Bringing home the lost vegetable sheep: A phylogenomic study of the Senecioneae genus *Haastia*.

A thesis submitted in partial fulfilment of the requirements for the Degree of Master of Science in Biology at the University of Canterbury.

By Kelvin McPherson.

University of Canterbury 2022

Acknowledgments:

Thank you to my supervisors, Illse Breitwieser, Pieter Pelser and Rob Smissen for the incredible support I was given. Through the many challenges and difficult times, I always felt confident I could rely on them.

Thank you to my amazing wife and family who believed in me and helped me and supported me throughout my thesis.

Thank you to Ines Schönberger and the rest of the Landcare staff for your guidance and welcoming attitude.

Thank you to the many university staff who taught and helped me.

Table of contents: Page

Abstract 3

1. Introduction 4

1.1 Senecioneae 4

1.1.1 Overview of Senecioneae 4

1.1.2 Overview of Brachyglottidinae 4

1.2 *Haastia* 4

1.2.1 Overview of *Haastia* 4

1.2.2 Pelser *et al.* 2007 5

1.3 Systematic challenges 6

1.3.1 Pelser *et al.* 2010 6

1.3.2 Taxon selection 7

1.3.3 Character selection 7

1.3.4 Sequencing 7

1.4 Aims and objectives 8

2 Methods 9

2.1 Sample selection 9

2.2 Character selection 9

2.3 Library preparation 10

2.4 Library enrichment 10

2.5 Data analysis 10

2.5.1 Assembly of nrDNA ITS regions from off-target sequences 10

2.5.2 Sequence trimming 10

2.5.3 Sequence assembly 10

2.5.4 Consensus sequence generation 11

2.6 Phylogenetic analysis 11

3 Results 12

3.1 Plastome phylogeny: Figure 6 12

3.2 ITS phylogeny: Figure 7 13

4 Discussion 13

4.1 Plastome phylogeny: Figure 6 13

4.2 ITS phylogeny: Figure 7 14

4.3 Congruence between figure 6 and figure 7 14

4.4 What this means for *Haastia* 14

4.5 What is next? 15

References 16

Abstract:

The genus *Haastia* has been shown to be polyphyletic and nested in the genus *Brachyglottis*. There is only low support for this in past phylogenies, however. This study aims to provide greater resolution and support of *Haastia's* placement among the sub-tribe Brachyglottidinae. We used more modern Next Generation Sequencing (NGS) techniques to generate a phylogenomic tree of the plastome. The results suggest that *Haastia* is monophyletic and sister to the core Brachyglottidinae clade, and not nested in *Brachyglottis.*

Further NGS of the Brachyglottidinae nuclear genome is suggested to provide further evidence for *Haastia's* delimitation and to provide greater resolution to other genera also nested in *Brachyglottis*.

1. Introduction

1.1 Senecioneae

1.1.1 Overview of Senecioneae

Senecioneae is the largest tribe within one of the largest families of flowering plants, Asteraceae, represented by over 3,000 species and about 150 genera (Nordenstam *et al.* 2009 ;Pelser et al. 2010). Senecioneae has a wide variety of plant morphologies and life-history strategies; these include, amongst others, sprawling vines, large shrubs, small herbs and trees(Barkley 1985). The tribe is found throughout nearly all continents, except Antarctica, and found in a broad range of habitats (Pelser *et al*. 2007). Senecioneae is characterised by having capitula with a commonly uniseriate involucre. This is a good character for distinguishing members of this tribe from those of other Asteraceae tribes (Pelser *et al*. 2007).

1.1.2 Overview of Brachyglottidinae

Senecioneae is divided into 5-7 subtribes (Nordenstam *et al.* 2009). Brachyglottidinae, also known as the *Brachyglottis* alliance, is a particularly morphologically diverse subtribe (Wagstaff and Breitwieser 2004; Breitwieser and Ward 2005; Pelser *et al*. 2007). Nearly the full spectrum of Senecioneae morphology is represented in Brachyglottidinae, and the same can be said for the range of environments the subtribe is found in (Kadereit and Jeffrey 2007). There are seven genera, with 56 species representing Brachyglottidinae (Kadereit 2007). The genera of Brachyglottidinae are *Brachyglottis* J.R.Forst. & G.Forst. (30 species), *Dolichoglottis* B.Nord. (2), *Haastia* Hook.f. (3)*, Traversia* Hook.f. (1), and *Urostemon* B.Nord. (1) being native to New Zealand, *Bedfordia* DC. (2) and *Centropapus* Hook.f. (1) being native to Australia, *Papuacalia* Veldkamp. (14) being native to New Guinea and *Acrisione* (Hook.et Arn.) B.Nord. (2) being native to Chile (Kadereit 2007; Pelser *et al*. 2007; Wagstaff and Breitwieser 2004).

1.2 *Haastia*

1.2.1 Overview of *Haastia*

Haastia is endemic to the highest reaches (Altitude 1,300m and up) of the New Zealand Southern Alps upon the scree and rocky slopes. This genus can survive due to its woolly leaves, which provide them cover from the elements, such as high winds, significant yearly variations in temperatures, snow and lack of shelter from the sun (Todd 1996). As the epithet of its scientific name suggests, *H. pulvinaris* has a round cushion-shaped growth form and can grow somewhat shrub-like in height but is most commonly prostrate. *Haastia recurva* (Hook.f.) and *H. sinclairii* (Hook.f.) are both completely prostrate and are not as compact as *H. pulvinaris.* The main difference between *H. sinclairii* and *H. recurva* is that the leaves of *H. recurva* are smaller, obovate and obtuse or subobtuse (20 x 10 mm), whereas those of *H. sinclairii* are larger, more oblong than obovate and acute or subacute (35 x 15 mm: Todd 1996, Fig 2). Each species of *Haastia* is composed of two varieties. In addition to the nomotypic varieties, these are *H. pulvinaris var. minor, H. sinclairii var. fulvida* and *H. sinclairii var. wallii*. The most recognisable differences are *H. pulvinaris var minor* has paler

and whiter leaves than *H. pulvinaris var. pulvinaris,* which are a more yellow-greyish colour, *H. sinclairii var. fulvida* has dark green leaves compared to the yellow-white leaves of *H. sinclairii var. sinclairii* due to its less dense tomentum, and *H. recurva var. wallii* has leaves that are greyish white whereas *H. recurva var. recurva* has leaves that are greyish-yellow to grey coloured (Allan 1961).

The genus *Haastia* and its three species were described in 1864 as obvious members of the family Asteraceae; however, its tribal position was less apparent (Wagstaff and Breitwieser 2002; Breitwieser and Ward 2005). Haastia was initially placed in the tribe Gnaphalieae (as Gnaphalioid Compositae) because the type species, *H*. *pulvinaris,* has a pulvinate growth form and a woolly tomentum. This inspired its common name “vegetable sheep”, which is also shared by some species of another genus in Gnaphalieae: *Raoulia* (Hook.f.), as some of its species also have those traits (Cockayne 1910; Fig 1). However, despite the shared pulvinate morphology, *Haastia* was recognised as not having some of the tribal characteristics. Hooker (1864) described *Haastia* as "A very singular and distinct genus, differing from the other Gnaphalioid Compositae in the tailless anthers". The affiliation with Gnaphalieae lasted until 1873 when Bentham & Hooker moved the genus to Astereae because they considered its style and anthers more diagnostic of that tribe. However, they did not mention why they considered them to be diagnostic of Astereae (Bentham & Hooker 1873). Later authors either placed *Haastia* in Gnaphalieae or Astereae until Wagstaff and Breitwieser (2002) researched the suitability of ITS data for resolving phylogenetic relationships by using New Zealand Asteraceae as their sample group. They included *H*. *pulvinaris* in their study and found it to be nested within Senecioneae instead of Gnaphalieae or Astereae (Wagstaff and Breitwieser 2002; see Figure 3). Breitwieser and Ward (2005) compared the morphological synapomorphies of the possible tribes *Haastia* could be affiliated with. They concluded that *Haastia* is different from Gnaphalieae because it does not have the unique two-layered sexine in the pollen wall found in this tribe, nor the absence of resin canals, caudate anthers, pollen with perforated columellae and papery involucral bracts. Compared to members of Astereae, *Haastia* does not have the unique adaxially glabrous style branches, anthers with radial endothecial tissue or radiate capitula with many slender ray florets (Breitwieser and Ward 2005). However, Breitwieser and Ward (2005) concluded that *Haastia* displays important morphological synapomorphies of Senecioneae, such as the uniseriate involucre around the capitulum.

1.2.2 Pelser *et al.* 2007

After the work by Wagstaff and Breitwieser (2004, 2002) and Breitwieser and Ward (2005), other systematic studies provided additional information about the evolutionary relationships of *Haastia*. In 2007 Pelser *et al*. attempted to infer a more complete ITS phylogeny of Senecioneae. Their results suggested that *Haastia* is a member of Brachyglottidinae. They also indicated that *Haastia* was nested in the genus *Brachyglottis* and that the latter genus is not monophyletic in its current delimitation. Furthermore, *Haastia* appeared to be polyphyletic. *B. perdicioides* appeared to be the closest relative of *H. sinclairii* and *H. recurva. H. sinclairii* and *H. recurva* were also more closely related to two more clades of *Brachyglottis* than they are to *H. pulvinaris*. Monophyly is crucial for a genus

as it shows that all species within the genus are each other's closest relatives. Therefore, the delimitation of the genera and their species should represent evolutionary patterns. Monophyly is the standard by which genera are delimited in modern taxonomy, and can be seen in most, if not all, modern taxonomic studies (Ashlock 1971). The results of Pelser et al. (2007) thus raising questions about the delimitation of both genera (Figure 4; Pelser *et al*. 2007).

1.3 Systematic challenges

More work is needed to discover if *Haastia* is monophyletic. The phylogeny in Pelser *et al.* (2007) and Wagstaff and Breitwieser (2004) helps to identify some of the challenges that need to be addressed to determine the phylogenetic placement of *Haastia* amongst Brachyglottidinae. In the remainder of this chapter, I will discuss the specific challenges and the reasons for the possible correct or incorrect phylogenies of *Haastia*.

1.3.1 Pelser *et al.* 2010

Pelser *et al.* 2010 also included *H. pulvinaris* in an ITS/ETS phylogeny of Senecioneae. However, it was now in a clade with Acrisione cymosa, Urostemon kirkii, Brachyglottis repanda, Traversia baccharoides, Papuacalia dindondl, Bedfordia arborescens and Centropappus brunonis. However, this clade had poor resolution, and therefore any new relationships between the Brachyglottidinae genera cannot be inferred. Pelser *et al.* (2010) investigated possible reasons for incongruence between the plastid and nuclear results in their 2007 paper. Although *Haastia* was not represented in the plastid data, all other Brachyglottidinae except *Traversia* was. Therefore, this paper may highlight some of the systematic challenges relating to *Haastia* as well. The two main reasons for incongruence that were theorised are hybridisation and incomplete lineage sorting (ILS). Hybridisation is well documented in Australasian Senecioneae, including *Haastia,* and has played a major role in the evolution and diversification of these taxa (Liew *et al*. 2018, Pelser *et al*. 2012, Todd 1996). Hybridisation can complicate phylogeny reconstruction, however. One of the issues that have been found in Australasian Senecioneae are chimeric ITS sequences and can therefore suggest the taxa comes from either lineage (Pelser *et al*. 2012). It is possible that hybridisation has occurred among *Haastia* species, among *Brachyglottis* species*,* or between both genera, although the latter is less likely as very few *Brachyglottis* species share their range with *Haastia* species. This study used DNA sequences of the plastid genome and compared it with nuclear data and found that there was incongruence amongst the datasets, although not so much within Brachyglottidinae. The two possibilities that were found to be feasible reasons for the incongruence are ILS and past hybridisation. It was discussed how taxa that undergo rapid species radiation are more likely to be subject to ILS, which is especially concerning to this study, as genera within Brachyglottidinae have had a history of rapid species radiation (Pelser *et al.* 2010; Wagstaff and Breitwieser. 2004). However, the research suggested that most of the incongruence in Senecioneae was due to hybridisation rather than ILS, and more importantly, that Brachyglottidinae was not amongst the taxa that may be incongruent due to ILS (Pelser *et al*. 2010)

1.3.2 Taxon selection

Pelser *et al.* (2007) were relatively thorough in their taxon selection of Senecioneae as a whole, although the emphasis of their study was on revising the tribe at a generic level, with a focus on the genus *Senecio* (L.). Although all *Haastia* species were represented, the focus on *Senecio* meant that some potentially informative Brachyglottidinae were not represented. Heath *et al.* (2008) reviewed a broad range of phylogenetic studies and was able to show that denser taxon sampling increased phylogenetic accuracy. Since it is unclear where the 3 *Haastia* species fall within Brachyglottidinae, it is important to add more taxa for the reason of increasing phylogenetic accuracy.

1.3.3 Character selection

In both Pelser *et al*. 2007 and 2010, branch support throughout the Brachyglottidinae clade is low (Figure 4). Pelser *et al.* 2007 and 2010 both used Bayesian inference posterior probabilities and maximum likelihood bootstrap values to display branch support. Posterior probabilities are shown to be more accurate but have an increased chance of erroneous values as they are less conservative than bootstrap values (Erixon *et al.* 2003). Posterior probabilities and bootstrap values are commonly used because if both show high support, then we can have higher confidence in what relationship the branch is inferring. The low branch support could suggest that the true phylogeny for Brachyglottidinae, and therefore *Haastia,* could be quite different to what either phylogeny suggests. However, the short branch lengths in these and previous phylogenies of Brachyglottidinae suggests that this subtribe has undergone recent species radiation of about five to two million years ago (Pelser *et al.* 2010; Wagstaff and Breitwieser 2004). If a clade has only recently radiated, then some DNA regions may not have accumulated enough variation to be informative for suggesting their phylogeny. This could indicate that the data set used by Pelser *et al*. (2007) does not have enough variation (Total length of 1,197 base pairs, 588 are variable, and 512 are parsimony informative) to inform a more correct phylogeny of *Haastia*.

1.3.4 Sequencing

So far, all phylogenies of *Haastia* have been obtained from ITS (Internal Transcribed Spacer) or trnK/matK sequences DNA sequences (Pelser *et al.* 2010, 2007; Wagstaff and Breitwieser 2004, 2002). These data were generated using the Sanger sequencing method, which was published by Fred Sanger in 1977 and became the primary method used for DNA sequencing for about thirty years (Schuster 2007). -However, Sanger sequencing did not generate enough meaningful variation to obtain well-supported clades within Brachyglottidinae (Pelser *et al.* 2010, 2007; Wagstaff and Breitwieser 2004, 2002).

Next-generation sequencing (NGS) is a newer alternative to Sanger sequencing. NGS differs from Sanger sequencing in that instead of sequencing individual genes, introns or intergenic spacers it sequences millions of small DNA fragments in parallel. These DNA fragments overlap to create contiguous sequences called contigs. Contigs are then used to comprise genomic rather than genetic data (Behjati and Tarpey 2013). Although Sanger sequencing can generate genomic data, it is extremely costly and time-consuming compared to NGS which can generate genomic data at a fraction of the cost and within a matter of weeks or

even days (Schuster 2007). For *Haastia* and other Brachyglottidinae, meaningful sequence variation is required to generate a better resolved and supported phylogeny. NGS could be used for this purpose with its genomic approach as it may find the variation needed. NGS is often used for other phylogenies, complementing or being an alternative to Sanger sequence data by using both phylogenetic and phylogenomic approaches (Hou *et al.* 2016; Fan *et al.* 2015).

1.4 Aims and objectives

I plan to address these challenges by using next-generation sequencing (NGS) methods instead of the Sanger sequencing approach used by Pelser et al. (2007), to potentially include a larger number of informative characters in my phylogenetic analyses. I will also expand the taxon sampling of Brachyglottidinae in this study. This is expected to result in a better resolved and supported phylogenetic hypothesis for an increased number of taxa that are closely related to the three *Haastia* species.

This study aims to contribute to a monophyletic generic delimitation of Brachyglottidinae by determining if *Haastia* is monophyletic and revising its delimitation if this is not the case. Its specific objectives are to:

1. Create a plastome phylogeny of *Haastia* using next-generation sequencing (NGS) methods and increased taxon sampling compared to previous studies. This will help resolve previously unsupported clades and result in a phylogenetic hypothesis that includes all *Haastia* species and its closest relatives.
2. Use this phylogeny to propose a revised generic classification of *Haastia* in which only monophyletic genera are recognised.

Even though this research is focused on *Haastia,* it may also provide insights to help more fully understand the phylogenetic relations of other closely related species or genera. This is especially important as some Brachyglottidinae species are of conservation concern, with one taxon listed as data deficient, three as nationally critical, one as nationally endangered, two as nationally vulnerable, one as risk and one as relict (de Lange *et al*. 2018). A more correct understanding delimitation of these Brachyglottidinae and their taxonomy is needed to guide conservation efforts more fully; this is because defining taxonomic units allows conservation planning and assessment for said units (Mace 2004). Even though *Haastia* is a New Zealand genus, the achievement of these objectives may also create a foundation for studies of Brachyglottidinae species outside of New Zealand; as this research may impact the placement of non-New Zealand taxa.

Using modern phylogenomic techniques, we can hopefully achieve these objectives and help bring home the lost (vegetable) sheep.

2. Methods

2.1 Sample selection

As per 1.3.2, I needed to achieve a higher level of taxon sampling so that I would be able to provide information on *Haastia's* placement relative to Brachyglottidinae, which was not included in Pelser *et al.* 2007. I also needed to include taxa from the previous phylogenies so my phylogeny would be comparable for analysis. Comparison to the previous phylogenies is important because they give the most recent placement of *Haastia* (Pelser *et al.* 2010, 2007; Wagstaff and Breitwieser 2004, 2002). All Brachyglottidinae ITS sequences from Pelser *et al.* 2007 were downloaded from GenBank to supplement the taxon sampling and make the analysis more comparable to the phylogeny in that study. I aimed to include at least one representative of each Brachyglottidinae genus in my analyses. Because nearly all other genera of this subtribe, including *Haastia*, were resolved as nested among *Brachyglottis* species in the Pelser *et al.* (2007) ITS phylogeny, I needed to ensure sufficient representation from the different *Brachyglottis* lineages identified in that study to avoid the effects of low taxon sampling (further detail in 1.3.2). To do this, I prioritised sequencing *Brachyglottis* species that were resolved as most closely related to *Haastia* species and those in phylogenetic positions that caused *Haastia* to be polyphyletic in Pelser *et al*. (2007).

I used 11 DNA samples that were previously extracted and used for Pelser *et al*. (2007). 21 samples used for my study were obtained from the DNA collection of Steve J. Wagstaff held at Manaaki Whenua Landcare Research, Lincoln. In addition, I newly extracted 8 DNA samples from herbarium specimens at CANU. For this, I used the Qiagen DNeasy mini kit with the same method as Pelser *et al*. 2007. All Brachyglottidinae ITS sequences from Pelser *et al.* 2007 were downloaded from GenBank to supplement the taxon sampling and make the analysis more comparable to the phylogeny in that study.

2.2 Character selection

The plastid genome is ideal for the purpose of resolving phylogenetic analyses among Brachyglottidinae taxa as it has been found to contain many informative characters that are useful across many taxonomic levels and has been used to resolve previously unresolved and unsupported phylogenies (Hollingsworth *et al.* 2016; Stull *et al.* 2013; Kane *et al*. 2012; Nock *et al*. 2011). Pelser *et al.* (2010) infers a plastid phylogeny of Senecioneae, including Brachyglottidinae, although without *Haastia* and *Traversia*. Sanger sequencing was used in Pelser *et al.* (2010) to obtain plastid reads of *ndhF* gene, the *trnL* intron, and the *psbA-trnH, 5' and 3' trnK*, and the *trnL-F* intergenic spacers of the plastid genome. Studies using sanger sequencing often used these short reads of the plastid genome, however, these were often supplementary to nuclear regions because the short plastid reads were too conserved and has less discriminatory power. However, when using a NGS genomic approach, the accumulative informative characters over the whole plastid genome make it sufficient for many phylogenetic purposes (Hollingsworth *et al.* 2016, 2011). Therefore, this study will use an NGS approach to obtain genomic plastid data.

2.3 Library preparation

A Qubit fluorometer was used to quantify DNA samples as required. DNA was amplified and barcoded using an Illumina TruSeq Nano DNA library prep kit, using the manufacturer's 2015 protocol. Most reactions were done with half scales to use less of the reagents. DNA sonication was done using the COVARIS machine, using waveguide 526 and rack 500609. Settings for sonication are as follows: Duration (s) 75, Peak power 40, Duty % factor 25, Cycles/burst 50 and Avg power 10. During the PCRnano step, the cycles were increased to 16 as the library preparation kept failing, but this gave usable results.

2.4 Library enrichment

Amplified and barcoded samples were then pooled into libraries of between 4 to 6 samples. Libraries were then run through targeted enrichment which uses oligonucleotide probes (also called baits). These baits are RNA strands that attach themselves to complementary DNA and RNA strands which then facilitates the hybridisation-based targeted enrichment, therefore enabling us to sequence mostly plastid DNA (Stull *et al*. 2013). Based on Stull et al. (2013), custom baits were used using the MyBaits (Arbor Biosciences) kit, protocol 5.1. These baits target angiosperm plastid genomes and was used at a hybridisation temperature of 50°C (Smissen and Scheele 2022). Libraries were sequenced on a MiSeq machine by Massey Genome Services.

2.5 Data analysis

2.5.1 Assembly of nrDNA ITS regions from off-target sequences

The ITS region is what all previous phylogenies of *Haastia* have included as well as 5' trnK/matK from the plastid. This would be advantageous to sample again because we can increase the taxon sampling and therefore help to alleviate impacts of low taxon sampling that may be present in the past phylogenies (Pelser *et al.* 2010, 2007; Wagstaff and Breitwieser 2004, 2002). The ITS region was able to be sequenced, along with the plastid genome as off-target sequences. Making this an affordable way to include this data set.

2.5.2 Sequence trimming

Adapter regions, where the indices attach to the DNA, can cause many downstream issues with data analysis. These were therefore trimmed from the sequences that were generated for this study using Trimmomatic (Bolger *et al*. 2014).

2.5.3 Sequence assembly

Plastid contigs were mapped to a *Dendrosenecio johnstonii* ((Oliv) B.Nord) plastome (GenBank NC037959) using BWA (Burrows-Wheeler Aligner; Li and Durbin 2009). If a sample's total reads have an average coverage of less than 10, then they are not used beyond this point. This is because it shows that not enough of the sample was sequenced for it to be useful for this analysis. Mapping contigs requires a closely related taxon so that contigs will be aligned in the correct order, creating a contiguous genomic sequence. Therefore, I chose this species because it is the most closely related species to the Brachyglottidinae, for which a fully annotated plastome is available on GenBank.

Dendrosenecio johnstonii is another member of Senecioneae. Therefore, I considered it sufficiently closely related for this purpose. A *Brachyglottis repanda* ITS sequence from Pelser *et al.* (2007) (GenBank AY554103) was used to map the newly generated ITS sequences with BWA.

2.5.4 Consensus sequence generation

For the plastid sequences, I used three consensus sequence thresholds (85%, 90% and 95%) using the Samtools consensus function (Danecek *et al*. 2021). The lower the consensus threshold, the fewer ambiguities, but a higher consensus threshold gives less chance for systematic errors, such as a misaligned patch of nucleotides changing the consensus sequence. Therefore, I determined which of the three consensus sequence thresholds resulted in the fewest ambiguities for each set of contigs without compromising it to systematic errors. For the ITS contigs, consensus sequences were generated using Geneious (version 10.2.6; https://www.geneious.com). This software was used instead of Samtools because the ITS region is relatively short (i.e. the ITS sequence for my samples is 651 base pairs long, whereas the plastome sequences are over 150,000 base pairs long) and ambiguities could therefore be resolved manually.

The consensus sequences obtained from the different samples were aligned with each other in Geneious using the MAFFT alignment add-on and using default parameters (Katoh 2013). I annotated the plastid sequences using the annotations provided with the *D. johnstonii* plastome sequence. I annotated the ITS sequences using a previously annotated *B. repanda* sequence (GenBank AY554103). I added mask annotations to areas of the plastid alignment that seemed to result from poor alignment, sequencing error or other systematic errors. These masked regions were not used for the subsequent phylogenetic analyses. This risks losing potentially phylogenetic informative characters but also helps to reduce erroneous characters that could mislead the phylogeny reconstruction. Every sample contained two copies of an inverted repeat region so one was removed from each plastid alignment (at position 125,917 – 150,608), as it adds no new informative characters and could add erroneous characters.

2.6 Phylogenetic analysis

Paup* (Version 4.0a; Swofford 2003) was used to generate trees with a maximum parsimony optimality criterion. A bootstrap analysis was used with 5,000 replicates, limited to 100,000,000 rearrangements and Maxtrees set to increase by 100 when the maximum number of trees to be saved has been reached. Other settings were left at default values. MrBayes (Version 3.2.7a; Huelsenbeck and Ronquist 2001) was used to generate trees using Bayesian inference. The Bayesian inference was run using the GTR + G model of nucleotide substitution and the monte carlo markov chain was set to run for 10000000 generations. The Bayes block reads as follows “lset nst=mixed rates=gamma; mcmc nruns=4 ngen=10000000;”. The run was made to stop once the standard deviation of split

frequencies reached 0.01. RAxML-ng (Kozlov *et al.* 2019) was used to generate trees with a maximum likelihood optimality criterion. The parameters that were set are datatype = DNA, DNA gtrcat = GTR, num patterns = 15000, ratehet model = +G, runtime = 5, select analysis = all, specify bootstraps = 1000, specify bscutoff = 0.3, specify bsmetric = fbp, specify typebootstraps = true and the rest of the parameters were set at the default. As mentioned in 1.3.3, using multiple optimality criteria for phylogeny reconstruction is a way to take the potential impact of their limitations on the phylogenetic conclusions into account. Therefore, if all optimality criteria support the same topology, the result can be more trusted.

The outgroup set for the three analyses was an unknown presumably *Senecio* species. This sample was sequenced because it was labelled as an *Acrisione,* but when blasted on Genbank, it came back as a *Senecio* (See fig 5). For a published work, this would need to be further investigated, or an alternative sample used, but for the purposes of this thesis I thought it would be fine. If I were to choose an alternative outgroup, I would choose a *Dolichoglottis* species because the genus has been shown to be the most distantly related Brachyglottidinae to *Haastia*. This was chosen because *Senecio* is also a member of Senecioneae (subtribe Senecioninae) but relatively distantly related to Brachyglottidinae and I wanted to see if my dataset would also support *Dolichoglottis's* placement seen in the previous phylogenies (Pelser *et al.* 2010, 2007; Wagstaff and Breitwieser 2004, 2002). Figtree (Version 1.4.4; http://tree.bio.ed.ac.uk/software/figtree/) was used to visualise the trees generated from all analyses. I compared the different trees manually to determine if the results were congruent. The standards I used for statistical support is as follows: Bayesian posterior probabilities (PP) ≥ 0.95 and maximum likelihood and parsimony bootstrap support (BS) >75%.

3. Results

3.1 Plastome phylogeny: Figure 6

35 out of 39 samples representing 18 species and 2 unknown taxa were used in the plastid phylogeny, as four samples did not meet the standard of having an average nucleotide coverage above 10. The length of the aligned reads for individual samples was between 150,430 and 150,743 characters. After removing ambiguous and poorly aligned areas and the aforementioned inverted repeat region, the length of the total plastid alignment was 127,704 characters. Of these, 125,972 characters are constant (98.64%), 886 variable

characters are parsimony-uninformative, and 846 variable characters are parsimony-informative.
All three phylogeny reconstruction methods (i.e. BI, MP, ML) resulted in congruent topologies, and a single consensus tree is therefore shown in Fig. 6 (separate trees shown in Fig. 6.1-6.3). *Dolichoglottis* is sister taxa to the other Brachyglottidinae genera. *Haastia* is recovered here as monophyletic with high support (1.00 posterior probability, 100% MP and ML bootstrap support). It is resolved as a sister clade to a core Brachyglottidinae clade (1.00/77/84). *H. recurva* and *H. sinclairii* form a well-supported clade (1.00/100/100). The well-supported core Brachyglottidinae clade (CBC; 1.00/100/100) consists of *Brachyglottis*, *Centropappus, Bedfordia, Acrisione, Urostemon, Traversia* and *Papuacalia.* CBC contains both well-supported and poorly-supported subclades. The terminal clades of CBC are all highly supported, except for the clade containing *Brachyglottis bidwillii* var. *viridis* and *Brachyglottis adamsii,* which has a high posterior probability, but low ML and MP bootstrap support (0.99/66/61).

3.2 ITS phylogeny: Figure 7

The same samples that were excluded from the plastid data set were also removed from the ITS data set, along sequences of a *Haastia recurva x sinclairii* sample, because the reads had an average nucleotide coverage of 60, 241 (the next highest in all my samples was 741) with only 0.7% being identical sites. It seemed that this could be the result of PCR product contamination, so I thought it would be safer to leave this sample out. Since I already have multiple other *Haastia recurva* and *Haastia sinclairii* representatives in this phylogeny, I judged that this would be the safest option for achieving accurate results. The ITS sequences used in Pelser *et al.* (2007) were added to this dataset, totalling 59 ITS sequences in the alignment. The length of the ITS alignment is 651 characters. Of these, 449 characters are constant (68. 97%), 79 variable characters are parsimony-uninformative, and 123 variable characters are parsimony-informative.

All three phylogeny reconstruction methods yielded congruent trees (Fig. 7; Separate trees shown in Fig. 7.1-7.3). Low support is seen throughout most of the branches, especially around the basal nodes. *Dolichoglottis* is sister to the rest of Brachyglottidinae, with low support (0.89/54/<50). *Haastia sinclairii* and *H. recurva* form a well-supported clade (1.00/95/88) but are now in a polytomy with two clades of *Brachyglottis* species containing *B*. *monroi, B. compacta, B. greyi, B. lagopus* and *B. perdiciodes.* This clade is sister to a *H. pulvinaris* clade (1.00/76/56).

4 Discussion

4.1 Plastome phylogeny: Figure 6

The plastome phylogeny had fewer Brachyglottidinae species represented than Pelser *et al*. (2007), with 18 species with 2 unknown taxa (An unknown Acrisione species and unknown possibly *Senecio* species), while Pelser *et al.* (2007) has 25 species (*B. bellidiodes, B. haastii* and *B. southlandica* have since been lumped into *B. lagopus* in Millar *et al.* (2018)). However, the sparser representation was sufficient for this plastome phylogeny. This is because at least one representative of every Brachyglottidinae genera was included and

Haastia was supported as sister to the core Brachyglottidinae clade (CBC). If *Haastia* was instead nested within this clade, then questions could be raised as to whether the species not represented in this phylogeny could be more closely related to *Haastia* and therefore change *Haastia's* phylogenetic position. *Haastia's* position as sister to the CBC is different to all previous phylogenies (Pelser *et al.* 2010, 2007; Wagstaff and Breitwieser 2004, 2002), which showed *Haastia* as polyphyletic nested in two positions within *Brachyglottis*.

Our analyses recover relationships within the CBC, similar to results from previous phylogenies (Pelser *et al.* 2010, 2007; Wagstaff and Breitwieser 2004, 2002). The various genera are again highly supported as distinct clades, except for *Brachyglottis*, in which the other genera in the CBC are nested. The support for the relationships between the genera in the CBC remains low; therefore, the relationships between the genera are still unknown. *Bedfordia* remains sister to *Centropapus*, which is unsurprising considering they are the only Australian representatives of Brachyglottidinae.

4.2 ITS phylogeny: Figure 7

The ITS phylogeny shows the same patterns as previous phylogenies (Pelser *et al.* 2010, 2007; Wagstaff and Breitwieser 2004, 2002). However, there was an increase in the representation of taxa. The newly represented species in Fig. 7 are *D. lyallii, Bedfordia arborescens, P. glossophylla, and P. mogrere.* These species each still group with congeneric species with high support. This gives additional confidence in the distinct lineages of these genera.

4.3 Congruence between figure 6 and figure 7

The plastome phylogeny (Fig. 6) and the ITS phylogeny (Fig. 7) are not congruent regarding their placement of *Haastia.* However, the incongruence arises from the aforementioned clade inferred in the ITS phylogenies that suggest *H. sinclairii* and *H. recurva* are sister taxa of two clades of *Brachyglottis.* This relationship is poorly supported; therefore, it cannot be reliably inferred that *H. pulvinaris* is not sister to *H. recurva* and *H. sinclairii* when considering the ITS phylogeny. However, as mentioned in 4.1, it is not only shown that *Haastia* is a single clade, but it is sister taxa to the CBC

4.4 What this means for *Haastia*

The robustly supported position of *Haastia* as monophyletic and sister to the CBC suggests that taxonomic changes based on earlier phylogenies would have been premature. For example, Kadereit *et al.* (2007) indicate that because *Haastia* is not monophyletic in what were the most recent phylogenies of the time (Pelser *et al.* 2007; Wagstaff and Breitwieser 2004), the genus should be split so that *H. sinclairii* and *H. recurva* are in a different genus to *H. pulvinaris.* Garnock-Jones (2014) recognises the polyphyly of *Haastia* and instead suggests two other solutions. One is to lump *B. compacta, B. greyii, B. haastii, B. lagopus, B. monroi* and *B. perdiciodes* into *Haastia.* The second solution is to lump *Haastia* into *Brachyglottis* instead (Garnock-Jones 2014). These three solutions from Kadereit *et al.* (2007) and Garnock-Jones (2014) would all ensure that the species of *Haastia* would be in a monophyletic genus. However, these solutions fail to consider the low support for *Haastia*

being polyphyletic and nested within *Brachyglottis* (Pelser *et al.* 2007). The plastome data suggests (Fig. 6) that *Haastia* can be retained as a monophyletic genus including *H. recurva, H. sinclairii* and *H. pulvinaris,* distinct from the rest of Brachyglottidinae. The data is consistent with earlier studies in suggesting several genera of Brachyglottidinae are nested within Brachyglottis as currently circumscribed suggesting taxonomic change at the generic level still needs to be considered. The suggestion for including all Brachyglottidinae into *Brachyglottis* would still organise the species of *Haastia* as monophyletic (Garnock-Jones 2014). However, with these results (Fig. 6), this would be an extreme solution, and therefore this decision would be premature before other evidence is brought forth.

Other evidence to support the retention of *Haastia* as a genus could be it's ecological niche and possibly some morphological features. *Haastia* remains the only alpine genus of Brachyglottidinae, except for some *B. lagopus* that grew near some *H. pulvinaris* in Alimarlock, birch ranch (CHR 215904). Morphological features that differ between *Haastia* and *Brachyglottis* are as follows: The growth form of *Haastia* is cushion or mat-forming, whereas *Brachyglottis* consists of Trees, shrubs or herbs. *Haastia* leaves are densely imbricate, spathulate, apically crenulated, covered with long hairs, or alternate, patent to recurved, oblong and tomentose, whereas *Brachyglottis* has leaves that are cauline or rosulate, sessile or petiolate, entire or dentate to lobed and pinnately veined. The capitula of *Haastia* are sessile and terminal, whereas the capitula of *Brachyglottis* are scapose, or several and corymbose, and radiate. The anthers of *Haastia* are ecaudate, whereas the anthers of *Brachyglottis* are caudate or sagittate. The style branches of *Haastia* are linear with an apical conical tuft of obtuse hairs, or slender and are dorsally papillate-hairy whereas *Brachyglottis* style branches are somewhat discrete. The cypselae of *Haastia* are compressed or subterete, whereas the *Brachyglottis* cypselae are ribbed. The pappus of *Haastia* are coarse and basally flattened, whereas *Brachyglottis* pappus is slender and usually persistent (Kadereit *et al.* 2007).

4.5 What is next?

The plastome phylogeny can show *Haastia* in a different position than previously suggested by genetic studies (Pelser *et al.* 2010, 2007; Wagstaff and Breitwieser 2004, 2002). However, this is only a gene tree that gives insight into the evolution of the plastome of these taxa. One of the reasons the plastome tree may not represent the species tree is due to past hybridisation. Brachyglottidinae have a history of hybridisation between even morphologically distinct taxa (Drury 1973). How this can confuse phylogenies is described in 1.3.1. Because of the known history of hybridisation and the impacts it has on constructing species trees, the plastome alone cannot give us the full picture of the evolutionary relationships between Brachyglottidinae. To generate the molecular species tree, NGS techniques must be used on the nuclear genome of *Haastia* and other Brachyglottidinae. This would provide many more informative characters for the delimitation of *Haastia* and all Brachyglottidinae. The comparatively short ITS and plastid sequences of previous phylogenies could not give enough resolution and support to infer the relationships between some of the species or most of the genera. This thesis has shown that NGS data can help to provide phylogenetic resolution and support for Brachyglottidinae. Analyses of

multiple nuclear DNA regions using NGS sequence capture methods could provide independent support for our plastome analyses, providing further evidence for the monophyly of *Haastia,* as shown in fig x, and can possibly bring greater resolution and support to the clades within the CBC (Grover *et al.* 2012). A potential method that could be used is described in Schmidt-Lebuhn and Bovill (2021) who used nuclear phylogenomic data on a group of Australian Gnaphalieae. However, there is a history of polyploidy in Brachyglottidinae (Millar *et al.* 2018) which could complicate analyses of nuclear loci if not accounted for. Polyploidy is not an issue with the plastome because it is inherited uniparentally and therefore does not re-combine as the bi-parentally nuclear genome does (Van de Peer *et al.* 2017). Because of the additional complications of nuclear data and the different evolutionary processes compared to the plastid data, the results could be incongruent and therefore could suggest alternative placements for the species and genus of *Haastia* and the rest of Brachyglottidinae.

References:

Allan H.H., 1961. Flora of New Zealand. Vol 1. *Government Printer, Wellington.*

Ashlock, P. D. (1971). Monophyly and Associated Terms. *Systematic Biology, 20(1),* pp.63–69.

Barkley, T.M., 1985. Generic boundaries in the Senecioneae. *Taxon*, *34*(1), pp.17-21.

Behjati, S. and Tarpey, P.S., 2013. What is next generation sequencing? *Archives of Disease in Childhood - Education & Practice Edition, 98*(6), pp.236–238

Bentham, G. and Hooker, J.D., 1873. Senecionidae. *Genera Plantarum, 2*(1), pp.434-453.

Bolger, A.M., Lohse, M., and Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics. 30*(15), pp.2114-2120.

Breitwieser, I. and Ward, J.M., 2005. Morphological evidence for the tribal position of Haastia (Asteraceae). *New Zealand Journal of Botany*, *43*(4), pp.767-777.

Cockayne, L., 1910. New Zealand plants and their story. *Government Printer, Wellington.*

Dawson, M.I., 2000. Index of chromosome numbers of indigenous New Zealand spermatophytes. *New Zealand Journal of Botany, 38*(1), pp.47-150

de Lange, P.J., Rolfe, J.R., Barkla, J.W., Courtney, S.P., Champion, P.D., Perrie, L.R., Beadel, S.M., Ford, K.A., Breitwieser, I., Schonberger, I., Hindmarsh-Walls, R., Heenan, P.B. and Ladley, K., 2018. Conservation status of New Zealand indigenous vascular plants, 2017. New Zealand Threat Classification Series 22. *Department of Conservation, Wellington.*

Denecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O, Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M. and Li, H., 2021. Twelve years of SAMtools and BCFtools. *Gigascience*, 10(2).

Drury, D.G., 1973. Nodes and leaf structure in the classification of some Australasian shrubby Senecioneae-Compositae. *New Zealand Journal of Botany, 11*(3), pp.525-554.

Erixon, P., Svennblad, B., Britton, T. and Oxelman, B., 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Systematic Biology, 52*(5), pp.665-673.

Fan, H., Ives, A.R., Surget-Groba, Y. and Cannon, C.H., 2015. An assembly and alignment-free method of phylogeny reconstruction from next-generation sequencing data. *BMC Genomics, 16*(1).

Garnock-Jones, P.J., 2014. Evidence-based review of the taxonomic status of New Zealand's endemic seed plant genera. *New Zealand Journal of Botany*, *52*(2), pp.163-212.

Grover, C.E., Salmon, A. and Wendel, J.F., 2012. Targeted sequence capture as a powerful tool for evolutionary analysis. *American Journal of Botany, 99,* pp.312-319.

Heath, T.A., Hedtke, S.M. and Hillis, D.M., 2008. Taxon sampling and the accuracy of phylogenetic analyses. *Journal of Systematics and Evolution, 46*(3), pp.239-257

Hollingsworth, P. M., Li, D. Z., van der Bank, M., & Twyford, A. D., 2016. Telling plant species apart with DNA: from barcodes to genomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1702).

Hollingsworth, P.M., Graham, S.W., Little, D.P., 2011. Choosing and Using a Plant DNA Barcode. *PLOS ONE 6*(5).

Hooker, J.D., 1864. Handbook of the New Zealand flora. *London, Reeve*.

Hou, C., Wikström, N., Strijk, J.S. and Rydin, C., 2016. Resolving phylogenetic relationships and species delimitations in closely related gymnosperms using high-throughput NGS, Sanger sequencing and morphology. *Plant Systematics and Evolution, 302*(9), pp.1345-1365.

Huelsenbeck, J.P., and Ronquist. F., 2001. MrBayes: Bayesian inference of phylogeny. *Biometrics* ,*17*, pp.754-755.

Kane, N., Sveinsson, S., Dempewolf, H., Yang, J.Y., Zhang, D., Engels, J.M.M. and Cronk, Q., 2012. Ultra-barcoding in cacao (*Theobroma* spp.; Malvaceae) using whole chloroplast genomes and nuclear ribosomal DNA. *American Journal of Botany*, *99*, pp.320-329.

Kadereit, J.W. and Jeffrey, C., 2007. Flowering plants: eudicots: astreales. Vol 8. *Springer Science and Business Media.*

Katoh, K. and Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, *30*(4), pp.772-780.

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B. and Stamatakis, A., 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, *35*(21), pp.4453-4455.

Li, H. and Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics, 25*, pp.1754-1760.

Liew, C., Memory, A.E., Oritz-Barrientos, D., de Lange, P.J. and Pelser, P.B., 2018. The delimitation and evolutionary history of the Australasian Lautusoid group of *Senecio* (Asteraceae: Senecioneae). *Taxon, 67*(1), pp.130-148

Mace, G. M., 2004. The role of taxonomy in species conservation*. Philosophical Transactions of the Royal Society B: Biological Sciences, 359*(1444), pp.711–719

Millar, T.R., Breitwieser, I., Pelser, P.B. and Smissen, R.D., 2018. A new classification of rosette-forming Brachyglottis (Asteraceae: Senecioneae) recognising a single species: Brachyglottis lagopus. *New Zealand Journal of Botany*, *56*(3), pp.237-263.

Nock, C.J., Waters, D.L., Edwards, M.A., Bowen, S.G., Rice, N., Cordeiro, G.M. and Henry, R.J., 2011. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnology Journal, 9* pp.328-333

Nordenstam, B., Pelser P.B., Kadereit, J.W. and Watson, L.E., 2009. Senecioneae. *Systematics, Evolution, and Biogeography of Compositae*, pp.503-525

Pelser, P.B., Abbott, R.J., Comes, P.H., Milton, J.J., Moller, M., Looseley, M.E., Cron, G.V., Barcelona, J.F., Kennedy, A.H., Watson, L.E., Barone, R.N., Hernandez, F. and Kadereit, J.W., 2012. The genetic ghost of invasion past: colonization and extinction revealed by historical hybridization in *Senecio*. *Molecular Ecology, 21,* pp.369-387.

Pelser, P.B., Kennedy, A.H., Tepe, E.J., Shidler, J.B., Nordenstam, B., Kadereit, J.W. and Watson, L.E., 2010. Patterns and causes of incongruence between plastid and nuclear Senecioneae (Asteraceae) phylogenies. *American Journal of Botany, 97*, pp.856-873.

Pelser, P.B., Nordenstam, B., Kadereit, J.W. and Watson, L.E., 2007. An ITS phylogeny of tribe Senecioneae (Asteraceae) and a new delimitation of *Senecio* L. *Taxon*, *56*(4), pp.1077-1104.

Ronquist, F., Teslenko, M.,van der Mark, P., Ayres, D., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. and Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, *61*, pp.539-542.

Schmidt-Lebuhn, A.N. and Bovill, J., 2021. Phylogenomic data reveal four major clades of Australian Gnaphalieae (Asteraceae). *TAXON*, *70*(5), pp.1020-1034.

Schuster, S.C., 2007. Next-generation sequencing transforms today's biology. *Nature Methods, 5*(1), pp.16-18.

Smissen, R.D., and Scheele, S.M., 2022. Plastid genome sequences are consistent with introduction of Phormium tenax to Norfolk Island and Chatham Islands by early East Polynesian voyagers but not vicariance explanations. *New Zealand Journal of Botany*, pp.1-19.

Stamatakis, A., 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics, 30*(9), pp.1312-1313

Stull, G.W., Duno de Stefano, R., Soltis, D.E. and Soltis, P.S., 2015. Resolving basal lamiid phylogeny and the circumscription of Icacinaceae with a plastome-scale data set. *American Journal of Botany, 102*, pp.1794-1813.

Stull, G.W., Moore, M.J., Mandala, V.S., Douglas, N.A., Kates, H.R., Qi, X., Brockington, S.F., Soltis, P.S., Soltis, D.E. and Gitzendanner, M.A., 2013. A targeted enrichment strategy for massively parallel sequencing of angiosperm plastid genomes. *Applications in Plant Sciences*, *1*(2).

Swofford, D.L., 2003. PAUP* Sinauer Associates. *Sunderland, MA*.

Theirs, B., Continuously updated. A global directory of public herbaria. http://sweetgum.nybg.org/science/ih/

Todd, M. J., 1996. Taxonomy of *Haastia*. *Unpublished Masters Thesis, UC.*

Van de Peer, Y., Mizrachi, E. and Marchal, K., 2017. The evolutionary significance of polyploidy. *Nature Reviews Genetics*, *18*(7), pp.411-424.

Wagstaff, S.J. and Breitwieser, I., 2004. Phylogeny and classification of *Brachyglottis* (Senecioneae, Asteraceae): An example of a rapid species radiation in New Zealand. *Systematic Botany, 29*(4), pp.1003-1010.

Wagstaff, S.J. and Breitwieser, I., 2002. Phylogenetic relationships of New Zealand Asteraceae inferred from ITS sequences. *Plant Systematics and Evolution*, *231*(1-4), pp.203-224.

Wilcox, T.P., Zwickl, D.J., Heath, T.A. and Hillis, D.M., 2002. Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. *Molecular Phylogenetics and Evolution, 25*(2), pp.361-371