# Phylogenetic implications of trans-specific chloroplast DNA sequence polymorphism in New Zealand Gnaphalieae (Asteraceae)

R. D. Smissen<sup>1</sup>, I. Breitwieser<sup>1</sup>, and J. M. Ward<sup>2</sup>

<sup>1</sup>Landcare Research, Lincoln, New Zealand <sup>2</sup>School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

Received December 2, 2003; accepted June 16, 2004 Published online: October 18, 2004 Springer-Verlag 2004

Abstract. The New Zealand endemic species of Gnaphalieae (Asteraceae) present a bewildering array of morphological forms recalcitrant to phylogenetic analysis using molecular or traditional characters. The chloroplast psbA-trnH intergenic spacer is highly variable in this group, with substitutions distinguishing 18 cpDNA haplotypes distributed among samples of Raoulia species and those of related genera. A subset of haplotypes was combined with sequences of the more slowly evolving trnL intron and trnL-trnF intergenic spacer region and subjected to median network and parsimony analyses. Although a number of chloroplast lineages were resolved, these do not correspond with taxonomic units, nor with groups supported by analysis of morphological characters, nor with groups indicated by nuclear rDNA ITS sequences. These data, considered in the context of morphological character state distribution in the group, suggest a complex set of interrelationships among extant species and genera as currently circumscribed. General implications for phylogeny reconstruction are discussed.

Key words: Raoulia, Asteraceae, phylogenetic incongruence, internal transcribed spacer, hybridisation, introgression, lineage sorting.

Phylogenetic studies using several chloroplast loci have become common since the advent of automated DNA sequencing. One reason plant molecular systematists rely heavily on the chloroplast genome is because it can generally be modelled as a single-copy, essentially nonrecombining locus (Clegg and Zurawski 1992). In theory it has the additional advantage, as a result of uniparental inheritance, of reduced coalescence times. However, these characteristics (non-recombining and uniparental inheritance) also place serious limitations on the use of cpDNA for phylogeny reconstruction (e.g. Neigel and Avise 1986, Avise 1994, Ennos et al. 1999, Avise 2000). These limitations relate to the implications of intraspecific and intra-population sequence polymorphism in cpDNA under different demographic scenarios of speciation and to the possibility of chloroplast introgression between species. One problem has been termed lineage sorting and recognises that gene lineages in extant species may coalesce on individuals prior to the divergence of the species concerned and therefore may not reflect organismal or taxic phylogeny (Avise 1994). The length of time

that multiple gene lineages will be maintained in populations is in part dependent on demographic parameters such as effective population size and population substructure (Neigel and Avise 1986).

A related problem is the introgression of genes between divergent species. Chloroplast DNA introgression has been widely reported both in studies of contemporary hybridisation (reviewed in Rieseberg and Brunsfeld 1992) and as an inference based on incongruence between nuclear and chloroplast DNA phylogenies (e.g. Soltis et al. 1991, Jansen et al. 1992, Comes and Abbott 2001). This latter class of studies implies that fixation of an introgressed chloroplast in the recipient species, at least in part of its range, may be common in at least some angiosperm groups. Chloroplast introgression has also been suggested on the basis of discrepancies between cpDNA haplotype distribution and morphological species boundaries (e.g. Steane et al. 1998).

By far the most widely used nuclear DNA sequences employed in plant systematics are those of the nuclear ribosomal genes and spacers (ITS1, ITS2, and ETS: Small et al. 2004). Unlike cpDNA these nrDNA regions are biparentally inherited and show recombination (at least in some taxa) as a coevolving gene family in high copy number (Hamby and Zimmer 1992, Baldwin et al. 1995, Alvarez and Wendel 2003). These facts are often put aside as in general ITS and ETS trees have considerable congruence with prior hypotheses about phylogenetic relationships based on morphology, biogeography or other data, and ITS and ETS sequences are generally treated in the same ways as those of non-recombining singlecopy genes in phylogenetic analyses (reviewed in Álvarez and Wendel 2003).

Where hybrid speciation or introgression is a factor in the evolution of a group, the behaviour of ITS and ETS sequences is likely to depart significantly from single-locus models. Studies examining the evolution of ITS sequences following crossing between parents with divergent sequences report that sequence capture of one parental sequence type

by an introgressed population is possible (Comes and Abbott 1999) as is the generation of novel recombinant types displaying a combination of parental types (e.g. Campbell et al. 1997). By contrast, discussion of the ETS in this context is largely absent from the literature. It seems reasonable to expect that introgression may contribute to a high level of homoplasy in ITS and ETS sequence data sets as well as possible incongruence with other data sets and prior hypotheses about phylogeny.

One example of an angiosperm group where ITS sequences show these properties is the Asteraceae genus complex consisting of Raoulia Hook.f., Leucogenes Beauverd, Rachelia J.M.Ward et Breitw., Anaphalioides (Benth.) Kirp. and New Zealand species currently ascribed to Ewartia Beauverd and Helichrysum Mill. corr. Pers. All of the approximately 52 species in this assemblage are endemic to New Zealand except for two species of *Anaphalioides*, which are endemic to New Guinea.

Numerous anecdotal reports of hybrids among these genera led to the hypothesis that they shared a recent common ancestry (Ward 1981, 1997) rather than resulting from several separate entries into New Zealand from different sources, as was commonly assumed due to their extreme morphological diversity and similarities to genera elsewhere (e.g. Anderberg 1991). Subsequently, the hybrid parentage of some of these morphologically intermediate plants (McKenzie 2001, Smissen et al. 2003, McKenzie et al. 2004) has been verified. Analysis of ITS sequences by Breitwieser et al. (1999) recovered this generic complex as a strongly supported clade (hereafter referred to as the ''New Zealand clade'') but with very little internal resolution. This lack of resolution mirrors the findings of non-molecular studies, where different character sets support different taxonomic groupings due to the reticulate nature of the character state distributions (see for example Breitwieser and Ward 1993, 2003; Ward 1993; Wilton 1997).

The extent of morphological diversity within the New Zealand clade does not appear to be the result of stabilised allopolyploidy. Polyploidy occurs in *Anaphalioides* (2 of 7 species: Glenny 1997), Leucogenes (2 of 4 species: Molloy 1995) and Raoulia (6 of c. 28 species: Dawson et al. 1993), but polyploids are almost always similar to extant diploid species.

Here we report the results of a survey of chloroplast psbA-trnH intergenic spacer, trnL-trnF intergenic spacer and trnL intron sequences from Raoulia and allied genera, focusing on the alpine cushion species linked to Leucogenes by their ITS sequences. We also report reanalysis of ITS sequences including important additional sampling.

## Material and methods

The plants. Chloroplast *psbA-trnH* sequences were sampled from a total of 14 individuals from seven populations of Raoulia eximia, ten individuals from six populations of *Helichrysum lanceolatum*, three individuals from different populations of R. grandi*flora*, four individuals from four populations of *.* buchananii, two individuals from different populations of Leucogenes grandiceps, two individuals from different populations of R. mammillaris, two individuals from the same population of Anaphalioides bellidioides, two individuals from the same population of R. rubra, and single specimens of A. mariae, A. trinervis, Ewartia sinclairii, Helichrysum filicaule, Leucogenes leontopodium, R. beauverdii, R. bryoides, R. goyenii, R. hectorii, R. hookeri, Raoulia sp. 'L', R. sp. 'M', R. tenuicaulis and R. youngii. Collection and voucher details are shown in Table 1. We also generated trnL-trnF intergenic spacer and trnL intron sequences from a subset of these samples including five individuals from four populations of Raoulia eximia, two individuals from different populations of Helichrysum lanceolatum, two individuals from different populations of R. grandiflora, and single individuals of Anaphalioides bellidioides, A. trinervis, Leucogenes grandiceps, L. leontopodium, R. sp. 'L', R. hookeri, and R. rubra. Representative sequences have been deposited in GenBank as detailed in Table 1.

To the ITS sequences already available (detailed in Breitwieser et al. 1999, Smissen et al. 2003) we added additional sequences for single specimens of R. hectorii, R. hookeri, R. mammil*laris*, R. sp.  $L$ , R. *youngii*, and three specimens of

R. grandiflora. We also increased our sampling of ITS sequences from R. eximia and R. buchananii to include additional populations. Sequence alignments are available from the corresponding author.

DNA extraction, PCR and sequencing. Genomic DNA was extracted by the CTAB method as described in Smissen et al. (2003) or by Dneasy Plant Mini Kit (Qiagen) following the manufacturer's instructions with the addition of a phenol chloroform extraction following elution of DNA where necessary. ITS sequences were amplified by PCR and sequenced according to Smissen et al. (2003). Amplification of psbA-trnH used PCR cycling conditions with an initial denaturation at 97  $\mathrm{^{\circ}C}$  for 120 s followed by 35 cycles of 94  $\rm{°C}$  for 30 s, 48  $\rm{°C}$  for 60 s, and 72  $\degree$ C for 2 min. Primer sequences were taken from the web page http://www.paeon.de/h1/ sang/dis4.html (psbAf 5' -GTTATGCATGAACG-TAATGCTC-3', trnHr 5'-CGCGCATGGTG-GATTCACAATC-3'). Sequencing used the same primers as PCR. Cycle sequencing was carried out using ABI BigDye 3.1 terminators and products were separated on an ABI 3730 Genetic Analyser. Sequences of  $trnL-trnF$  and  $trnL$  intron were generated using the primers of Taberlet et al. (1991) and PCR conditions as for the *psbA-trnH* detailed above. ITS sequences were generated as described in Smissen et al. (2003).

Data analysis. All New Zealand Gnaphalieae cpDNA sequences analysed in this study were readily aligned by eye. Likewise, the ITS sequences of members of the New Zealand endemic clade were readily aligned by eye. Ingroup ITS sequences were aligned with outgroup ITS sequences using ClustalX (Thompson et al. 1997) and subsequent visual inspection. Median networks were constructed using splits generated via Hadamard conjugation by Spectronet (Huber et al. 2002). The option of Hadamard conjugation was employed as Spectronet is unable to generate splits directly from data columns with more than two states and thus information is lost. An alternative strategy available in Spectronet, reducing sequence data to purine/pyrimidine  $(r/y)$  information at these sites, is not justified for chloroplast sequences, as these generally do not show a transition/transversion bias. Shortest parsimony trees for ITS sequences were found using PAUP4.0b10 heuristic search with TBR branch swapping, steepest descent and multrees in effect, and a simple addition sequence. Searches for shortest trees for ITS







sequences using TBR branch swapping and 1000 random addition replicates, with multrees but without steepest descent in effect (see http:// paup.csit.fsu.edu/problems.html), were also conducted. DELTRAN character optimisation was used as PAUP4.0b10 reports incorrect branch lengths for outgroup rooted trees for our data with ACCTRAN optimisation (see http://paup.csit.fsu.edu/problems.html). Bootstrap values were calculated from 1000 bootstrap replicates resampling the original number of characters using full heuristic search, multrees and steepest descent in effect, collapsing branches if their minimum length was zero. Alignment gaps were treated as missing data for ITS sequence analysis. Heteroplasmy, where detected in ITS sequences, was treated as ambiguity for the purposes of analysis. For the psbA-trnH analysis alignment gaps were coded as presence/ absence characters and included in analyses. Analysis of chloroplast DNA sequences for New Zealand endemic Gnaphalieae together with those of Euchiton audax, Argyrotegium mackayi, Pterygopappus lawrencei, Ewartia catipes and several Craspedia samples (Smissen, unpublished sequences) was also undertaken using similar methods to those described above.

## Results

Eighteen haplotypes were identified from psbA–trnH sequences. Distribution of these among the species sampled is shown in Table 2 and relationships inferred among them by median network analysis are shown in Fig. 1. Nucleotide substitutions are apparent at a total of 19 sites and four alignment gaps attributable to insertion/deletion events (indels) unique to haplotypes were detected. Alignment gaps range in size from 1 to 36 base pairs in length. Two characters support conflicting splits between samples and this is reflected in the box linking haplotypes C1, C2, C3 and C4 in the median network. Maximum nucleotide divergence among the haplotypes (between A5 and C6) was 3.8%. The maximum haplotype divergence detected within a species was 3.1%. This was between the A5 and C5 haplotypes detected in R. buchananii.

The trnL intron and trnL-trnF intergenic spacer appear to have slower rates of

Haplotype	Species occurrence	Specimens
A1	Raoulia rubra	CHR 546183, CHR 546184
	Leucogenes leontopodium	CHR510025
A2	Leucogenes grandiceps	CHR514141, CANU38729
A3	Anaphalioides bellidioides	CANU38726, CANU38727
A <sub>4</sub>	Anaphalioides mariae	CHR492179
A <sub>5</sub>	Anaphalioides trinervis	CHR569864
	Raoulia buchananii	CHR 546188, CHR546190
A <sub>6</sub>	Helichrysum lanceolatum	CHR569865, CHR569866,
A7	Helichrysum lanceolatum	CHR565241
A8	Helichrysum lanceolatum	CHR569867, CHR569871
A <sub>9</sub>	Helichrysum lanceolatum	CHR569868, CHR569869, CHR569870,
		CHR569872, CHR569873
A10	Ewartia sinclairii	<b>CANU37227</b>
B1	Raoulia grandiflora	Lagoon saddle (unvouchered)
B2	Raoulia hectorii	<b>CANU38722</b>
	Raoulia eximia	CHR546180, CHR546182
C1	Raoulia sp. 'L'	<b>CANU38725</b>
C <sub>2</sub>	Raoulia bryoides	CHR502568
	Raoulia eximia	CHR546147
	Raoulia sp. 'M'	<b>CANU38721</b>
C <sub>3</sub>	Raoulia tenuicaulis	unvouchered <sup>a</sup>
	Raoulia hookeri	CHR569861
C <sub>4</sub>	Raoulia beauverdii	<b>CANU38724</b>
	Helichrysum filicaule	CHR569874
C <sub>5</sub>	Raoulia buchananii	CHR565559, CHR 569758
	Raoulia eximia	CHR546162, CHR546168, CHR546177,
		CHR546178, CHR546154, CHR546160,
		CHR546164, CHR569875
	Raoulia goyenii	CHR560933
	Raoulia grandiflora	<b>CANU38728</b>
	Raoulia mammillaris	CHR569860, CANU38629
	Raoulia youngii	<b>CANU38723</b>
C6	Raoulia eximia	CHR546146, CHR546151, CHR509743

Table 2. Distribution of cpDNA haplotypes among plant samples

<sup>a</sup>The DNA extract used for this sample is the same as that used by Glenny and Wagstaff (1997) for which no voucher was made.

nucleotide substitution than the psbA-trnH intergenic spacer in this group (maximum divergence between the sampled haplotypes of 1.2% for each locus). Combining psbA-trnH sequences with those of the *trn*L intron and trnL-trnF intergenic spacer region for a subset of 12 haplotypes gives a total of 27 variable sites, 8 of which are potentially informative about phylogenetic relationships among haplotypes. A median network generated for these data (Fig. 2) shows slightly improved resolution of relationships over that for the psbAtrnH sequences alone. An unrooted strict consensus of shortest trees from a parsimony analysis for these data is also presented (Fig. 3). We explored possible outgroup rooting with Australian Gnaphalieae and New Zealand Gnaphalieae previously evidenced to be outside the New Zealand endemic clade (Ward 1993, Breitwieser et al. 1999, Smissen et al. 2003, Breitwieser and Ward 2003). Different outgroups grouped with different parts of the



Fig. 1. Median network for *psbA-trnH* haplotypes of *Raoulia* and allied genera. "Boxes" represent conflicting signals in the data. Branch lengths are drawn proportional to the number of differences between sequences. The species from which each haplotype has been sampled are listed. Species shown in bold are those from which non-monophyletic groups of haplotypes have been sampled



ingroup (attributable to homoplasy) giving no clear picture as to where the root should be placed (not shown). Consequently we have



Fig. 2. Median network for combined trnL intron, trnL-trnF intergenic spacer and psbA-trnH intergenic spacer sequences of selected haplotypes. Labels refer to haplotypes shown in Fig. 1 and Table 2. Branch lengths are drawn proportional to the number of differences between sequences

Fig. 3. Unrooted strict consensus of four most parsimonious trees for combined trnL intron, trnLtrnF intergenic spacer, and psbA-trnH intergenic spacer sequences. Labels refer to haplotypes shown in Fig. 1 and Table 2. The most parsimonious trees contributing to this consensus had a Consistency Index of 0.967, Retention Index of 0.970 and Rescaled Consistency Index of 0.937

limited our conclusions to those that can be drawn from un-rooted networks.

Of the eight species for which more than one individual has been sampled (Anaphalioides bellidioides, Leucogenes grandiceps, Raoulia eximia, R. mammillaris, R. rubra, R. buchananii, R. grandiflora, Helichrysum lanceolatum), three possess non-monophyletic haplotype combinations. Two haplotypes were sampled from R. buchananii and these are separated by nine substitutions in Fig. 1 (haplotype A5 and C5 respectively). Four haplotypes were observed in R. eximia: C6 is so far unique to this species, C5 is shared with five other Raoulia species, B2 is shared with R. hectorii, and C2 is shared with R. bryoides and  $R$ . sp. 'M'. Two individuals of  $R$ . grandiflora had the C5 haplotype shared by six Raoulia species while the third had the unique B1 haplotype related to the B2 haplotype shared by R. eximia and R. hectorii. Of the remaining three species, only two individuals of R. rubra, from the same population, have been sequenced. Raoulia rubra is endemic to a small area in the North Island (Smissen et al. 2003) and probably has little cpDNA diversity. Only a small part of the geographic range of L. grandiceps is represented by the two individuals sampled, so little can be read from the fact that they display haplotypes identical at the psbA-trnH locus. Some diversity has been sampled from the widespread Helichrysum lanceolatum with four closely related haplotypes so far identified (A6-9). Relationships between these *H. lanceolatum* haplotypes and the related A10 haplotype (sampled from Ewartia sinclairii) are unresolved (Fig. 1).

Heuristic search on ITS sequences returns 22958 trees of 291 steps (Consistency Index 0.659, Retention Index 0.832 and Rescaled Consistency Index 0.548 excluding uninformative characters). These 22958 trees condense to a single topology if branches with a minimum length of zero (i.e. no unambiguous character support) are collapsed (Fig. 4). Searches with random addition sequence did not report any trees shorter than 291 steps nor any of 291 steps not agreeing with Fig. 4. Although PAUP4.0b10 reported finding more shortest trees in the random addition search (23374) than in the simple addition sequence search, the program appears to be saving too many trees (see http://paup.csit.fsu.edu/problems.html) as filtering the saved trees for ''best score'' retains only 22958 trees. The single condensed tree for the most part confirms the previously reported pattern of relationships (Smissen et al. 2003). A significant new result compared with previously published analyses (Breitwieser et al. 1999, Smissen et al. 2003) is the exclusion of R. grandiflora from a wellsupported group containing the remaining sampled species of *Raoulia* and allied genera (93% bootstrap). Although Fig. 4 depicts R. grandiflora as the sister clade to a clade comprising the New Zealand endemic clade and sampled species of Argyrotegium, Euchiton, Australian species of Ewartia, and Pterygopappus lawrencei, bootstrap support for this clade is essentially absent (52%). The ITS data can therefore be viewed as consistent with *R. grandiflora* being sister group to the New Zealand endemic clade. There is little resolution within the New Zealand endemic clade, but three sub-groupings are noteworthy. Firstly, sequences from the species of Anaphalioides group weakly together (bootstrap 57%). Secondly, the ITS sequences of South Island alpine cushion species R. eximia, R. buchananii, and R. goyenii group with the sequence of Leucogenes leontopodium (bootstrap 64%). Thirdly, the sequences from a group of morphologically similar lowland mat-forming Raoulia species (R. hookeri, R. monroi, and R. tenuicaulis) form a well-supported clade (bootstrap 91%).

### **Discussion**

cpDNA haplotype distribution. Median network and parsimony analysis of combined psbA-trnH, trnL intron, and trnL-trnF sequences produced very similar results, differing only in the capacity of median networks to display conflicting signal, and allow conflicting phylogenetic signal to be distinguished



1 change

Fig. 4. Most parsimonious tree for ITS sequences. Numbers above branches are bootstrap percentages from 1000 replicates (where >50). Note that the ''New Zealand endemic clade'' includes the Papua New Guinea endemic Anaphalioides mariae. Country of origin is New Zealand unless otherwise noted (Aust=Australia, PNG=Papua New Guinea). Sequences of *Craspedia* and *Pycnosorus* have been defined as outgroups for this graph. Branch lengths are drawn proportional to the inferred number of substitutions. Species in bold are samples new to this study, others are detailed in Smissen et al. (2003) or Breitwieser et al. (1999)

easily from lack of phylogenetic signal. We present the parsimony network (Fig. 3) to demonstrate the loss of information in parsimony networks relative to the median network, and to demonstrate the general agreement over groups recovered. Combining psbA-trnH sequences with those of the trnLtrnF intergenic spacer and trnL intron provides improved resolution (compare Fig. 1 and Fig. 2) but adds little additional conflict, suggesting that data from these different regions of the chloroplast genome are congruent and, therefore, that the apparent pattern of chloroplast haplotype sharing observed across species is not the product of misleading homoplastic changes in the trnH-psbA intergenic spacer.

Even with the limited sampling reported here, it is clear that significant cpDNA polymorphism in New Zealand Gnaphalieae is common, and that attempts to recover a species phylogeny from cpDNA sequences are likely to be highly misleading. It should be noted that we took no steps to increase the probability of sampling introgressed individuals. In fact, in the case of  $R$ . *eximia*, we specifically tried to avoid hybrid populations. Likewise, we had no reason to expect R. buchananii specimens displaying the A5 haplotype shared with Anaphalioides to be unusual in any way.

However, some haplotype sharing observed in our study is consistent with expectations of relationships based on morphology and, in some cases, ITS sequences. The three alpine cushion species linked by similar morphology and ITS sequences, R. eximia, R. goyenii and R. buchananii, all share haplotype C5, although other haplotypes were also sampled from both R. eximia and R. buchananii. A fourth alpine cushion species, R. rubra, is morphologically very similar to the other three but has a distinct ITS sequence type and has cpDNA haplotype A1. Two lowland matforming Raoulia species, R. hookeri and R. tenuicaulis, have identical trnH-psbA sequences and have ITS sequences differing by a single unambiguous substitution. A very similar ITS sequence (differing from the R. tenuicaulis sequence only by the presence of an ambiguous site) is also shared by the morphologically similar species R. *monroi*, for which no cpDNA sequence is available. The Anaphalioides species sampled all share group A haplotypes, although these were also found in some other species.

Although numerous hybrid combinations are known among the species within the New Zealand clade that were sampled in this study, as well as other species for which sequence data are unavailable, it is not possible to explain all the haplotype sharing observed among morphologically dissimilar species on the basis of contemporary hybridisation. However, R. bryoides and R. eximia are sympatric in the St Arnaud Range where R. eximia samples display the C2 haplotype shared with R. bryoides. Occasional putative hybrids have been collected elsewhere where the two species grow together (Allan 1961, Ward 1997). Our samples of Raoulia rubra and Leucogenes leontopodium share the A1 haplotype. These species were both collected from Mt Holdsworth (Wellington, North Island) where they occur in sympatry, and putative hybrids displaying intermediate morphology and additive ITS sequences have been collected (Allan 1961, Smissen et al. 2003). The three South Island alpine cushion species R. buchananii, R. eximia, and R. goyenii possess an entirely different set of cpDNA haplotypes from R. rubra and Leucogenes, but share ITS sequence characters with  $L$ . *leontopodium* that are not shared by  $R$ . rubra or *L. grandiceps*. No hybrids between any of these three South Island alpine cushion species and *L. leontopodium* are known (but there is little or no overlap in their ranges). Further complicating matters, L. grandiceps is widespread in the South Island and does hybridise with at least R. eximia (McKenzie 2001), and, on Stewart Island, R. goyenii (Allan 1961, I.B. pers. observation). It is likely that the alpine cushion group of Raoulia species and the species of *Leucogenes* are linked by past and possibly ongoing reticulation, but in the absence of knowledge of ancestral ranges and the direction of hypothetical

sequence introgression, specific hypotheses accounting for the sequences we have observed would be highly speculative.

Haplotype A5 was sampled from Anaphalioides trinervis and Raoulia buchananii. No hybrids between this pair of species are known but hybrids of A. trinervis and A. bellidioides are common. Anaphalioides bellidioides is reported to hybridise with a range of Raoulia, Helichrysum and Leucogenes species (Glenny 1997). Also, experimental crosses between A. trinervis and other New Zealand Gnaphalieae, including species of Raoulia and Leucogenes, set seed (McKenzie 2001). Likewise, hybrids between R. eximia and R. hectorii (which share the B2 haplotype) are unknown, but the two are sympatric at Mt Potts, and the combination is not unlikely.

The distribution of cpDNA polymorphism among species of New Zealand Gnaphalieae reported here is difficult to explain other than through introgression among species after species divergence. However, the level of polymorphism found and the seemingly random association of haplotypes with species precludes specific simple hypotheses of introgression events. Given that haplotypes coexisting in species are widely separated in the network (Fig. 1), an alternative explanation based on maintenance of ancestral polymorphism would imply that divergence of species occurred after the evolution of most of the haplotype diversity observed (i.e. that the hypothetical ancestor possessed haplotypes indistinguishable from extant haplotypes at the psbA–trnH locus). A rapidly expanding population can be expected to maintain greater haplotype diversity than static or shrinking populations (Avise 2000), as lineage extinction is comparatively rare under such circumstances. Therefore, if speciation events producing most of the extant species of the New Zealand clade occurred in rapid succession, or even contemporaneously, and occurred after a period of major population growth in their common ancestor (perhaps following dispersal to New Zealand) this level of polymorphism might conceivably have been developed. How-

ever, such a hypothetical ancestral species would have needed both a high level of interpopulation gene-flow and a high level of local differentiation to produce the extant pattern of morphological and cpDNA diversity seen today. Given these sorts of constraints, the differences between the two scenarios (ancestral polymorphism with local differentiation and inter-population gene-flow, and introgression among divergent species) start to become quantitative rather than qualitative and at least partly semantic. Importantly, both involve reticulation among diverging populations. Whether these are viewed as species, races or populations depends on the degree of divergence and definitions of the categories.

Other studies of cpDNA polymorphism in vascular plants have revealed strong associations of haplotypes with geographical areas, either as phylogeographic structure within a species or across species boundaries, perhaps as a result of co-occurring glacial refugia (Steane et al. 1998, Trewick et al. 2002, Petit et al. 2004). No such pattern associating haplotypes with geographical areas is evident from our data.

ITS sequence types. Few clades within the New Zealand endemic clade are well supported by the ITS sequence analysis. However, the grouping of Anaphalioides species accords with morphology and is not strongly contradicted by the cpDNA data. As discussed above, R. monroi, R. tenuicaulis, and R. hookeri are morphologically similar and their grouping in the ITS tree is therefore not unexpected. As discussed in Smissen et al. (2003) the grouping of South Island alpine cushion species  $R.$  eximia,  $R.$  buchananii and  $R.$  govenii is consistent with morphology, but inclusion of Leucogenes leontopodium in this clade is not. The groupings of R. hectorii with R. sp.  $'L$  and the grouping of Rachelia glaria with Leucogenes grandiceps are weakly supported by ITS sequences and by their morphological similarity, whereas the pairing of Helichrysum filicaule with Raoulia youngii is weakly supported by ITS sequences and strongly contradicted by morphology.

Importantly, the well-supported branch separating the outgroup taxa and Raoulia grandiflora from the New Zealand endemic clade including the rest of Raoulia in the shortest ITS tree is not reflected in the cpDNA evidence. Also, phenetic and cladistic analyses of morphology, leaf anatomy and flavonoid profiles do not provide any support for this separation (Ward 1993; Breitwieser and Ward 1993, 2003). Overall morphology suggests a close relationship between R. grandiflora and R. youngii. Of the two cpDNA haplotypes observed in R. grandiflora, one is so far unique but nested well within the diversity of the rest of the group (grouped with a haplotype shared by R. hectorii and R. eximia); the other is shared with at least five other Raoulia species.

However, this well-supported branch in the ITS tree, and the ITS sequence data in general, should not be ignored. Firstly, these data seem to indicate a longer history of the New Zealand clade than previously evidenced (Breitwieser et al. 1999). Because ITS sequences are biparentally inherited and potentially recombining, their divergence presumably requires a low level or absence of interbreeding. The ITS separation between R. grandiflora and the New Zealand endemic clade strongly suggests that diversified lineages were present from an early stage in the history of the group in New Zealand (conflicting with explanations of cpDNA diversity relying exclusively on incomplete lineage sorting). Secondly, this result necessitates a re-evaluation of the conclusion of Breitwieser et al. (1999) that the New Zealand clade has evolved in New Zealand following a single dispersal event from Australia. ITS sequences provide no unambiguous evidence as to whether R. grandiflora is sister to the rest of Raoulia and allies, or originated within a distinct clade (as shown in Fig. 4). Thirdly, the ITS sequence analysis provides a temporal context for understanding the evolution of diversity in New Zealand Gnaphalieae. This analysis suggests an early divergence of R. grandiflora from the rest of the group, a later radiation generating most of the morphological and taxonomic diversity, and a second phase of species radiations evidenced for the South Island alpine cushion Raoulia species, lowland mat-forming Raoulia species, and *Anaphalioides* species. Reticulation may have played a role in any of these phases.

Evolutionary consequences of reticulation. Although we have presented evidence for cpDNA introgression between endemic New Zealand Gnaphalieae, there is as yet no firm evidence of nuclear gene introgression, and in particular no evidence of adaptively significant introgression. Thus, while reticulation seems to be a good explanation for the failure of DNA sequences to recover phylogenetic structure, the importance of reticulation in adaptive evolution is still a matter for speculation. However, Kim and Rieseberg (1999) have shown that nuclear gene introgression is not only possible, but has probably resulted in the formation of new taxa in Helianthus L. Also, given the lack of congruent hierarchical structure in morphological and chemical data sets previously analysed in the New Zealand Gnaphalieae (Breitwieser and Ward 1993, Ward 1993), it is tempting to speculate that these characters have also been subject to reticulate evolution. Future investigations of New Zealand Gnaphalieae should focus on nuclear gene introgression in selected hybridising species using multi-locus approaches such as AFLP or simple-sequencerepeat markers.

A cautionary note for phylogenetic studies. Regardless of the importance of reticulation in adaptive evolution, the high degree of intraspecific cpDNA polymorphism apparent from our limited sampling has important implications for the use of cpDNA sequences for phylogeny estimation in Asteraceae. The large within-species divergences (up to  $3.1\%$ ) reflect, to a large part, the high substitution rate of the psbA-trnH intergenic spacer region of the chloroplast DNA molecule; approximately three-fold lower divergences were found in the trnL intron and trnL-trnF intergenic spacer. Rapid species radiations, such as those that appear to have occurred in New Zealand Gnaphalieae, may be a common

pattern in angiosperms (Bateman 1999). A DNA-sequence-based phylogenetic study of Australian Gnaphalieae (Bayer et al. 2002) also indicates a rapid radiation of genera. Moreover, patterns of transpecific chloroplast polymorphism have been reported for Senecio sect. Senecio (Senecioneae: Asteraceae) by Comes and Abbott (1999, 2001) on the basis of restriction fragment length polymorphism. The concurrent appearance of Asteraceae pollen in the fossil records of several continents (Graham 1996) in the late Oligocene to early Miocene is consistent with a rapid spread and diversification of Asteraceae on a worldwide scale in the ancient past. Given these facts, it seems reasonable to assume that at least some of the major extant lineages of Asteraceae are descended from rapid species radiations similar to those seen in the New Zealand flora and elsewhere in the world today.

It is clearly important that before relationships among extant lineages are inferred from chloroplast phylogenies, the possibility that lineage sorting of ancestral cpDNA polymorphism or chloroplast introgression has influenced chloroplast lineage distribution be evaluated. Neigel and Avise (1986) conclude from modelling experiments that ''phylogenetic distributions of mtDNA can lack concordance with species boundaries when species are recently separated''. Resolving this lack of concordance requires sufficient generational time for matrilineal coalescence in each species, and sufficient reproductive isolation to prevent exchange of maternally inherited DNA between the new species. If further speciations occur before both of these conditions are met then maternally inherited DNA trees may lack concordance with species phylogeny, unaffected by the passage of further time (see hypothetical example, Fig. 5).

In the hypothetical example depicted in Fig. 5 it is assumed that nucleotide substitutions occur such that the true chloroplast phylogeny can be estimated (Fig. 5b). The true species tree for these hypothetical species is (AB)(CD). Because the divergence of



Fig. 5. Fate of chloroplast lineages during a hypothetical species radiation illustrating potential consequences of introgression and lineage sorting. Adapted from Avise (2000, p. 65). b. Expected chloroplast DNA tree for taxa A, B, C, and D given sufficient sampling of nucleotide sites. Relationships between chloroplast lineages are correctly estimated, but these do not correspond with the order of divergence of the species

species C and D occurred relatively soon after the divergence of their parent species Y from X, a subset of the chloroplast lineages present in Z was retained and subsequently sorted so that the cpDNA supports the tree  $((AB)C)D$ . The branch supporting the A+  $B+C$  clade is relatively short, reflecting the shallow divergence of chloroplast lineages at the time of speciation. Conversely, the divergence of species A and B occurred after lineage sorting had left species X with only one of the chloroplast lineages inherited from Z, and after reproductive isolation between species X and Y was strong enough to prevent chloroplast introgression. As a result the branch supporting the  $A+B$  clade is relatively long. Species A appears twice in the cpDNA tree as a result of polymorphism resulting from reticulation early in the divergence of species A and B. Given sufficient further time, in the absence of further chloroplast introgression, lineage sorting will occur and each of A and B will lose all but one of the chloroplast lineages shown.

As in the hypothetical example in Fig. 5, lineages descending from ancestors that evolved during rapid radiations, whose relationships are likely to be obscured by lineage sorting or introgression effects, will be grouped by relatively short branches in phylogenetic trees derived from cpDNA. Except in exceptional cases where inferences about ancestral demographics can be made, we suggest cpDNA clades grouped by branch lengths shorter than distances between haplotypes cooccurring in species should not be used to infer species phylogeny.

Importantly, these caveats apply not only to interspecific comparisons (Wendel and Doyle 1998), but also to any lineages descending from ancient rapid radiations (Avise and Wollenberg 1997). In the past, short internal branches in cpDNA trees that are likely to be misleading if used to infer species phylogeny will have tended to attract little statistical support as measured by bootstrapping, and little decay analysis (Bremer 1994) support because of their short length. Thus a conservative view of the tree will have placed little confidence in these branches. However, as multigene studies analysing larger numbers of nucleotides are undertaken, shorter branches (as measured by substitutions per site rather than absolute numbers of changes) in chloroplast phylogenetic trees will be estimated with statistical confidence (as measured by tests including bootstrapping).

Where short branches in cpDNA trees are independently supported (for example by morphological or nuclear DNA data) then it may be reasonable to place confidence in them. However, where independent support is lacking, short internal branches in phylogenetic trees estimated with the aid of very long cpDNA sequences (especially multiple chloroplast genes) should not be considered reliable estimators of species phylogeny, regardless of statistical significance, unless they exceed the level of contemporary intraspecific cpDNA variation. It should also be noted that parsimony optimisations may tend to overestimate the lengths of internal branches in trees (relative to external branches) in a process similar to the comparatively well understood phenomenon of longbranch attraction (Felsenstein 1978). Maximum likelihood optimisations are probably more appropriate for this comparison (Smissen unpublished).

Here we have shown that levels of intraspecific cpDNA sequence polymorphism can be sufficient that attempts to estimate phylogenetic relationships among well-differentiated species are completely confounded.

The authors thank Robert McKenzie for plant material and discussion, and David Glenny, Peter Lockhart and Phil Novis for field collections. Previous versions of the manuscript benefited from review by Christine Bezar, Peter B. Heenan, Steven J. Wagstaff and two anonymous reviewers.

### References

- Allan H. H. (1961) Flora of New Zealand, Vol. 1. Government Printer, Wellington.
- Alvarez I., Wendel J. F. (2003) Ribosomal ITS sequences and plant phylogenetic inference. Molec. Phylogenet. Evol. 29: 417–434.
- Anderberg A. (1991) Taxonomy and phylogeny of the tribe Gnaphalieae (Asteraceae). Opera Bot. 104: 1–195.
- Avise J. C. (1994) Molecular markers. Natural history and evolution. Chapman and Hall, New York.
- Avise J. C. (2000) Phylogeography: The history and formation of species. Harvard University Press, Cambridge, MA.
- Avise J. C., Wollenberg K. (1997) Phylogenetics and the origin of species. Proc. Natl. Acad. Sci. USA 94: 7748–7755.
- Baldwin B. G., Sanderson M. J., Porter J.M., Wojciechowski M. F., Campbell C. S., Donoghue M. J. (1995) The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. Ann. Missouri Bot. Gard. 82: 247–277.
- Bateman R. M. (1999) Integrating molecular and morphological evidence of evolutionary relationships. In: Hollingsworth P. M., Bateman R. M., Gornall R. J. (eds.) Molecular systematics and plant evolution. Taylor and Francis, London, pp. 432–471.
- Bayer R. J., Greber D. G., Bagnall N. H. (2002) Phylogeny of Australian Gnaphalieae (Asteraceae) based on chloroplast and nuclear sequences, the trnL intron, trnL/trnF intergenic spacer, matK, and ETS. Syst. Bot. 27: 801-814.
- Breitwieser I., Ward J. M. (1993) Systematics of New Zealand Inuleae (Compositae-Asteraceae)- 3 Numerical phenetic analysis of leaf anatomy and flavonoids. New Zealand J. Bot. 31: 43–58.
- Breitwieser I., Ward J. M. (2003) Phylogenetic relationships and character evolution in New Zealand and selected Australian Gnaphalieae (Compositae) inferred from morphological and anatomical data. Bot. J. Linn. Soc. 141: 183–203.
- Breitwieser I., Glenny D. S., Thorne A., Wagstaff S. J. (1999) Phylogenetic relationships in Australasian Gnaphalieae (Compositae) inferred from ITS sequences. New Zealand J. Bot. 37: 399–412.
- Bremer K. (1994) Branch support and tree stability. Cladistics 10: 295–304.
- Campbell C. S., Wojciechowski M. F., Baldwin B. G., Alice L. A., Donoghue M. J. (1997) Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex. Molec. Biol. Evol. 14: 81–90.
- Clegg M. T., Zurawski G. (1992) Chloroplast DNA and the study of plant phylogeny: present status and future prospects. In: Soltis P. S., Soltis D. E., Doyle J. J. (eds.) Molecular systematics of plants. Chapman and Hall, New York, pp. 252– 279.
- Comes H. P., Abbott R. J. (1999) Reticulate evolution in the Mediterranean complex of Senecio sect. Senecio: uniting phylogenetic and population level approaches. In: Hollingsworth

P. M., Bateman R. M., Gornall R. J. (eds.) Molecular systematics and plant evolution. Taylor and Francis, London, pp. 171–198.

- Comes H. P., Abbott R. J. (2001) Molecular phylogeography, reticulation, and lineage sorting in Mediterranean Senecio sect. Senecio (Asteraceae). Evolution 55: 1943–1962.
- Dawson M. I., Ward J. M., Groves B. E., Hair J. B. (1993) Contributions to a chromosome atlas of the New Zealand flora. 32. Raoulia (Inuleae-Compositae (Asteraceae). New Zealand J. Bot. 31: 97–106.
- Ennos R. A., Sinclair W. T., Hu X-S., Langdon A. (1999) Using organelle markers to elucidate the history, ecology and evolution of plant populations. In: Hollingsworth P. M., Bateman R. M., Gornall R. J. (eds.) Molecular systematics and plant evolution. Taylor and Francis, London, pp. 432–471.
- Felsenstein J. (1978) Cases in which parsimony or compatibility will be positively misleading. Syst. Zool. 27: 401–410.
- Glenny D. (1997) A revision of the genus Anaphalioides (Asteraceae: Gnaphalieae). New Zealand J. Bot. 35: 451–477.
- Glenny D., Wagstaff S. J. (1997) Evolution and biogeography of New Zealand Anaphalis DC. (Asteraceae: Gnaphalieae) inferred from rDNA sequences. New Zealand J. Bot. 35: 441–449.
- Graham A. (1996) A contribution to the geologic history of the Compositae. In: Hind D. J. N., Beentje J. J. (eds.) Compositae: Systematics. Proceedings of the International Composiate Conference, Kew, 1994. Vol 1. Royal Botanic Gardens, Kew, pp. 123–140.
- Hamby R. K., Zimmer E. A. (1992) Ribosomal RNA as a phylogenetic tool in plant systematics. In: Soltis P. S., Soltis D. E., Doyle J. J. (eds.) Molecular systematics of plants. Chapman and Hall, New York, pp. 252–279.
- Huber K. T., Langton M., Penny D., Moulton V., Hendy M. (2002) Spectronet: a package for computing spectra and median networks. Appl. Bioinformatics 1: 159–161.
- Jansen R. K., Michaels H. J., Wallace R. S., Kim K. –J., Keeley S. C., Watson L. E., Palmer J. D. (1992) Chloroplast DNA variation in the Asteraceae: Phylogenetic and evolutionary implications. In: Soltis P. S., Soltis D. E., Doyle J. J. (eds.) Molecular systematics of plants. Chapman and Hall, New York, pp. 252–279.

R. D. Smissen et al.: Chloroplast polymorphism in New Zealand Gnaphalieae 53

- Kim S.-C., Rieseberg L. H. (1999) Genetic architecture of species differences in annual sunflowers: Implications for adaptive trait introgression. Genetics 153: 965–977.
- McKenzie R. J. (2001) Intergeneric hybridisation in the New Zealand Gnaphalieae (Compositae). Unpublished PhD thesis. University of Canterbury, New Zealand.
- McKenzie R. J., Ward J. M., Lovis J. D., Breitwieser I. (2004) Morphological evidence for natural intergeneric hybridisation in the New Zealand Gnaphalieae (Compositae): Anaphalioides bellidioides  $\times$  Ewartia sinclairii. Bot. J. Linn. Soc. 145: 59–75.
- Molloy B. P. J. (1995) Two new species of Leucogenes (Inuleae: Asteraceae) from New Zealand, and typification of L. grandiceps. New Zealand J. Bot. 33: 53–63.
- Neigel J. E., Avise J. C. (1986) Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: Karlin S., Nevo E. (eds.) Evolutionary processes and theory. Academic Press, Orlando, pp. 515–534.
- Petit R. J., Bodénès C., Ducousso A., Roussel G., Kremer A. (2004) Hybridisation as a mechanism of invasion in oaks. New Phytol. 161: 151–164.
- Rieseberg L. H., Brunsfeld S. J. (1992) Molecular evidence and plant introgression. In: Soltis P. S., Soltis D. E., Doyle J. J. (eds.) Molecular systematics of plants. Chapman and Hall, New York, pp. 151–176.
- Small R. L. Cronn R. C., Wendel J. F. (2004) Use of nuclear genes for phylogeny reconstruction in plants. Austral Syst. Bot. 17: 145–170.
- Smissen R. D., Breitwieser I., Ward J. M., McLenachan P. A., Lockhart P. J. (2003) Use of ISSR profiles and ITS-sequences to study the biogeography of alpine cushion plants in the genus Raoulia (Asteraceae). Plant Syst. Evol. 239: 79–94.
- Soltis D. E., Soltis P. S., Collier T. G., Edgerton M. L. (1991). Chloroplast DNA variation within and among genera of the Heuchera group (Saxifragaceae): evidence for chloroplast transfer and paraphyly. Amer. J. Bot. 78: 1091–1112.
- Steane D. A., Byrne M., Vaillancourt R. E., Potts B. M. (1998) Chloroplast DNA polymorphism signals complex interspecific interactions in Eucalyptus (Myrtaceae). Austral. Syst. Bot. 11: 25–40.
- Taberlet P., Gielly G., Pautou G., Bouvet J. (1991) Universal primers for the amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17: 1105–1109.
- Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F., Higgins D. G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl. Acids Res. 24: 4876– 4882.
- Trewick S. A., Morgan-Richards M., Russell S. R., Henderson S., Rumsey F. J., Pintér I., Barrett J. A., Gibby M., Vogel J. C. (2002) Polyploidy, phylogeography and Pleistocene refugia of the rockfern Asplenium ceterach: evidence from cpDNA. Mol. Ecol. 11: 2003–2012.
- Ward J. M. (1981) Numerical phenetics and the classification of Raoulia (Gnaphaliinae-Compositae). Unpublished PhD thesis. University of Canterbury, New Zealand.
- Ward J. M. (1993) Systematics of New Zealand Inuleae (Compositae-Asteraceae). A numerical phenetic study of Raoulia and allied genera. New Zealand J. Bot. 31: 29–42.
- Ward J. M. (1997) Raoulia and its hybrids. In: Sheppard J. S. (ed.) Southern Alpines '96 Conference Proceedings. Christchurch, New Zealand, pp. 40–44.
- Ward J. M., Breitwieser I., Flann C. A. (2003) Argyrotegium, a new genus of Gnaphalieae (Compositae). New Zealand J. Bot. 41: 603–611.
- Wendel J. F., Doyle J. J. (1998) Phylogenetic incongruence: Window into genome history and molecular evolution. In: Soltis P., Doyle J. (eds.) Molecular systematics of plants II. Chapman and Hall, New York, pp. 265-296.
- Wilton A. D. (1997) An evolutionary investigation of the New Zealand Inuleae (Compositae): stem anatomy and flowering phenology. Unpublished PhD thesis, University of Canterbury, New Zealand.

Addresses of the authors: Rob Smissen (e-mail: smissenr@landcareresearch.co.nz), Ilse Breitwieser Landcare Research, P.O. Box 69, Lincoln 8152, New Zealand. Josephine Ward School of Biological Sciences, University of Canterbury, Private Bag 4800, Christchurch, New Zealand.