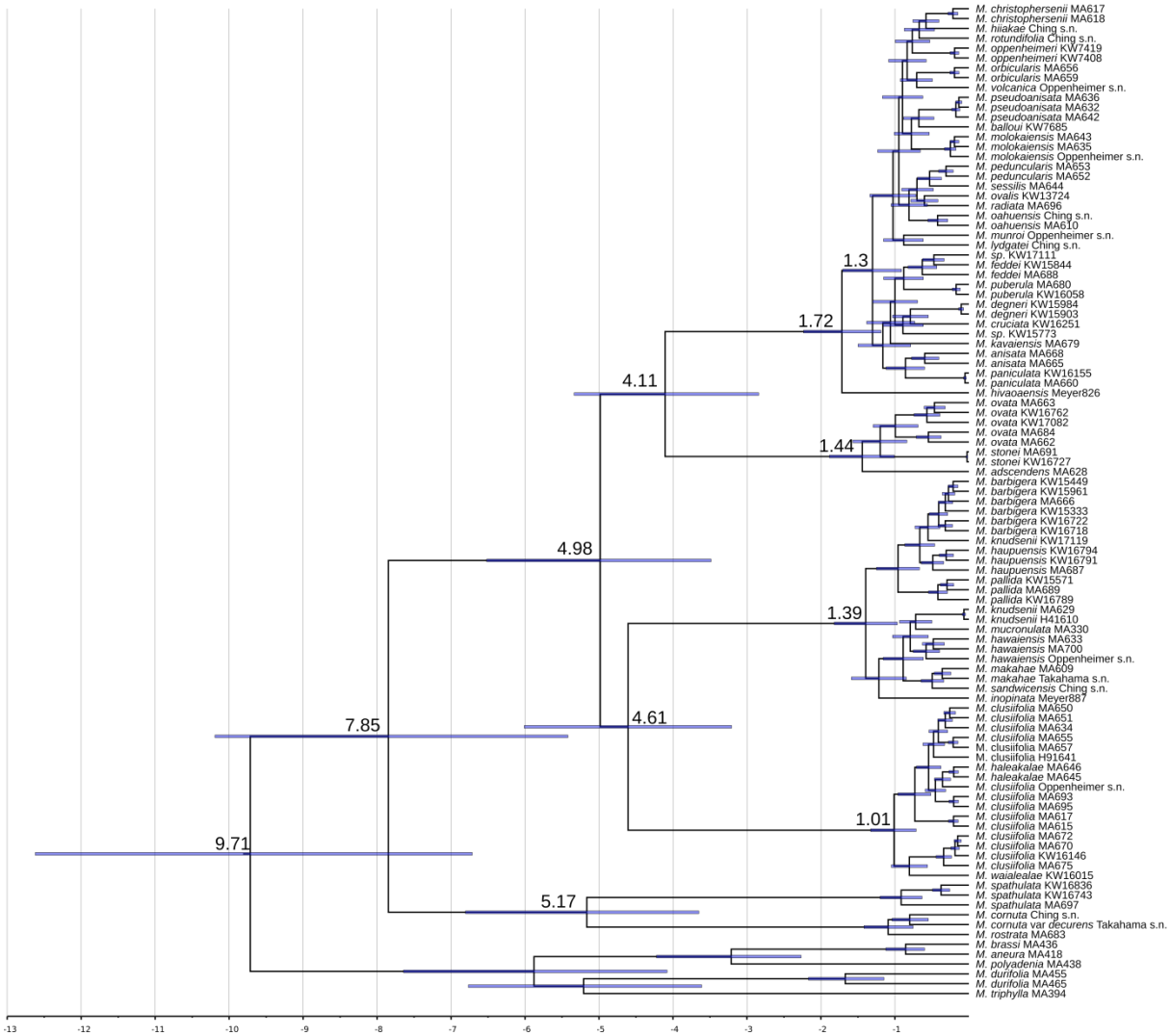
**Supplemental Table 4.3** | Habitat occupation for Hawaiian *Melicope* species investigated in MuSSE analysis. *represents a total of six closely related Marquesan species. **represents the Maui populations of *M. knudsenii*.

| species | habitat |
|---|---|
| *M. adscendens* | mesic |
| *M. anisata* | mesic-wet |
| *M. balloui* | wet |
| *M. barbigera* | mesic |
| *M. christophersenii* | wet |
| *M. clusiifolia* | mesic-wet-bog |
| *M. cornuta* | mesic |
| *M. cruciata* | wet |
| *M. degeneri* | wet |
| *M. feddei* | wet-bog |
| *M. haleakalae* | wet |
| *M. haupuensis* | mesic |
| *M. hawaiiensis* | dry-mesic |
| *M. hiiakae* | wet |
| *M. hivaoaensis** | wet |
| *M. inopinata* | wet |
| *M. kavaiensis* | wet-bog |
| *M. knudsenii* | dry-mesic |
| *M. knudsenii*** | dry-mesic |
| *M. lydgatei* | mesic |
| *M. makahae* | mesic |
| *M. molokaiensis* | mesic-west |
| *M. mucronulata* | mesic |
| *M. munroi* | mesic-wet |
| *M. oahuensis* | mesic-wet |
| *M. oppenheimeri* | wet |
| *M. orbicularis* | mesic-wet |
| *M. ovalis* | wet |
| *M. ovata* | mesic |
| *M. pallida* | mesic |
| *M. paniculata* | wet |
| *M. peduncularis* | mesic |
| *M. pseudoanisata* | mesic-wet |
| *M. puberula* | wet-bog |
| *M. radiata* | dry-mesic-wet |
| *M. rostrata* | mesic-wet |
| *M. rotundifolia* | mesic-wet |
| *M. sandwicensis* | mesic |
| *M. sessilis* | wet |
| *M. spathulatha* | mesic-wet |
| *M. stonei* | mesic |
| *M. volcanica* | mesic-wet |
| *M. waialealae* | wet-bog |

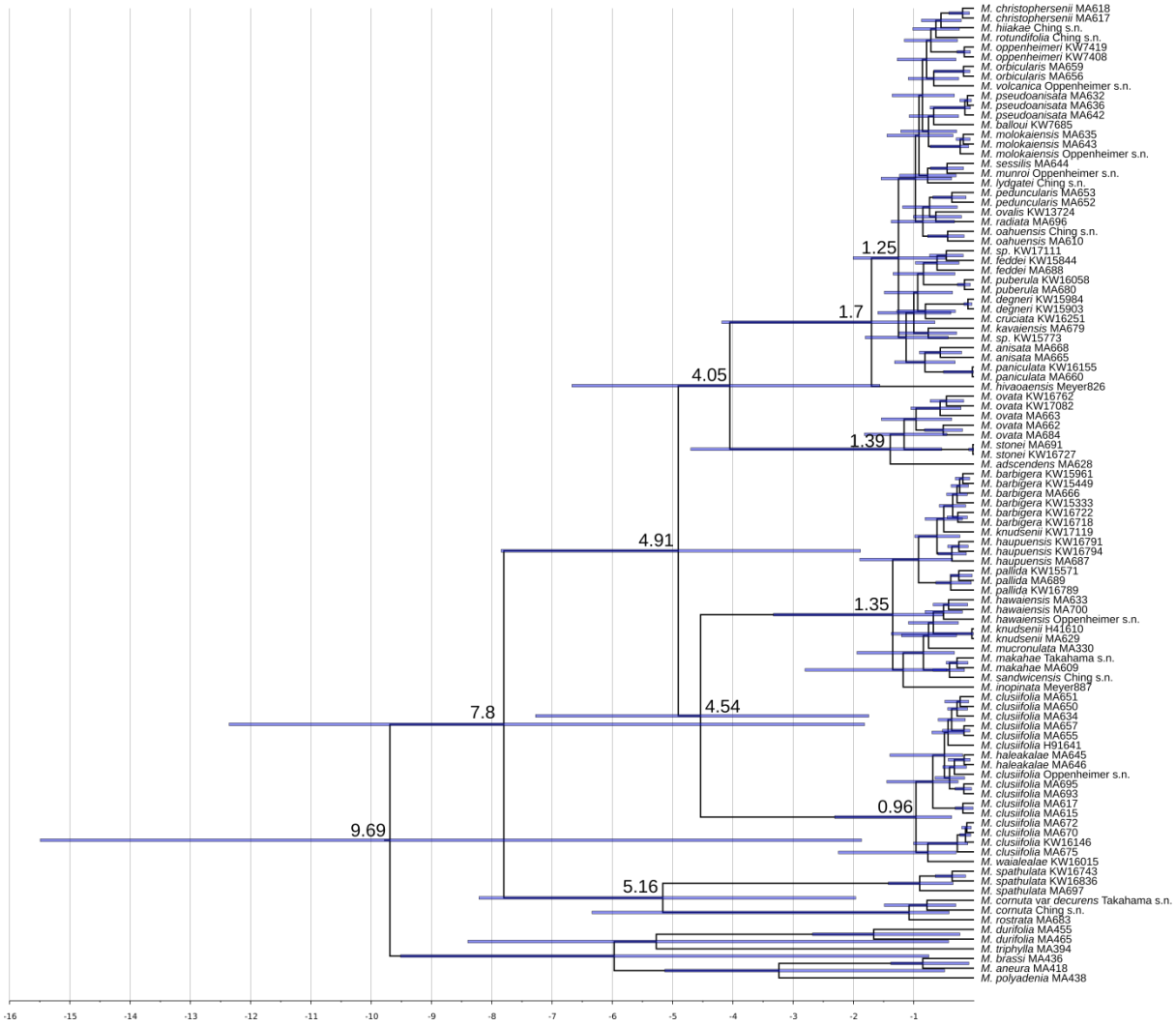
| | |
|---|---|
| *M. sp. (wawraeana-like)* | mesic |

**Supplemental Table 4.4** | Results of multiple linear regression models (LM) to infer the impact of dataset size on inferred node ages for 11 focal nodes (Figure 4.2).

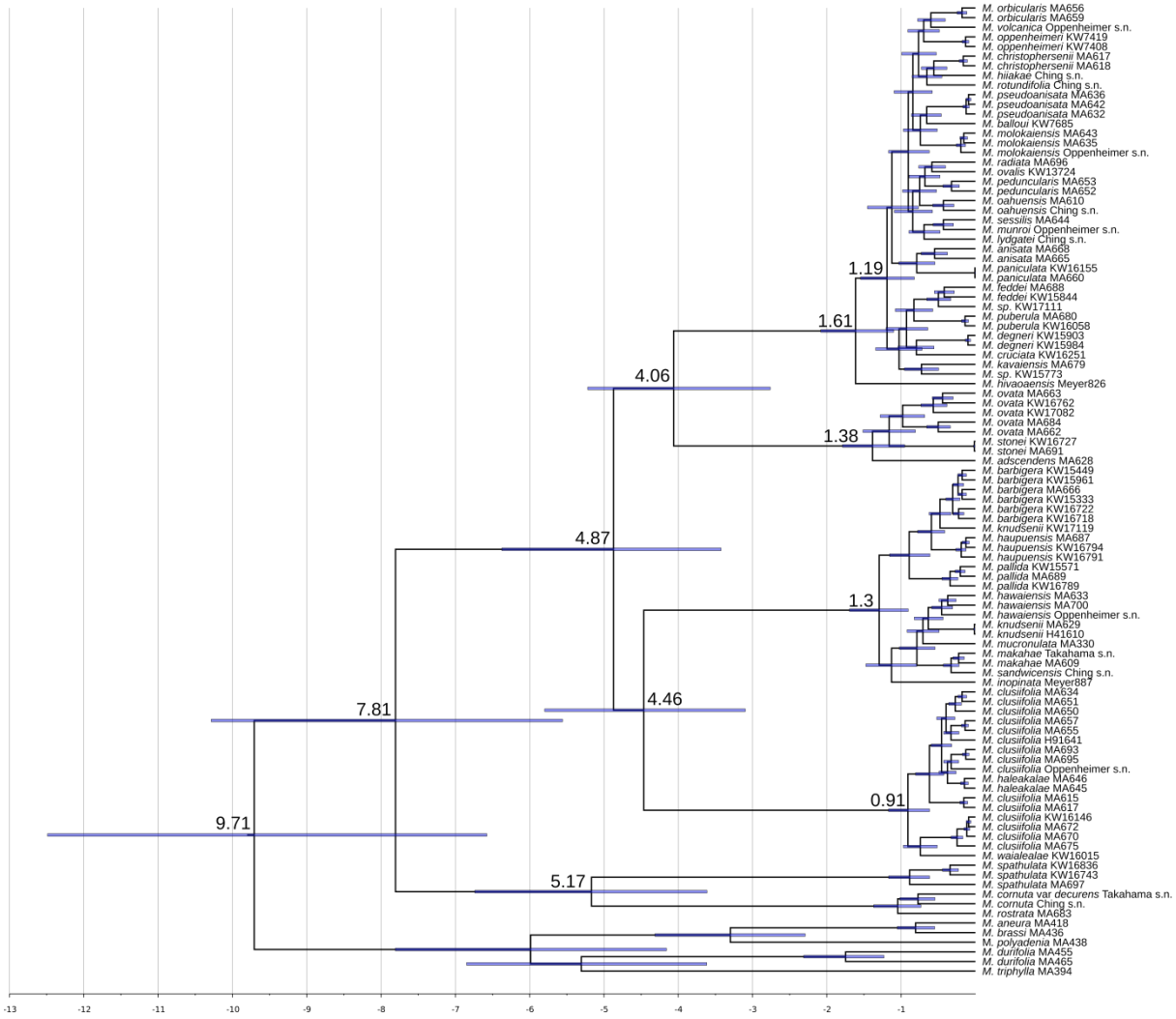
| | multiple $R^2$ | $F_{1,2}$ | p |
|---|---|---|---|
| **node 1** | 0.001458 | 0.00292 | 0.9618 |
| **node 2** | 0.5853 | 2.823 | 0.235 |
| **node 3** | 0.5243 | 2.204 | 0.275 |
| **node 4** | 0.769 | 6.657 | 0.1231 |
| **node 5** | 0.8852 | 15.42 | 0.0592 |
| **node 6** | 0.5842 | 2.81 | 0.2356 |
| **node 7** | 0.9798 | 97.02 | 0.01015 |
| **node 8** | 7697 | 6.685 | 0.1227 |
| **node 9** | 0.9731 | 33.17 | 0.0288 |
| **node 10** | 0.9613 | 49.68 | 0.0195 |
| **node 11** | 0.9703 | 65.29 | 0.01497 |



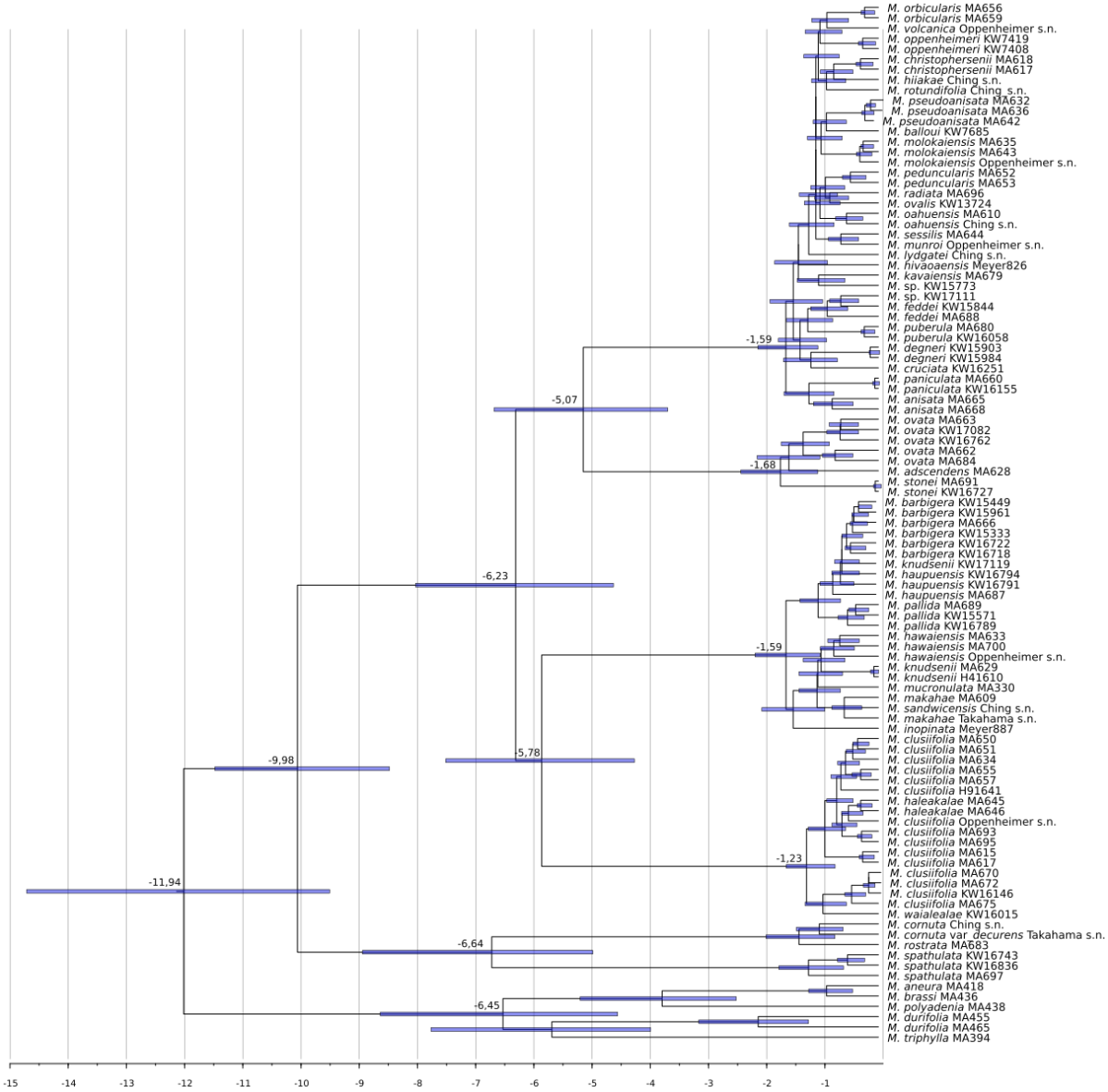

**Supplemental Figure 4.1** | Divergence times in Hawaiian *Melicope* as inferred by BEAST in the min50 dataset. The maximum clade credibility consensus tree is shown with the credibility intervals of estimated ages displayed as light blue bars. Mean ages of nodes corresponding the origin of major lineages are shown.



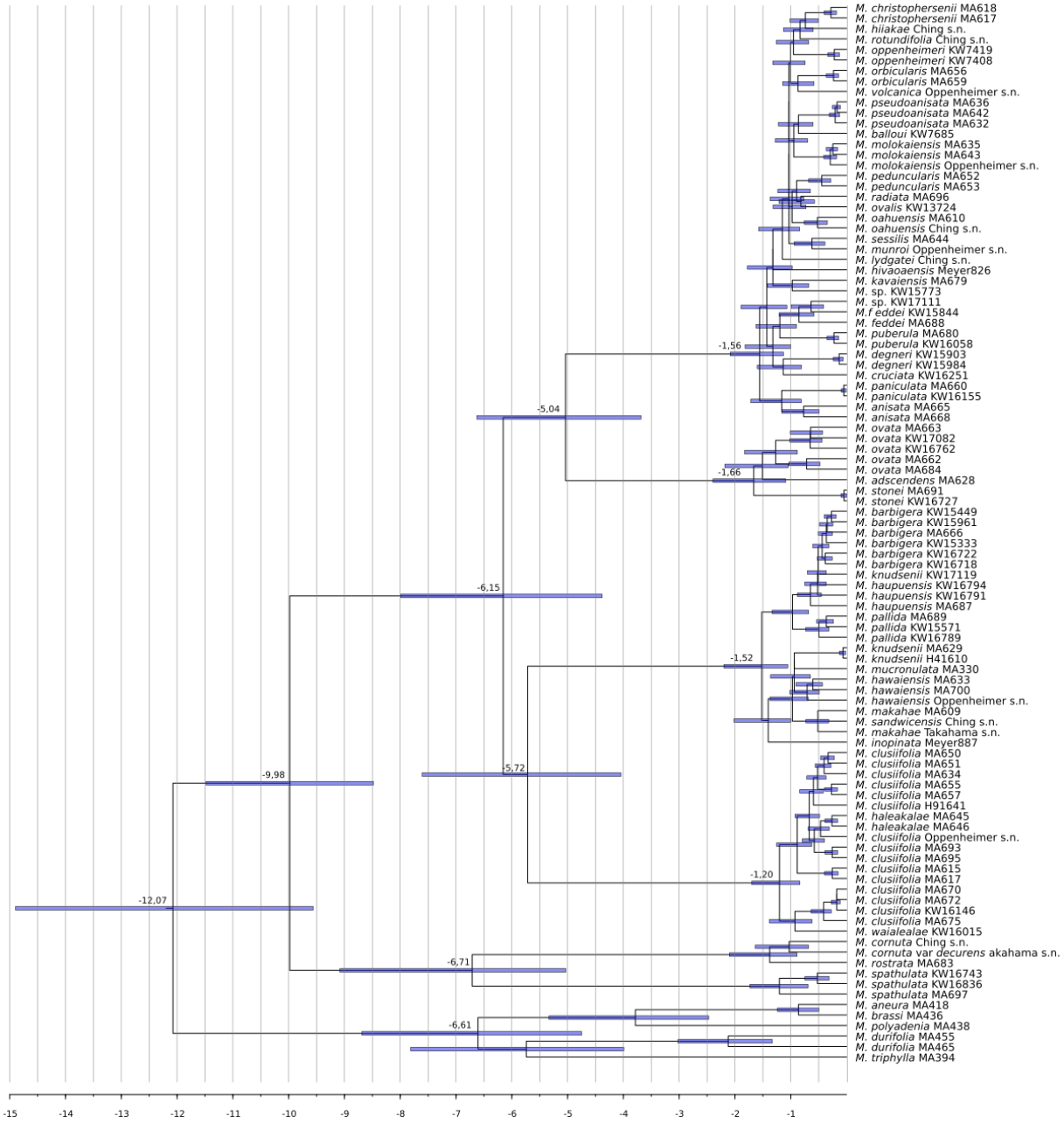

**Supplemental Figure 4.2** | Divergence times in Hawaiian *Melicope* as inferred by BEAST in the min67 dataset. The maximum clade credibility consensus tree is shown with the credibility intervals of estimated ages displayed as light blue bars. For nodes corresponding the origin of major lineages mean ages are shown in addition to the 95% credibility interval.



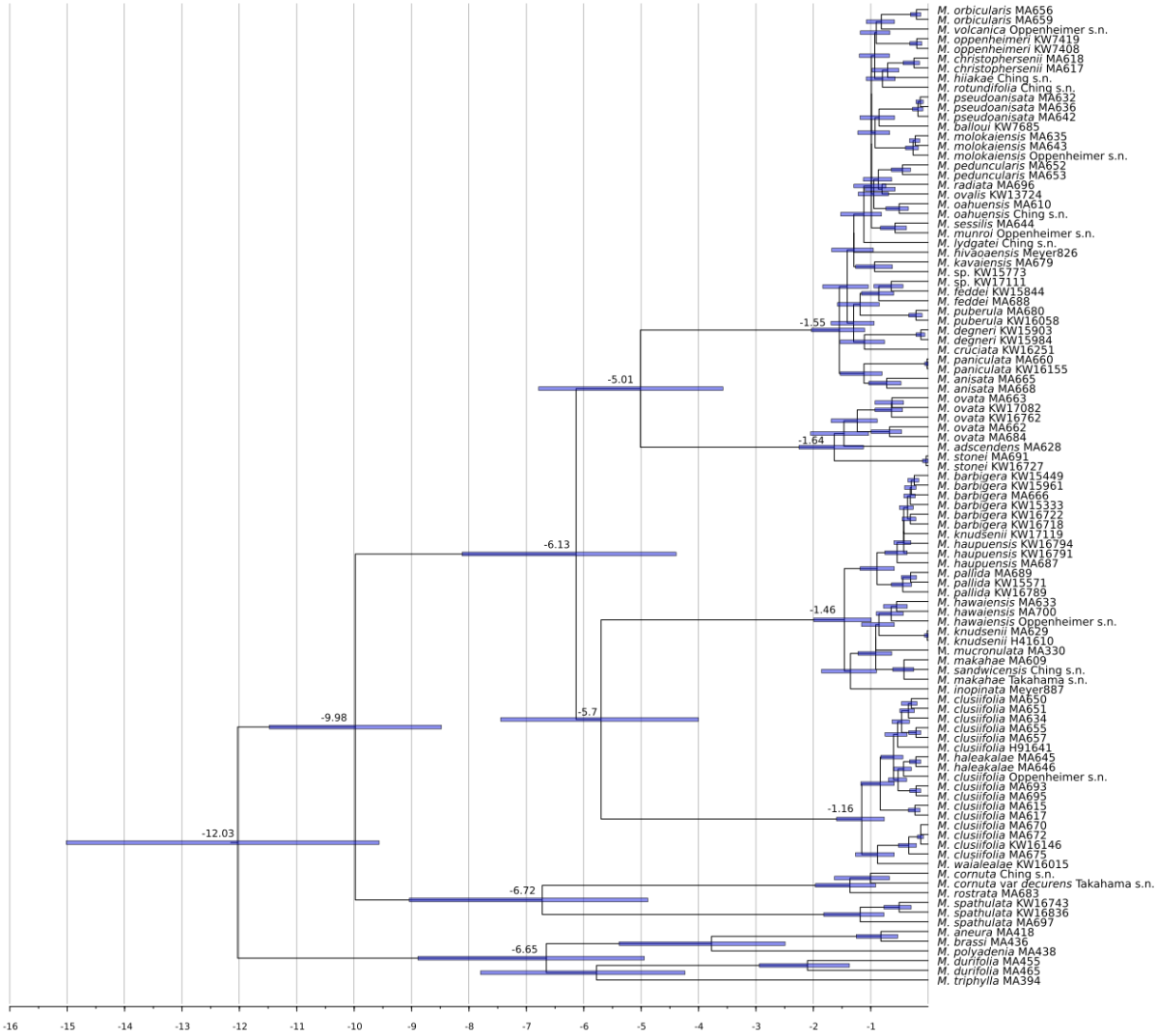

**Supplemental Figure 4.3** | Divergence times in Hawaiian *Melicope* as inferred by BEAST in the min85 dataset. The maximum clade credibility consensus tree is shown with the credibility intervals of estimated ages displayed as light blue bars. For nodes corresponding the origin of major lineages mean ages are shown in addition to the 95% credibility interval.



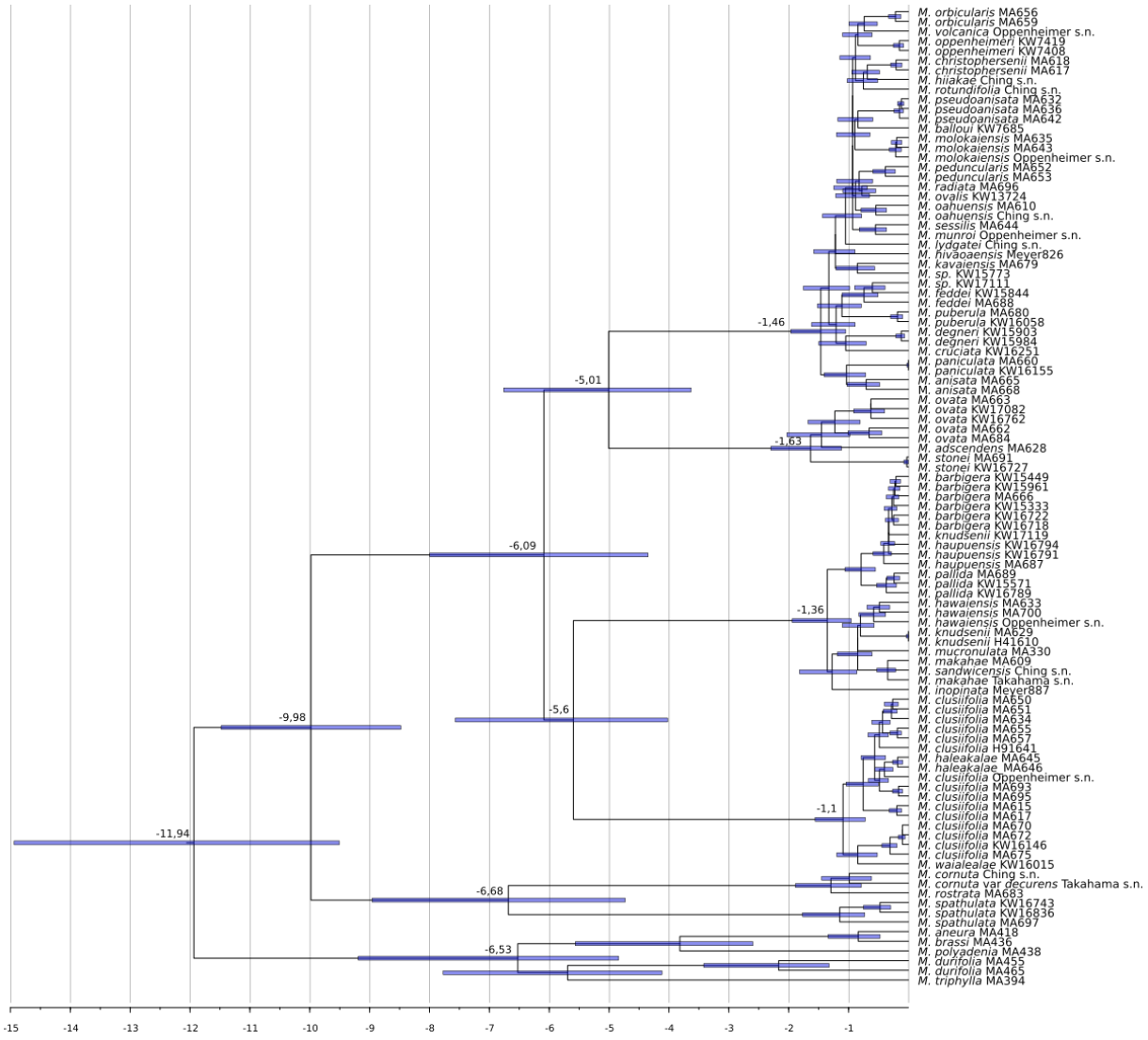

**Supplemental Figure 4.4** | Divergence times in Hawaiian *Melicope* as inferred by LSD2 with IQ-Tree in the min32 dataset. The Maximum Likelihood tree is shown with the credibility intervals of estimated ages displayed as light blue bars. For nodes corresponding the origin of major lineages mean ages are shown in addition to the 95% credibility interval.



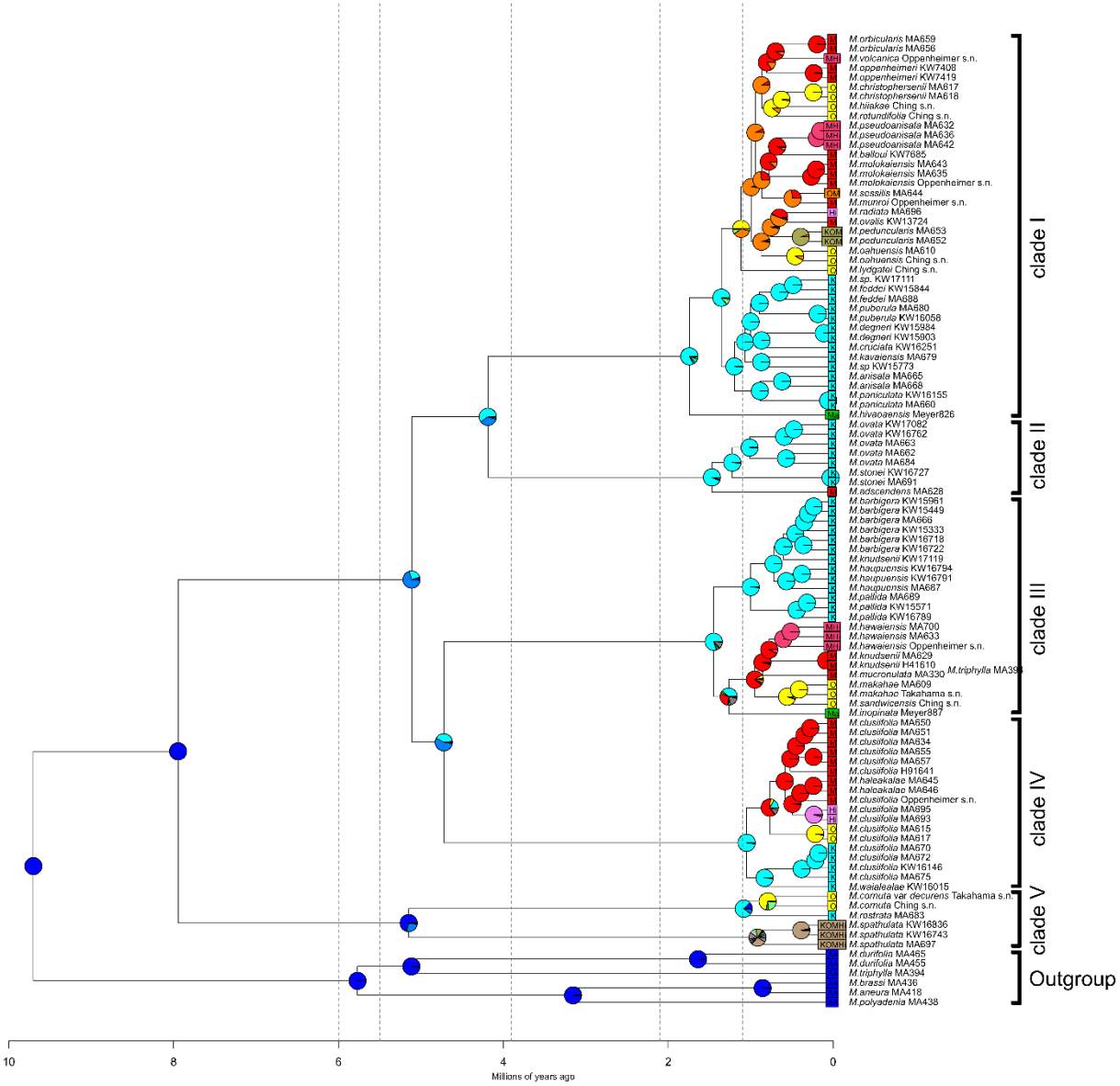

**Supplemental Figure 4.5** | Divergence times in Hawaiian *Melicope* as inferred by LSD2 with IQ-Tree in the min50 dataset. The Maximum Likelihood tree is shown with the credibility intervals of estimated ages displayed as light blue bars. For nodes corresponding the origin of major lineages mean ages are shown in addition to the 95% credibility interval.



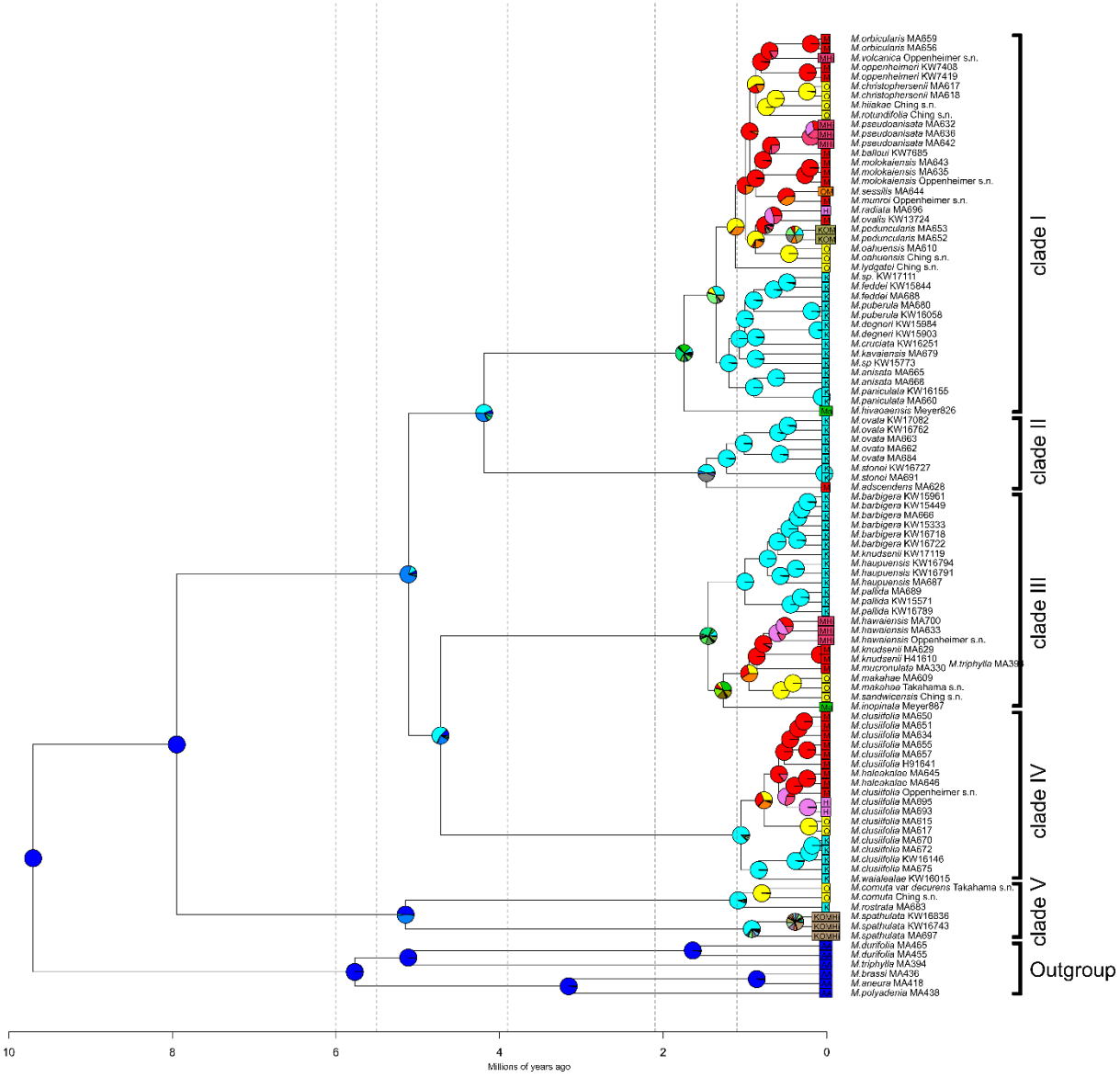

**Supplemental Figure 4.6** | Divergence times in Hawaiian *Melicope* as inferred by LSD2 with IQ-Tree in the min67 dataset. The Maximum Likelihood tree is shown with the credibility intervals of estimated ages displayed as light blue bars. For nodes corresponding the origin of major lineages mean ages are shown in addition to the 95% credibility interval.



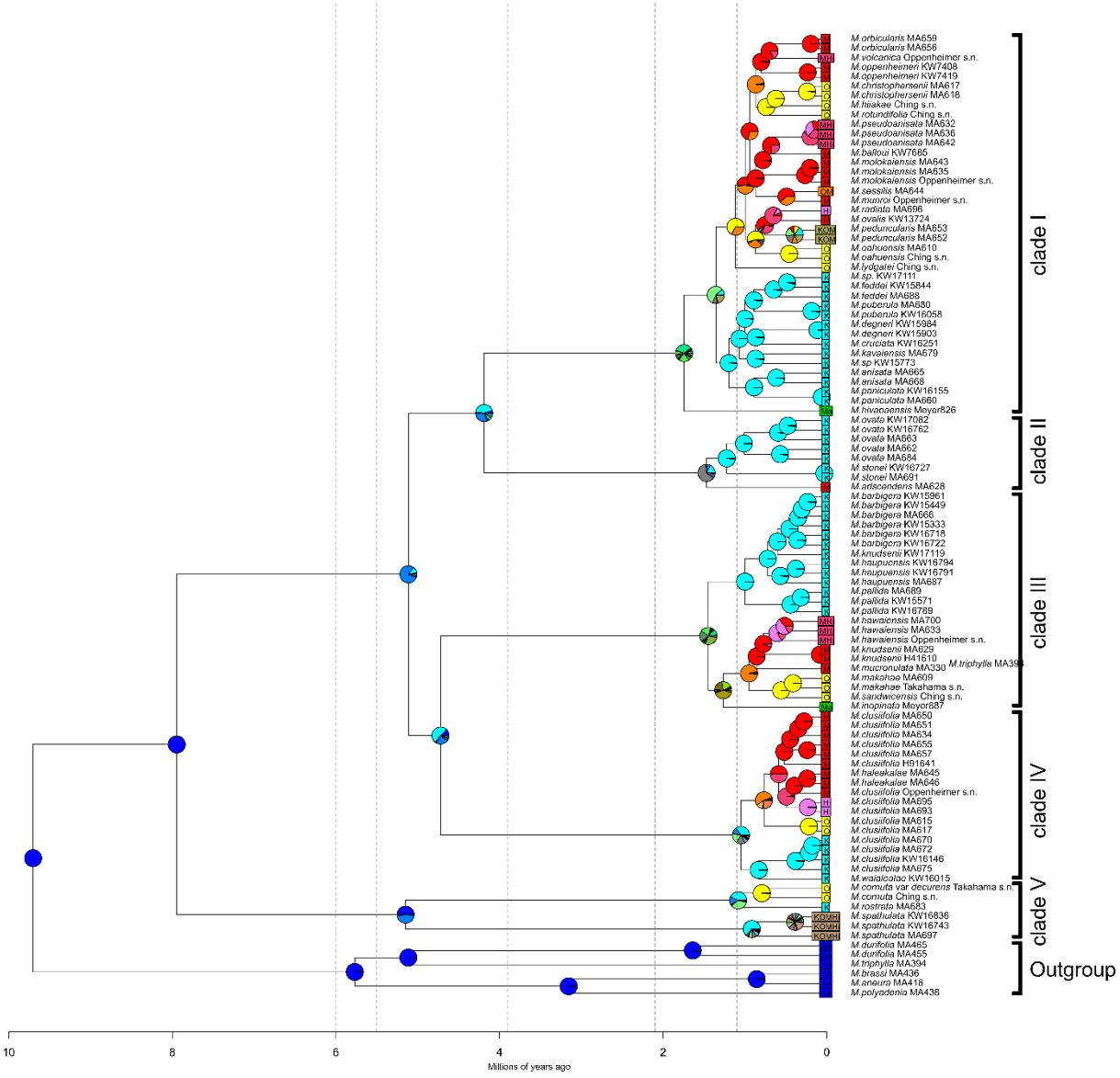

**Supplemental Figure 4.7** | Divergence times in Hawaiian *Melicope* as inferred by LSD2 with IQ-Tree in the min85 dataset. The Maximum Likelihood tree is shown with the credibility intervals of estimated ages displayed as light blue bars. For nodes corresponding the origin of major lineages mean ages are shown in addition to the 95% credibility interval.



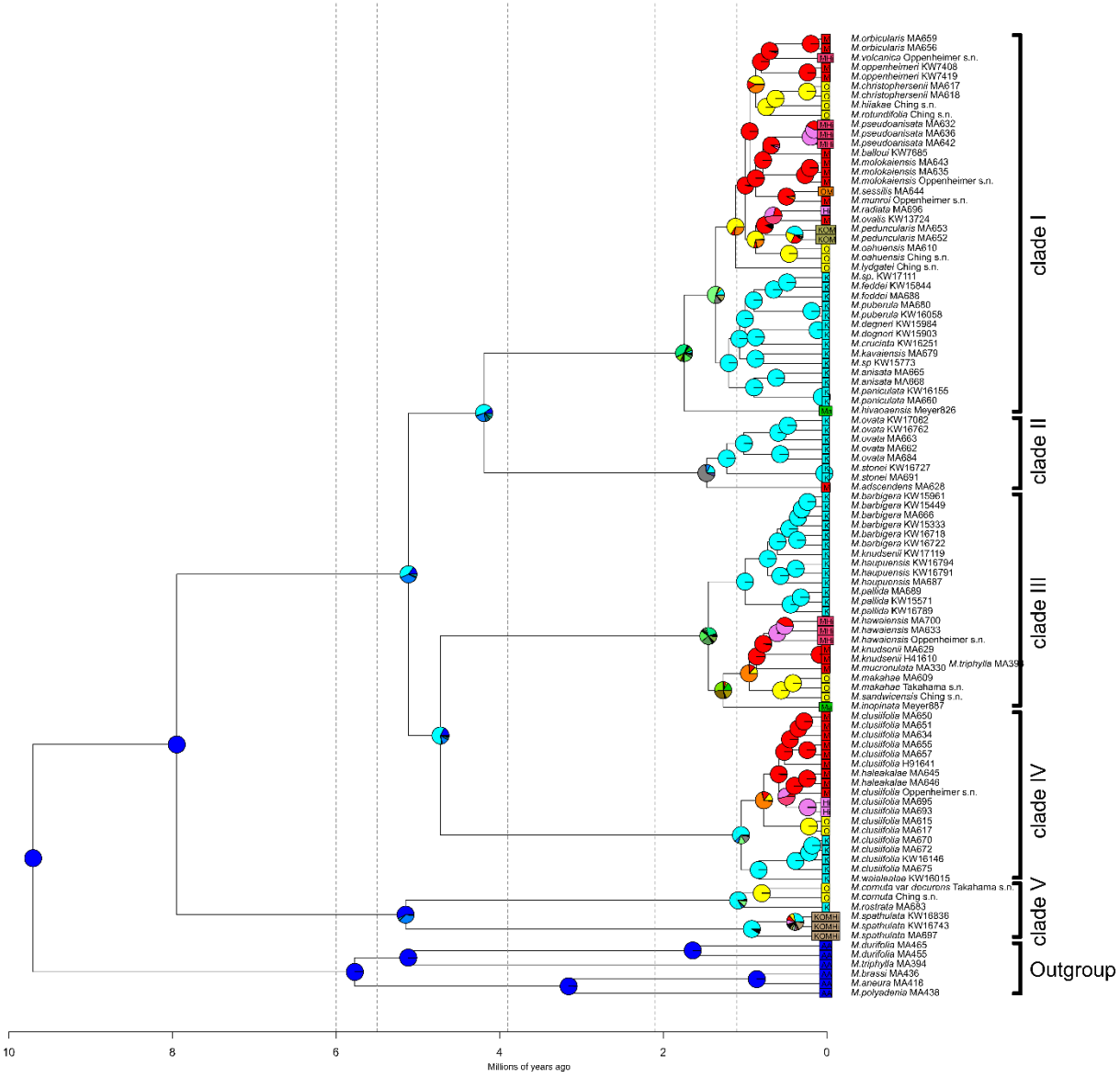

**Supplemental Figure 4.8** | Ancestral area reconstruction according to the BAYAREALIKE model implemented in BioGeoBEARS. The origin of the Islands is indicated in the phylogeny as vertical, dashed lines.



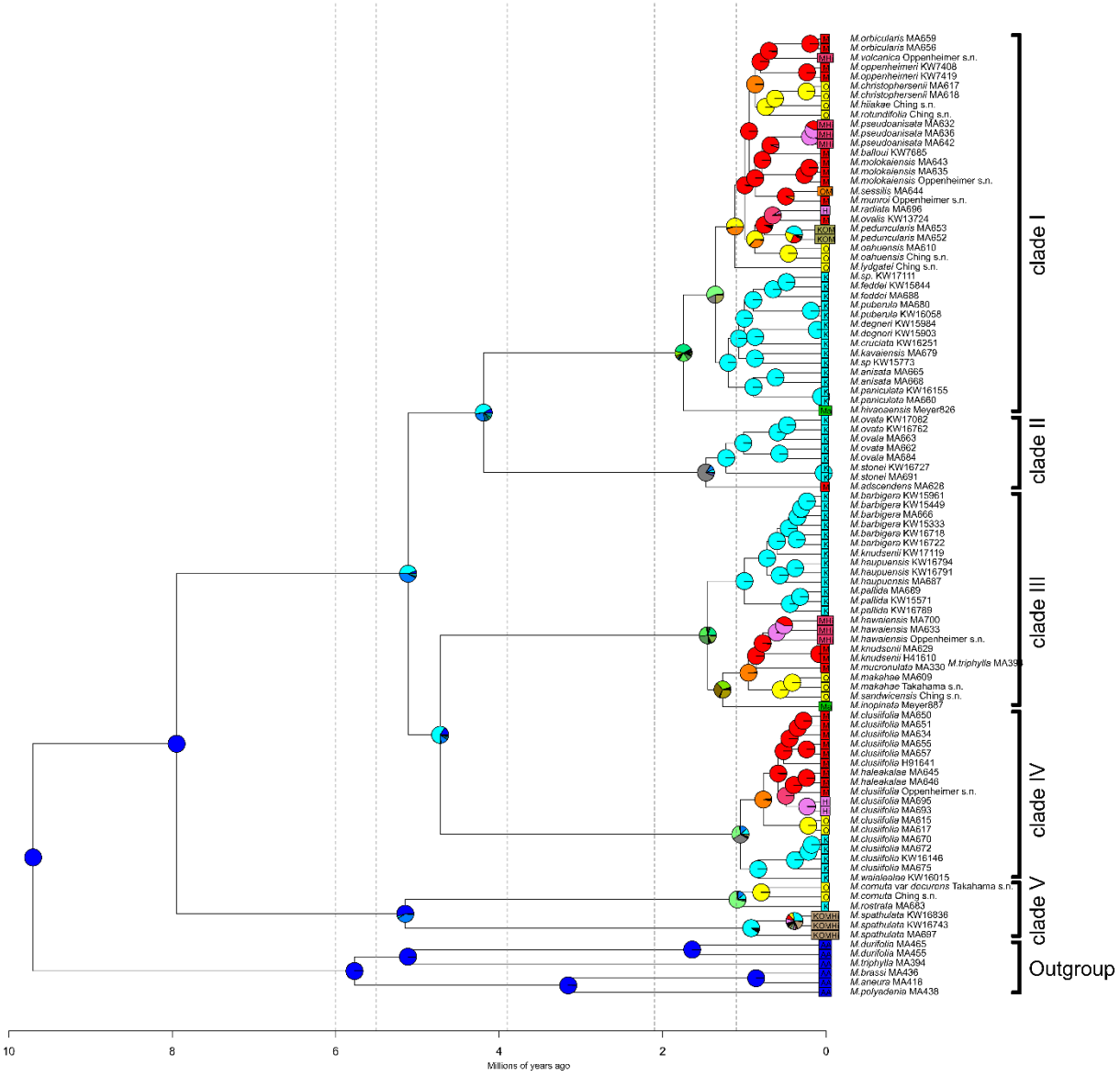

**Supplemental Figure 4.9** | Ancestral area reconstruction according to the DEC+J model implemented in BioGeoBEARS. The origin of the Islands is indicated in the phylogeny as vertical, dashed lines.



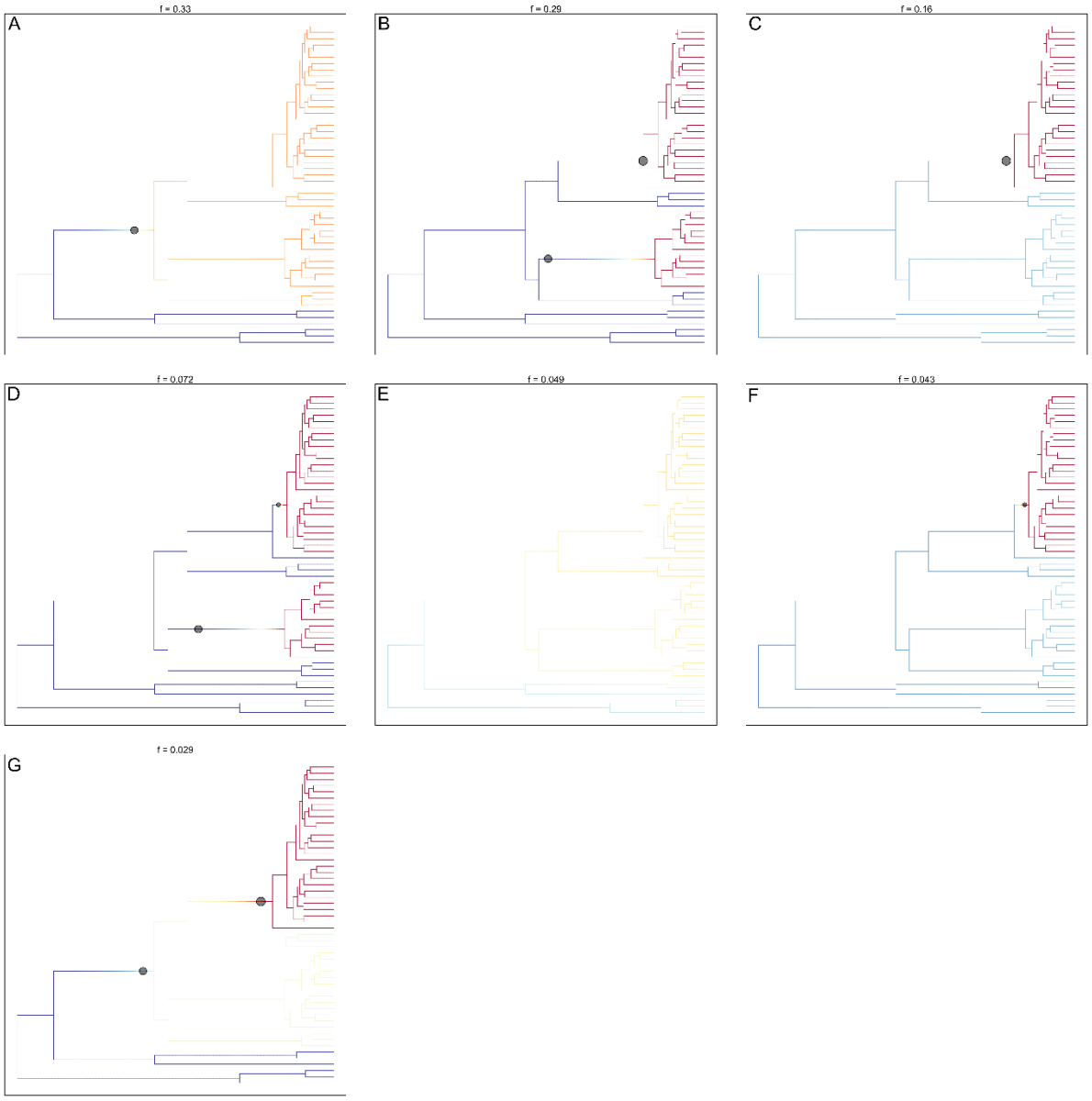

**Supplemental Figure 4.10** | Ancestral area reconstruction according to the DEC model implemented in BioGeoBEARS. The origin of the Islands is indicated in the phylogeny as vertical, dashed lines.



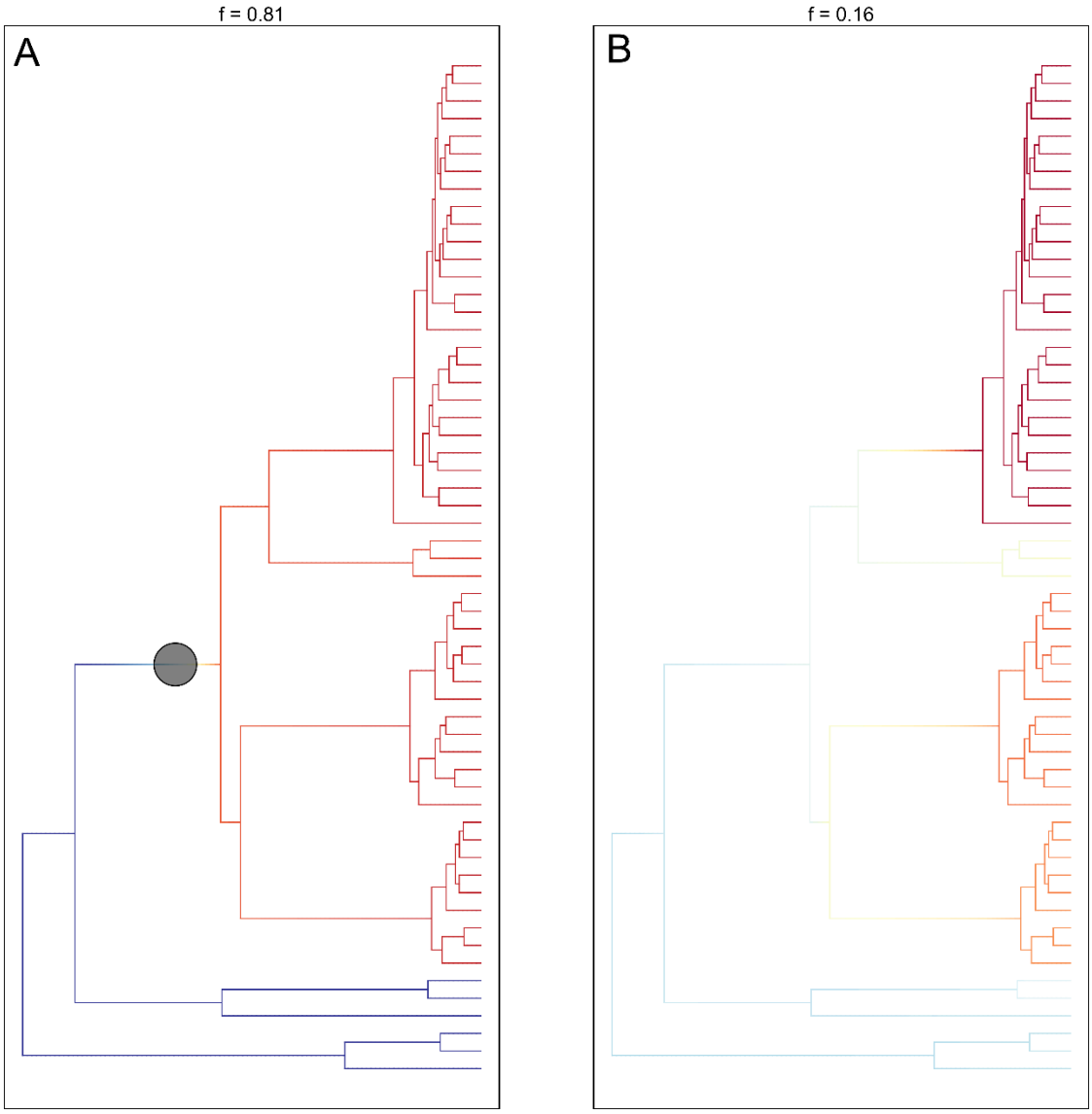

**Supplemental Figure 4.11** | Ancestral area reconstruction according to the DIVALIKE+J model implemented in BioGeoBEARS. The origin of the Islands is indicated in the phylogeny as vertical, dashed lines.




**Supplemental Figure 4.12** | Ancestral area reconstruction according to the DIVALIKE model implemented in BioGeoBEARS. The origin of the Islands is indicated in the phylogeny as vertical, dashed lines.




**Supplemental Figure 4.13** | Distinct shift configurations in the credible shift set resolved by BAMM analysis of the DA1 dataset ordered by frequency (f). Warm colours represent higher relative diversification rates, cold colours lower, relative diversification rates. Grey circles represent lineage-specific shifts in relative diversification rate.




**Supplemental Figure 4.14** | Distinct shift configurations in the credible shift set resolved by BAMM analysis of the DA2 dataset ordered by frequency (f). Warm colours represent higher relative diversification rates, cold colours lower, relative diversification rates. Grey circles represent lineage-specific shifts in relative diversification rate.

# 5. | Discussion

Na Pali Coast, Northern Kaua‘i. Photograph: Claudia Paetzold

## 5.1 Phylogeny and Spatio-temporal evolution of Hawaiian *Melicope*

Elucidating relationships in Hawaiian *Melicope* has been challenging in the past. Datasets generated by Sanger-Sequencing of a small number of genomic regions did not contain sufficient variation to resolve the majority of relationships (Harbaugh et al., 2009; Appelhans et al., 2014a, 2014b). The RAD-seq datasets herein are several magnitudes larger compared to previous efforts both in the number of base pairs and in the number of informative sites (Table 3.2; Harbaugh et al., 2009; Appelhans et al., 2014a, 2014b) and offer an unprecedented resolution of species-level relationships in the lineage.

Divergence time estimation using the RAD-seq dataset (chapter 4) resulted in a crown age of 7.9 mya (6.7-12.9 mya) for the extant Hawaiian *Melicope* (Figure 4.1). This estimate confirms previous results dating the origin of the lineage prior to the rise of the current high islands (Appelhans et al., 2018b). Consequently, the ancestor of the radiation most likely colonized either Necker (ca. 10.3 mya), the Twin Banks (9.6 mya) or, less likely, Nihoa (ca 7.5 mya) (Garcia et al., 2015). The colonizer originated from the Australasian region and is hypothesized to have arrived via stepping stone dispersal of other Pacific Island systems, though the resolution of relationships is not sufficient to support that claim (Appelhans et al., 2018b; Price and Wagner, 2018).

Chromosome numbers for Hawaiian *Melicope* are uniformly n = 18 or 2n = 36 (chapter 2), with only one reported deviation in extant species (Guerra, 1984). These values represent the base chromosome number in the entire Rutacean subfamily Amyridioideae dating back to an ancient polyploidization event at least 70 MA (Appelhans et al., 2012). There is no evidence to suggest that the ancestor of the Hawaiian lineage was a recent polyploid.

The colonization of the archipelago was seemingly not followed by immediate large-scale diversification. The established colonizer persisted for a period of ca. 2 MA in the archipelago but did not radiate, before eventually colonizing Kaua‘i. Lag times between colonization and onset of diversification have been connected to decreased initial competitiveness due to the low population numbers during the establishment of new arrivals (Gillespie, 2004). Yet the majority of island adaptive radiations are characterized by an early burst of speciation soon after colonization (e.g. Baldwin and Sanderson, 1998; Lindqvist et al., 2003; Gillespie, 2004). The exceptions to that early burst pattern are the lineages that colonized the archipelago prior to the emergence of the current main islands, Hawaiian Lobeliads (Givnish et al., 2009) and drosophilids (Russo et al., 1995) and now apparently *Melicope*. Considering that the leeward refugial islands were short-lived, low in elevation and already in decline at the time of the *Melicope* arrival (Clague, 1996; Price and Clague, 2002), the majority of

ecological niches were likely saturated. Consequently, the population size of the established colonists might have remained small until colonization of the young Kaua‘i provided ecological opportunity to diversify.

Extant Hawaiian *Melicope* are divided into five, fully supported main clades (Figure 3.2). Ancestral area reconstruction in chapter 4 suggests that Kaua‘i was colonized twice; once by the ancestor to clade V (*Platydesma*) and a second time by the ancestor to clades I-IV (Figure 4.2). That taxon subsequently diverged into four lineages comparatively rapidly within less than 1 MA after arrival (Figure 4.2).

The relationships of clade III are incongruent across different RAD-seq datasets (Figure 3.2, Supplemental Figures 3.1-3.4) with quartet sampling QD scores indicating ILS as a less likely source than horizontal gene transfer (Figure 3.3; Pease et al., 2018). Partitioned *D*-statistics tests (Eaton and Ree, 2013) with each of the clades I-IV designated as donor lineages in all possible combinations were conducted to identify signals of ancient introgression events (chapter 3). The results were statistically significant for $D_1$ and $D_2$ in all performed tests, while $D_{12}$ produced significant results exclusively in tests with clades III and IV designated as donor lineages (Table 3.3). This indicates that two separate introgressive hybridization events occurred between these four ancestral species, one involving the ancestors to clades I and IV, while the second involved the ancestors to clades II and III (Figure 3.2). Significant values for $D_{12}$ represent the presence of alleles by the progenitor of clades I and II, which the two descending species shared and subsequently introduced into the ancestral species of clades III and IV, respectively. The introgressed information might even have resulted in the apparent morphological connection between the clades. All species in clades I and IV have syncarpous fruits, while clades II and III are characterized by apocarpous fruits (Stone et al., 1999). In general, some caution is warranted as *D*-statistics tests are sensitive to confounding signals from multiple introgression events due to phylogenetic non-independence of tests (Eaton et al., 2015). In addition, I could not sample *M. elliptica* for my thesis, which was resolved as an independent lineage and sister to clade I in a previous analysis (Appelhans et al., 2014b). As this taxon is suspected to be involved in hybridization events itself (Price and Wagner, 2004), incorporating the species is crucial to exclude the possibility of the results inferred herein resulting from a ghost-lineage effect.

As currently standing, the introgressive hybridization events are associated with the time frame of the subaerial shield building of Kaua‘i (Figure 1.1; Clague and Sherrod, 2014). The older islands in the mountain chain, Necker, Twin Banks, and Nihoa, were eroded substantially or completely at this time, while Kaua‘i experienced the phase of most active volcanism, accompanied by lava flows, earthquakes, and landslides, which might have caused tsunamis (Moore et al., 1989; Price and Clague, 2002;

Clague and Sherrod, 2014; Denlinger and Morgan, 2014). Catastrophic events have been suggested to promote hybridization events though the exact reasons are unknown (Stuessy et al., 2014). These events might force species to survive in refugial, suboptimal habitats thus increasing close proximity of each other.

The divergence of the five main lineages was followed by a period of little apparent diversification lasting 3 MA. The currently recognized 54 species evolved during a phase of rapid diversification in the last 1.5 MA (Figure 4.1). The burst in divergence was accompanied by the colonization of the younger islands in the archipelago, and the Marquesas Islands (Figure 4.2).

The estimated diversification rates for Hawaiian *Melicope*, especially during the last <2 MA (Figure 4.3), render the previous interval of seemingly low divergence unlikely. Instead, the "handle-and-broom" shape of the tree with long basal branches and bushy tips characterizes a high extinction fraction, i.e. a high rate of extinction relative to speciation (Crisp and Crone, 2009). The estimated rates of species turnover indicate that the majority of diversity in the lineage is extinct and the extant species descend from a small number of surviving lineages. As already mentioned, the Hawaiian Islands are geologically highly active regularly experiencing large scale catastrophic events (Moore et al., 1989; Price and Clague, 2002; Denlinger and Morgan, 2014). Changing sea levels during glacial cycles repeatedly reshaped or discarded entire habitat zones (Price and Elliott-Fisk, 2004) while tropical cyclones increased frequency and intensity during warming phases (Montaggioni, 2005; Fedorov et al., 2010). All of these factors combined with changing ecological opportunity driven by island ontogeny (Whittaker et al., 2010) are liable to cause substantial amounts of extinction in lineages with high degrees of single-island endemism and species adapted to narrow ecological niches, like Hawaiian *Melicope* (Stone et al., 1999). Unfortunately, we currently lack evolutionary models that can adequately estimate extinction from genomic data alone (Sanmartín and Meseguer, 2016).

The overall biogeographical pattern in Hawaiian *Melicope* seems to be represented by a per-clade progression rule (Wagner and Funk, 1995). In each clade, islands were colonized in the order of their age (Figure 4.2). However, the pattern does not reflect the younger islands becoming successively available. Both O'ahu and Maui Nui were already in post-shield stage when they were colonized. It rather reflects the relative proximity of neighboring islands as destinations for dispersers. *Melicope*, including the Hawaiian radiation, shows adaptions to bird dispersal (Stone et al., 1999; Hartley, 2001). At least in prehistoric times, the Hawaiian Islands had a rich endemic avifauna providing a wide palette of possible dispersers (Pratt et al., 2009). However, most forest birds have adapted to island life by evolving a sedentary behavior and a

reduced, low-cost flight apparatus in response to competitive release. The majority of native bird species are single-island endemics that lost their inclination and ability to travel larger distances and thus find the water body between the main islands a barrier (Pratt, 2009). Hence, the colonization of subsequently younger islands in the chain is more likely a function of distance and might be contingent on non-standard vectors or migration, e.g. birds blown away by tropical storms.

In most cases, the colonization of a new island resulted in a speciation event (Figure 4.2). However, the majority of species are the product of intra-island diversification subsequent to colonization and in one case from diversification following back-colonization of an older island (clade I, Figure 4.2). In general, however, back-colonizations are infrequent and not associated with speciation events, rather they result in a species being more widespread.

Presently, it seems no general pattern describes patterns of divergence in Hawaiian *Melicope*. Rather every major clade seems to represent a unique pattern, which in turn is associated with a specific diversification rate. Clade V (*Platydesma*) shows the lowest diversification rate in the entire lineage. The taxa in this clade are characterized by perfect flowers, whereas they are functionally dioecious in all other clades (Stone et al., 1999; Appelhans et al., 2017). The shift to hermaphroditism occurred only once and likely relatively early in the history of the lineage during the relegation to refugial islands (Figure 4.1). When population sizes were low, the shift to hermaphroditism might have represented a short-term advantage but would have possibly made the population susceptible to inbreeding-depression (Sakai et al., 1995), which resulted in low diversification rates. A shift from dioecy to perfect flowers following colonization of Hawaii has also been noted for *Rhus* L. (Anacardiaceae; Sakai et al., 1995), which is represented by only one endemic species that colonized the archipelago around 13.5 mya (Yi et al., 2004). In either lineage, the shift to hermaphroditism might have enabled the survival of bottlenecks at the cost of future diversification.

Both, clade II and IV show comparatively low diversification rates and species numbers. Clade IV might represent a widespread, morphologically variable ecological generalist, with *M. clusiifolia* occupying a wide range of habitats and elevational ranges and the remaining two species more specialized (Figure 4.4; Stone et al., 1999). A generalist strategy would result in low diversification rates and raise the question of how gene flow is maintained across the islands. Caution is required for this assessment, as the clade requires taxonomic revision (see below). Species in clade II are exclusively adapted to mesic habitats and have mostly small elevational ranges. However, this pertains also to clade III, which shows comparatively high rates of diversification. Quartet sampling QD scores (Figure 3.3), as well as morphological observations in the field (K. Wood, personal communication), indicate

that hybridization events occur with some frequency. Repeated hybridization events can increase gene flow between species or result in hybrid speciation, and thus increase diversification rates.

Finally, clade I is the most species-rich clade in the entire lineage and shows the highest diversification rates (Figure 4.3). One subclade comprises approximately half the diversity and is strictly endemic to Kauaʻi while the second subclade comprises all species occurring on the younger islands (Figure 4.2). Species in this clade show a wide range of habitat preferences, including a high frequency of habitat shifts and elevational range expansions/contractions between species (Figure 4.4, Table 4.2). There are also some instances of species-to-species-matching, where species of different islands occupy similar habitats and show similar morphology; e.g. *M. molokaiensis, M. oahuensis, M. hiiakae, M. ovalis* (Figure 4.4; Stone et al., 1999). Diversification in this clade might be represented by a pattern of repeated specialization to small, open ecological niches with species adapting to different habitats and height ranges.

## 5.2 Taxonomic implications of RAD-seq phylogeny

Phylogenetic relationships based on RAD-seq datasets highlights the necessity for a taxonomic revision in Hawaiian *Melicope*. Of the five informally used Stone's sections in the lineage, I only could confirm only two as monophyletic groups, the former genus *Platydesma* and Stone's section *Pelea* (Figure 3.2). Considering the distinctiveness of the *Platydesma* lineage, especially the palmoid habit, hermaphroditic flowers and staminal tube (Wagner et al., 1999b) and its early divergence (Figure 4.1), the group should receive formal recognition. The three remaining Stone's sections, which comprise the vast majority of the lineage are each non-monophyletic. Stone's sections *Cubicarpa* and *Megacarpa* are paraphyletic to each other and all taxa belonging to either were resolved in clade I with no apparent pattern to the divergence between the two fruit morphologies (Figure 3.2). The main characteristic delimitating the two groups is the degree of carpel connation, which is described as "up to 2/3 of their length" for *Megacarpa* and "nearly to completely" for *Cubicarpa* (Stone et al., 1999). The results I present in chapter 3 suggest the separation between these two Stone's sections is artificial and they should be merged as carpel connation seems a continuous character rather than two discrete states. Stone's section *Apocarpa* is divided into two different lineages with the majority of taxa resolved in clade *Apocarpa 1* (clade III, Figure 3.2) while only three species included herein comprise the *Apocarpa 2* clade (clade II, Figure 3.2). These three species do share some morphological traits, noticeably the sprawling, shrubby habit (except *M. stonei*), a glabrous exo- and endocarp and the few-flowered inflorescences.

Additionally, these three taxa occur exclusively in mesic habitats (Stone et al., 1999; Wood et al., 2017). However, neither one of these characteristics, nor their combination, is unique to this lineage. Further studies of the characteristics of all *Apocarpa* species are required to identify trait patterns characterizing these clades. These research efforts need to include samples of *M. elliptica*, which I could not sample for my thesis. *M. elliptica* represents the type species of Stone's section *Apocarpa*. Results from Sanger sequencing suggest the species represents a third apocarpous lineage within the Hawaiian *Melicope* and may be closely related to clade I (Appelhans et al., 2014b). Price and Wagner (2004) hypothesized that the individuals of *M. elliptica* might represent hybrids, as they grew near several species of clade I. Inclusion of *M. elliptica* samples is required to determine, whether one of the two *Apocarpa* clades resolved will have to be renamed or even both. The revision of the lineage must also necessarily include the Marquesan species of *Melicope*. My results confirm previous results suggesting that the Hawaiian Islands are the origin of two independent colonization events to the Marquesas Islands, one resulting in a local adaptive radiation of seven species (Hartley, 2001; Appelhans et al., 2018b) represented here by *M. hivaoaensis* (clade I; Figure 3.2). Inclusion of these altogether eight Marquesan species in further research is necessary to conclusively identify phenotypic traits characterizing clades in Hawaiian *Melicope*.

Some species require taxonomic revision as well. In this thesis, I sampled 84% of the 54 currently described Hawaiian species, 24 of which I could include with multiple samples. Of these 24 multi-sampled species four were resolved as non-monophyletic: *M. clusiifolia*, *M. haupuensis, M. knudsenii*, and the *M. feddei*. *Melicope knudsenii* is the only species in the entire lineage occurring on non-neighboring islands, Kauaʻi and Maui (Stone et al., 1999). Previous phylogenetic efforts in Hawaiian *Melicope* provided evidence that the populations of *M. knudsenii* actually represent up to three distinct species (Appelhans et al., 2014b). Subsequently, one population occurring on the island of Kauaʻi was described as a new species (Wood et al., 2017). In this thesis, I could confirm that the populations of *M. knudsenii* on Kauaʻi and Maui represent two distinct taxa as indicated in Appelhans et al. (2014b). The individuals are all members of clade III, with the Maui populations sister to *M. hawaiensis* and the populations from the type location on Kauaʻi type population sister to *M. barbigera*. Either relationship receives high statistical support (Figure 3.3). However, the results of quartet sampling indicate the possibility of reticulate evolution in the Maui population, possibly involving *M. barbigera* (Figure 3.3). This could result in the reconstructed close relationship between the two taxa. Further studies are required to definitively elucidate the relationship of the Maui population of *M. knudsenii* before it can be resurrected as a separate species during taxonomic revision.

*Melicope haupuensis* is polyphyletic (clade III, Figure 3.2) and the only sample in my thesis with incongruent constellations among the individuals based on different

datasets (Figure 3.2, Supplemental Figures 3.1-3.4). Several individuals were observed in the field showing intermediate phenotypes between *M. haupuensis* and *M. barbigera* (personal observation K. Wood), which might be of hybrid origin. The results of quartet sampling support the indication of extensive reticulate evolution, as several nodes show skewed discord (Figure 3.3; Pease et al., 2018). However, the taxon sampling in my thesis is not sufficient to address this question further. A thorough population-level sampling of the taxa in question, including *M. knudsenii* as sister to *M. barbigera,* will be required to address the possibility, extent, and direction of horizontal gene transfer in this clade.

Finally, *M. clusiifolia* is paraphyletic with respect to both *M. haleakalae* and *M. waialealae* (Figure 3.2). *Melicope clusiifolia* is the most widespread species in the entire lineage, occurring on all main islands, and characterized by large morphological plasticity. The most recent taxonomic treatment synonymized all previous attempts at subdividing this species (St. John, 1944; Stone, 1969) considering the variability to represent continuously varying rather than discrete characters. In the same spirit, *M. haleakalae*, differing from *M. clusiifolia* mainly by persistent sepals and *M. waialealae* differing mainly in leaf shape (Stone et al., 1999) might be considered representing further morphological variability. Thus, Stone's section *Pelea* might only comprise one highly plastic species. On the other hand, the relationships of these three taxa might indicate speciation in progress. In this case, the deep nesting of *M. haleakalae* within *M. clusiifolia* would represent a case of progenitor-derivative-speciation (Crawford, 2010). Finally, there is a clear geographical signal for the diversification in this group (Figure 4.2), which would indicate limited gene flow among islands. Extensive sampling of populations of Stone's section *Pelea* and subsequent morphological and molecular research will be required to definitively elucidate the taxonomy of this clade.

*M. wawraeana* might be the most cryptic taxon within the lineage. My sampling included two individuals from Kauaʻi that correspond closely, but not entirely to the type-morphology of the species. The core populations of the species are situated on Oʻahu and show some degree of morphological variability. In "non-core" populations this variability is extended further and resulted in synonymization of several species names (Stone et al., 1999). However, one of these synonyms, *M. hiiakae,* was tentatively resurrected by Hartley and Stone (1989) and treated as a separate species ever since (Wagner et al., 1999a; Imada et al., 2011; Wood et al., 2016). My results provide evidence for this resurrection as the taxon is resolved as sister to *M. christophersenii* in clade I and not closely related to the two *wawraeana*-like individuals (Figure 3.2). This indicates that other non-core populations might represent different taxa as well. The *M. wawraeana*-like populations from Kauaʻi sampled herein are cryptic as well. One of these *M.* sp. (specimen KW17111) is

resolved in a monophyletic clade with the individuals representing *M. feddei* (Figure 3.2). This indicates that at least one of the populations from Kaua'i might actually represent *M. feddei* instead of *M. wawraeana.* The second *M.* sp. (specimen KW15733) is closely related to *M. kavaiensis* and both taxa show rogue behavior related to the Marquesan sample *M. hivaoaensis* (Figure 3.3, Supplemental Figures 3.1-3.4), which might reflect a hybridization event. In addition to putative hybridization events in the *M. wawraeana* complex (chapter 3, Imada et al., 2011), there is also evidence for polyploidy in some individuals (Guerra, 1984). Further detailed morphological, molecular and karyological studies are required to determine the status and evolutionary history of this potentially artificial taxon (Stone et al., 1999).

Finally, the sampling also included two species with samples corresponding to both, the type morphology and deviating morphotypes, *M. barbig*era and *M. ovata*. The deviating morphotypes are resolved as monophyletic sister groups to the samples representing the type morphology, which are also resolved as monophyletic (Figure 3.2). This again indicates speciation including phenotypic changes. Further studies are required to ascertain whether the speciation events are in process or sufficiently advanced to recognize these deviating morphotypes as separate species or subspecies.

### 5.3 Hawaiian *Melicope* and island adaptive radiation

The results presented in the previous three chapters offer the opportunity to evaluate the existing theories about the evolution of adaptive radiations on oceanic islands. We will find that Hawaiian *Melicope* support several hypotheses and negate others.

The ancestor to Hawaiian *Melicope* did not show any trait hypothesized to characterize successful island colonizers. Instead, the taxon showed traits characterizing already established island lineages. Woodiness is the ancestral state of the entire Rutaceae family, which comprises only a small number of herbaceous taxa (Kubitzki et al., 2011). *Melicope* is an exclusively woody genus with no herbaceous representatives (Hartley, 2001). Several hypotheses have been proposed for the evolution of secondary woodiness in island settings. According to the competition hypothesis, herbaceous colonists gain an advantage when they are growing taller than their competitors eventually evolving secondary woody structures to support their ever taller growth (Darwin, 1859). The longevity hypothesis, suggested by Wallace (1878), states that woodiness provides a longer lifespan, which in turn increases the chances for reproduction when pollinators might be scarce. Finally, the moderate insular climate hypothesis states that, since island climates are typically more mild and moist compared to the climate in the respective source areas, plants can grow throughout the whole year and thus woodiness is promoted (Carlquist, 1974). However, a woody colonist would profit from the advantages of the habit

right away, instead of evolving it presumably with a time delay and thus even having a competitive advantage. Frequent shifts to woodiness are likely owed to the high frequency of herbaceous colonizers profiting from the mentioned advantages.

Herbs often have a higher dispersal ability compared to trees and thus represent the majority of island colonizers (Darwin, 1859; Carlquist, 1966a, 1966b, 1966c). Based on observations of Pacific island floras, Carlquist (1966a, 1966b, 1966c) connected the higher dispersability of herbaceous taxa to lightweight, small seeds that are easy to swallow or transport, and the germination typically occurring in open areas with direct access to sunlight. On the other hand, trees generally show larger seed sizes, because they comprise the embryo and additional nutrient tissues enabling growth in a forest understorey with limited light exposure after germination (Carlquist, 1966a). However, the presence of nutrient tissues in herbaceous and woody taxa and its impact on dispersal to oceanic islands has never been empirically studied. Establishment of colonizers is characterized by a loss of dispersibility putatively in order to avoid loss of propagules at sea or in ecologically unfavorable regions on islands (Darwin, 1842; Carlquist, 1966b, 1966c, 1974). The phenomenon is best investigated in Pacific Asteraceae (Carlquist, 1966b); Hawaiian *Bidens*, for example, show the reduction of attaching structures on seeds thus reducing their dispersal range (Carlquist, 1966c). Many other island plants have increased seed or fruit sizes so as to be swallowed less easily or transported only over short distances due to their weight (Carlquist, 1966b, 1966c; Price and Wagner, 2004). For *Melicope* detailed, empirical studies regarding dispersal ranges are lacking. While seed sizes are highly variable in *Melicope*, there is no observable trend for a reduction of seed size in Hawaiian *Melicope* or *Melicope* island species compared to mainland representatives of the genus (Stone et al., 1999; Hartley, 2001). *Melicope* is adapted to bird dispersal, yet the Hawaiian avifauna is characterized by a high degree of endemicity and many species balking at habitat barriers (Pratt, 2009). If such restricted-range bird species serve as dispersers for Hawaiian *Melicope*, seed size in terms of dispersal ability might have not been under strong selective pressure.

Dioecy is a common feature in island floras with e.g. 12-13% of the New Zealand flora (Webb and Kelly, 1993) and at least 14.7% (Sakai et al., 1995) of the Hawaiian flora dioecious, compared to only 4% of Angiosperms worldwide (Yampolski and Yampolski, 1922). In *Melicope,* dioecy is present in two of the four sections, *Lepta* and *Pelea* (Hartley, 2001). The shift from hermaphroditism to dioecy occurred several times within the genus and characterizes several island lineages including the Hawaiian one, for which dioecy is the ancestral state (Hartley, 2001; chapter 2). However, the dioecious colonizer to Hawaiian *Melicope* is not necessarily an exception, as approximately 10% of the current species diversity in Hawaii descends from dioecious colonizers, while 31.8% of dioecious species evolved autochthonously

in the archipelago (Sakai et al., 1995). In New Zealand, only five shifts to dioecy were reconstructed in the native flora and most of the dioecious species descend from dimorphic ancestors (Lloyd, 1985). This suggests that the selective advantage of outcrossing in island systems is high enough to both, promote the shift to dimorphism in hermaphroditic colonizers (Baker, 1955), and to outweigh the disadvantages of being a dioecious colonizer requiring at least two individuals for establishment (Carlquist, 1974). On the other hand, strong evidence for self-incompatibility was only found in one Hawaiian radiation so far, the Hawaiian silversword/tarweed alliance (Asteraceae, Madiinae; Carr et al., 1986). Although not many Hawaiian lineages have been empirically tested for self-incompatibility, the few existing results suggest a lack thereof. Hence, the causality of the evolution and success of dioecy in island radiations is not yet established (Sakai et al., 1995). In the general terms of the island syndrome, we can surmise that it is characterized by a comparatively higher frequency of dioecy per se, but that a shift in breeding system, regardless of whether it occurs prior to colonization or autochthonously on the islands is not strictly required to derive successful radiations. A world-wide sister-lineage comparison revealed dioecy is connected to lower diversification rates on the family and genus level compared to monomorphism (Heilbuth, 2000). The genus *Melicope* represents a case where shifts in breeding system occur within a genus, and thus an organismic level explicitly not tested, yet. Further research is required to infer whether dioecy or self-incompatibility is characterizing either adaptive radiations as a whole or oceanic island lineages in particular and to establish causality.

Another feature of the island syndrome, especially in plants, is polyploidy. In the Hawaiian flora 88% of native plants are polyploid (Carr, 1998), for New Zealand the number ranges around 63% (Murray et al., 2005). In addition, several species-rich Angiosperm families are characterized by a whole-genome duplication (WGD) event at their origin, which is linked to the evolution of key traits, e.g. the composite flower in Asteraceae (Schranz et al., 2012). Considering that the Angiosperm radiation as a whole is characterized by two ancient polyploidization events (Jiao et al., 2011; Amborella Genome Project et al., 2013), polyploidy is suggested to represent a key feature characterizing major adaptive radiations (Schranz et al., 2012). Since all Angiosperms might be considered polyploid (Jiao et al., 2011), commonly only the latest polyploidization event in a taxons history is considered (Leitch and Bennett, 2004). There are two advantages to polyploidy, that promote adaptive radiation: heterosis (in case of an allopolyploid), gene redundancy, which allows neo- or subfunctionalization of duplicated genes or masking of deleterious mutations (Comai, 2005). In the danthonioid grasses, the frequency of long-distance dispersal events across oceanic barriers is significantly increased in polyploid lineages compared to diploid ones (Linder and Barker, 2014). The association reflects the

effects on establishment, rather than the likelihood of dispersal events *per se* (Linder and Barker, 2014). Polyploidy can disrupt self-incompatibility systems and thus increase the chances for successful establishment of a colonizer from a single propagule (Miller and Venable, 2000; Mable, 2004). The genomic plasticity conveyed by a fixed higher heterozygosity, heterosis and gene redundancy, may increase the evolvability compared to diploid taxa thus enabling subsequent adaption to a wide range of ecological niches (Doyle et al., 2008).

Chromosome counts and DNA content of Hawaiian and extra-Hawaiian *Melicope* (chapter 2) show a chromosome number of 2n=36 shared by all Hawaiian species, with the possible exception of one specimen of *M. wawraeana.* The 2n=36 chromosome configuration is ancestral to the entire subfamily Amyridioideae (Kubitzki et al., 2011; Morton and Telmer, 2014) and originates from a WGD event about 70 mya (Figure 2.1; Appelhans et al., 2012). The MCRA of Hawaiian *Melicope* was thus a palaeopolyploid that has likely not undergone an additional polyploidization event prior to the colonization of the islands. The DNA content in the nuclei of Hawaiian *Melicope* is equal to or even smaller than that of other Rutaceae species with n=9 (Figure 2.3), suggesting substantial diploidization (Dodsworth et al., 2016). Post-WGD diploidization was suggested to be crucial to counter the negative effects of a polyploid genome, especially dosage effects and selection barriers. Thus, the causal agent for the positive effects of WGDs may not the actual duplication event itself, but the downsizing following it (Dodsworth et al., 2016). This suggestion receives support from the observation that the majority of polyploidization events result in extinction rather than diversification (Mayrose et al., 2011; Arrigo and Barker, 2012) and provides an explanation for the apparent lag-time between a WGD event and the onset of diversification in lineages that prevail (Dodsworth et al., 2016).

The hypothesis that diploidization subsequent to polyploidization events conditions taxa for diversification raises the expectation that diversification rates be elevated in other amyridioid genera, characterized by palaeopolyploidy. *Melicope* and *Zanthoxylum* are the largest genera in Rutaceae, comprising ca. 230 species each (Kubitzki et al., 2011). *Zanthoxylum* has a wider distributional range, occurring in world-wide in tropical regions and extends to subtropical and temperate regions in Asia and North America (Beurton, 1994). Both genera comprise widely distributed, woody species, show shifts in breeding system between functionally unisexual and perfect flowers as well as similar fruits adapted to bird dispersal (Hartley, 2001; Kubitzki et al., 2011). In addition, chromosome numbers in *Zanthoxylum* range from 2n=32 to 2n=136 (-144) (Kiehn and Lorence, 1996; Kubitzki et al., 2011) showing extensive aneuploidy and dysploid variation (Stace et al., 1993) illustrating the effects of diploidization. Consequently, the diversification rates in *Melicope* and *Zanthoxylum*

should at least be equal if not higher in *Zanthoxylum*, considering the putatively additive effects of repeated WGD events and the wider distribution potentially increasing access to ecological opportunity. Surprisingly, diversification rates in *Melicope* are considerably higher than in *Zanthoxylum* (Appelhans et al., 2014b) in general and on the Hawaiian Islands. In the archipelago, *Zanthoxylum* is represented by a monophyletic group of only four morphologically diverse species with a crown-group age of 11.8 MA (6.9-17.5 MA) (Appelhans et al., 2014b, 2018a) compared to the 54 species in *Melicope* (Wood et al., 2016, 2017; Appelhans et al., 2017). In addition, many other polyploid, native Hawaiian lineages are represented by low species numbers as well and have not radiated (Carr, 1998; Wagner et al., 1999b).

This suggests that polyploidy *per se* is not necessarily the trait characterizing adaptive radiations. For only two Hawaiian Angiosperm radiations, the originating polyploidization events were reconstructed from molecular data; the colonists to Hawaiian mints and the Hawaiian silversword Alliance were neoallopolyploids; i.e. the result of an interspecific hybridization event in the respective source areas (Barrier et al., 1999; Lindqvist and Albert, 2002; Lindqvist et al., 2003; Roy et al., 2015). The majority of the positive effects of polyploidization pertain especially to allopolyploids, e.g. heterosis and increased fixed heterozygosity (Comai, 2005). So far, plant adaptive radiations resulting from polyploidization events have generally been linked to a hybrid origin (reviewed in Schranz et al., 2012).

Hybridization events as a putative catalyst to adaptive radiation have garnered increased attention in the era of genomic studies (e.g. Seehausen, 2004, 2013; Abbott et al., 2013; Marques et al., 2019). Partitioned *D*-statistics and quartet sampling on RAD-seq data in chapter 3 indicate several putative introgressive hybridization events in Hawaiian *Melicope* (Figure 3.3, Table 3.3). Past and present reticulate evolution on the Hawaiian Islands has also been shown for *Cyrtandra* (Johnson et al., 2019; Kleinkopf et al., 2019) and *Schiedea* (Willyard et al., 2011). Interspecific hybridization is quite common with an estimated 10% of animal and 30% of plant species regularly hybridizing (Mallet, 2005). Besides hybrid speciation (Rieseberg, 1997), one frequent outcome of hybridization is the introgression of alleles from one species into another by backcrossing. This process might combine selectively favored alleles that arose under different ecological conditions (Abbott et al., 2013). Mutation events are rare; statistically, there are only $10^{-8}$ – $10^{-9}$ per base and generation (Abbott et al., 2013). Speciation events during adaptive radiation often occur too rapidly to evolve novel adaptions by mutation. Instead, adaptive traits are mostly the result of existing genetic variation (Barrett and Schluter, 2008). Hybridization can provide an abundant source of genetic variation and numerous different combinations of adaptive alleles. Further research regarding the effect of hybridization on adaptive radiation is required to identify its significance. Genomic studies have provided evidence for hybridization in many lineages, where there was no or limited

indication for it before. Hawaiian *Melicope* are illustrating that issue, as analysis of few genomic loci did not provide an indication for hybridization (Appelhans et al., 2014b) but analysis of thousands of genomic loci (chapter 3) did. However, combinatorial analyses are necessary to provide evidence for introgressed alleles being under selection and for providing an adaptive advantage (Suarez-Gonzalez et al., 2018). This requires extensive genomic information but has been accomplished in *Populus* (Suarez-Gonzalez et al., 2016) and the Lake Victoria cichlid fishes (Meier et al., 2017).

### Patterns of Diversification

Diversification rates in Hawaiian *Melicope* are generally high (Table 4.2), comparable to those inferred for other adaptively radiating clades, e.g. orchids (Givnish et al., 2015). *Melicope* species in the Hawaiian islands are adapted to four habitat types: dry, mesic, wet forests and bog habitats (Stone et al., 1999). Each habitat type has a significant impact on species diversification possibly reflecting its average lifetime on an island. Mesic and wet habitat types represent the majority of habitats on the archipelago. Within them, speciation rates are comparatively low, but extinction rates are even lower by an order of magnitude, resulting in a positive net diversification and thus a steady gain of species numbers in the lineage in these habitats. In addition, the character transition rates are highest for a shift from either dry or wet to mesic habitats (Table 4.2), which is likely due to a large extent to the frequent shifts for habitat adaption in species of clade I (Figure 4.4). In more extreme habitats on the islands, dry ranges, and bogs, the inferred net diversification rate is negative, resulting in a net loss of species (Table 4.2). However, this diversification rate results from elevated speciation rates, which are up to 40x higher than for mesic and wet habitats. Consequently, extinction rates for the two extreme habitat types are considerably high as well (Table 4.2). These rates represent a high amount of species turnover associated with bog and dry habitat ranges, and possibly a shorter average time of prevalence per species due to the transitory nature of these habitat types, which are strongly affected by island orogeny and climate changes.

Mesic and wet ranges come into existence comparatively early in the existence of an island and persist in large areas for almost all its lifetime. That is not to mean, that these ranges are static; they shift with regards to area, dissection and elevational range due to the island's growth, subsidence, glacial cycles, and wind currents. However, these areas are present in substantial amounts for nearly the entire lifespan of the island. Extreme habitat types are impacted by these environmental and orogenic factors more substantially. Bog habitats do not initially exist on young islands. The formation of bogs requires a long-term sequence of events; the primary succession of rainforest, erosion, and depletion of nutrients from the soil; combined with steady high amounts of rainfall. Both, their formation and their persistence are

vulnerable to decreases in water supply (Mueller-Dombois and Boehmer, 2013). Yet, the islands have experienced repeated periods of drier conditions during glacial maxima (Price and Elliott-Fisk, 2004). During these periods bog habitats have likely shrunk or disappeared from some mountain ranges entirely before expanding again during wetter conditions. In general, this habitat type is becoming available later than the others, comparatively short-lived and likely highly competitive, thus resulting in high species turnover.

The highest speciation and extinction rates were estimated for dry habitats (Table 4.2). On the Hawaiian Islands, these are generally associated with lowland, leeward areas (Wagner et al., 1999a). These ranges are transient in the early stages of island formation. At the end of the shield stage, when the island moves away from the mantel plume, subsidence occurs rapidly (Clague, 1996; Clague and Sherrod, 2014) resulting in the submergence of initial lowland areas. Lowland, dry habitats are most extensive during glacial maxima. Large areas of land emerge when sea levels drop and the conditions are generally cooler and much drier (Price and Elliott-Fisk, 2004). Glacial cycles result in repeated extension and contraction of dry, lowland areas and result in high rates of species turnover. The high estimated rates of extinctions are an echo of the effects of the current interglacial period setting in with a rapid rise of sea levels beginning only 19.000 years ago (Lambeck and Chappell, 2001). However, diversification rates for dry habitats might be contorted by human influence. Lowland regions on the Hawaiian Islands have been exceptionally impacted by anthropogenic land use due to settlements and agriculture. Many of the species adapted to these areas are already extinct or at high risk for extinction (Sakai et al., 2002). The inferred rates for extinction might reflect this anthropogenic impact.

On the other hand, the trait "habitat" as used here is a very broad concept mainly relating to the water regime and only one possible trajectory for adaption. In the Asteraceae genus *Encelia* Adans. adaption to light intensity and temperatures along an elevational gradient resulted in different degrees of leaf pubescence (Ehleringer and Clark, 1988). Differences in leaf hairiness, especially along the midrib, is also known from *Melicope* (Stone et al., 1999; Hartley, 2001). However, despite our current understanding of the geology of the islands, detailed data regarding environmental factors such as light intensity or soil properties are limited. Reconstructing the adaptive landscape (Schluter, 2000) of Hawaiian *Melicope* requires fine-scaled data about these ecological factors, detailed distributional maps for the lineage linking species to specific ecological conditions, as well as corresponding phenotypic traits to identify adaption. For example, the difference in fruit morphology throughout the lineage (Stone et al., 1999) an the gene flow necessary to sustain widespread species (see 5.1) might indicate adaption to different vector species. Currently lacking knowledge about pollinator identity, availability, and specificity, as well as mating barriers, might be crucial to understand the realization of hybridization and adaption

to local insect communities. Flowers of Hawaiian *Melicope* produce nectar, but are generally small and inconspicuously colored suggesting a general strategy of pollination by insects with no obvious morphological specialization (Hartley, 2001). However, this theory remains without empirical evidence. With the hermaphroditic *Platydesma* lineage showing floral traits (copious nectar production and connate stamens) interpreted as adaptions to bird pollination (Appelhans et al., 2017) and the wealth of secondary metabolites present in Rutaceae (Kubitzki et al., 2011) in general and *Melicope* in particular, diversification driven by adaptions to biotic factors is likely but untested.

For Hawaiian *Melicope* I estimated the highest diversification rates in clade I (Figure 4.3) suggesting one key to diversification in Hawaiian *Melicope* might be flexibility, i.e. evolvability. The ability to frequently shift habitat type, elevational range (Figure 4.4) and possibly other characters, as required to exploit open ecological niches, might create an abundance of small-niche specialists, each with the capacity to spawn new species to compete for other niches. This evolvability would result in a lineage that is highly adapted to the continually emerging and changing habitats typical for the geologically active environment on oceanic islands.

Many oceanic lineages present a different picture with high initial rates of diversification (Whittaker and Fernández-Palacios, 2010) and higher rates of diversification on younger islands (Borregaard et al., 2017). Either of these patterns is not directly observable in Hawaiian *Melicope* (Figures 4.2, 4.3). In Hawaiian *Melicope* the majority of diversification seems limited to the comparatively short period of the recent ca. 1.5 MA (Figure 4.1) to the island of Kauaʻi. The low observed rates of initial variation are quite possibly an artifact of the previously discussed high amounts of species turnover. As species are adapted to narrow ecological niches, the local disappearance of said niche results in the extinction of a species or in a shift to a new niche. Considering that taller islands tend to support more habitat types than lower ones (Hobohm, 2000), diversity on Kauaʻi was possibly even greater in the past. The observed low diversity results from the high background extinction rates and extant diversity resulting from the few lineages leaving descendants to the present. High rates of extinction might also explain – at least partially – the comparatively low diversity on younger islands and their recent colonization compared to island ages (Figure 4.2). This would require entire established lineages on the younger islands to have gone extinct and be replaced by the comparatively arrivals of the most wave of arrivals. According to the general dynamic model of island biogeography (Whittaker et al., 2007, 2008), all islands except for Hawaiʻi have reached or surpassed their maximum carrying capacity in terms of the number of species, so broom-and-handle shaping of the tree and the comparable youth of diversity might also represent "extinction-based-saturation" of niches (MacArthur and Wilson, 1967; Sax and

Gaines, 2008). In this case, the higher diversity on Kaua‘i might be explained by some degree of allopatric speciation in dissected habitats representing identical niches. Alternatively, the high extinction rates correspond to a mass extinction event in the recent past (1-2 mya), which impacted O‘ahu and Maui more severely than Kaua‘i. However, evidence for such an event would likely be provided by fossils, which are lacking for Hawaiian *Melicope*, or by comparison across Hawaiian lineages.

On the other hand, the differences in species diversity on the islands in the archipelago might correspond to a founder effect. In island settings, bottlenecks occur on four different levels: the colonization of the archipelago itself, the subsequent colonization of individual islands, the colonization of habitat patches on islands (Whittaker and Fernández-Palacios, 2010), and the climate changes associated with glacial cycles reducing some habitat regimes drastically in size (Price and Clague, 2002). If evolvability is an important trait in Hawaiian *Melicope* the colonization of the younger islands might have resulted in local populations with a reduced genetic regarding adaptive traits, which would result in a lower evolutionary opportunity, (Glor, 2010), increased vulnerability to climatic effects and consequently lower species numbers.

Unfortunately, several difficulties exist for determining speciation and extinction rates through time from phylogenies using current methods (reviewed in Sanmartín and Meseguer, 2016). The handle-and-broom shape of the Hawaiian *Melicope* phylogeny can be a result from three different speciation and extinction signatures: a high background extinction rate across the entire tree, a mass extinction event sometime around ca. 1.5 mya and a true shift in diversification rates in the lineage (Crisp and Crone, 2009; Antonelli and Sanmartín, 2011). Distinguishing between these three scenarios, requires additional information, such as paleontological data (Sanmartín and Meseguer, 2016). Unfortunately, there are no applicable fossils for *Melicope* or other Hawaiian endemic lineages to date. However, a mass extinction event would have impacted a range of Hawaiian taxa and its signature should, therefore, be found in all their phylogenies. Genomic studies on Hawaiian lineages are scarce but increasing (Izuno et al., 2016, 2017; Jennings et al., 2016; Welch et al., 2016; Kleinkopf et al., 2019), offering the possibility to compare a number of resolved phylogenies in the near future. Considering the high volcanic activity of the Hawaiian mantle plume and its consequences, e.g. lava flows, slope failures, and tsunamis (McMurtry et al., 2004; Clague and Sherrod, 2014), frequent meteorological events, e.g. hurricanes, in addition to the effects of island ontogeny (Figure 1.1; Whittaker and Fernández-Palacios, 2010) and repeated glacial cycles (Price and Elliott-Fisk, 2004; Fernández-Palacios et al., 2016), makes a general high background extinction rate the most likely. Yet, while modeling approaches have recently become more sophisticated, the majority are unable to estimate speciation and extinction rate heterogeneously across time and clades without prior assumptions (Sanmartín and

Meseguer, 2016). The closest models to achieve this are the class of SSE methods, which associate rates to the evolution of a trait character. While I have employed one of these methods in chapter 4 to investigate the effects of habitat adaption to diversification (see above), the application of these methods is computationally and data-intensive. The application of these methods to *Melicope* is currently limited by the restricted availability of ecological data as well as the necessity of taxonomic revision to accurately investigate morphological traits.

Subsequently, research efforts have to be extended to the *Melicope* genus as a whole, preferably using genomic methods to resolve species relationships, reticulate evolution, adaption, and speciation. *Melicope* distributed throughout Australasia with local adaptive radiations on nearly all Pacific island systems and mainland areas. The island radiations are typically monophyletic and younger than the Hawaiian radiation (Hartley, 2001; Appelhans et al., 2018b). The framework of adaptive radiation on oceanic islands research currently tends to focus on few iconic lineages and to overlook colonizers that have not produced either endemic species or substantial variation (Warren et al., 2015). In the *Acronychia-Melicope* clade, there are several widespread species, e.g. *M. triphylla,* which occurs throughout Malesia and on several Pacific islands (Hartley, 2001). There are also several instances of multiple colonizations of island systems, e.g. Lord Howe Island was colonized three times independently (Appelhans et al., 2018b) and the Marquesas Islands were colonized twice from Hawaiian ancestry (this thesis). Thus the *Acronychia-Melicope* clade represents an ideal case study for island adaptive radiation with Hawaiian *Melicope* representing the tip of the proverbial iceberg and an ideal subset to establish methods and research questions.

## 5.4 Big data in plant systematics – quo vadis?

Application of phylogenomic datasets to systematics has provided insights into a large number of recalcitrant relationships across the tree of life. Amongst others, it revealed that horizontal gene transfer is very common in multicellular organisms (e.g. Abbott et al., 2016; Gallardo, 2017). Phylogenomic approaches provide an unprecedented amount of information and thus the ability to resolve bursts of speciation at both recent and deep evolutionary time scales (e.g. Wanke et al., 2017). However, the sheer amount of data creates and reveals methodological and computational challenges regarding matrix assembly, species tree reconstruction, and algorithmic complexity.

The assembly of HTS reads into a matrix is either reference-based (reads are mapped to a reference genome) or *de novo* (reads are compared to each other). The process is generally computationally intensive, especially in the case of *de novo* assembly since

it requires an all-to-all read comparison. A variety of software and pipelines exist for either approach. In the family Rutaceae, sequencing and annotation of entire genomes has so far only been performed for crop species from the genus *Citrus* (Wu et al., 2013; Xu et al., 2013). Unfortunately, these genomes are not suitable as a reference for *Melicope.* The two genera are members of different subfamilies in Rutaceae, which diverged ca. 55-70 mya (Appelhans et al., 2012; Morton and Telmer, 2014; Koenen et al., 2016). In addition, *Citrus* is characterized by a base chromosome number of n=9, compared to n=18 in *Melicope* (see chapter 2). Since a closely related reference genome was not available for *Melicope*; I assembled RAD-seq reads *de novo* using the software *ipyrad* (Eaton, 2014) in chapter 3. However, all available software implements the same paradigm of assembling sequences into contigs purely by sequence similarity. In RAD-seq assembly that paradigm is asserted twice, first when reads are clustered into putative genomic loci within each individual sample (in-sample-clustering) and then when consensus sequences called from in-sample loci are clustered across samples (between sample clustering) into a matrix (Catchen et al., 2011; Eaton, 2014; Ree and Hipp, 2015). Putative genomic loci are characterized by a user-specified maximum level of divergence; in *ipyrad* that is implemented as a minimum clustering threshold for the aligned sequences in a locus (Eaton, 2014). The selection of appropriate clustering thresholds is thus crucial to avoid introducing systematic errors to the analysis (Mastretta-Yanes et al., 2015), possibly resulting in the resolution of erroneous relationships (Misof et al., 2014). Unfortunately for the majority of taxonomic groups, reliable estimates for genomic divergence within and between genomes are lacking. Consequently, most RAD-seq studies employ a range of clustering thresholds, estimate tree topologies from the resulting alignments and determine the best assembly parameters by maximizing average statistical support, the number of SNPs or the recovery of specific relationships (e.g. Leaché et al., 2015; Suchan et al., 2017). Strategies to determine which loci are most likely assembly artifacts have been proposed (Shen et al., 2017), and could potentially be used to determine the best assembly. However, that approach might be computationally very demanding when a thorough exploration of parameter space becomes necessary. Few suggestions on identifying the optimal assembly parameters *a priori* exist, e.g. by using technical replicates (Mastretta-Yanes et al., 2015) or exploring the effects of clustering thresholds (Paris et al., 2017). In addition, in most cases the same threshold is employed for both in-sample- and between-sample-clustering (Leaché et al., 2015b; Suchan et al., 2017), implicitly assuming that the divergence between alleles in the genome is identical to the divergence between genomes of different species. This assumption may be biologically realistic in a population genomics approach (for which most of these methods have been developed), but unlikely when sampling divergent species.

For the assembly of the RAD-seq dataset of Hawaiian *Melicope,* I have attempted to reduce assembly artifacts and the introduction of systematic error by adapting the approach of Paris et al. (2017) to a phylogenetic framework (chapter 3). I have iterated over core clustering parameters in the *ipyrad* pipeline, separately for in-sample and between-sample clustering, to obtain an optimal parameter set for the assembly process.

However, even the optimization of the assembly process cannot provide a dataset of entirely accurately assembled loci. In *Citrus* L. (Rutaceae) the observed heterozygosity on chromosome 6 varies more than 3 fold between 500bp windows (Wu et al., 2014). Highly conserved loci would benefit from a very strict clustering threshold and are at risk of over-merging already. On the contrary, highly variable loci will be under-merged still. During between-sample-clustering, the issue of variable divergence of loci is exacerbated by the divergence between species. We now deal with varying genetic divergence between regions of one genome as well as different divergence times between loci of different taxa. As of now, there is no *de novo* clustering algorithm which can account for rates of evolution differing between genomic loci. Subsequent assembly and filtering steps will likely recognize some of these loci. However, a small amount of assembly error will likely remain even in the most optimal dataset. An 'optimal' dataset must therefore currently be one, which reduces the assembly error to a point that it is overpowered by the true signal.

All in all, the assembly process is vulnerable to systematic errors when assembly parameters are miss-specified. The RAD-seq assembly process would benefit from algorithms, which are aware of different levels of divergence and implement clustering parameter optimization into the assembly process. These could use a deep learning approach to identify bins of clusters requiring the same clustering threshold. These bins would also represent divergence rates of fractions of the genomes in question, and thus provide an additional layer of phylogenetic information. Alternatively, algorithms might incorporate genomic signal into the clustering process (Mendizabal-Ruiz et al., 2018).

In Hawaiian *Melicope*, the relationships of clade III were resolved incongruent between datasets and inference methods. Concatenated BI (CA-BI), ML (CA-ML) and BEAST analyses resolved clade II as sister to clade IV and the entire lineage as a sister group to a lineage comprising clades I and II with maximum PP and high ML-NBS support, but some discord revealed by QD scores for four of the datasets (Figure 3.2, Supplemental Figures 3.1, 3.2, 3.4). Relationships as resolved on the CA-BI and CA-ML of the min67 matrix place clade III as the only sister to clades I + II, and clade IV as sister to the lineage comprising clades I + II and III (Supplemental Figure 3.3). The SVD-Quartets analysis of the five datasets resulted in a third alternative topology,

where clade II is sister to clade III and that lineage is sister clades I + IV (Supplemental Figures 3.1-3.9). While the ultimate reason for the detected discord is likely the ancient introgression event (see 5.3), the question which topology best describes the relationships of the major lineages in Hawaiian *Melicope* remains.

The performance and evaluation of different species-tree inference methods and models is a matter of ongoing research especially with regards to the scalability to large genomic datasets with thousands or hundreds of thousands of loci (e.g. Chou et al., 2015; Roch and Warnow, 2015; Mirarab et al., 2016; Molloy and Warnow, 2018; Nute et al., 2018). Concatenated analysis of multiple genomic regions has been suggested to possibly resolve erroneous relationships with high statistical support (Kubatko and Degnan, 2007). Concatenation does not model the effects of heterogeneity in the evolutionary history of individual genomic regions resulting in gene-tree/species-tree discord. Accordingly, the concatenation approach is statistically inconsistent in the presence of incomplete lineage sorting (Kubatko and Degnan, 2007). As typically only a limited number of sites are informative for specific nodes in a topology (Palmer et al., 2019), a 'handful' of loci can drive the resolution of erroneous relationships with high support (Shen et al., 2017). Short branches are especially susceptible to this effect, as many genomic loci fail to coalesce between speciation events and thus provide contentious information (Kumar et al., 1990). The Multispecies Coalescent (MSC) was developed to explicitly model gene-tree heterogeneity when resolving species-tree relationships (Degnan and Rosenberg, 2009). A range of methods has been developed for species tree estimation from multi-locus datasets under the MSC model (reviewed in Warnow, 2017). These methods fall into three categories: (a) co-estimation of gene trees and species trees, as implemented for example in BEAST (Heled and Drummond, 2010). Sampling both species trees and gene trees simultaneously is computationally very demanding, even when employing heuristics (Wang and Nakhleh, 2018). Thus the method is currently computationally not feasible for RAD-seq sized datasets. (b) summary methods. The majority of available coalescent-based methods developed to date aim to estimate the species tree from existing gene trees by applying a summary statistic, e.g. ASTRAL (Mirarab et al., 2016), BUCKy (Larget et al., 2010), and MP-EST (Liu et al., 2010). Most summary methods have been proven statistically consistent under the MSC (Warnow, 2017), provided a large sample of true gene trees is provided (Mirarab et al., 2014a). In genome sized dataset, the likelihood of all loci containing substantial phylogenetic signal is low resulting in a high amount of gene tree estimation error (GTEE) (Bayzid et al., 2015). For analyzing RAD datasets, where the majority of individual loci comprise only limited phylogenetic information (Ree and Hipp, 2015), the vulnerability of MSC methods to GTEE is especially pronounced. In addition, the number of loci in RAD matrices and the resulting number of locus trees require tremendous computational efforts, which is currently hardly feasible.

Binning approaches (Bayzid and Warnow, 2013; Mirarab et al., 2014a; Zimmermann et al., 2014; Bayzid et al., 2015) offer computational improvements but their impact on accuracy is highly variable and currently poorly understood (Streicher et al., 2018). (c) site-based coalescent methods. These methods aim to bypass the calculation of gene-trees by analyzing individual site patterns to estimate the species tree. The method SNAPP (Bryant et al., 2012) employs a Bayesian MCMC algorithm that is computationally highly expensive and only feasible for a limited number of samples in the dataset (Yoder et al., 2013). A computationally feasible site-based coalescent method is SVD-Quartets, which computes quartet topologies and summarizes them into a species tree topology using a heuristic (Chifman and Kubatko, 2014).

The majority of coalescent-based methods have been proven to be statistically consistent under the MSC model, curiously with the exception of SVD-Quartets, where the mathematical proof is still outstanding (Roch et al., 2019). However, statistical consistency does not necessarily translate to the accuracy of inferred relationships. In comparative studies on simulated and empirical data, CA-ML performed at least as well and frequently better than coalescent-based methods under a variety of proportions of ILS and gene tree estimation error (Chou et al., 2015; Mirarab et al., 2016; Molloy and Warnow, 2018; Nute et al., 2018; Streicher et al., 2018; Palmer et al., 2019). The assumptions posed for statistical consistency under the MSC are not biologically realistic conditions (Roch et al., 2019); the MSC requires that individual loci are independent, that there is no recombination and that gene tree discord is exclusively caused by ILS (Degnan and Rosenberg, 2009). Statistical consistency under the MSC is the proof that the respective method will converge on the true species tree topology with probability 1 as the number of genes and the number of sites per gene both increase without upper bound (Roch et al., 2019). When the number of sites per locus has an upper bound, irrespective of its size, all of the hitherto tested MSC models as well as fully partitioned ML analysis become statistically inconsistent and can even become positively misleading, i.e. converging on an erroneous species tree topology with high probability, with long-branch attraction effects suggested as the ultimate cause (Roch et al., 2019). These simulations did not include site-based methods, however, SVD-Quartets has been proven to be less accurate than summary and co-estimation MSC models as well as CA-ML under all levels of ILS, GTEE (Roch and Warnow, 2015; Nute et al., 2018). In addition, it is vulnerable to missing data, especially when the distribution is non-random (Schmidt-Lebuhn et al., 2017; Nute et al., 2018). Concatenated analysis, on the other hand, generally resolves relationships with high accuracy under all simulated and empirically tested conditions of GTEE, ILS and missing data, despite not being statistically consistent under the MSC (e.g. Roch and Warnow, 2015; Eaton et al., 2017; Nute et al., 2018).

The limits of concatenated analyses of large phylogenomic datasets are poorly understood (Molloy and Warnow, 2018) and research has so far focused solely on concatenated Maximum Likelihood (CA-ML) inference. In this thesis, I have also employed concatenated Bayesian (CA-BI, chapter 3) as well as BEAST (chapter 4) analysis on RAD matrices for Hawaiian *Melicope*. The performance of either approach, with regards to robustness to ILS, GTEE, and HGT has not been formally assessed, yet. For CA-ML approaches it was suggested concatenation results in a synergistic effect where the combination of loci is more accurate than the proverbial sum of the signals provided by individual regions (Rivers et al., 2016; Palmer et al., 2019). This synergistic effect or 'hidden support' provides robustness to conflicting signals originating from incongruent gene trees, GTEE and missing data to the point that correct species relationships are inferred even when the majority of individual loci support an alternate, erroneous relationship (Rivers et al., 2016). The increased robustness to conflict might be imparted by the high-level distribution of phylogenetic signal for individual nodes versus the low-level distribution of conflict (Nute et al., 2018). In addition, Quartet sampling detected overall less discord in relationships resolved by on concatenated analysis. Accordingly, I have accepted the topology resolved by analysis of the concatenated analyses (Figure 3.2) as the more likely relationship.

However, as an ancestral introgression event was inferred as the most likely cause for conflicting relationships of clade III (chapter 3, Figure 3.3), phylogenetic reconstruction methods modeling HGT are required to definitively confirm the relation of the major clades. Phylogenetic Networks aim to explicitly incorporate reticulate evolution by allowing horizontal edges in addition to bifurcating ones (Huson et al., 2010). With the increasing recognition of the prevalence of hybridization events across the tree of life (Mallet et al., 2016), algorithms for inferring explicit phylogenetic networks have recently been advanced from a proof-of-concept state (Huson et al., 2010) to fully implemented software. As the evolution of individual loci within the network is tree-like, the MSC has been extended to phylogenetic networks (Yu et al., 2011, 2012) and recently network inference in a Bayesian framework has been implemented (Zhu et al., 2018). However, the relationship between phylogenetic trees and networks is very complex, and many problems are computationally expensive (Zhu et al., 2016). The computation of explicit networks, inferring simultaneously a rooted network with edges representing evolutionary distances and the number of reticulate events within that network is an NP-hard problem, so heuristics and algorithmic constraints, e.g. considering the number of reticulation events, have to be enforced (Huson et al., 2010). Still, phylogenetic network space is considerably larger than that of phylogenetic trees on the same number of taxa (Wen et al., 2018). Consequently, the majority of the current methods can only handle a small number of taxa and genomic

loci and do not scale to phylogenomic datasets (Zhu and Nakhleh, 2018; Elworth et al., 2019; Zhu et al., 2019). However, scalability is substantially improved in a recently proposed divide-and-conquer method, which drastically reduces computation times (Zhu et al., 2019), might provide an algorithm applicable in Hawaiian *Melicope* in the near future.

Streicher et al. (2018) posed the relevant question of whether researchers will have the data determine the analytical method or if the method should determine the data to be chosen for analysis. Currently, computational restrictions related to dataset size and complexity are often one of the deciding factors in method choice. This particular situation is pragmatic, but undoubtedly not ideal. However, the increased awareness of model assumptions and computational challenges provided by studies like the one presented in this thesis serve to highlight current challenges and to drive future development.

## 5.5 Conclusion and Prospects

Hawaiian *Melicope* represent a great model to study island adaptive radiation. The application of RAD-seq has resulted in unprecedented resolution of relationships in this mesmerizing lineage allowing for efficient testing of evolutionary hypotheses.

The lineage colonized the archipelago before the origin of the current high islands, and the oldest high island, Kauaʻi was colonized nearly upon its emergence. The younger islands were colonized repeatedly. However, the only direct evidence we find is for very recent colonizations of the younger islands, long after they had emerged (Figure 4.2). In general, diversification in the lineage is characterized by extinction just as much as by adaption to continuously changing habitats and speciation events (Figure 4.4, Table 4.2). The occurrence of some widespread species closely related to narrow endemics (Figure 4.4) indicates shifts in either pollinator or dispersal vectors; although in general there is no detailed knowledge of pollinator or disperser identity. A shift in breeding system is displayed by the members of clade V (*Platydesma*) from a dioecious ancestor to hermaphroditism in the extant species. The shift to monomorphism was estimated to have occurred around the time when Kauaʻi was first colonized by the clade. Two ancient introgression events were indicated involving the ancestors to the remaining four clades of Hawaiian *Melicope.* In addition, the RAD-seq datasets and the observation of morphologically intermediate individuals in the field, indicate additional, recent hybridization events. All in all occasional hybridization events might be crucial mechanisms to allow for the combination of adaptive traits from existing genetic variation during rapid speciation (Barrett and Schluter, 2008; Abbott et al., 2013).

The limited sampling for populations of species suspected in recent and current hybridization events in this thesis as well as the nature of RAD-seq data limits definitive assessment of the frequency, direction, and impact of introgression events on the evolution of Hawaiian *Melicope.* In addition, a taxonomic revision is required to investigate phenotypic trait evolution. In order to address the issues raised by this thesis, I am currently drafting a project aiming to employ a TE approach to Hawaiian *Melicope* using a custom bait set designed by myself. Sampling will be expanded to include additional populations of species resolved as non-monophyletic, species suspected to be involved in hybridization events, e.g. *M. barbigera* and *M. haupuensis* (Figure 3.3), as well as the entire Marquesan radiation and additional populations of widespread Hawaiian taxa. Using the resulting gene alignments, I will revise the question of phylogenetic inference methods with regards to accuracy and computational feasibility. In addition, loci resulting from the ancient hybridization event (Figure 3.2) will be identified and evaluated for conferring adaptive traits. Phasing of alleles will allow the inference of paternal lineages to recent and ongoing hybridization events, potentially adding further insights into the evolution of adaptive traits.

# References

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., et al. (2013). Hybridization and speciation. *Journal of Evolutionary Biology* 26, 229–246. doi:10.1111/j.1420-9101.2012.02599.x.

Abbott, R. J., Barton, N. H., and Good, J. M. (2016). Genomics of hybridization and its evolutionary consequences. *Molecular Ecology* 25, 2325–2332. doi:10.1111/mec.13685.

Aberer, A. J., Kobert, K., and Stamatakis, A. (2014). ExaBayes: Massively parallel Bayesian tree inference for the whole-genome era. *Molecular Biology and Evolution* 31, 2553–2556. doi:10.1093/molbev/msu236.

Aberer, A. J., Krompass, D., and Stamatakis, A. (2013). Pruning rogue taxa improves phylogenetic accuracy: An efficient algorithm and webservice. *Systematic Biology* 62, 162–166. doi:10.1093/sysbio/sys078.

Allison, P. A., and Bottjer, D. J. eds. (2011). *Taphonomy: Bias and process through time*. Dordrecht, Netherlands: Springer. doi:10.1007/978-90-481-8643-3_1.

Amborella Genome Project, Albert, V. A., Barbazuk, W. B., dePamphilis, C. W., Der, J. P., Leebens-Mack, J., et al. (2013). The *Amborella* genome and the evolution of flowering plants. *Science* 342, 1241089–1241089. doi:10.1126/science.1241089.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., and Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* 17, 81–92. doi:10.1038/nrg.2015.28.

Andrews, S. (2010). FastQC: A quality control tool from high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Antonelli, A., and Sanmartín, I. (2011). Mass extinction, gradual cooling, or rapid radiation? Reconstructing the spatiotemporal evolution of the ancient Angiosperm genus *Hedyosmum* (Chloranthaceae) using empirical and simulated approaches. *Systematic Biology* 60, 596–615. doi:10.1093/sysbio/syr062.

Appelhans, M. S., Keßler, P. J. A., Smets, E., Razafimandimbison, S. G., and Janssens, S. B. (2012). Age and historical biogeography of the pantropically distributed Spathelioideae (Rutaceae, Sapindales). *Journal of Biogeography* 39, 1235–1250. doi:10.1111/j.1365-2699.2012.02686.x.

Appelhans, M. S., Reichelt, N., Groppo, M., Paetzold, C., and Wen, J. (2018a). Phylogeny and biogeography of the pantropical genus *Zanthoxylum* and its closest relatives in the proto-Rutaceae group (Rutaceae). *Molecular Phylogenetics and Evolution* 126, 31–44. doi:10.1016/j.ympev.2018.04.013.

Appelhans, M. S., Smets, E., Razafimandimbison, S. G., Haevermans, T., van Marle, E. J., Couloux, A., et al. (2011). Phylogeny, evolutionary trends and classification of the

*Spathelia–Ptaeroxylon* clade: morphological and molecular insights. *Annals of Botany* 107, 1259–1277. doi:10.1093/aob/mcr076.

Appelhans, M. S., Wen, J., Duretto, M., Crayn, D., and Wagner, W. L. (2018b). Historical biogeography of *Melicope* (Rutaceae) and its close relatives with a special emphasis on Pacific dispersals. *Journal of Systematics and Evolution* 56, 576-599. doi:10.1111/jse.12299.

Appelhans, M. S., Wen, J., and Wagner, W. L. (2014a). A molecular phylogeny of *Acronychia*, *Euodia*, *Melicope* and relatives (Rutaceae) reveals polyphyletic genera and key innovations for species richness. *Molecular Phylogenetics and Evolution* 79, 54–68. doi:10.1016/j.ympev.2014.06.014.

Appelhans, M. S., Wen, J., Wood, K. R., Allan, G. J., Zimmer, E. A., and Wagner, W. L. (2014b). Molecular phylogenetic analysis of Hawaiian Rutaceae (*Melicope*, *Platydesma* and *Zanthoxylum* ) and their different colonization patterns. *Botanical Journal of the Linnean Society* 174, 425–448. doi:10.1111/boj.12123.

Appelhans, M. S., Wood, K. R., and Wagner, W. L. (2017). Reduction of the Hawaiian genus *Platydesma* into *Melicope* section *Pelea* (Rutaceae) and notes on the monophyly of the section. *PhytoKeys* 91, 125–137. doi:10.3897/phytokeys.91.21363.

Arrigo, N., and Barker, M. S. (2012). Rarely successful polyploids and their legacy in plant genomes. *Current Opinion in Plant Biology* 15, 140–146. doi:10.1016/j.pbi.2012.03.010.

Asker, S., and Jerling, L. (1992). *Apomixis in plants*. Boca Raton, USA: CRC Press.

Baele, G., Li, W. L. S., Drummond, A. J., Suchard, M. A., and Lemey, P. (2013). Accurate Model Selection of Relaxed Molecular Clocks in Bayesian Phylogenetics. *Molecular Biology and Evolution* 30, 239–243. doi:10.1093/molbev/mss243.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3, e3376. doi:10.1371/journal.pone.0003376.

Baker, H. G. (1955). Self-Compatibility and establishment after "long-distance" dispersal. *Evolution* 9, 347–349. doi:10.1111/j.1558-5646.1955.tb01544.x.

Baldwin, B. G. (1998). "Evolution in the endemic Hawaiian Compositae." in *Evolution and speciation of island plants*, eds. T. F. Stuessy and M. Ono (Cambridge, UK: Cambridge University Press), 49–73.

Baldwin, B. G. (2005). Origin of the serpentine-endemic *Layia discoidea* from the widespread *L. glandulosa* (Compositae). *Evolution* 59, 2473–2479. doi:10.1554/05-147.1.

Baldwin, B. G., Kyhos, D. W., Dvorak, J., and Carr, G. D. (1991). Chloroplast DNA evidence for a North American origin of the Hawaiian silversword alliance (Asteraceae). *Proceedings of the National Academy of Sciences* 88, 1840–1843. doi:10.1073/pnas.88.5.1840.

Baldwin, B. G., and Sanderson, M. J. (1998). Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences* 95, 9402–9406. doi:10.1073/pnas.95.16.9402.

Baldwin, B. G., and Wagner, W. L. (2010). Hawaiian Angiosperm radiations of North American origin. *Annals of Botany* 105, 849–879. doi:10.1093/aob/mcq052.

Ballard, H. E., and Sytsma, K. J. (2000). Evolution and biogeography of the woody Hawaiian Violets (*Viola*, Violaceae): arctic origins, herbaceous ancestry and bird dispersal. *Evolution* 54, 1521–1532.

Barido-Sottani, J., Bošková, V., Plessis, L. D., Kühnert, D., Magnus, C., Mitov, V., et al. (2018). Taming the BEAST—A Community Teaching Material Resource for BEAST 2. *Systematic Biology* 67, 170–174. doi:10.1093/sysbio/syx060.

Barrett, R. D. H., and Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in Ecology & Evolution* 23, 38–44. doi:10.1016/j.tree.2007.09.008.

Barrett, S. C. H., Emerson, B., Mallet, J. W., Clarke, B. C., and Grant, P. R. (1996). The reproductive biology and genetics of island plants. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 351, 725–733. doi:10.1098/rstb.1996.0067.

Barrier, M., Baldwin, B. G., Robichaux, R. H., and Purugganan, M. D. (1999). Interspecific hybrid ancestry of a plant adaptive radiation: allopolyploidy of the Hawaiian silversword alliance (Asteraceae) inferred from floral homeotic gene duplications. *Molecular Biology and Evolution* 16, 1105–1113. doi:10.1093/oxfordjournals.molbev.a026200.

Bayly, M. J., Holmes, G. D., Forster, P. I., Cantrill, D. J., and Ladiges, P. Y. (2013). Major clades of Australasian Rutoideae (Rutaceae) based on *rbcL* and *atpB* Sequences. *PLoS ONE* 8, e72493. doi:10.1371/journal.pone.0072493.

Bayzid, M. S., Mirarab, S., Boussau, B., and Warnow, T. (2015). Weighted statistical binning: Enabling statistically consistent genome-scale phylogenetic analyses. *PLoS ONE* 10, e0129183. doi:10.1371/journal.pone.0129183.

Bayzid, M. S., and Warnow, T. (2013). Naive binning improves phylogenomic analyses. *Bioinformatics* 29, 2277–2284. doi:10.1093/bioinformatics/btt394.

Bennett, M. D., and Leitch, I. J. (2007). Plant DNA C-values database (release 3.0). Accessed 01.10.2019. Available at: https://cvalues.science.kew.org/

Beurton, C. (1994). Gynoecium and perianth in *Zanthoxylum* s.l. (Rutaceae). *Plant Systematic Evolution* 189, 165–191. doi:10.1007/BF00939724.

Böhle, U. R., Hilger, H. H., and Martin, W. F. (1996). Island colonization and evolution of the insular woody habit in *Echium* L. (Boraginaceae). *Proceedings of the National Academy of Sciences* 93, 11740–11745. doi:10.1073/pnas.93.21.11740.

Borgmann, E. (1964). Anteil der Polyploiden in der Flora des Bismarcksgebirges von Ostneuguinea. *Zeitschrift für Botanik* 55, 118–173.

Borregaard, M. K., Amorim, I. R., Borges, P. A. V., Cabral, J. S., Fernández-Palacios, J. M., Field, R., et al. (2017). Oceanic island biogeography through the lens of the general dynamic model: assessment and prospect. *Biological Reviews* 92, 830–853. doi:10.1111/brv.12256.

Bramwell, D. (1976). "The endemic flora of the Canary Islands; distribution, relationships and phytogeography." in *Biogeography and ecology in the Canary Islands*, ed. G. Kunkel (Dordrecht: Springer Netherlands), 207–240. doi:10.1007/978-94-010-1566-0_6.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., and RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution* 29, 1917–1932. doi:10.1093/molbev/mss086.

Cariou, M., Duret, L., and Charlat, S. (2013). Is RAD-seq suitable for phylogenetic inference? An in silico assessment and optimization. *Ecology and Evolution* 3, 846–852. doi:10.1002/ece3.512.

Carlquist, S. (1966a). The biota of long-distance dispersal. I. Principles of dispersal and evolution. *The Quarterly Review of Biology* 41, 247–270. doi:10.1086/405054.

Carlquist, S. (1966b). The biota of long-distance dispersal. II. Loss of dispersibility in Pacific Compositae. *Evolution* 20, 30-48. doi: 10.2307/2406147

Carlquist, S. (1966c). The biota of long-distance dispersal. III. Loss of dispersibility in the Hawaiian flora. *Brittonia* 18, 310-335. doi:10.2307/2805148.

Carlquist, S. (1967). The biota of long-distance dispersal. V. Plant dispersal to Pacific islands. *Bulletin of the Torrey Botanical Club* 94, 129–162. doi:10.2307/2484044.

Carlquist, S. (1969). Wood anatomy of Lobelioideae (Campanulaceae). *Biotropica* 1, 47-72. doi:10.2307/2989761.

Carlquist, S. (1974). *Island Biology*. New York, NY: Columbia University Press.

Carman, J. G. (1997). Asynchronous expression of duplicate genes in Angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. *Biological Journal of the Linnean Society* 61, 51–94. doi:10.1111/j.1095-8312.1997.tb01778.x.

Carr, G. D. (1978). Chromosome numbers of Hawaiian flowering plants and the significance of cytology in selected taxa. *American Journal of Botany* 65, 236–242. doi:10.1002/j.1537-2197.1978.tb06061.x.

Carr, G. D. (1998). "Chromosome evolution and speciation in Hawaiian flowering plants." in *Evolution and speciation of island plants*, eds. T. F. Stuessy and M. Ono (Cambridge, UK: Cambridge University Press), 97–119.

Carr, G. D., Powell, E. A., and Kyhos, D. W. (1986). Self-Incompatibility in the Hawaiian Madiinae (Compositae): An exception to Baker's Rule. *Evolution* 40, 430–434. doi:10.2307/2408823.

Carracedo, J. C. (1999). Growth, structure, instability and collapse of Canarian volcanoes and comparisons with Hawaiian volcanoes. *Journal of Volcanology and Geothermal Research* 94, 1–19. doi:10.1016/S0377-0273(99)00095-5.

Cássia-Silva, C., and Sales, L. P. (2019). The imprints left by historical contingency on marsupials' life-history traits. *Journal of Zoology* 307, 149–158. doi:10.1111/jzo.12629.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., and Postlethwait, J. H. (2011). *Stacks* : Building and genotyping loci *de novo* from short-read sequences. *Genes|Genomes|Genetics* 1, 171–182. doi:10.1534/g3.111.000240.

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A., and Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology* 22, 3124–3140. doi:10.1111/mec.12354.

Chase, M. W., Morton, C. M., and Kallunki, J. A. (1999). Phylogenetic relationships of Rutaceae: a cladistic analysis of the subfamilies using evidence from *RBC* and *ATP* sequence variation. *American Journal of Botany* 86, 1191–1199. doi:10.2307/2656983.

Chifman, J., and Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics* 30, 3317–3324. doi:10.1093/bioinformatics/btu530.

Choi, J. Y., Purugganan, M., and Stacy, E. A. (2019). Divergent selection and primary gene flow shape incipient speciation of a riparian tree on Hawaii Island. *Molecular Biology and Evolution* msz259. doi:10.1093/molbev/msz259

Chou, J., Gupta, A., Yaduvanshi, S., Davidson, R., Nute, M., Mirarab, S., et al. (2015). A comparative study of SVDquartets and other coalescent-based species tree estimation methods. *BMC Genomics* 16, S2. doi:10.1186/1471-2164-16-S10-S2.

Clague, D. A. (1996). "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,* eds. A. Keast and S.C. Miller (Berlin, Germany: Balogh Scientific Books), 35–50.

Clague, D. A., Braga, J. C., Bassi, D., Fullagar, P. D., Renema, W., and Webster, J. M. (2010). The maximum age of Hawaiian terrestrial lineages: geological constraints from Kōko Seamount. *Journal of Biogeography* 37, 1022–1033. doi:10.1111/j.1365-2699.2009.02235.x.

Clague, D. A., and Sherrod, D. R. (2014). Growth and degradation of Hawaiian volcanoes. *U.S. Geological Survey Professional Papers* 1801, 97-146.

Cockayne, L., and Allan, H. H. (1934). An annotated list of groups of wild hybrids in the New Zealand flora. *Annals of Botany* 48, 1-55.

Collins, F. S., Morgan, M., and Patrinos, A. (2003). The human genome project: Lessons from large-scale biology. *Science* 300, 286–290. doi:10.1126/science.1084564.

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nature Reviews Genetics* 6, 836–846. doi:10.1038/nrg1711.

Cowie, R. H., and Holland, B. S. (2006). Dispersal is fundamental to biogeography and the evolution of biodiversity on oceanic islands. *Journal of Biogeography* 33, 193–198. doi:10.1111/j.1365-2699.2005.01383.x.

Crawford, D. J. (2010). Progenitor-derivative species pairs and plant speciation. *Taxon* 59, 1413–1423. doi:10.1002/tax.595008.

Crawford, D. J., Lowrey, T. K., Anderson, G. J., Bernardello, G., Santos-Guerra, A., and Stuessy, T. F. (2009). "Genetic diversity in Asteraceae endemic to oceanic islands: Baker's Law and polyploidy." in *Systematics, evolution, and biogeography of Compositae,* ed. V. Funk (Vienna, Austria: International Association for Plant Taxonomy), 139–151.

Crawford, D. J., Sang, T., Stuessy, T. F., Kim, S.-C., and Silva O., M. (1998). "*Dendroseris* (Asteraceae: Lactuceae) and *Robinsonia* (Asteraceae: Senecioneae) on the Juan Fernandez Islands: similarities and differences in biology and phylogeny." in *Evolution and speciation of island plants*, eds. T. F. Stuessy and M. Ono (Cambridge, UK: Cambridge University Press), 97–120.

Crisp, M. D., and Crone, L. G. (2009). Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies. *Evolution* 63, 2257–2265.

Cronk, Q. C. B., Kiehn, M., Wagner, W. L., and Smith, J. F. (2005). Evolution of *Cyrtandra* (Gesneriaceae) in the Pacific Ocean: the origin of a supertramp clade. *American Journal of Botany* 92, 1017–1024. doi:10.3732/ajb.92.6.1017.

Currat, M., Ruedi, M., Petit, R. J., and Excoffier, L. (2008). The hidden side of invasions: Massive introgression by local genes. *Evolution* 62, 1908–1920. doi:10.1111/j.1558-5646.2008.00413.x.

da Silva, M. F. das G. F., Gottlieb, O. R., and Ehrendorfer, F. (1988). Chemosystematics of the Rutaceae: Suggestions for a more natural taxonomy and evolutionary interpretation of the family. *Plant Systematic Evolution* 161, 97–134. doi:10.1007/BF00937293.

Darwin, C. (1842). *Journal of researches into the geology and natural history of the various countries visited by the voyage of H.M.S. "Beagle", under the command of Captain Fitzroy, R. N. from 1832-1836.* London, UK: Henry Colborn.

Darwin, C. (1859). *On the origin of species by means of natural selection, or the preservation of the favoured races in the struggle for life.* London, UK: John Murray.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., and Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* 12, 499–510. doi:10.1038/nrg3012.

Degnan, J. H., and Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution* 24, 332–340. doi:10.1016/j.tree.2009.01.009.

Denlinger, R. P., and Morgan, J. K. (2014). Instability of Hawaiian volcanoes. *U.S. Geological Survey Professional Papers* 1801, 149-176.

Díaz-Arce, N., Arrizabalaga, H., Murua, H., Irigoien, X., and Rodríguez-Ezpeleta, N. (2016). RAD-seq derived genome-wide nuclear markers resolve the phylogeny of tunas. *Molecular Phylogenetics and Evolution* 102, 202–207. doi:10.1016/j.ympev.2016.06.002.

Dobzhansky, T. (1937). *Genetics and the origin of species.* New York: Columbia Univ. Press.

Dodsworth, S., Chase, M. W., and Leitch, A. R. (2016). Is post-polyploidization diploidization the key to the evolutionary success of Angiosperms?: Diploidization in polyploid Angiosperms. *Botanical Journal of the Linnean Society* 180, 1–5. doi:10.1111/boj.12357.

Doležel, J., and Göhde, W. (1995). Sex determination in dioecious plants *Melandrium album* and *M. rubrum* using high-resolution flow cytometry. *Cytometry* 19, 103–106. doi:10.1002/cyto.990190203.

Doyle, J. J., Flagel, L. E., Paterson, A. H., Rapp, R. A., Soltis, D. E., Soltis, P. S., et al. (2008). Evolutionary genetics of genome merger and doubling in plants. *Annual Review of Genetics* 42, 443–461. doi:10.1146/annurev.genet.42.110807.091524.

Drummond, A. J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7, 214. doi:10.1186/1471-2148-7-214.

Dulin, M. W., and Kirchoff, B. K. (2010). Paedomorphosis, secondary woodiness, and insular woodiness in plants. *The Botanical Review* 76, 405–490. doi:10.1007/s12229-010-9057-5.

Dunbar-Co, S., Wieczorek, A. M., and Morden, C. W. (2008). Molecular phylogeny and adaptive radiation of the endemic Hawaiian *Plantago* species (Plantaginaceae). *American Journal of Botany* 95, 1177–1188. doi:10.3732/ajb.0800132.

Durand, E. Y., Patterson, N., Reich, D., and Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution* 28, 2239–2252. doi:10.1093/molbev/msr048.

Eaton, D. A. R. (2014). PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30, 1844–1849. doi:10.1093/bioinformatics/btu121.

Eaton, D. A. R., Hipp, A. L., González-Rodríguez, A., and Cavender-Bares, J. (2015). Historical introgression among the American live oaks and the comparative nature of tests for introgression. *Evolution* 69, 2587–2601. doi:10.1111/evo.12758.

Eaton, D. A. R., and Ree, R. H. (2013). Inferring phylogeny and introgression using RADseq data: An example from flowering plants (*Pedicularis*: Orobanchaceae). *Systematic Biology* 62, 689–706. doi:10.1093/sysbio/syt032.

Eaton, D. A. R., Spriggs, E. L., Park, B., and Donoghue, M. J. (2017). Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology* 66, 399–412. doi:10.1093/sysbio/syw092.

Eggens, F., Popp, M., Nepokroeff, M., Wagner, W. L., and Oxelman, B. (2007). The origin and number of introductions of the Hawaiian endemic *Silene* species (Caryophyllaceae). *American Journal of Botany* 94, 210–218. doi:10.3732/ajb.94.2.210.

Ehleringer, J. R., and Clark, C. (1988). "Evolution and adaptation in *Encelia* (Asteraceae)." in *Plant evolutionary biology*, eds. L. D. Gottlieb and S. K. Jain (Dordrecht: Springer Netherlands), 221–248. doi:10.1007/978-94-009-1207-6_9.

Ehrendorfer, F. (1980). "Polyploidy and distribution." in *Polyploidy: Biological relevance*, ed. W. H. Lewis (Boston, MA: Springer US), 45–60. doi:10.1007/978-1-4613-3069-1_3.

Elworth, R. A. L., Ogilvie, H. A., Zhu, J., and Nakhleh, L. (2019). "Advances in computational methods for phylogenetic networks in the presence of hybridization." in *Bioinformatics and phylogenetics*, ed. T. Warnow (Basel, Switzerland: Springer International Publishing), 317–360. doi:10.1007/978-3-030-10837-3_13.

Emerson, B. C. (2002). Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular Ecology* 11, 951-966.

Engler, A. (1931). "Cneoraceae, Rutaceae." in *Die natürlichen Pflanzenfamilien*, eds. A. Engler and K. Prantl (Leipzig, Germinay: W. Engelmann), 184–359.

Fedorov, A. V., Brierley, C. M., and Emanuel, K. (2010). Tropical cyclones and permanent El Niño in the early Pliocene epoch. *Nature* 463, 1066–1070. doi:10.1038/nature08831.

Fernández-Mazuecos, M., Mellers, G., Vigalondo, B., Sáez, L., Vargas, P., and Glover, B. J. (2018). Resolving recent plant radiations: Power and robustness of genotyping-by-sequencing. *Systematic Biology* 67, 250–268. doi:10.1093/sysbio/syx062.

Fernández-Palacios, J. M., Rijsdijk, K. F., Norder, S. J., Otto, R., Nascimento, L. de, Fernández-Lugo, S., et al. (2016). Towards a glacial-sensitive model of island biogeography. *Global Ecology and Biogeography* 25, 817–830. doi:10.1111/geb.12320.

FitzJohn, R. G. (2014). Diversitree: comparative phylogenetic analyses of diversification in *R*. *Methods in Ecology and Evolution* 6, 1084–1092. doi:10.1111/j.2041-210X.2012.00234.x.

Floyd, A. G. (1989). *Rainforest trees of mainland South-Eastern Australia.* Melbourne, Australia: Inkata Press.

Forest, F. (2009). Calibrating the Tree of Life: fossils, molecules and evolutionary timescales. *Annals of Botany* 104, 789–794. doi:10.1093/aob/mcp192.

Francisco-Ortega, J., Santos-Guerra, A., Hines, A., and Jansen, R. K. (1997). Molecular evidence for a Mediterranean origin of the Macaronesian endemic genus *Argyranthemum* (Asteraceae). *American Journal of Botany* 84, 1595–1613. doi:10.2307/2446622.

Francisco-Ortega, J., Santos-Guerra, A., Kim, S.-C., and Crawford, D. J. (2000). Plant genetic diversity in the Canary Islands: a conservation perspective. *American Journal of Botany* 87, 909–919. doi:10.2307/2656988.

Fresnillo, B., and Ehlers, B. K. (2008). Variation in dispersability among mainland and island populations of three wind-dispersed plant species. *Plant Systematics and Evolution* 270, 243–255. doi:10.1007/s00606-007-0615-1.

Frith, H. J., Crome, F. H. J., and Wolfe, T. O. (1976). Food of fruit-pigeons in New Guinea. *Emu* 76, 49–58. doi:10.1071/mu9760049.

Gadagkar, S. R., Rosenberg, M. S., and Kumar, S. (2005). Inferring species phylogenies from multiple genes: Concatenated sequence tree versus consensus gene tree. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 304, 64–74. doi:10.1002/jez.b.21026.

Gallardo, M. H. (2017). "Phylogenetics, reticulation and evolution." in *Phylogenetics*, ed. I. Y. Abdurakhmonov (London, UK: InTech), 47-58. doi:10.5772/intechopen.68564.

Garcia, M. O., Smith, J. R., Tree, J. P., Weis, D., Harrison, L., and Jicha, B. R. (2015). "Petrology, geochemistry, and ages of lavas from Northwest Hawaiian Ridge volcanoes," in *The origin, evolution, and environmental impact of oceanic large igneous provinces* eds. C. R. Neal, W. W. Sager, T. Sano, and E. Erba, 1–25.

Gatesy, J., and Springer, M. S. (2014). Phylogenetic analysis at deep timescales: Unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. *Molecular Phylogenetics and Evolution* 80, 231–266. doi:10.1016/j.ympev.2014.08.013.

Geist, D., Snell, H., Snell, H., Goddard, C., and Kurz, M. (2014). "Palaeogeography of the Galápagos Islands and biogeographical implications." in *The Galápagos: A natural laboratory for the earth sciences*, eds. K. S. Harpp, E. Mittelstaedt, N. d'Ozouville, and D. W. Graham (Washington, D.C., USA: American Geophysical Union), 145–166.

Gillespie, R. (2004). Community assembly through adaptive radiation in Hawaiian spiders. *Science* 303, 356–359. doi:10.1126/science.1091875.

Givnish, T. J. (1997). "Adaptive radiation and molecular systematics: issues and approaches," in *Molecular evolution and adaptive radiation*, eds. T. J. Givnish and K. J. Systma (Cambridge, UK: Cambridge University Press), 1–54.

Givnish, T. J. (1998). "Adaptive plant evolution on islands: classical patterns, molecular data, and new insights." in *Evolution on islands*, ed. P. R. Grant (Oxford, UK: Oxford University Press), 281–304.

Givnish, T. J. (2015). Adaptive radiation versus 'radiation' and 'explosive diversification': why conceptual distinctions are fundamental to understanding evolution. *New Phytologist* 207, 297–303. doi:10.1111/nph.13482@10.1111.

Givnish, T. J., Bean, G. J., Ames, M., Lyon, S. P., and Sytsma, K. J. (2013). Phylogeny, floral evolution, and inter-island dispersal in Hawaiian *Clermontia* (Campanulaceae) based on ISSR variation and plastid spacer sequences. *PLoS ONE* 8, e62566. doi:10.1371/journal.pone.0062566.

Givnish, T. J., Millam, K. C., Mast, A. R., Paterson, T. B., Theim, T. J., Hipp, A. L., et al. (2009). Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). *Proceedings of the Royal Society B: Biological Sciences* 276, 407–416. doi:10.1098/rspb.2008.1204.

Givnish, T. J., Spalink, D., Ames, M., Lyon, S. P., Hunter, Zuluaga, A., et al. (2015). Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proceedings of the Royal Society B: Biological Sciences* 282, 20151553. doi:10.1098/rspb.2015.1553.

Glor, R. E. (2010). Phylogenetic insights on adaptive radiation. *Annual Review of Ecology, Evolution, and Systematics* 41, 251–270. doi:10.1146/annurev.ecolsys.39.110707.173447.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E. M., Brockman, W., et al. (2009). Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature Biotechnology* 27, 182–189. doi:10.1038/nbt.1523.

Grant, P. R. (1998). "Patterns on islands and microevolution." in *Evolution on islands,* ed. P. R. Grant (Oxford, UK: Oxford University Press), 1–17.

Grant, P. R., and Grant, B. R. (2002). Adaptive radiation of Darwin's finches: Recent data help explain how this famous group of Galápagos birds evolved, although gaps in our understanding remain. *American Scientist* 90, 130–139.

Grant, P. R., and Grant, B. R. (2008). *How and why species multiply: The radiation of Darwin's finches*. Princeton, USA: Princeton University Press.

Grant, V. (1963). *The origin of adaptations*. New York, USA: Columbia University Press.

Groppo, M., Pirani, J. R., Salatino, M. L. F., Blanco, S. R., and Kallunki, J. A. (2008). Phylogeny of Rutaceae based on two noncoding regions from cpDNA. *American Journal of Botany* 95, 985–1005. doi:10.3732/ajb.2007313.

Guerra, M. dos S. (1984). New chromosome number in Rutaceae. *Plant Systematics and Evolution* 146, 13–30. doi:10.1007/BF00984051.

Hair, J. B. (1966). Biosystematics of the New Zealand flora, 1945–1964. *New Zealand Journal of Botany* 4, 559–595. doi:10.1080/0028825X.1966.10430184.

Hammer, Ø., Harper, D. A. T., and Ryan, P. D. (2001). PAST: Palaeontological statistics software package for education and data analysis. *Palaeontologica Electronica* 4, 9.

Harbaugh, D. T. (2008). Polyploid and hybrid origins of Pacific island sandalwoods (*Santalum*, Santalaceae) inferred from low-copy nuclear and flow cytometry data. *International Journal of Plant Sciences* 169, 677–685. doi:10.1086/533610.

Harbaugh, D. T., and Baldwin, B. G. (2007). Phylogeny and biogeography of the sandalwoods (*Santalum*, Santalaceae): repeated dispersals throughout the Pacific. *American Journal of Botany* 94, 1028–1040. doi:10.3732/ajb.94.6.1028.

Harbaugh, D. T., Wagner, W. L., Allan, G. J., and Zimmer, E. A. (2009). The Hawaiian Archipelago is a stepping stone for dispersal in the Pacific: an example from the plant genus *Melicope* (Rutaceae). *Journal of Biogeography* 36, 230–241. doi:10.1111/j.1365-2699.2008.02008.x.

Hartley, T. G. (2001). On the taxonomy and biogeography of *Euodia* and *Melicope* (Rutaceae). *Allertonia* 8, 1–328.

Hartley, T. G., and Stone, B. C. (1989). Reduction of *Pelea* with new combinations in *Melicope* (Rutaceae). *Taxon* 38, 119–123. doi:10.2307/1220910.

Harvey, M. G., Smith, B. T., Glenn, T. C., Faircloth, B. C., and Brumfield, R. T. (2016). Sequence capture versus Restriction site associated DNA sequencing for shallow systematics. *Systematic Biology* 65, 910–924. doi:10.1093/sysbio/syw036.

Havran, J. C., Sytsma, K. J., and Ballard, H. E. (2009). Evolutionary relationships, interisland biogeography, and molecular evolution in the Hawaiian violets (*Viola*: Violaceae). *American Journal of Botany* 96, 2087–2099. doi:10.3732/ajb.0900021.

Heads, M. (2011). Old taxa on young islands: A critique of the use of island age to date island-endemic clades and calibrate phylogenies. *Systematic Biology* 60, 204–218. doi:10.1093/sysbio/syq075.

Heilbuth, J. C. (2000). Lower species richness in dioecious clades. *The American Naturalist* 156, 221–241. doi:10.1086/303389.

Heled, J., and Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* 27, 570–580. doi:10.1093/molbev/msp274.

Higgins, S. I., Nathan, R., and Cain, M. L. (2003). Are long-distance dispersal events in plants usually caused by nonstandard means of dispersal? *Ecology* 84, 1945–1956. doi:10.1890/01-0616.

Hipp, A. L., Eaton, D. A. R., Cavender-Bares, J., Fitzek, E., Nipper, R., and Manos, P. S. (2014). A framework phylogeny of the American Oak clade based on sequenced RAD data. *PLoS ONE* 9, e93975. doi:10.1371/journal.pone.0093975.

Ho, S. Y. W., and Duchêne, S. (2014). Molecular-clock methods for estimating evolutionary rates and timescales. *Molecular Ecology* 23, 5947–5965. doi:10.1111/mec.12953.

Hobohm, C. (2000). Plant species diversity and endemism on islands and archipelagos, with special reference to the Macaronesian Islands. *Flora* 195, 9–24. doi:10.1016/S0367-2530(17)30943-X.

Hodel, R. G. J., Chen, S., Payton, A. C., McDaniel, S. F., Soltis, P., and Soltis, D. E. (2017). Adding loci improves phylogeographic resolution in red mangroves despite increased missing data: comparing microsatellites and RAD-Seq and investigating loci filtering. *Scientific Reports* 7, e17598. doi:10.1038/s41598-017-16810-7.

Hohmann, N., Wolf, E. M., Lysak, M. A., and Koch, M. A. (2015). A time-calibrated road map of Brassicaceae species radiation and evolutionary history. *The Plant Cell* 27, 2770–2784. doi:10.1105/tpc.15.00482.

Höhna, S., and Freyman, W. A. (2016). *RevGadgets: Process output generated by RevBayes. R package version 1.0.0.* https://rdrr.io/github/revbayes/RevGadgets/.

Höhna, S., Landis, M. J., Heath, T. A., Boussau, B., Lartillot, N., Moore, B. R., et al. (2016). *RevBayes*: Bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Systematic Biology* 65, 726–736. doi:10.1093/sysbio/syw021.

Holzmeyer, L., Duretto, M., Crayn, D., Hörandl, E., Heslewood, M., Jayanthan, J., et al. (2015). Phylogeny of *Acronychia* (Rutaceae) and first insights into its historical biogeography and the evolution of fruit characters. *PLoS ONE* 10, e0136296. doi:10.1371/journal.pone.0136296.

Hsu, C.-C. (1968). Preliminary chromosome studies on the vascular plants of Taiwan (II). *Taiwania* 11, 11–27.

Huang, H., and Knowles, L. L. (2009). What is the danger of the anomaly zone for empirical phylogenetics? *Systematic Biology* 58, 527–536. doi:10.1093/sysbio/syp047.

Huson, D. H., Rupp, R., and Scornavacca, C. (2010). *Phylogenetic networks: Concepts, algorithms and applications*. Cambridge, UK: Cambridge University Press.

Imada, C., Clifford, P., and Lau, J. Q. C. (2011). *2010 Rare plant survey, O'ahu forest national wildlife refuge, Waipi'o O'ahu. 1st edition*. Honolulu, USA: Bishop Museum Press.

Innis, G. J. (1989). Feeding ecology of fruit pigeons in subtropical rainforests of South-Eastern Queensland. *Wildlife Research* 16, 365–394. doi:10.1071/wr9890365.

Izuno, A., Hatakeyama, M., Nishiyama, T., Tamaki, I., Shimizu-Inatsugi, R., Sasaki, R., et al. (2016). 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. doi:10.1007/s10265-016-0822-3.

Izuno, A., Kitayama, K., Onoda, Y., Tsujii, Y., Hatakeyama, M., Nagano, A. J., et al. (2017). The population genomic signature of environmental association and gene flow in an ecologically divergent tree species *Metrosideros polymorpha* (Myrtaceae). *Molecular Ecology* 26, 1515–1532. doi:10.1111/mec.14016.

Jeffroy, O., Brinkmann, H., Delsuc, F., and Philippe, H. (2006). Phylogenomics: the beginning of incongruence? *Trends in Genetics* 22, 225–231. doi:10.1016/j.tig.2006.02.003.

Jennings, H., Wallin, K., Brennan, J., Valle, A. D., Guzman, A., Hein, D., et al. (2016). Inbreeding, low genetic diversity, and spatial genetic structure in the endemic Hawaiian lobeliads *Clermontia fauriei* and *Cyanea pilosa* ssp. *longipedunculata*. *Conservation Genetics* 17, 497–502. doi:10.1007/s10592-015-0785-2.

Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., et al. (2011). Ancestral polyploidy in seed plants and Angiosperms. *Nature* 473, 97–100. doi:10.1038/nature09916.

Johnson, M. A., Clark, J. R., Wagner, W. L., and McDade, L. A. (2017). A molecular phylogeny of the Pacific clade of *Cyrtandra* (Gesneriaceae) reveals a Fijian origin, recent diversification, and the importance of founder events. *Molecular Phylogenetics and Evolution* 116, 30–48. doi:10.1016/j.ympev.2017.07.004.

Johnson, M. A., Pillon, Y., Sakishima, T., Price, D. K., and Stacy, E. A. (2019). Multiple colonizations, hybridization and uneven diversification in *Cyrtandra* (Gesneriaceae) lineages on Hawai‘i Island. *Journal of Biogeography* 46, 1178–1196. doi:10.1111/jbi.13567.

Jordan, S., Simon, C., Foote, D., and Englund, R. A. (2005). Phylogeographic patterns of Hawaiian *Megalagrion* damselflies (Odonata: Coenagrionidae) correlate with Pleistocene island boundaries. *Molecular Ecology* 14, 3457–3470. doi:10.1111/j.1365-294X.2005.02669.x.

Kagawa, K., and Takimoto, G. (2018). Hybridization can promote adaptive radiation by means of transgressive segregation. *Ecology Letters* 21, 264–274. doi:10.1111/ele.12891.

Keeley, S. C., and Funk, V. A. (2011). “Origin and evolution of Hawaiian endemics: new patterns revealed by molecular phylogenetic studies.” in *The biology of island floras*, eds. D. Bramwell and J. Caujape-Castells (Cambridge, UK: Cambridge University Press), 57–88. doi:10.1017/CBO9780511844270.005.

Kiehn, M. (2005). Chromosome numbers of hawaiian Angiosperms: New records and comments. *Pacific Science* 59, 363–377. doi:10.1353/psc.2005.0036.

Kiehn, M., and Lorence, D. H. (1996). Chromosome counts on Angiosperms cultivated at the National Tropical Botanical Garden, Kaua‘i, Hawai‘i. *Pacific Science* 50, 317–323.

Kim, S. C., Crawford, D. J., Francisco-Ortega, J., and Santos-Guerra, A. (1996). A common origin for woody *Sonchus* and five related genera in the Macaronesian Islands: molecular evidence for extensive radiation. *Proceedings of the National Academy of Sciences* 93, 7743–7748. doi:10.1073/pnas.93.15.7743.

Kleinkopf, J. A., Roberts, W. R., Wagner, W. L., and Roalson, E. H. (2019). Diversification of Hawaiian *Cyrtandra* (Gesneriaceae) under the influence of incomplete lineage sorting and hybridization. *Journal of Systematics and Evolution* 57, 561-578. doi:10.1111/jse.12519.

Knope, M. L., Morden, C. W., Funk, V. A., and Fukami, T. (2012). Area and the rapid radiation of Hawaiian *Bidens* (Asteraceae). *Journal of Biogeography* 39, 1206–1216. doi:10.1111/j.1365-2699.2012.02687.x.

Koenen, E. J. M., Clarkson, J. J., Pennington, T. D., and Chatrou, L. W. (2016). Recently evolved diversity and convergent radiations of rainforest mahoganies (Meliaceae) shed new light on the origins of rainforest hyperdiversity. *New Phytologist* 207, 327–339. doi:10.1111/nph.13490.

Kozlov, A. M., Aberer, A. J., and Stamatakis, A. (2015). ExaML version 3: a tool for phylogenomic analyses on supercomputers. *Bioinformatics* 31, 2577–2579. doi:10.1093/bioinformatics/btv184.

Krastel, S., Schmincke, H.-U., Jacobs, C. L., Rihm, R., Bas, T. P. L., and Alibés, B. (2001). Submarine landslides around the Canary Islands. *Journal of Geophysical Research: Solid Earth* 106, 3977–3997. doi:10.1029/2000JB900413.

Kubatko, L. S., and Degnan, J. H. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology* 56, 17–24. doi:10.1080/10635150601146041.

Kubitzki, K., Kallunki, J. A., Duretto, M., and Wilson, P. G. (2011). "Rutaceae." in *Flowering plants. Eudicots: Sapindales, Cucurbitales, Myrtaceae - The families and genera of vascular plants*, ed. K. Kubitzki (Berlin, Germany: Springer), 276–356.

Kumar, S., Filipski, A. J., Battistuzzi, F. U., Kosakovsky Pond, S. L., and Tamura, K. (2012). Statistics and truth in phylogenomics. *Molecular Biology and Evolution* 29, 457–472. doi:10.1093/molbev/msr202.

Kumar, V., Karunaratne, V., Sanath, M. R., Meegalle, K., and MacLeod, J. K. (1990). Two fungicidal phenylethanones from *Euodia lunu-ankenda* root bark. *Phytochemistry* 29, 243–245. doi:10.1016/0031-9422(90)89042-8.

Lambeck, K., and Chappell, J. (2001). Sea level change through the last glacial cycle. *Science* 292, 679–686. doi:10.1126/science.1059549.

Lammers, T. G. (1988). Chromosome numbers and their systematic implications in Hawaiian Lobelioideae (Campanulaceae). *American Journal of Botany* 75, 1130–1134.

Landis, M. J. (2017). Biogeographic dating of speciation times using paleogeographically informed processes. *Systematic Biology* 66, 128–144. doi:10.1093/sysbio/syw040.

Landis, M. J., Freyman, W. A., and Baldwin, B. G. (2018). Retracing the Hawaiian silversword radiation despite phylogenetic, biogeographic, and paleogeographic uncertainty. *Evolution* 72, 2343–2359. doi:10.1111/evo.13594.

Landis, M. J., Matzke, N. J., Moore, B. R., and Huelsenbeck, J. P. (2013). Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology* 62, 789–804. doi:10.1093/sysbio/syt040.

Larget, B. R., Kotha, S. K., Dewey, C. N., and Ané, C. (2010). BUCKy: Gene tree/species tree reconciliation with Bayesian concordance analysis. *Bioinformatics* 26, 2910–2911. doi:10.1093/bioinformatics/btq539.

Le Friant, A., Harford, C. L., Deplus, C., Boudon, G., Sparks, R. S. J., Herd, R. A., et al. (2004). Geomorphological evolution of Montserrat (West Indies): importance of flank collapse and erosional processes. *Journal of the Geological Society* 161, 147–160. doi:10.1144/0016-764903-017.

Leaché, A. D., Banbury, B. L., Felsenstein, J., de Oca, A. Nieto-Montes, and Stamatakis, A. (2015a). Short tree, long tree, right tree, wrong tree: New acquisition bias corrections for inferring SNP phylogenies. *Systematic Biology* 64, 1032–1047. doi:10.1093/sysbio/syv053.

Leaché, A. D., Chavez, A. S., Jones, L. N., Grummer, J. A., Gottscho, A. D., and Linkem, C. W. (2015b). Phylogenomics of phrynosomatid lizards: Conflicting signals from sequence capture versus restriction site associated DNA sequencing. *Genome Biology and Evolution* 7, 706–719. doi:10.1093/gbe/evv026.

Lecaudey, L. A., Schliewen, U. K., Osinov, A. G., Taylor, E. B., Bernatchez, L., and Weiss, S. J. (2018). Inferring phylogenetic structure, hybridization and divergence times within Salmoninae (Teleostei: Salmonidae) using RAD-sequencing. *Molecular Phylogenetics and Evolution* 124, 82–99. doi:10.1016/j.ympev.2018.02.022.

Leitch, I. J., and Bennett, M. D. (2004). Genome downsizing in polyploid plants. *Biological Journal of the Linnean Society* 82, 651–663. doi:10.1111/j.1095-8312.2004.00349.x.

Lens, F., Davin, N., Smets, E., and del Arco, M. (2013). Insular woodiness on the Canary Islands: A remarkable case of convergent evolution. *International Journal of Plant Sciences* 174, 992–1013. doi:10.1086/670259.

Lerner, H. R. L., Meyer, M., James, H. F., Hofreiter, M., and Fleischer, R. C. (2011). Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. *Current Biology* 21, 1838–1844. doi:10.1016/j.cub.2011.09.039.

Linder, H. P., and Barker, N. P. (2014). Does polyploidy facilitate long-distance dispersal? *Annals of Botany* 113, 1175–1183. doi:10.1093/aob/mcu047.

Lindqvist, C., and Albert, V. A. (2002). Origin of the Hawaiian endemic mints within North American *Stachys* (Lamiaceae). *American Journal of Botany* 89, 1709–1724. doi:10.3732/ajb.89.10.1709.

Lindqvist, C., Motley, T. J., Jeffrey, J. J., and Albert, V. A. (2003). Cladogenesis and reticulation in the Hawaiian endemic mints (Lamiaceae). *Cladistics* 19, 480–495. doi:10.1111/j.1096-0031.2003.tb00384.x.

Liu, L., Wu, S., and Yu, L. (2015). Coalescent methods for estimating species trees from phylogenomic data. *Journal of Systematics and Evolution* 53, 380–390. doi:10.1111/jse.12160.

Liu, L., and Yu, L. (2011). Estimating species trees from unrooted gene trees. *Systematic Biology* 60, 661–667. doi:10.1093/sysbio/syr027.

Liu, L., Yu, L., and Edwards, S. V. (2010). A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. *BMC Evolutionary Biology* 10, 302. doi:10.1186/1471-2148-10-302.

Lloyd, D. G. (1985). Progress in understanding the natural history of New Zealand plants. *New Zealand Journal of Botany* 23, 707–722. doi:10.1080/0028825X.1985.10434239.

Lorence, D. H., and Perlman, S. (2007). A new species of *Cyrtandra* (Gesneriaceae) from Hawai‘i, Hawaiian Islands. *Novon* 17, 357–361.

Losos, J. B. (1992). The evolution of convergent structure in caribbean *Anolis* communities. *Systematic Biology* 41, 403–420. doi:10.1093/sysbio/41.4.403.

Losos, J. B., and Mahler, D. L. (2010). "Adaptive radiation: The interaction of ecological opportunity, adaption and speciation." in *Evolution since Darwin: The first 150 years*, eds. M. A. Bell, D. J. Futuyama, W. F. Eanes, and J. S. Levinton (Sunderland, MA, USA: Sinauer Associates Inc.), 381–420.

Losos, J. B., and Ricklefs, R. E. (2009). Adaptation and diversification on islands. *Nature* 457, 830–836. doi:10.1038/nature07893.

Ma, X.-F., and Gustafson, J. P. (2005). Genome evolution of allopolyploids: a process of cytological and genetic diploidization. *Cytogenetic and Genome Research* 109, 236–249. doi:10.1159/000082406.

Mable, B. K. (2004). Polyploidy and self-compatibility: is there an association? *New Phytologist* 162, 803–811. doi:10.1111/j.1469-8137.2004.01055.x.

MacArthur, R. H., and Wilson, E. O. (1967). *The theory of island biogeography*. Princeton, USA.: Princeton University Press.

Maddison, W. P. (2006). Confounding asymmetries in evolutionary diversification and character change. *Evolution* 60, 1743–1746. doi:10.1111/j.0014-3820.2006.tb00517.x.

Maddison, W. P., Midford, P. E., and Otto, S. P. (2007). Estimating a Binary Character's Effect on Speciation and Extinction. *Systematic Biology* 56, 701–710. doi:10.1080/10635150701607033.

Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology & Evolution* 20, 229–237. doi:10.1016/j.tree.2005.02.010.

Mallet, J., Besansky, N., and Hahn, M. W. (2016). How reticulated are species? *BioEssays* 38, 140–149. doi:10.1002/bies.201500149.

Mann, H. (1866). Revision of the Rutaceae of the Hawaiian Islands. *Proceedings of the Boston Society of Natural History* 10, 312–319.

Marques, D. A., Meier, J. I., and Seehausen, O. (2019). A combinatorial view on speciation and adaptive radiation. *Trends in Ecology & Evolution* 34, 531–544. doi:10.1016/j.tree.2019.02.008.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12. doi:10.14806/ej.17.1.200.

Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., et al. (2013). Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research* 23, 1817–1828. doi:10.1101/gr.159426.113.

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., and Emerson, B. C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and *de novo* assembly optimization for population genetic inference. *Molecular Ecology Resources* 15, 28–41. doi:10.1111/1755-0998.12291.

Matschiner, M. (2019). *Species tree inference with SNP data, GitHub repository*. https://github.com/mmatschiner/tutorials/species_tree_inference_with_snp_data/

Matzke, N. J. (2013). *BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R scripts. R package version 1.1.2.* https://github.com/nmatzke/BioGeoBEARS/.

Matzke, N. J. (2014). Model Selection in Historical Biogeography Reveals that Founder-Event Speciation Is a Crucial Process in Island Clades. *Systematic Biology* 63, 951–970. doi:10.1093/sysbio/syu056.

Mayer, C., Sann, M., Donath, A., Meixner, M., Podsiadlowski, L., Peters, R. S., et al. (2016). BaitFisher: A software package for multispecies target DNA enrichment probe design. *Molecular Biology and Evolution* 33, 1875–1886. doi:10.1093/molbev/msw056.

Mayr, E. (1942). *Systematics and the origin of species from the viewpoint of a zoologist*. New York: Columbia University Press

Mayrose, I., Zhan, S. H., Rothfels, C. J., Magnuson-Ford, K., Barker, M. S., Rieseberg, L. H., et al. (2011). Recently formed polyploid plants diversify at lower rates. *Science* 333, 1257–1257. doi:10.1126/science.1207205.

McArthur, R. H., and Wilson, E. O. (1967). *Island biogeography*. Princeton, USA.: Princeton University Press.

McMullen, C. K. (1987). Breeding systems of selected Galápagos Islands Angiosperms. *American Journal of Botany* 74, 1694–1705. doi:10.1002/j.1537-2197.1987.tb08770.x.

McMurtry, G. M., Watts, P., Fryer, G. J., Smith, J. R., and Imamura, F. (2004). Giant landslides, mega-tsunamis, and paleo-sea level in the Hawaiian Islands. *Marine Geology* 203, 219–233. doi:10.1016/S0025-3227(03)00306-2.

McNab, B. K. (1994). Energy conservation and the evolution of flightlessness in birds. *The American Naturalist* 144, 628–642. doi:10.1086/285697.

Medeiros, A. C. (2004). "Phenology, reproductive potential, seed dispersal and predation, and seedling establishment of three invasive plant species in a Hawaiian rain forest." Dissertation. University of Hawaii at Manoa.

Meier, J. I., Marques, D. A., Mwaiko, S., Wagner, C. E., Excoffier, L., and Seehausen, O. (2017). Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature Communications* 8, 1–11. doi:10.1038/ncomms14363.

Mendizabal-Ruiz, G., Román-Godínez, I., Torres-Ramos, S., Salido-Ruiz, R. A., Vélez-Pérez, H., and Morales, J. A. (2018). Genomic signal processing for DNA sequence clustering. *PeerJ* 6, e4264. doi:10.7717/peerj.4264.

Meng, C., and Kubatko, L. S. (2009). Detecting hybrid speciation in the presence of incomplete lineage sorting using gene tree incongruence: A model. *Theoretical Population Biology* 75, 35–45. doi:10.1016/j.tpb.2008.10.004.

Meredith, R. W., Janečka, J. E., Gatesy, J., Ryder, O. A., Fisher, C. A., Teeling, E. C., et al. (2011). Impacts of the Cretaceous terrestrial revolution and KPg extinction on Mammal diversification. *Science* 334, 521–524. doi:10.1126/science.1211028.

Miller, J. S., and Venable, D. L. (2000). Polyploidy and the evolution of gender dimorphism in plants. *Science* 289, 2335–2338. doi:10.1126/science.289.5488.2335.

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., and Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site-associated DNA (RAD) markers. *Genome Research* 17, 240–248. doi:10.1101/gr.5681207.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* 37, 1530–1534. doi:10.1093/molbev/msaa015.

Mirarab, S., Bayzid, M. S., Boussau, B., and Warnow, T. (2014a). Statistical binning enables an accurate coalescent-based estimation of the avian tree. *Science* 346, 1250463 doi:10.1126/science.1250463.

Mirarab, S., Bayzid, M. S., and Warnow, T. (2016). Evaluating summary methods for multilocus species tree estimation in the presence of incomplete lineage sorting. *Systematic Biology* 65, 366–380. doi:10.1093/sysbio/syu063.

Mirarab, S., Reaz, R., Bayzid, M. S., Zimmermann, T., Swenson, M. S., and Warnow, T. (2014b). ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics* 30, i541–i548. doi:10.1093/bioinformatics/btu462.

Mirarab, S., and Warnow, T. (2015). ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* 31, i44–i52. doi:10.1093/bioinformatics/btv234.

Misof, B., Meusemann, K., von Reumont, B. M., Kück, P., Prohaska, S. J., and Stadler, P. F. (2014). A priori assessment of data quality in molecular phylogenetics. *Algorithms for Molecular Biology* 9, 22. doi:10.1186/s13015-014-0022-4.

Molloy, E. K., and Warnow, T. (2018). To include or not to include: The impact of gene filtering on species tree estimation methods. *Systematic Biology* 67, 285–303. doi:10.1093/sysbio/syx077.

Montaggioni, L. F. (2005). History of Indo-Pacific coral reef systems since the last glaciation: Development patterns and controlling factors. *Earth-Science Reviews* 71, 1–75. doi:10.1016/j.earscirev.2005.01.002.

Moore, B. R., Höhna, S., May, M. R., Rannala, B., and Huelsenbeck, J. P. (2016). Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. *Proceedings of the Royal Society* 113, 9569–9574. doi:10.1073/pnas.1518659113.

Moore, J. G., and Clague, D. A. (1992). Volcano growth and evolution of the island of Hawaii. *Geological Society of America Bulletin* 104, 1471–1484.

Moore, J. G., Clague, D. A., Holcomb, R. T., Lipman, P. W., Normark, W. R., and Torresan, M. E. (1989). Prodigious submarine landslides on the Hawaiian Ridge. *Journal of Geophysical Research* 94, 17465. doi:10.1029/JB094iB12p17465.

Morawetz, W. (1986). Remarks on karyological differentiation patterns in tropical woody plants. *Plant Systematics and Evolution* 152, 49–100. doi:10.1007/BF00985351.

Morden, C. W., and Harbin, S. C. (2013). Evolution and biogeographic origins of the endemic Hawaiian genus *Hesperomannia* (Asteraceae). *Pacific Science* 67, 219–235. doi:10.2984/67.2.6.

Morgan, J. R. (1996). *Hawai'i: a unique geography*. Honolulu, USA: Bess Press.

Morton, C. M., and Telmer, C. (2014). New subfamily classification for the Rutaceae. *Annals of the Missouri Botanical Garden* 99, 620–641. doi:10.3417/2010034.

Mueller-Dombois, D., and Boehmer, H. J. (2013). Origin of the Hawaiian rainforest and its transition states in long-term primary succession. *Biogeosciences* 10, 5171–5182. doi:10.5194/bg-10-5171-2013.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (1986). Specific enzymatic amplification of DNA *in vitro*: The polymerase chain reaction. *Cold Spring Harbor Symposium on Quantitative Biology* 51, 263–273. doi:10.1101/SQB.1986.051.01.032.

Murray, B. G., De Lange, P. J., and Ferguson, A. R. (2005). Nuclear DNA variation, chromosome numbers and polyploidy in the endemic and indigenous grass flora of New Zealand. *Annals of Botany* 96, 1293–1305. doi:10.1093/aob/mci281.

Nathan, R., Schurr, F. M., Spiegel, O., Steinitz, O., Trakhtenbrot, A., and Tsoar, A. (2008). Mechanisms of long-distance seed dispersal. *Trends in Ecology & Evolution* 23, 638–647. doi:10.1016/j.tree.2008.08.003.

NCBI Resource Coordinators (2018). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research* 46, D8–D13. doi:10.1093/nar/gkx1095.

Neall, V. E., and Trewick, S. A. (2008). The age and origin of the Pacific islands: a geological overview. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363, 3293–3308. doi:10.1098/rstb.2008.0119.

Nepokroeff, M., Sytsma, K. J., Wagner, W. L., and Zimmer, E. A. (2003). Reconstructing ancestral patterns of colonization and dispersal in the Hawaiian understory tree genus *Psychotria* (Rubiaceae): A comparison of parsimony and likelihood approaches. *Systematic Biology* 52, 820–838. doi:10.1080/10635150390251072.

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evololution* 32, 268–274. doi:10.1093/molbev/msu300.

NTBG (2019). The herbarium catalogue. Assessed 03.09.2019. Available at: https://ntbg.org/database/herbarium.

Nute, M., Chou, J., Molloy, E. K., and Warnow, T. (2018). The performance of coalescent-based species tree estimation methods under models of missing data. *BMC Genomics* 19, 286. doi:10.1186/s12864-018-4619-8.

Ono, M., and Masuda, Y. (1981). Chromosome numbers of some endemic species of the Bonin Islands II. *Ogasawara Research* 4, 1–24.

Otto, F. (1990). "DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA." in *Methods in cell biology: Flow Cytometry*, eds. Z. Darzynkiewicz and H. A. Crissman (London, UK: Academic Press), 105–110. doi:10.1016/S0091-679X(08)60516-6.

Paetzold, C., Kiehn, M., Wood, K. R., Wagner, W. L., and Appelhans, M. S. (2018). The odd one out or a hidden generalist: Hawaiian *Melicope* (Rutaceae) do not share traits associated with successful island colonization. *Journal of Systematics and Evolution* 56, 621–636. doi:10.1111/jse.12454.

Paetzold, C., Wood, K. R., Eaton, D. A. R., Wagner, W. L., and Appelhans, M. S. (2019). Phylogeny of Hawaiian *Melicope* (Rutaceae): RAD-seq resolves species relationships and reveals ancient introgression. *Frontiers in Plant Science* 10, e1074. doi:10.3389/fpls.2019.01074.

Palmer, M., Venter, S. N., McTaggart, A. R., Coetzee, M. P. A., Wyk, S. V., Avontuur, J. R., et al. (2019). The synergistic effect of concatenation in phylogenomics: the case in *Pantoea*. *PeerJ* 7, e6698. doi:10.7717/peerj.6698.

Pancho, J. V. (1971). IOPB chromosome number reports XXXIV. *Taxon* 20, 794–795.

Pannell, J. R. (2015). Evolution of the mating system in colonizing plants. *Molecular Ecology* 24, 2018–2037. doi:10.1111/mec.13087.

Paris, J. R., Stevens, J. R., and Catchen, J. M. (2017). Lost in parameter space: a road map for *stacks*. *Methods in Ecology and Evolution* 8, 1360–1373. doi:10.1111/2041-210X.12775.

Pease, J. B., Brown, J. W., Walker, J. F., Hinchliff, C. E., and Smith, S. A. (2018). Quartet sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany* 105, 385–403. doi:10.1002/ajb2.1016.

Pease, J. B., and Hahn, M. W. (2015). Detection and polarization of introgression in a five-taxon phylogeny. *Systematic Biology* 64, 651–662. doi:10.1093/sysbio/syv023.

Pelser, P. B., Kennedy, A. H., Tepe, E. J., Shidler, J. B., Nordenstam, B., Kadereit, J. W., et al. (2010). Patterns and causes of incongruence between plastid and nuclear Senecioneae (Asteraceae) phylogenies. *American Journal of Botany* 97, 856–873. doi:10.3732/ajb.0900287.

Percy, D. M., Garver, A. M., Wagner, W. L., James, H. F., Cunningham, C. W., Miller, S. E., et al. (2008). Progressive island colonization and ancient origin of Hawaiian *Metrosideros* (Myrtaceae). *Proceedings of the Royal Society B: Biological Sciences* 275, 1479–1490. doi:10.1098/rspb.2008.0191.

Poon, W.-S., Shaw, P.-C., Simmons, M. P., and But, P. P.-H. (2007). Congruence of molecular, morphological, and biochemical profiles in Rutaceae: a cladistic analysis of the subfamilies Rutoideae and Toddalioideae. *Systematic Botany* 32, 837–846. doi:10.1600/036364407783390692.

Porter, D. M. (1979). "Endemism and evolution in Galápagos Islands vascular plants." in *Plants and islands*, ed. D. Bramwell (London, UK: Academic Press), 225–258.

Porter, D. M. (1984). "Endemism and evolution in terrestrial plants." in *Galápagos*, ed. R. Perry (Oxford, UK: Pergamon Press), 85–99.

Pratt, T. K. (2009). "Origins and evolution." in *Conservation biology of Hawaiian forest birds: implications for island avifauna*, eds. T. K. Pratt, C. T. Attkinson, P. C. Banko, J. D. Jacobi, and B. L. Woodworth (New Haven, USA: Yale University Press), 3–24.

Pratt, T. K., Attkinson, C. T., Banko, P. C., Jacobi, J. D., and Woodworth, B. L. eds. (2009). *Conservation biology of Hawaiian forest birds: implications for island avifauna*. New Haven, USA: Yale University Press.

Price, J. P. (2004). Floristic biogeography of the Hawaiian Islands: influences of area, environment and paleogeography. *Journal of Biogeography* 31, 487–500. doi:10.1046/j.0305-0270.2003.00990.x.

Price, J. P., and Clague, D. A. (2002). How old is the Hawaiian biota? Geology and phylogeny suggest recent divergence. *Proceedings of the Royal Society B: Biological Sciences* 269, 2429–2435. doi:10.1098/rspb.2002.2175.

Price, J. P., and Elliott-Fisk, D. (2004). Topographic history of the Maui Nui complex, Hawai'i, and its implications for biogeography. *Pacific Science* 58, 27–45. doi:10.1353/psc.2004.0008.

Price, J. P., and Wagner, W. L. (2004). Speciation in Hawaiian Angiosperm lineages: Cause, consequence, and mode. *Evolution* 58, 2185–2200. doi:10.1111/j.0014-3820.2004.tb01597.x.

Price, J. P., and Wagner, W. L. (2018). Origins of the Hawaiian flora: Phylogenies and biogeography reveal patterns of long-distance dispersal. *Journal of Systematics and Evolution* 56, 600–620. doi:10.1111/jse.12465.

Price, J. R. (1963). "The distribution of alkaloids in the Rutaceae." in *Chemical plant taxonomy.* (London, UK: Academic Press), 429–452.

Puritz, J. B., Matz, M. V., Toonen, R. J., Weber, J. N., Bolnick, D. I., and Bird, C. E. (2014). Demystifying the RAD fad. *Molecular Ecology* 23, 5937–5942. doi:10.1111/mec.12965.

Pyle, D. M. (2015). "Sizes of volcanic eruptions," in *The encyclopedia of volcanoes. Second edition*, ed. H. Sigurdsson (Amsterdam, the Netherlands: Academic Press), 257–264. doi:10.1016/B978-0-12-385938-9.00013-4.

R Core Team (2019). *R: A language and environment for statistical computing.* Available at: https://www.R-project.org/.

Rabosky, D. L. (2014). Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS ONE* 9, e89543. doi:10.1371/journal.pone.0089543.

Rabosky, D. L., Donnellan, S. C., Grundler, M., and Lovette, I. J. (2014a). Analysis and visualization of complex macroevolutionary dynamics: An example from Australian Scincid lizards. *Systematic Biology* 63, 610–627. doi:10.1093/sysbio/syu025.

Rabosky, D. L., Grundler, M., Anderson, C., Title, P., Shi, J. J., Brown, J. W., et al. (2014b). BAMMtools: an *R* package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution* 5, 701–707. doi:10.1111/2041-210X.12199.

Rattenbury, J. A. (1957). Chromosome numbers in New Zealand Angiosperms. *Transactions of the Royal Society of New Zealand* 84, 936–938.

Ree, R. H., and Hipp, A. L. (2015). "Inferring phylogenetic history from restriction site associated DNA (RADseq)." in *Next-Generation sequencing in plant systematics,* eds. E. Hörandl and M. S. Appelhans (Königstein, Germany: Koeltz Scientific Books), 181–204.

Ree, R. H., Moore, B. R., Webb, C. O., and Donoghue, M. J. (2005). A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59, 2299–2311. doi:10.1111/j.0014-3820.2005.tb00940.x.

Ree, R. H., and Smith, S. A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* 57, 4–14. doi:10.1080/10635150701883881.

Rieseberg, L. H. (1997). Hybrid origins of plant species. *Annual Review of Ecology and Systematics* 28, 359–389. doi:10.1146/annurev.ecolsys.28.1.359.

Rivers, D. M., Darwell, C. T., and Althoff, D. M. (2016). Phylogenetic analysis of RAD-seq data: examining the influence of gene genealogy conflict on analysis of concatenated data. *Cladistics* 32, 672–681. doi:10.1111/cla.12149.

Roch, S., Nute, M., and Warnow, T. (2019). Long-branch attraction in species tree estimation: Inconsistency of partitioned likelihood and topology-based summary methods. *Systematic Biology* 68, 281–297. doi:10.1093/sysbio/syy061.

Roch, S., and Warnow, T. (2015). On the robustness to gene tree estimation error (or lack thereof) of coalescent-based species tree methods. *Systematic Biology* 64, 663–676. doi:10.1093/sysbio/syv016.

Roderick, G. K. (1997). Herbivorous insects and the Hawaiian silversword alliance: Coevolution or Cospeciation? *Pacific Science* 51, 440-449

Ronquist, F. (1997). Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. *Systematic Biology* 46, 195–203. doi:10.1093/sysbio/46.1.195.

Roy, T., Cole, L. W., Chang, T.-H., and Lindqvist, C. (2015). Untangling reticulate evolutionary relationships among New World and Hawaiian mints (Stachydeae, Lamiaceae). *Molecular Phylogenetics and Evolution* 89, 46–62. doi:10.1016/j.ympev.2015.03.023.

Rubin, B. E. R., Ree, R. H., and Moreau, C. S. (2012). Inferring phylogenies from RAD sequence data. *PLoS ONE* 7, e33394. doi:10.1371/journal.pone.0033394.

Russo, C. A., Takezaki, N., and Nei, M. (1995). Molecular phylogeny and divergence times of drosophilid species. *Molecular Biology and Evolution* 12, 391–404. doi:10.1093/oxfordjournals.molbev.a040214.

Sakai, A. K., Wagner, W. L., Ferguson, D. M., and Herbst, D. R. (1995). Origins of dioecy in the Hawaiian flora. *Ecology* 76, 2517–2529. doi:10.2307/2265825.

Sakai, A. K., Wagner, W. L., and Mehrhoff, L. A. (2002). Patterns of endangerment in the Hawaiian flora. *Systematic Biology* 51, 276–302. doi:10.1080/10635150252899770.

Salichos, L., and Rokas, A. (2013). Inferring ancient divergences requires genes with strong phylogenetic signals. *Nature* 497, 327–331. doi:10.1038/nature12130.

Salichos, L., Stamatakis, A., and Rokas, A. (2014). Novel information theory-based measures for quantifying incongruence among phylogenetic trees. *Molecular Biology and Evolution* 31, 1261–1271. doi:10.1093/molbev/msu061.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* 74, 5463–5467. doi:10.1073/pnas.74.12.5463.

Sanmartín, I., and Meseguer, A. S. (2016). Extinction in phylogenetics and biogeography: From timetrees to patterns of biotic assemblage. *Frontiers in Genetics* 7, 35. doi:10.3389/fgene.2016.00035.

Sax, D. F., and Gaines, S. D. (2008). Species invasions and extinction: The future of native biodiversity on islands. *Proceedings of the National Academy of Sciences* 105, 11490–11497. doi:10.1073/pnas.0802290105.

Schluter, D. (2000). *The ecology of adaptive radiation*. New York, USA: Oxford University Press.

Schmidt-Lebuhn, A. N., Aitken, N. C., and Chuah, A. (2017). Species trees from consensus single nucleotide polymorphism (SNP) data: Testing phylogenetic approaches with simulated and empirical data. *Molecular Phylogenetics and Evolution* 116, 192–201. doi:10.1016/j.ympev.2017.07.018.

Schranz, E. M., Mohammadin, S., and Edger, P. P. (2012). Ancient whole genome duplications, novelty and diversification: the WGD radiation lag-time model. *Current Opinion in Plant Biology* 15, 147–153. doi:10.1016/j.pbi.2012.03.011.

Seehausen, O. (2004). Hybridization and adaptive radiation. *Trends in Ecology & Evolution* 19, 198–207. doi:10.1016/j.tree.2004.01.003.

Seehausen, O. (2006). African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences* 273, 1987–1998. doi:10.1098/rspb.2006.3539.

Seehausen, O. (2013). Conditions when hybridization might predispose populations for adaptive radiation. *Journal of Evolutionary Biology* 26, 279–281. doi:10.1111/jeb.12026.

Seo, T.-K. (2008). Calculating bootstrap probabilities of phylogeny using multilocus sequence data. *Molecular Biology and Evolution* 25, 960–971. doi:10.1093/molbev/msn043.

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52, 591–611. doi:10.2307/2333709.

Shen, X.-X., Hittinger, C. T., and Rokas, A. (2017). Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nature Ecology & Evolution* 1, 0126. doi:10.1038/s41559-017-0126.

Silvestro, D., Zizka, G., and Schulte, K. (2014). Disentangling the effects of key innovations on the diversification of Bromelioideae (Bromeliaceae). *Evolution* 68, 163–175. doi:10.1111/evo.12236.

Simpson, G. G. (1953). *The major features of evolution.* New York, USA: Columbia University Press.

Smith-White, S. (1954). Chromosome numbers in the Boronieae (Rutaceae) and their bearing on the evolutionary development of the tribe in the Australian flora. *Australian Journal of Botany* 2, 287–303. doi:10.1071/bt9540287.

Solís-Lemus, C., Yang, M., and Ané, C. (2016). Inconsistency of species tree methods under gene flow. *Systematic Biology* 65, 843–851. doi:10.1093/sysbio/syw030.

St. John, H. (1944). Diagnoses of Hawaiian species of *Pelea* (Rutaceae) - Hawaiian plant studies 13. *Lloydia* 7, 265–274.

Stace, H. M., Armstrong, J. A., and James, S. H. (1993). Cytoevolutionary patterns in Rutaceae. *Plant Systematics and Evolution* 187, 1–28. doi:10.1007/BF00994089.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi:10.1093/bioinformatics/btu033.

Staudigel, H., and Koppers, A. A. P. (2015). "Seamounts and island building." in *The encyclopedia of volcanoes. Second edition*, ed. H. Sigurdsson (Amsterdam: Academic Press), 405–421. doi:10.1016/B978-0-12-385938-9.00022-5.

Stebbins, G. L. (1951). *Variation and evolution in plants.* New York: Columbia University Press.

Stegemann, S., Keuthe, M., Greiner, S., and Bock, R. (2012). Horizontal transfer of chloroplast genomes between plant species. *Proceedings of the National Academy of Sciences* 109, 2434–2438. doi:10.1073/pnas.1114076109.

Stevens, P. F. (2001). Angiosperm phylogeny website. Version 14, July 2017 [and more or less continuously updated since]. Assessed 01.11.2019. Available at: h*ttp://www.mobot.org/MOBOT/research/APweb/.*

Stone, B. C. (1969). *The genus Pelea A. Gray. (Rutaceae, Evodiinae.) A taxonomic monograph. 3rd edition*. Berlin, Germany: J. Cramer Verlag.

Stone, B. C., Wagner, W. L., and Herbst, D. R. (1999). "Rutaceae," in *Manual of the flowering plants of Hawai'i. Revised edition.*, eds. W. L. Wagner, D. R. Herbst, and S. H. Sohmer (Honolulu, USA: University of Hawaii Press and Bishop Museum Press), 1174–1216.

Streicher, J. W., Miller, E. C., Guerrero, P. C., Correa, C., Ortiz, J. C., Crawford, A. J., et al. (2018). Evaluating methods for phylogenomic analyses, and a new phylogeny for a major frog clade (Hyloidea) based on 2214 loci. *Molecular Phylogenetics and Evolution* 119, 128–143. doi:10.1016/j.ympev.2017.10.013.

Stroud, J. T., and Losos, J. B. (2016). Ecological opportunity and adaptive radiation. *Annual Review of Ecology, Evolution, and Systematics* 47, 507–532. doi:10.1146/annurev-ecolsys-121415-032254.

Stuessy, T. F., and Crawford, D. J. (1998). "Chromosomal stasis during speciation in Angiosperms of oceanic islands." in *Evolution and speciation of island plants*, eds. T. F. Stuessy and M. Ono (Cambridge, UK: Cambridge University Press), 307–324.

Stuessy, T. F., Greimler, J., and Dirnbôck, T. (2005). "Landscape modification and impact on specific and genetic diversity in oceanic islands." in *Plant diversity and complexity patterns: Local, regional, and global dimensions: Proceedings of an international symposium held at the Royal Danish Academy of Sciences and Letters in Copenhagen, Denmark, 25-28 May, 2003*, eds. I. Friis, H. Balslev, and K. D. videnskabernes selskab (Copenhagen, Denmark: Kgl. Danske Videnskabernes Selskab), 89–102.

Stuessy, T. F., Marticorena, C., Rodriguez R., R., Crawford, D. J., and Silva O., M. (1992). Endemism in the vascular flora of the Juan Fernández Islands. *Aliso: A Journal of Systematic and Evolutionary Botany* 13, 297–307.

Stuessy, T. F., Takayama, K., López-Sepúlveda, P., and Crawford, D. J. (2014). Interpretation of patterns of genetic variation in endemic plant species of oceanic islands. *Botanical Journal of the Linnean Society* 174, 276–288. doi:10.1111/boj.12088.

Suarez-Gonzalez, A., Hefer, C. A., Christe, C., Corea, O., Lexer, C., Cronk, Q. C. B., et al. (2016). Genomic and functional approaches reveal a case of adaptive introgression from *Populus balsamifera* (balsam poplar) in *P. trichocarpa* (black cottonwood). *Molecular Ecology* 25, 2427–2442. doi:10.1111/mec.13539.

Suarez-Gonzalez, A., Lexer, C., and Cronk, Q. C. B. (2018). Adaptive introgression: a plant perspective. *Biology Letters* 14, 20170688. doi:10.1098/rsbl.2017.0688.

Suchan, T., Espíndola, A., Rutschmann, S., Emerson, B. C., Gori, K., Dessimoz, C., et al. (2017). Assessing the potential of RAD-sequencing to resolve phylogenetic relationships within species radiations: The fly genus *Chiastocheta* (Diptera: Anthomyiidae) as a case study. *Molecular Phylogenetics and Evolution* 114, 189–198. doi:10.1016/j.ympev.2017.06.012.

Suchan, T., Pitteloud, C., Gerasimova, N. S., Kostikova, A., Schmid, S., Arrigo, N., et al. (2016). Hybridization capture using RAD probes (hyRAD), a new tool for performing genomic analyses on collection specimens. *PLoS ONE* 11, e0151651. doi:10.1371/journal.pone.0151651.

Swofford, D. L. (2002). *Phylogenetic analysis using parsimony (*and other methods)*. Sunderland, MA, USA: Sinauer Associates.

Swofford, D. L. (2018). *PAUP*: phylogenetic analysis using parsimony (and other methods), version 4.a165. at https://people.sc.fsu.edu/~dswofford/ paut_test/.*

To, T.-H., Jung, M., Lycett, S., and Gascuel, O. (2016). Fast Dating Using Least-Squares Criteria and Algorithms. *Systematic Biology* 65, 82–97. doi:10.1093/sysbio/syv068.

Twyford, A. D., and Ennos, R. A. (2012). Next-generation hybridization and introgression. *Heredity* 108, 179–189. doi:10.1038/hdy.2011.68.

Volz, E. M., and Frost, S. D. W. (2017). Scalable relaxed clock phylogenetic dating. *Virus Evolution* 3. doi:10.1093/ve/vex025.

Wagner, C. E., Keller, I., Wittwer, S., Selz, O. M., Mwaiko, S., Greuter, L., et al. (2013). Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology* 22, 787–798. doi:10.1111/mec.12023.

Wagner, W. L., and Funk, V. A. eds. (1995). *Hawaiian biogeography: Evolution on a hot spot archipelago.* Washington, D.C., USA: Smithsonian Institution Press

Wagner, W. L., Bruegmann, M. M., Herbst, D. R., and Lau, J. Q. C. (1999a). Hawaiian vascular plants at risk: 1999. *Bishop Museum Occasional Papers* 60, 1–58.

Wagner, W. L., Herbst, D. R., and Sohmer, S. H. (1999b). *Manual of the flowering plants of Hawai'i. Revised edition*. Honolulu, USA: University of Hawaii Press.

Wagner, W. L., Weller, S. G., and Sakai, A. K. (2005). Monograph of *Schiedea* (Caryophyllaceae subfam. Alsinoideae). *Systematic Botany Monographs* 72.

Wallace, A. R. (1878). *Tropical nature, and other essays*. London, UK: Macmillan and Company.

Wang, Y., and Nakhleh, L. (2018). Towards an accurate and efficient heuristic for species/gene tree co-estimation. *Bioinformatics* 34, i697–i705. doi:10.1093/bioinformatics/bty599.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* 10, 57–63. doi:10.1038/nrg2484.

Wanke, S., Granados Mendoza, C., Müller, S., Paizanni Guillén, A., Neinhuis, C., Lemmon, A. R., et al. (2017). Recalcitrant deep and shallow nodes in *Aristolochia* (Aristolochiaceae) illuminated using anchored hybrid enrichment. *Molecular Phylogenetics and Evolution* 117, 111–123. doi:10.1016/j.ympev.2017.05.014.

Warnow, T. (2017). *Computational phylogenetics: An introduction to designing methods for phylogeny estimation*. Cambridge, UK: Cambridge University Press.

Warren, B. H., Simberloff, D., Ricklefs, R. E., Aguilée, R., Condamine, F. L., Gravel, D., et al. (2015). Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters* 18, 200–217. doi:10.1111/ele.12398.

Waterman, P. G. (1975). Alkaloids of the Rutaceae: their distribution and systematic significance. *Biochemical Systematics and Ecology* 3, 149–180. doi:10.1016/0305-1978(75)90019-8.

Webb, C. J., and Kelly, D. (1993). The reproductive biology of the New Zealand flora. *Trends in Ecology & Evolution* 8, 442–447. doi:10.1016/0169-5347(93)90007-C.

Welch, A. J., Collins, K., Ratan, A., Drautz-Moses, D. I., Schuster, S. C., and Lindqvist, C. (2016). The quest to resolve recent radiations: Plastid phylogenomics of extinct and endangered Hawaiian endemic mints (Lamiaceae). *Molecular Phylogenetics and Evolution* 99, 16–33. doi:10.1016/j.ympev.2016.02.024.

Wen, D., Yu, Y., Zhu, J., and Nakhleh, L. (2018). Inferring phylogenetic networks using *PhyloNet*. *Systematic Biology* 67, 735–740. doi:10.1093/sysbio/syy015.

Wendel, J. F. (2015). The wondrous cycles of polyploidy in plants. *American Journal of Botany* 102, 1753–1756. doi:10.3732/ajb.1500320.

Wenny, D. G., Sekercioglu, Ç. H., Cordeiro, N. J., Rogers, H. S., and Kelly, D. (2016). "Seed dispersal by fruit-eating birds." in *Why birds matter: Avian ecological function and ecosystem services*, eds. Ç. H. Sekercioglu, D. G. Wenny, and C. J. Whelan (Chicago, USA: University of Chicago Press), 107–146.

Wetterstrand, K. A. (2019). DNA sequencing costs: Data from the NHGRI genome sequencing program (GSP). Accessed 01.10.2019. Available at www.genome.gov/sequencingcostsdata.

Whelan, F., and Kelletat, D. (2003). Submarine slides on volcanic islands - a source for mega-tsunamis in the Quaternary. *Progress in Physical Geography: Earth and Environment* 27, 198–216. doi:10.1191/0309133303pp367ra.

Whittaker, R. J. (1998). *Island biogeography: Ecology, evolution and conservation. 1st edition*. Oxford, UK: Oxford University Press.

Whittaker, R. J., and Fernández-Palacios, J. M. (2010). *Island biogeography - Ecology, evolution, and conservation*. *2nd edition*. Oxford, UK: Oxford University Press.

Whittaker, R. J., Ladle, R. J., Araújo, M. B., Fernández-Palacios, J. M., Delgado, J. D., and Arévalo, J. R. (2007). The island immaturity – speciation pulse model of island evolution: an alternative to the "diversity begets diversity" model. *Ecography* 30, 321–327. doi:10.1111/j.0906-7590.2007.04761.x.

Whittaker, R. J., Triantis, K. A., and Ladle, R. J. (2008). A general dynamic theory of oceanic island biogeography: A general dynamic theory of oceanic island biogeography. *Journal of Biogeography* 35, 977–994. doi:10.1111/j.1365-2699.2008.01892.x.

Whittaker, R. J., Triantis, K. A., and Ladle, R. J. (2010). "A general dynamic theory of oceanic island biogeography - Extending the McArthur-Wilson theory to accommodate the rise and fall of volcanic islands," in *The theory of island biogeography revisited* (Princeton, New Jersey, USA.: Princeton University Press), 88–142.

Willyard, A., Wallace, L. E., Wagner, W. L., Weller, S. G., Sakai, A. K., and Nepokroeff, M. (2011). Estimating the species tree for Hawaiian *Schiedea* (Caryophyllaceae) from multiple loci in the presence of reticulate evolution. *Molecular Phylogenetics and Evolution* 60, 29–48. doi:10.1016/j.ympev.2011.04.001.

Wood, K. R., Appelhans, M. S., and Wagner, W. L. (2016). *Melicope oppenheimeri*, section *Pelea* (Rutaceae), a new species from West Maui, Hawaiian Islands: with notes on its ecology, conservation, and phylogenetic placement. *PhytoKeys*, 51–64. doi:10.3897/phytokeys.69.8844.

Wood, K. R., Appelhans, M. S., and Wagner, W. L. (2017). *Melicope stonei*, section *Pelea* (Rutaceae), a new species from Kauaʻi, Hawaiian Islands: with notes on its distribution, ecology, conservation status, and phylogenetic placement. *PhytoKeys*, 119–132. doi:10.3897/phytokeys.83.13442.

Wu, G. A., Prochnik, S., Jenkins, J., Salse, J., Hellsten, U., Murat, F., et al. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nature Biotechnology* 32, 656–662. doi:10.1038/nbt.2906.

Wu, Z.-Y., Monro, A. K., Milne, R. I., Wang, H., Yi, T.-S., Liu, J., et al. (2013). Molecular phylogeny of the nettle family (Urticaceae) inferred from multiple loci of three genomes and extensive generic sampling. *Molecular Phylogenetics and Evolution* 69, 814–827. doi:10.1016/j.ympev.2013.06.022.

Xu, Q., Chen, L.-L., Ruan, X., Chen, D., Zhu, A., Chen, C., et al. (2013). The draft genome of sweet orange (*Citrus sinensis*). *Nature Genetics* 45, 59–66. doi:10.1038/ng.2472.

Yampolski, C., and Yampolski, H. (1922). Distribution of sex forms in phanerogamic flora. *Bibliotheca Genetica* 3, 1–62.

Yang, Y., Morden, C. W., Sporck-Koehler, M. J., Sack, L., Wagner, W. L., and Berry, P. E. (2018). Repeated range expansion and niche shift in a volcanic hotspot archipelago: Radiation of C4 Hawaiian *Euphorbia* subgenus *Chamaesyce* (Euphorbiaceae). *Ecology and Evolution* 8, 8523–8536. doi:10.1002/ece3.4354.

Yi, T., Miller, A. J., and Wen, J. (2004). Phylogenetic and biogeographic diversification of *Rhus* (Anacardiaceae) in the Northern hemisphere. *Molecular Phylogenetics and Evolution* 33, 861–879. doi:10.1016/j.ympev.2004.07.006.

Yoder, J. B., Briskine, R., Mudge, J., Farmer, A., Paape, T., Steele, K., et al. (2013). Phylogenetic signal variation in the genomes of *Medicago* (Fabaceae). *Systematic Biology* 62, 424–438. doi:10.1093/sysbio/syt009.

Yu, Y., Degnan, J. H., and Nakhleh, L. (2012). The probability of a gene tree topology within a phylogenetic network with applications to hybridization detection. *PLoS Genetics* 8, e1002660. doi:10.1371/journal.pgen.1002660.

Yu, Y., Than, C., Degnan, J. H., and Nakhleh, L. (2011). Coalescent histories on phylogenetic networks and detection of hybridization despite incomplete lineage sorting. *Systematic Biology* 60, 138–149. doi:10.1093/sysbio/syq084.

Zhou, W., Ji, X., Obata, S., Pais, A., Dong, Y., Peet, R., et al. (2018). Resolving relationships and phylogeographic history of the *Nyssa sylvatica* complex using data from RAD-seq

and species distribution modeling. *Molecular Phylogenetics and Evolution* 126, 1–16. doi:10.1016/j.ympev.2018.04.001.

Zhu, J., Liu, X., Ogilvie, H. A., and Nakhleh, L. K. (2019). A divide-and-conquer method for scalable phylogenetic network inference from multilocus data. *Bioinformatics* 35, i370–i378. doi:10.1093/bioinformatics/btz359.

Zhu, J., and Nakhleh, L. (2018). Inference of species phylogenies from bi-allelic markers using pseudo-likelihood. *Bioinformatics* 34, i376–i385. doi:10.1093/bioinformatics/bty295.

Zhu, J., Wen, D., Yu, Y., Meudt, H. M., and Nakhleh, L. (2018). Bayesian inference of phylogenetic networks from bi-allelic genetic markers. *PLoS Computational Biology* 14, e1005932. doi:10.1371/journal.pcbi.1005932.

Zhu, J., Yu, Y., and Nakhleh, L. (2016). In the light of deep coalescence: revisiting trees within networks. *BMC Bioinformatics* 17, 415. doi:10.1186/s12859-016-1269-1.

Zimmer, E. A., and Wen, J. (2015). Using nuclear gene data for plant phylogenetics: Progress and prospects II. Next-gen approaches. *Journal of Systematics and Evolution* 53, 371–379. doi:10.1111/jse.12174.

Zimmermann, T., Mirarab, S., and Warnow, T. (2014). BBCA: Improving the scalability of *BEAST using random binning. *BMC Genomics* 15, S11. doi:10.1186/1471-2164-15-S6-S11.

# List of Publications

Oelschlägel B, Nuss M, von Tschirnhaus M, **Pätzold C**, Neinhuis C, Dötterl S & Wanke S (2014) The betrayed thief: the unique strategy of *Aristolochia rotunda* to deceive its pollinators. *New Phytologist* 206(1)

Appelhans MS, Reichelt N, Groppo M, **Paetzold C** & Wen J (2018) Phylogeny and biogeography of the pantropical genus *Zanthoxylum* and its closest relatives in the proto-Rutaceae group (Rutaceae). *Molecular Phylogenetics and Evolution*, doi: 10.1016/j.ympev.2018.04.013

**Paetzold C**, Kiehn M, Wood KR, Wagner WL & Appelhans MS. (2018). The odd one out or a hidden generalist: Hawaiian *Melicope* (Rutaceae) do not share traits associated with successful island colonization. *Journal of Systematics and Evolution*. 56(6): 621-636. doi: 10.1111/jse.12454 (impact factor: 3.657)

**Paetzold C**, Wood KR, Eaton DAR, Wagner WL & Appelhans MS (2019). Phylogeny of Hawaiian *Melicope* (Rutaceae): RAD-Seq resolves species relationships and reveals ancient introgression. *Frontiers of Plant Science*, doi: 10.3389/fpls.2019.0107 (impact factor: 4.106)

Tomasello S, Karbstein K, Hodač L, **Paetzold C &** Hörandl E (2020). Phylogenomics unravels Quaternary vicariance and allopatric speciation patterns in temperate-montane plant species: A case study on the *Ranunculus auricomus* species complex. *Molecular Ecology* 29, doi: 10.1111/mec.15458

Appelhans M, **Paetzold C**, Wood KR & Wagner WL (2020). RADseq resolves the phylogeny of Hawaiian *Myrsine* L. (Primulaceae) and provides evidence for hybridization. *Journal of Systematics and Evolution*, doi: 10.1111/jse.12668

**Paetzold C**, Wood KR, Wagner WL & Appelhans M (submitted). Historical biogeography and diversification of Hawaiian *Melicope* (Rutaceae): flexibility is key. Submitted to *New Phytologist*

# Thesis Declarations

Declaration of the author's contribution to manuscripts with multiple authors.

## The odd one out or a hidden generalist: Hawaiian *Melicope* (Rutaceae) do not share traits associated with successful island colonization.

CP and MSA conceived and designed this study. MSA, CP, and KRW collected leaf material for the Flow Cytometry and data analysis. MK collected meristematic tissue for chromosome counts. CP conducted Flow Cytometry analysis. MK conducted chromosome counts. CP drafted this article and all authors contributed to writing and editing.

## Phylogeny of Hawaiian *Melicope* (Rutaceae): RAD-seq resolves species relationships and reveals ancient introgression

MSA, CP, and WLW conceived and designed the study. MSA, CP, and KRW collected the samples. CP carried out the laboratory work and performed all analyses. DARE provided valuable input for the analyses. CP drafted the article and all authors contributed to writing and editing.

## Historical Biogeography and diversification of Hawaiian *Melicope* (Rutaceae): flexibility is key.

MSA, CP, and WLW conceived and designed the study. CP performed all analyses and drafted the article. All authors contributed to writing and editing.

# Acknowledgments

I would like to thank Dr. Marc Appelhans for placing his trust in me to carry out this research, for his excellent supervision, inspiring discussion and insightful advice throughout these last years.

My sincerest thanks to Prof. Elvira Hörandl for helpful discussions and mentorship throughout this project, for providing opportunities to broaden my horizons and become involved in some of the department's projects. I would also like to express my gratitude to Prof. Alexander Schmidt for agreeing to be my second reviewer and for his helpful comments during the project's duration.

Fieldwork and species identification would have been impossible without the help of Ken Wood, Susann Ching, Steve Perlman, Jonathan Price, Kobe Togikawa, Kerri Fay, and Hank Oppenheimer. I would like to thank the Hawai‘i State Department of Land and Natural Resources (Permits: P-242, KPI2017-102, ODF-051316R, and MDF-092216A) for the permission to collect plants in forest reserves on Kaua‘i, O‘ahu, Maui, and Hawai‘i (Big Island); the Nature Reserve for the permission to collect plants at the Waikamoi Preserve on Maui; the Puu Kukui Watershed Preserve for the permission to collect along the Puu Kukui Trail; and the United States Fish & Wildlife Service for the permission to export samples (Permit: MA96221B-O) and C. Dang for help with collection permits.

This project was financially supported by the German Science Foundation (DFG; Grant AP 251/3-1 to MA).

I would also like to thank all my colleagues at the department for their help and advice for all kinds of issues and the many memories of tea breaks and cookies. A special thanks to Dr. Birthe Barke und Gabi Ließmann for many an entertaining lunch filled with silly and serious conversation. Thank you also to my friends in Bad Dürrenberg and Dresden, who lend long-distance helping hands when needed.

Finally, for unwavering support and love in all my endeavors, I am eternally grateful to my parents, without whom I would not have made it to this point.