

RESEARCH ARTICLE

Open Access

Morphological differentiation despite gene flow in an endangered grasshopper

Eddy J Dowle^{1,2*}, Mary Morgan-Richards¹ and Steven A Trewick¹

Abstract

Background: Gene flow is traditionally considered a limitation to speciation because selection is required to counter the homogenising effect of allele exchange. Here we report on two sympatric short-horned grasshoppers species in the South Island of New Zealand; one (*Sigaus australis*) widespread and the other (*Sigaus childi*) a narrow endemic.

Results: Of the 79 putatively neutral markers (mtDNA, microsatellite loci, ITS sequences and RAD-seq SNPs) all but one marker we examined showed extensive allele sharing, and similar or identical allele frequencies in the two species where they co-occur. We found no genetic evidence of deviation from random mating in the region of sympatry. However, analysis of morphological and geometric traits revealed no evidence of morphological introgression.

Conclusions: Based on phenotype the two species are clearly distinct, but their genotypes thus far reveal no divergence. The best explanation for this is that some loci associated with the distinguishing morphological characters are under strong selection, but exchange of neutral loci is occurring freely between the two species. Although it is easier to define species as requiring a barrier between them, a dynamic model that accommodates gene flow is a biologically more reasonable explanation for these grasshoppers.

Background

Although taxonomy implies abrupt disjunctions between biological entities, we know that speciation usually involves non-instantaneous change [1,2]. The existence of hybrids and the implications of hybridisation have long intrigued evolutionists [3], however the incorporation of gene flow into speciation models has only recently gained acceptance [4-6].

Genetic introgression occurs when two genetically distinct populations come into contact enabling individuals from each to interbreed. When this occurs through secondary contact the process has usually been regarded as hybridisation [7,8], however, broader definitions of hybridisation accommodate the continuum from normal intraspecific mating to rare interspecies exchange [9]. The fertility of resulting offspring mediates gene flow between populations. This situation underpins the popular biological species concept [10], but the frailty and circularity of the *a priori* assumption that species are always reproductively isolated is readily demonstrated [4,11]. Stark reminders that

hybridisation is not a valid test of species status come from observations that it is also a potent force in plant speciation that can result in the rapid formation of distinct and reproductively isolated taxa [12,13].

Since the 1960's, direct measures of variable genetic loci have provided strong evidence that genomes are not unitary and exchange of loci between populations may be uneven [5,9,14,15]. Where gene flow between somewhat distinct genomes is not contained by the formation of hybrid zones [8] or abrupt speciation [16], it can have numerous outcomes; reinforcement of reproductive barriers, the evolution of a new species, the loss of one or both parental species, limited adaption due to homogenization, or provide a means to pass adaptive traits between populations [6-8,17,18].

Gene flow may therefore result in species with mosaic genomes, comprised of alleles from different ancestral populations, which has been described as a potentially important evolutionary mechanism for the formation of many animal species [12,13,15]. Indeed, allelic leakage may be fairly persistent where gene flow is mediated not by extrinsic geophysical barriers, but by locus-specific selection [9,19-23]. Empirical data showing the maintenance of incipient species in the face of ongoing gene flow

* Correspondence: Eddy.Dowle@cawthron.org.nz

¹Ecology Group, IAE, Massey University, Private Bag 11222, Palmerston North 4442, New Zealand

²Coastal and Freshwater Group, Cawthron Institute, Nelson, New Zealand

between populations are gradually accumulating, aided by increasingly sophisticated genetic tools [24–30].

Historically, one of the most informative animal groups in this field of study have been Orthoptera and in particular grasshoppers [14,31–35]. Here we report on flightless New Zealand short-horned grasshoppers (Orthoptera: Acrididae). Most of the fifteen New Zealand species, in four endemic genera, occupy subalpine native grasslands above the tree line [36]. Prior to the arrival of humans in New Zealand (~1260 AD), the landscape was mostly dense forest [37–39]. Grasshopper habitat was therefore mostly in the mountain ranges of the South Island, although a few species occur at lower altitude in areas with semi-arid climate or braided river-beds (*Brachaspis robustus*, *Sigauss minutus* and *Sigauss childi* [40–42]).

The species *Sigauss australis* appears, on the basis of mtDNA sequence data, to encompass several narrow endemics and one widespread species [42,43]. Typical *Sigauss australis* are relatively large (adult females ~26 mm) and abundant in South Island subalpine grasslands between 1000 and 1800 m asl. Sympatric with this widespread species is the microendemic *Sigauss childi*, which is restricted to a low-lying, semi-arid region of about 100 km² around the town of Alexandra (Central Otago) (Figure 1). *Sigauss australis* is also present in this region, but the two species are readily distinguished by their appearance. An intriguing feature of *S. childi* is that their colour patterns appear to be specific to the substrate on which individuals are found. Colour patterns within *S. childi* range from almost white or grey on quartz pebbles, brown and red on schist gravels, to green and black like the tumbling lichen (*Chondropsis semiviridis*) that grows on rocks in some areas of Central Otago. Although, inferences of camouflage are subjective they support the conjecture that these grasshoppers are under selection by visual predators. *Sigauss australis* are more boldly patterned, often with longitudinal stripes, and tend to be colour-pattern variable within locations.

Sigauss childi, although morphologically distinct, could not be distinguished from *S. australis* using mtDNA data [42]. Perhaps *S. childi* evolved recently and has retained ancestral mtDNA haplotypes (Incomplete Lineage Sorting), or has exchanged genetic information since diverging [43]. Or perhaps divergence has occurred and been maintained despite gene flow. Genetic exchange between populations might be experienced at different rates across the genome; selection could operate on some loci to limit local exchange of alleles even when net (genome wide) gene flow continues. These alternatives make different predictions about the pattern of morphological and genetic character sharing (Figure 2). In order to understand the evolution of this system we applied six types of data; morphology, mtDNA sequencing, microsatellite genotyping, multi-copy nuclear sequencing, single nucleotide polymorphisms (SNP) and spatial position. We used these putatively independent

data to contrast species integrity as characterised by morphology (subject to natural selection) and neutral characters that allowed us to test the stability of species delimitation, assess the extent and evenness of gene flow and thus gain an understanding of where these grasshopper populations are in the speciation continuum.

Results

Morphology

Discriminant analysis of traditional morphological data collected from species description traits of 169 adult grasshoppers revealed strong support for the two entities *Sigauss childi* and *S. australis* (Table 1). All 28 adult *S. childi* were readily separated from the remaining sample containing *S. australis*.

Principal component analysis also revealed two distinct groups; *S. childi* vs. all other *S. australis* specimens in the analysis with these groups further subdivided into males and females (Figure 3). The first four components of the PCA accounted for >95% of the variation. Thus the morphometric data based on traditional taxonomic characters suggest there are just two morphological entities: *S. childi* and *S. australis*.

Geometric analysis of the grasshopper pronotum gave a similar result to that of traditional morphology. PCA analysis of the 14 landmark measurements showed two major groupings (Figure 4a); one composed of *S. australis* and the other composed of *S. childi*. All *S. childi* individuals grouped together due to their distinctive pronotum shape. Variation within *S. childi* was likely due to their extremely cryptic shape formed by the 'broken' edge of the pronotum resulting in little uniformity within species. The *S. childi* mean was significantly different from *S. australis* over the entire range ($P < 0.001$ T-square: 1101.9925) and *S. australis* in the area of sympatry ($P < 0.001$ T-square: 935.5225).

The PCA of pronotum shape from those grasshoppers sampled within the area of sympatry was scrutinised for evidence of hybrids (Figure 4a). However, within the area of sympatry, not only did the means of *S. australis* and *S. childi* differ ($P < 0.001$ T-square 1291.4541) but there was no overlap between the two forms. Thus we found no evidence of morphological intermediates in adults or juveniles.

Mitochondrial DNA sequence

Partial Cytochrome oxidase subunit I was sequenced from 59 grasshoppers and combined with 46 previously published sequences [43] (GenBank EF544487–EF544562). A total of 66 haplotypes were identified in the alignment (519 bp) representing 105 individuals (72 *S. australis* and 33 *S. childi*). Six haplotype clusters were identified that were each geographically restricted within the range of *Sigauss australis* (Figure 4b). The MtDNA clusters did not

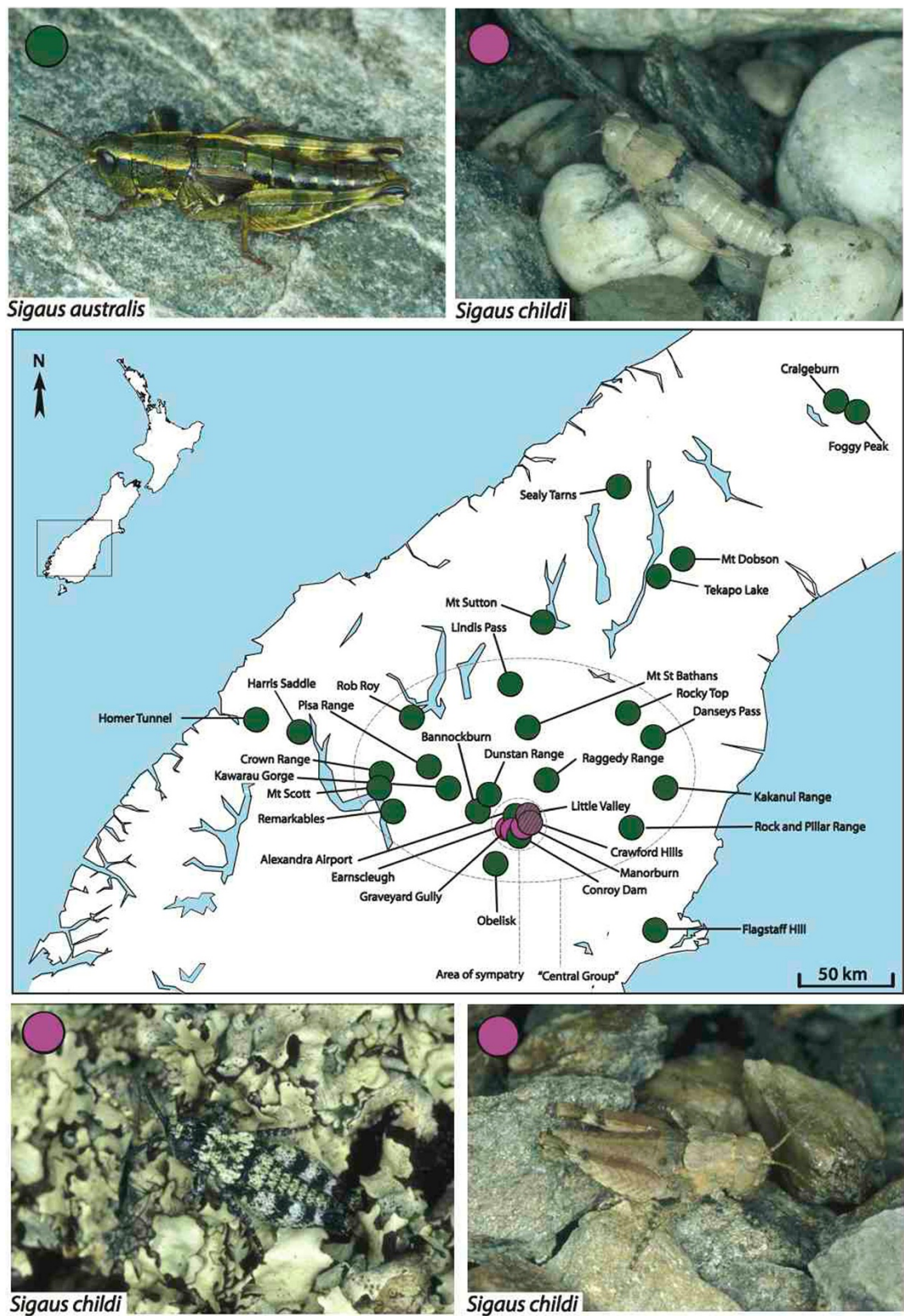


Figure 1 (See legend on next page.)

(See figure on previous page.)

Figure 1 Sample Map. Sample locations in South Island, New Zealand, of the *Sigauss australis* complex grasshoppers used in this study. The two main species *Sigauss australis* (green) and *Sigauss childi* (pink) are morphologically very different; *S. childi* tends to be smaller and more camouflaged to its local habitat than *S. australis*. The 'Central Group' and 'Area of sympatry' defined here are used to analyse subsets of the specimens. Hatched circles represent locations with both species present.

correspond to current taxonomic groups or morphological types. MtDNA sequences from *S. childi* fell in several parts of the *Sigauss australis* complex phylogeny (Figure 4b). NETWORK analysis revealed the extent of sharing between *S. australis* and *S. childi* (Figure 5). As a result of interspecific sharing, some *S. childi* have haplotypes more similar to haplotypes in *S. australis* than other *S. childi*. No single mtDNA clade can be confidently construed as being primarily associated with *S. childi* ancestry. The high genetic diversity detected at Alexandra appears to be a

result of the meeting of three distinct mtDNA clades that otherwise have separate ranges.

No evidence for isolation by distance was detected among the mtDNA diversity within the *S. australis* complex ($p = 0.5410$). Although the mtDNA diversity within clades also did not fit a model of isolation by distance there was a non-significant positive relationship ($p = 0.1410$) and the power of this test was limited by smaller within-clade sample sizes. There was no statistical support for population genetic differentiation between *S. childi* and *S.*

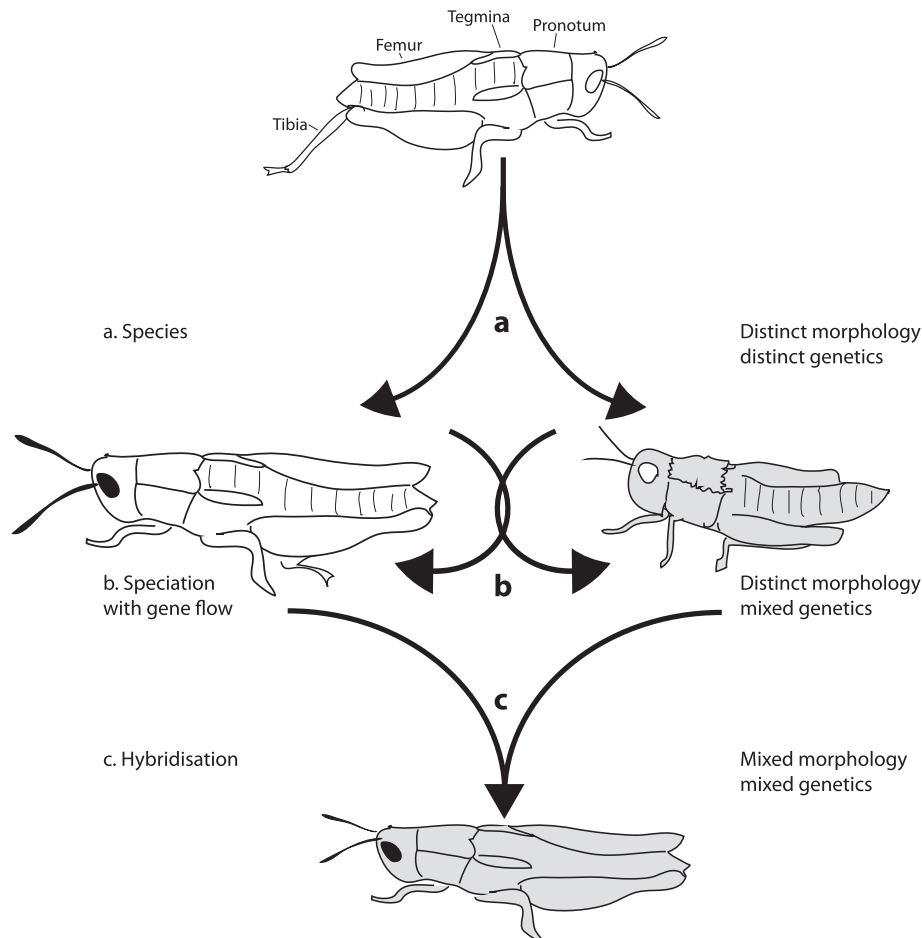


Figure 2 Alternative hypotheses. Alternative hypotheses to explain the relationship between morphological differentiation and gene flow in this study of grasshoppers in New Zealand. Two morphologically defined species exist that have a common ancestor and may share identical alleles due to descent, but as distinct species derived alleles are also expected. (a) An abrupt speciation event is expected to result in accumulation of distinct traits and genotypes, including at neutral loci. (b) Sharing of neutral genetic alleles might be maintained by ongoing gene flow whilst alleles at some loci are subject to selection. Morphological difference is an observable expression of genetic loci under diverging selection. (c) Hybridisation or reticulation is expected to result in individuals with intermediate forms.

Table 1 Discriminant analysis

Summary of classification with cross-validation		
True group	Predicted group	
	<i>S. australis</i>	<i>S. childi</i>
<i>S. australis</i>	134	0
<i>S. childi</i>	0	28
Total N	134	28
N correct	134	28
Proportion	1.000	1.000

Discriminant analysis with cross-validation using the character states employed in traditional species diagnosis for adult grasshoppers of the *Siga* complex in South Island, New Zealand. *Siga childi* (in bold) individuals were correctly grouped together. The total squared difference between the two groups was 274.

australis from either the central group $F_{CT} = 0.09175$ ($P = 0.09677$), or from the area of sympatry $F_{CT} = 0$ ($P = 1$) (Figure 1).

Microsatellites

The three microsatellite loci surveyed each had between 16 and 18 alleles. No evidence of linkage disequilibrium was detected and Hardy-Weinberg expectations were met in the majority of population samples. A positive relationship between geographic distance and genetic differentiation (pairwise F_{ST}) supported a model of isolation by distance for *Siga australis* (Figure 1) (57 individuals; $P = 0.0108$). Analysis of *S. australis* microsatellite data (excluding *S. childi* samples) using STRUCTURE [44] showed

evidence of extensive gene flow among populations. The optimum ΔK was $K = 2$ (Figure 4c), which is consistent with a grouping of populations in Central Otago (Figure 1 and 4b).

Analysis of microsatellite data from 57 *S. australis* and 13 *S. childi* individuals resolved the same geographic subdivision of genetic variation: $K = 2$ (Figure 4c). There was no support for $K = 3$, which was unexpected given that the data encompassed two morphologically distinct species and spatial structure had already been indicated. It is important to note that although the microsatellite dataset covers a similar geographical range to that of the mtDNA dataset there is little similarity in the genetic structure detected. To reduce the possible influence of uneven sample size of the two species we restricted the data to include sampling only from the area of sympatry (Figure 1). STRUCTURE [44] analysis found no support for genetic partitioning within these data (i.e. $K = 1$), contrary to the expectation that the two morphologically defined species would represent discrete genetic units (Figure 4c).

As with the mtDNA data we sought evidence of genetic structure concordant with taxonomy and morphology using analysis of the correlation of genotypes between species by grouping the samples according to the morpho-species *S. australis* and *S. childi* and estimating F_{CT} . No significant genetic differentiation between the morpho-species was found within the central group: $F_{CT} = 0.01780$ ($P = 0.18573$), or the area of sympatry $F_{CT} = -0.07424$ ($P = 0.65494$), although these samples were not all taken from the same generation. This result was consistent with the inference from STRUCTURE.

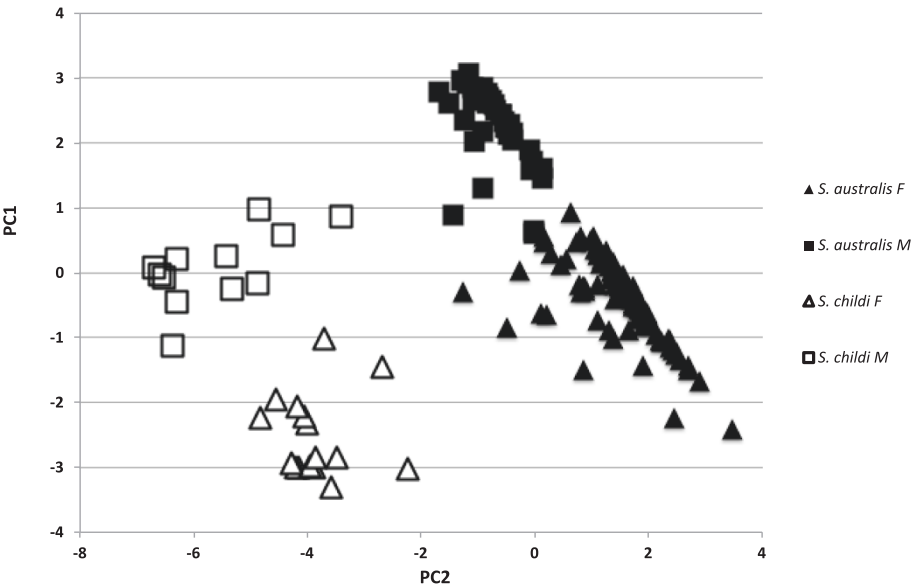


Figure 3 PCA traditional species diagnostics. Principle component analysis using morphological character states used in traditional species diagnostics of adult *Siga* grasshoppers from South Island New Zealand. Morphologically, *Siga childi* is readily separated from *S. australis*, F = females (triangles); M = males (squares).

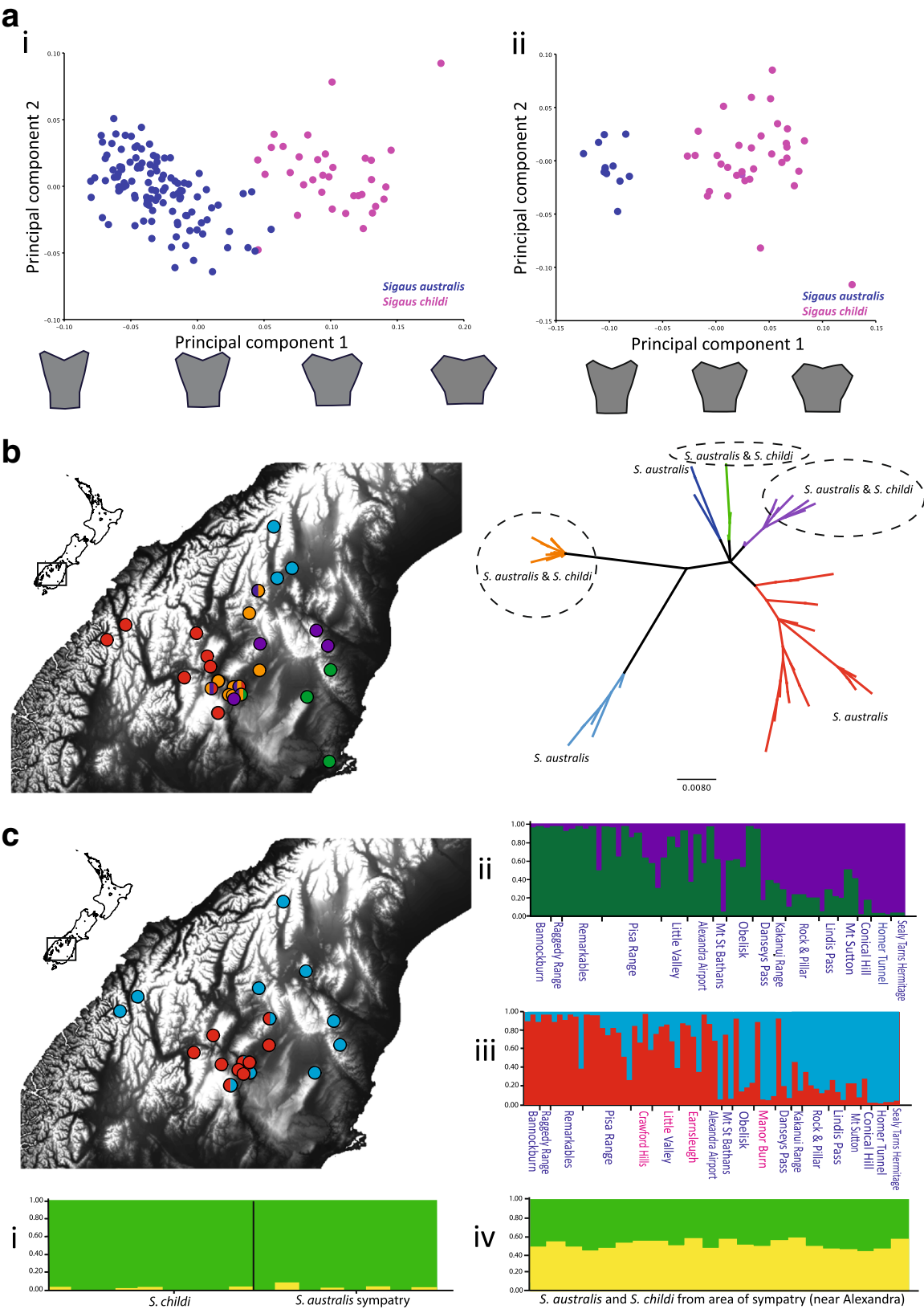


Figure 4 (See legend on next page.)

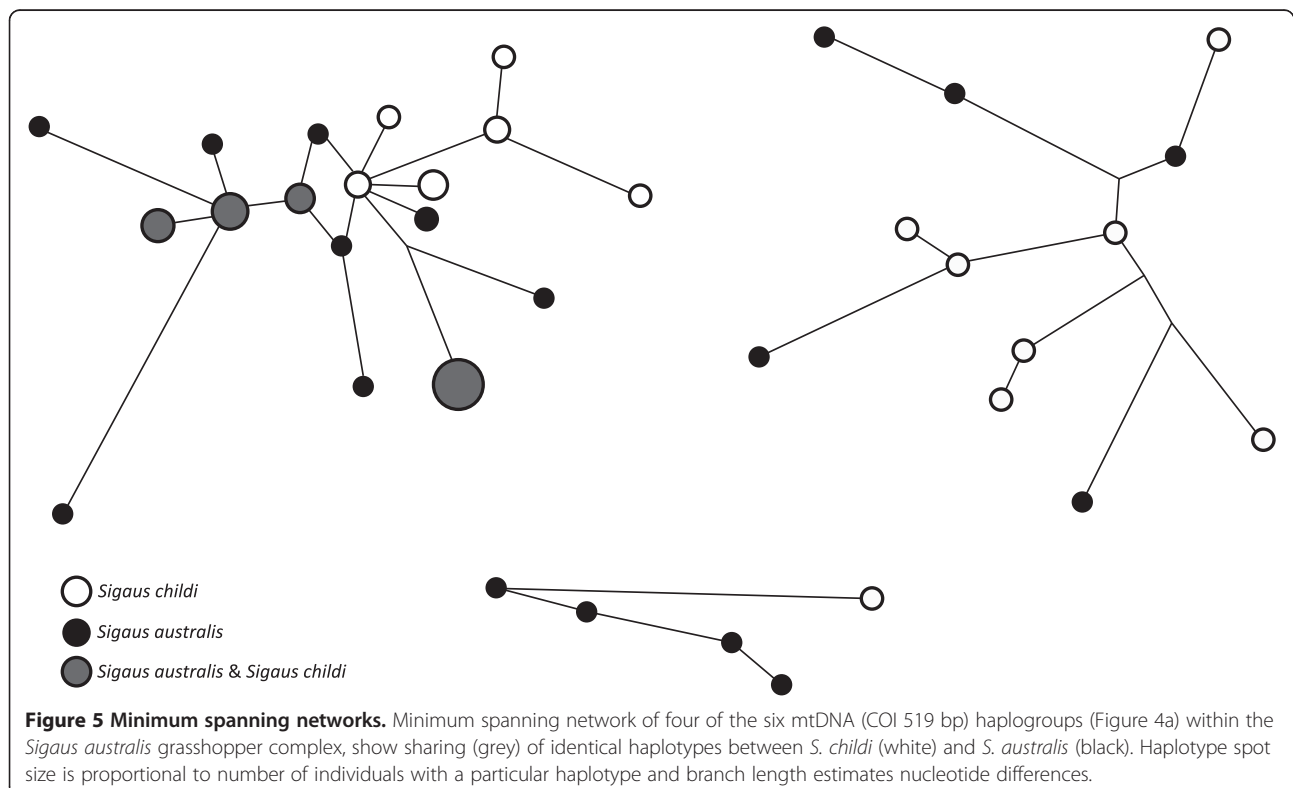
(See figure on previous page.)

Figure 4 Genetic and morphological structure within *Siga* grasshoppers in South Island New Zealand. (a) Variation in the shape of the pronotum using digital imagery of *Siga* grasshopper pronotum shape (PCA analysis from MORPHOJ): (i) PCA for both species from all areas, with pronotum shape changes indicated along the PC1 axis. The two major groupings comprise *S. childi* separated from *S. australis* (ii) PCA for just individuals from the area of sympatry (Figure 1), with pronotum shape changes indicated along the PC1 axis. Within the area of sympatry there was no clear evidence of morphological intermediates. (b) Neighbour Joining tree of mtDNA haplotypes (COI, 519 bp) and a distribution map showing the spatial distribution of haplogroups. The circled clades (dashed circles) are used in the Network analysis (Figure 5). (c) Genetic structure genotype data (STRUCTURE analysis) for the *Siga* grasshoppers: (i) Results from the 74 RAD-seq SNPs of the two species in sympatry, *S. australis* and *S. childi* K = 2 (ii) Microsatellite genotypes from *S. australis* populations at K = 2; (iii) Microsatellite genotypes from *S. australis* and *S. childi* individuals (colours shown on map) at K = 2; (iv) Microsatellite genotypes from *S. australis* and *S. childi* individuals from the area of sympatry (see Figure 1) at K = 2.

Nuclear sequencing

We amplified and sequenced the ITS region (706 bp including 5.8S, ITS1 and 2) from 40 grasshoppers. Some of the grasshoppers had unambiguous single ITS sequence but many had more than one ITS sequence, consistent with these grasshoppers being heterozygotes of mixed ancestry. Of twenty-five grasshoppers (15 *S. childi* and 10, *S. australis*) collected near the township of Alexandra, 16 (11 *S. childi*, 5 *S. australis*) had more than one sequence which differed by the presence of an INDEL approximately 100 bp from the ITS1 forward primer. Sequences of ITS2 from these individuals were unambiguous except at single nucleotide polymorphic sites (SNPs), confirming that these grasshoppers carried more than one ITS sequence per genome. There were 16 SNPs in the set of unambiguous sequences. However the presence of an INDEL near the start of ITS1 meant grasshoppers with more than one

sequence had only 13 observable SNPs. Only one of the 15 grasshoppers from outside the Alexandra area (a specimen from Raggedy Range) appeared to have more than one ITS sequence per genome that involved the large INDEL. However, many individuals (mostly collected from the northern part of the species' range) had an independent 8 bp insertion that occurred in all their copies of ITS. When more than one ITS sequence was detected in a grasshopper DNA, we found that ambiguity could parsimoniously be explained by combinations of unambiguous (single copy) sequences that we separately identified in other grasshoppers. We detected only one copy of ITS in a genome-wide survey of *Siga australis*, suggesting that there was only one family of ITS in these grasshoppers; sequence variation within this family occurs where individuals have recently exchanged genetic material.



Rad-Seq SNPs

The Illumina sequencing provided 9,789,323 forward reads of 100 bp for the 30 grasshoppers, of which 8,934,377 were retained after quality checks in process_radtags.pl (part of the Stacks package). These comprised between 3,743 and 978,246 reads per individual with a total of 30,439 loci. Three individuals (2 *S. australis* and 1 *S. childi*) were removed due to low coverage. From these data sets we identified 8,958 loci that occurred in >2 individuals and these were subjected to further selection based upon coverage per population. The Lindis population was removed from subsequent analysis as several individuals failed to produce data of sufficient quality (likely due to poor DNA quality) and most of the putative-loci resolved were not represented in the other samples. Of 8,958 loci, 74 were retained as they occurred in $\geq 50\%$ of the two populations (*Sigauss childi* and *Sigauss australis* in sympatry). The relatively low proportion of loci that were represented across both population samples was due to insufficient representation of their very large genome despite our protocol involving quantification and compensation for this. Nevertheless we obtained ample data for our purpose.

The distribution of pairwise (*S. childi*/*S. australis*) F_{ST} values for each of the 74 putative-loci revealed the high frequency of low scores expected in the absence of significant structure (Figure 6a). A test for deviations from expected frequencies of neutral loci in BayeScan indicated that one marker may have been subject to diversifying selection, $\log(PO) > 0$ alpha 0.878 (Figure 6b). A BLAST search of the sequence containing this SNP did not result in any matches to known sequences on Genbank. Mean population pairwise F_{ST} was low (0.025), with a confidence interval that effectively included zero (CI 2.5% 0.001, CI 97.5% 0.053), providing little evidence that these samples represent more than one population, with random mating. Population differentiation estimated with STRUCTURE suggested extensive sharing of genetic material among populations, with no species structure detected (Figure 4ci).

Table 2 Migration rates

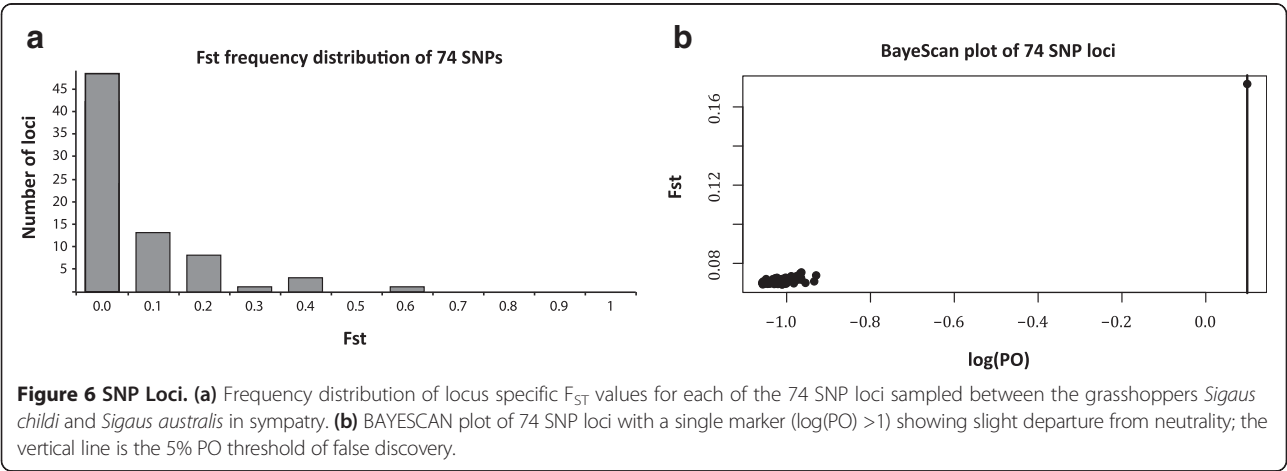
All loci	Mean	0.025	0.975	Nm
θ_1	0.09632	0.08413	0.10907	
θ_2	0.08723	0.07547	0.09907	
$M2 > 1$	2046.7	1927.3	2051.0	49.3 (2- > 1)
$M1 > 2$	3010.8	2874.7	3010.7	65.6 (1- > 2)

Extensive gene flow between the grasshoppers *Sigauss australis*₁ and *S. childi*₂ in sympatry was revealed using MIGRATE-N with 74 RAD-seq SNP makers. Theta θ is an estimate of population size, $\theta = 4N_e\mu$ in the SNPs, where N_e is population size and μ is mutation rate, population size was generally large as is expected for grasshoppers. Mutation scaled migration rates (M) were converted into Nm (number of migrants per generation) via $\theta_1M_{2 \rightarrow 1} = 4Nm_1$. The results show extensive gene flow in both directions.

This was confirmed by analysis in MIGRATE-N, which indicated extensive gene flow between the species in the zone of sympatry (Table 2).

Discussion

We found morphological support for two distinct entities consistent with their existing taxonomic treatment as species: *Sigauss australis* and *Sigauss childi*. *Sigauss australis* has a comparatively wide geographic range that can be subdivided into a number of phylogeographically distinct mtDNA haplogroups. *Sigauss childi* is nested within *S. australis* in terms of habitat, geographic range and genetic diversity; in stark contrast to its clear morphological distinction. Despite occurring in sympatry, no phenotypic intermediates were detected. We found no evidence of genetic partitioning in putatively-neutral mtDNA sequence, microsatellite, ITS sequence loci or SNP data. None of the mtDNA diversity detected within *Sigauss australis* is concordant with current taxonomic subdivision. Microsatellite allele frequencies within *S. childi* are indistinguishable from those within sympatric *S. australis* suggesting recent (and on-going) gene flow. The SNP data show no population structure and extensive gene flow between the two species in sympatry, with one marker showing some sign of diversifying selection. The



presence of more than one ITS1-ITS2 sequence within the genomes of single grasshoppers is also consistent with recent gene flow near the township of Alexandra. Concerted evolution normally results in homogenisation of variation in the rRNA cassette containing ITS [45].

A possible explanation for this lack of genetic differentiation is that *S. childi* is a phenotypic variant of *S. australis* generated by different local environmental conditions. Extreme plasticity is known to occur in other grasshoppers such as the locust (*Locusta migratoria*) which has two distinct life history strategies that are partitioned in time [46]. Indeed, *S. childi* is restricted to a limited lowland habitat in contrast with the subalpine environment experienced by most *S. australis*. The lowland conditions of Central Otago have been described as semiarid [47,48], but despite altitudinal separation from subalpine areas, these habitats are climatically similar in terms of their extreme day/night and seasonal temperature cycles. More pertinent is the fact that *S. australis* and *S. childi* are sympatric in the lowland semiarid environment of Central Otago. They occur at the same places at the same times with, for example, specimens of both morphotypes used in our analysis collected as adults within metres of each other in Little Valley, Alexandra on the same day. These circumstances are inconsistent with an interpretation of phenotypic differences being driven by environmental induced plasticity. Although we cannot exclude this possibility, a novel type of micro-environmental control of grasshopper development would need to be invoked.

Alternatively, the observed morphological divergence but lack of genetic structure is consistent with strong character specific selection in the presence of high levels of gene flow or incipient speciation (Figure 2b). Under both such circumstances neutral genetic markers, such as those examined, may not detect population structuring [22]. Contemporary introgression can be difficult to distinguish from incomplete lineage sorting [49,50], however the sharing of identical mtDNA haplotypes and sharing of alleles across neutral nuclear loci in these two species suggests that very recent and/or ongoing reticulation is more likely. Gene flow is expected to homogenise variation between species and it has traditionally been considered that speciation is unlikely to proceed in the presence of gene flow. Many models of speciation have emphasized partitioning of populations by some extrinsic process (e.g. allopatry) as a prerequisite [10,51]. However, if selection on particular loci is sufficiently intense, the effects of gene flow could be mitigated. Models that accommodate permeability of putative species boundaries and acknowledge that selection can be locus-specific rather than genome wide are not new [5,14,20,52], and empirical data that demonstrate this process are emerging [30,53-55].

Different evolutionary responses to selection could explain the observed morphological distinction. *Sigauss childi*, is a small, highly cryptic species of grasshopper suggesting it has been or is under selection from visual predators. In contrast, *Sigauss australis*, is a larger grasshopper, usually with more striking colour markings that are more easily observed against the substrate, suggesting a different mode of predator avoidance. Although the two species appear to share the same neutral alleles; they remain morphologically distinct with a single genetic marker showing some evidence of diversifying selection between the two species. All individuals examined fell into one of two morphological groups and no specimens could be classed as morphologically intermediate within the area of sympatry even when juveniles were examined. The apparent lack of morphological intermediates (hybrids) suggests that selection, even in today's highly modified environment, is intense. It should be noted, however, that F1 phenotypes are often not intermediate between parents and this might also explain our observations [56]. The conservation status of *S. childi* limited our sample sizes and precluded any observation of mating behaviour and juvenile colouration and survival where the species are sympatric. Further examination might reveal morphological intermediates.

Situations where speciation and selection are most likely to be observable in nature are those with high environmental heterogeneity, temporal instability and/or novel environments [57,58]. Not surprisingly many examples of contemporary speciation in action therefore come from anthropogenic settings [59,60], and this may be relevant to these *Sigauss* grasshoppers. The South Island of New Zealand was settled by Polynesian colonists starting about 800 years ago, and this was accompanied by episodes of scrub and forest fire [61,62]. Expansion of grass and herbs following reduction of forest that may previously have formed a habitat barrier between the alpine and lowland grasshoppers may have facilitated population mixing. The area now shared by *S. childi* and *S. australis* was further modified by European introduction of plants and grazing animals, and mining practices in the last 150 years. These changes could have facilitated increased gene flow, but were analogous to the effects of Pleistocene climate cycling. Disentangling their respective influence on the grasshoppers is not simple [63].

The taxonomic status of these species is problematic, as traditional methods cannot resolve the conflicting information from morphology and genetics resulting from the process of evolution. Although *S. childi* is not genetically isolated from *S. australis* it is morphologically well differentiated, and in our relatively small SNP dataset we were able to find one marker possibly under selection. Models of speciation with gene-flow predict a continuum from partially isolated populations to reproductive isolation

[5,23]. In the *Siga* system divergence seems to be at an early stage; *Siga australis* and *Siga childi* do not appear to be losing morphological distinction, but our neutral genetic data does show extensive gene flow. The absence of any morphological hybrids suggests selection is intense, removing relatively conspicuous intermediates and holding these two species apart. This may provide them the opportunity to diverge at other loci.

Conclusions

The findings of our study endorse Charles Darwin's original dynamic view of speciation [3], but are contrary to those expected from a more restrictive but popular view of species as reproductively isolated units [10]. There is a grand irony that while for many, genetic methods are seen as tools for testing species status (e.g. DNA barcoding), genetic data are actually the key to revealing that speciation is not clear cut [2,20,23,54,55]. In our study we found that the only characters that reliably distinguished species were morphological, whilst 78 neutral genetic markers showed that distinct morphotypes do not correspond to genetically isolated units.

Methods

We collected grasshoppers by hand when they were active during the New Zealand summer season (December–March, between 1995 and 2009). Sampling included all recognised members of the *S. australis* complex (*S. australis*, *S. childi*, *S. obelisci* and *S. homerensis*) from their full geographic range (Figure 1). As already noted *S. australis* is widespread in subalpine habitat with a few populations extending down to low elevation (~300 m asl) areas in some locations. *Siga childi* occurs only at a single low elevation site in Central Otago where it is sympatric with *S. australis*. Due to the legal protection given to the endangered *S. childi*, sample sizes were limited and sampling spanned more than one overlapping generation (c.f. usual assumptions of population genetic models). Sampling from different generations, which are already overlapping is, however, not likely to increase the similarity of population allele frequencies, and therefore we do not consider this will have hindered any of the analyses. *Siga obelisci* and *S. homerensis* are each recorded from single subalpine locations within the range of *S. australis*. The identity of *S. obelisci* and *S. homerensis* specimens were confirmed by Simon Morris (pers comm. to SAT). Individuals were preserved by freezing or in 95% ethanol and identified following Bigelow [36], Morris [64], and Jamieson [40].

Morphology

Morphological data were collected for all adult *Siga australis* complex grasshoppers in two ways. The first used the traditional species diagnostic characteristics,

although we note that much of the information used to distinguish some of these taxa has been geographic location and altitude [65,66]. The exception is *Siga childi*, for which the sinuous caudal margin of the pronotum and tegminal size are diagnostic [40]. Male genitalia are taxonomically informative for some grasshopper species but consistent differences have not been reported among species in this complex. For instance, male genitalia in *S. homerensis* and *S. obelisci*, are each described as being near identical to *S. australis* and *S. "remarkables"*, which is a synonym of *S. australis* [64–66]. Thus male genitalia can be interpreted as being variable within *S. australis*, but uninformative for species delimitation [36,40,64]. Because of this absence of diagnostic information and in light of previous genetic information [43] *S. homerensis* and *S. obelisci* are incorporated into *S. australis* here.

The traditional species diagnostic characters for the grasshoppers were examined and measured with the aid of a dissecting microscope. Four metrics were recorded for each grasshopper using callipers accurate to 0.01 mm; maximum pronotum width, mid-line pronotum length, femur length and body length. In addition, six characters with discrete states were examined; sex, length of tegmina (see below), shape of pronotum posterior margin (sinuous or concave), cuticle rugosity (rugose or smooth), shading on pronotum posterior margin (pigmented or not), and shape of pronotum lateral margins (irregular or smooth). For most characters, the alternative and intermediate states were coded as 2, 0 or 1 respectively. The length of tegmina was classified by reference to the number of abdominal tergites across which they extended; not beyond the pronotum (as in many *S. childi*) coded 0, not beyond first abdomen segment coded as 1, and so forth to a maximum of 4 (no tegmina ever reached beyond the posterior margin of the 5th tergite).

Adults were distinguished by the tegmina concealing the relictual hind wing, which is the case only in the last instar. Juveniles were excluded from this morphometric analysis. The data were analysed using Discriminant Analysis and Principle component analysis (PCA) approaches implemented in MINITAB 15 [67]. The discriminant analysis with cross validation tested whether character information could be used to group individuals into their *a priori* categories: *S. australis* and *S. childi*. A PCA was applied to all the morphological characters and the scores saved. PCA required no *a priori* grouping, allowing us to determine whether the data could be partitioned into taxonomically meaningful groups based solely on the documented morphological character states.

As an alternative to the traditional taxonomic characters, we tested for shape differences of the pronotum among species using geometric analysis. This method is more powerful and avoids any circularity that could arise from

using traditional species characteristics as the sole morphological traits analysed. Much of the taxonomy in the *Siga* genus relies on the pronotum shape, but descriptions are often vague, based on discrete states and inferred from few individuals making species identification difficult [36,65,66]. Using two digital images of the pronotum of each of 147 individuals (113 *S. australis*, 34 *S. childi*) that were obtained with the aid of a dissecting microscope we tested whether shape variation could be detected from metric data. Using IMAGEJ [68], 14 landmarks were identified around the perimeter of the dorsal surface of the pronotum on each image of each grasshopper and measured. The landmarks were selected to maximise variation among individuals. These measurements were analysed using MORPHOJ [69]. A Procrustes fit aligned by principal axes was performed to eliminate size differences before a Procrustes ANOVA was used to examine the error of image capture. This analysis revealed that the error arising from image capture variation was biologically irrelevant: mean squares for image capture was 32 times smaller than the variation found between individual grasshoppers. Juveniles, adults and both sexes were included in the analyses and tested to confirm they did not partition in the results. Principal component and discriminant analyses with cross validation were each performed on the averaged value for each individual from all four species (34 *S. childi* and 113 *S. australis*). These analyses were also separately applied to the *S. childi* and *S. australis* individuals collected within the area of sympatry (34 *S. childi* 12 *S. australis*) (Figure 1).

Mitochondrial DNA sequence

In order to improve genealogical resolution in relation to taxonomy and geography, we supplemented existing published (genbank EF544523-EF544562) mtDNA Cytochrome Oxidase Subunit I (COI) data for the *Siga australis* grasshopper complex [43]. Tissue was dissected from femora of recently collected grasshoppers and DNA extracted using a salting-out method [70,71]. DNA from specimens preserved for more than one year, was extracted using incubation at 55°C with Proteinase K and a CTAB buffer (2% Hexadecyltrimethylammonium bromide, 100 mmol/L Tris-HCl pH8.0, 1.4 mol/L NaCl, 20 mmol/L EDTA), followed by a combined phenol/chloroform/isoamyl alcohol (25:24:1) cleanup. Extractions were eluted in water and diluted as necessary for PCR reactions. Primers C1-J-2195 and LI-N-3014 [72] were used to target the 3' portion of COI. Polymerase chain reactions (PCRs) were performed in 10 µl volumes using ABgene Red Hot Taq (Thermo Fisher Scientific). Thermocycling conditions were 94°C for three minutes; 94°C for 45 seconds, 52°C for 45 seconds and 72°C for 75 seconds repeated 36 times; followed by a 2 minute final extension. Cycle sequencing

used Perkin Elmer BigDye 3.1 chemistry following the manufacturer's protocols analysed on an ABI Prism 377 DNA sequencer (Applied Biosystems, Inc., Foster City, California). Sequences were checked using SEQUENCHER version 4.10.1 (Gene Codes) and aligned with existing data using SeAl version 2.0 and GENEIOUS PRO version 5.3.4 [73,74].

GENEIOUS was used to estimate a neighbour-joining tree for all the lineages. NETWORK version 4.5.1.6. [75] was used to estimate haplotype networks within clades. To test for a correlation between genetic and geographic distance (expected under a model of isolation by distance), Mantel tests [76] were performed using ISOLATION BY DISTANCE WEB SERVICE version 3.16 [77] with 10,000 randomizations to assess the significance of distance correlations. Distance by distance analysis was applied to all data and separately to data within haplo-groups.

A standard AMOVA was used to test for significant genetic differences based on the estimate of genetic partitioning among groups (F_{CT}) using ARLEQUIN version 3.5.1.2 [78]. The first run tested *S. childi* against all populations in the central group (as indicated in Figure 1) and the second included only those individuals from the area of sympatry (Figure 1).

Microsatellites

To examine population structure using nuclear loci we developed primers to amplify microsatellite loci using a modified enriched microsatellite library protocol (Additional file 1).

Screening of fifty microsatellite loci revealed three that were polymorphic and amplified consistently among a subset of DNA samples from the target taxa. The loci were checked for large allele dropout, stuttering, and null alleles using 1000 randomisations in MICROCHECKER version 2.2.3 [79]. Not all populations had sufficient sampling to analyse in MICROCHECKER and within the area of sympatry the *S. childi* and *S. australis* populations were treated as a single population for this purpose. Although there was evidence of null alleles in some of the loci within some of the populations (*S. childi* sympatry and *S. australis* sympatry), this is unlikely to influence the detection of genetic differentiation [80]. To test a hypothesis of isolation by distance, geographical distances (km) among pairs of *S. australis* population samples were linearly regressed against their pairwise F_{ST} estimates. Nineteen populations had sufficient sampling for this analysis (Figure 1). Mantel testing [76] was performed using ISOLATION BY DISTANCE WEB SERVICE version 3.16 [77] with 10,000 randomizations to assess the significance of distance correlations.

Population structure was assessed without *a priori* groupings using STRUCTURE version 2.3.4 [44]. First we looked for evidence of population structure in the data from *S.*

australis samples only, secondly among all samples from all areas sampled for the complex, and thirdly among all samples collected in the area of sympatry (*S. childi* and *S. australis*) and central group *S. australis* (Figure 1). The analyses were run using an admixture model with correlated allele frequency, 100,000 generations of burn-in followed by 100,000 generations, and the number of groups (K) set from 1 to 20 (10 replicates each). The optimum value of K was found using the ΔK method except for K = 1, which was determined by examination of the bar-plots and structure harvester [81]. Charts were averaged over the 10 replicates and re-drawn using CLUMPP and distruct [82,83]. We sought evidence of genetic differentiation concordant with morphology using the populations within the central group subset identified by STRUCTURE (Figure 1).

A standard AMOVA was used to test for significant genetic differences based on the estimate of genetic partitioning among groups (F_{CT}) using ARLEQUIN version 3.5.1.2 [78]. The first run tested *S. childi* against all populations in the central group and the second only those *S. childi* and *S. australis* individuals from the area of sympatry (Figure 1).

Nuclear sequence

Nuclear sequences representing the internal transcribed spacers (ITS1 and ITS2) of the rRNA cluster and the intervening rRNA 5.8S gene were obtained using the primers ITS4 and ITS5 [84]. PCR conditions and sequencing followed standard protocols as above. Sequences were aligned using GENEIOUS PRO version 5.3.4 [74] and checked by eye. Sequences were generated for all grasshoppers from the area of sympatry of *S. childi* and *S. australis* (the Alexandra region). Alignment and comparison of unambiguous with ambiguous sequences allowed us to identify the most likely combinations of sequences that gave the observed heterozygotes (S3). Where sequence variants differed by single nucleotide substitutions we could identify and resolve the polymorphism. Where sequence variation involved INDELs the resulting length polymorphism was evident by abrupt onset of sustained nucleotide ambiguity at the INDEL position, but sequencing in both directions allowed identification of the combination of sequences involved. To examine the number of families of ITS per grasshopper genome we interrogated a DNA dataset generated by high throughput sequencing. Genomic DNA from a single *Sigauss australis* individual was sequenced on an Illumina Hi-Seq 2000 (Beijing Genomics Institute) resulting in >1GB of sequence. The sequence was de-novo assembled via VELVET [85] with mapping performed using BOWTIE version 2 [86] and the results viewed in TABLET version 1.12.09.03 [87]. The resulting contigs were blasted to Genbank and all matches to ITS were selected, aligned, mapped back and, checked for copy number.

Rad-Seq SNPs

Single nucleotide polymorphic (SNP) anonymous nuclear markers were generated using high throughput sequencing, with individual DNA fragments coded so we could identify individual grasshopper genotypes. The double digest Rad-Seq protocol [88] was applied with minor modifications. We estimated genome size to help us optimise the selection of endonucleases and sequencing coverage. To do this we used flow-cytometry on a FACSCalibur system and CellQuest software (BD Biosciences, San Jose, CA, USA), following staining of cells with Propidium Iodide and reference to an internal control (chicken or locust). Our estimates of the *Sigauss* genome were approximately 11.9 pg (consistent with estimates of other short-horned grasshopper species (<http://www.genomesize.com>)). In light of this information we used the restriction enzymes PstI and BamHI to digest the whole genomic DNA extracted from 30 grasshoppers (10 *S. childi*, 10 *S. australis* from Alexandra, 10 *S. australis* from Lindis). The DNA fragments were tagged with DNA sequences that identified each individual before size selecting at 300-400 bp and pooling as per [84].

Data were generated using an Illumina Hi-Seq (New Zealand Genomics Limited), and sorted using the STACKS version 0.99992 pipeline [89]. Settings for coverage and sites per read were adjusted iteratively. Read coverage settings vary in the literature [88], so we initially ran trials ranging from 7 to 30 reads, but found no alteration in the results. We report results using an optimum coverage of 15 reads per individual (excluding all stacks with a lower coverage), a maximum of two mismatches between reads for a single individual as well as allowing four mismatches between primary and secondary reads within stacks. We allowed the program to remove any potentially spurious highly repetitive stacks. In cstacks we allowed 3 mismatches between samples when generating the SNP set (-m 15 -N2 -M4 -n3 -t). We restricted our analysis to a single SNP per putative locus (always the first), thus avoiding potential problems of non-independence between markers. Data file conversion for programs was performed using PGDSPIDER version 2.0.4.0 [90]. Population pairwise F_{ST} as calculated for each putative-locus across all loci in STACKS, and an AMOVA was run in GENODIVE version 2.0b24 [91] to determine F_{ST} across populations. STRUCTURE version 2.3.4 [44] was used to estimate population differentiation using an admixture model with correlated allele frequency. A burnin of 100,000 generations was followed by 10 replications of 100,000 generations with the number of groups (K) set from 1 to 3. The optimum value of K was found from ΔK method, via structure harvester, except for K = 1, which was determined by examination of the bar-plots [81]. Charts were averaged over the 10 replicates and re-drawn using CLUMPP and distruct [82,83]. STRUCTURE was run using one SNP per

putative-locus (read), and each putative-locus appeared in both *S. australis* and *S. childi* populations, and occurred in $\geq 50\%$ of the individuals.

Gene flow between the two populations in sympatry was estimated using MIGRATE-N version 3.5.1 [92,93], although algorithms that test for gene flow are often not ideal for situations where gene flow is very high, which is likely in this case. MIGRATE-N was implemented with the Bayesian inference strategy. Initial runs involved only half the markers as we optimised settings. The starting values for θ and M were generated initially from F_{ST} with subsequent runs using the resulting θ and M values. The uniform prior distributions were used for both parameters with slice sampling; one long chain was run recording every 5 steps after a burnin of 50,000 with a static heating scheme with five chains. Four runs were conducted on half the data before three final runs on all loci were undertaken using the starting values for θ and M of the previous run. BAYESCAN version 2.01 [90,94,95] was used to examine the individual markers for evidence of selection using the default settings. Prior odds of a neutral model were 10 times more likely than the model with selection at a locus. This prior was tested further by changing it to one, without any identifiable change in results. The alpha value was used to determine the direction of selection with a positive value suggesting diversifying selection and a negative value suggesting balancing selection. Results were viewed in R version 3.0.0 using an FDR of 0.05 [96] and markers with evidence of selection were subjected to a BLAST search via NCBI [97].

Data archiving

All sanger sequenced data (KM576255-KM576292) and raw Illumina reads (bioproject-261083 <http://www.ncbi.nlm.nih.gov/bioproject/261083>) are available on genbank. ITS sequence data from *de novo* assembly is available on genbank (KM576254). ITS data indicating SNP and indel variation is available in Additional file 1. Data used in the geometric morphometrics analysis is provided in Additional file 2 along with the SNP/microsatellite set used in this paper (Additional file 3).

Ethics

This study did not require ethical approval.

Additional files

Additional file 1: Extended Methods. Further details on methodology. Includes details on the microsatellite discovery methodology, along with primer sequences. As well as details on sample number/location at each analysis. As well as a table indicating SNP and indel variations/positions in ITS sequencing.

Additional file 2: Geometric Morphometrics. Procrustes fit values for geometric morphometrics.

Additional file 3: SNP and microsatellites. Spreadsheet of SNPs and microsatellites values across populations used in this study.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

EJD, SAT and MMR conceived and designed experiments and collected grasshoppers. EJD and SAT carried out molecular analysis. EJD carried out morphometric analysis. EJD, SAT and MMR drafted and edited paper. All authors read and approved the final manuscript.

Acknowledgements

The authors have no conflict of interest to declare. We thank: Craig Wilson (New Zealand Department of Conservation), Simon Morris, Julia Goldberg, Ted Trewick and Bee Trewick for their collection help; Vaughan Symonds for allowing us to use his microsatellite library protocol; Trish McLenachan for all her help and support in the lab; Lara Shepherd for all her microsatellite advice; Fran Wolber for instructions and use of flow cytometry equipment. Elizabeth Daly, Briar Smith, Simon Hills and Mariana Bulgarella of the phoenix group (evolves.massey.ac.nz) for their help with analysing results and reviewing the manuscript. The manuscript benefited from comments from a number of anonymous reviewers. This work was funded in part by Department of Conservation Taxonomic Units Fund (contract 4076) and Massey University Vice-Chancellor's Scholarship.

Received: 21 July 2014 Accepted: 29 September 2014

Published online: 16 October 2014

References

- Wright S: Character change, speciation, and the higher taxa. *Evolution* 1982, **36**(3):427–443.
- Mallet J, Meyer A, Nosil P, Feder JL: Space, sympatry and speciation. *J Evol Biol* 2009, **22**(11):2332–2341.
- Darwin C: *On the Origin of Species by Means of Natural Selection*. London: Murray; 1859.
- Mallet J: Mayr's view of Darwin: was Darwin wrong about speciation? *Biol J Linn Soc* 2008, **95**(1):3–16.
- Wu C-I: The genic view of the process of speciation. *J Evol Biol* 2001, **14**(6):851–865.
- Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJ, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R, Butlin RK, Dieckmann U, Eroukhmanoff F, Grill A, Cahan SH, Hermansen JS, Hewitt G, Hudson AG, Jiggins C, Jones J, Keller B, Marczewski T, Mallet J, Martinez-Rodriguez P, Möst M, Mullen S, Nichols R, Nolte AW, Parisod C, Pfennig K, et al.: Hybridization and speciation. *J Evol Biol* 2013, **26**(2):229–246.
- Morgan-Richards M, Smitten RD, Shepherd LD, Wallis GP, Hayward JJ, Chan CH, Chambers GK, Chapman HM: A review of genetic analyses of hybridisation in New Zealand. *J R Soc NZ* 2009, **39**(1):15–34.
- Harrison RG: *Hybrid Zones and the Evolutionary Process*. New York: Oxford University Press; 1993.
- Harrison RG: The language of speciation. *Evolution* 2012, **66**(12):3643–3657.
- Mayr E: *Animal Species and Evolution*. Cambridge: Belknap Press of Harvard University Press; 1963.
- Mallet J: A species definition for the modern synthesis. *Trends Ecol Evol* 1995, **10**(7):294–299.
- Xu S: Phylogenetic analysis under reticulate evolution. *Mol Biol Evol* 2000, **17**(6):897–907.
- Susnik S, Weiss S, Odak T, Delling B, Treer T, Snoj A: Reticulate evolution: ancient introgression of the Adriatic brown trout mtDNA in softmouth trout *Salmo obtusirostris* (Teleostei: Salmonidae). *Biol J Linn Soc* 2007, **90**(1):139–152.
- Barton NH, Hewitt GM: Analysis of hybrid zones. *Annu Rev Ecol Syst* 1985, **16**:113–148.
- Nosil P, Feder JL: Genomic divergence during speciation: causes and consequences. *Philos Trans R Soc Lond B Biol Sci* 2012, **367**(1587):332–342.
- Soltis P, Soltis D: The role of hybridization in plant speciation. *Annu Rev Plant Biol* 2009, **60**(1):561–588.

17. Dowling TE, Secor CL: **The role of hybridization and introgression in the diversification of animals.** *Annu Rev Ecol Syst* 1997, **28**:593–619.
18. Servedio MR, Noor MAF: **The role of reinforcement in speciation: theory and data.** *Annu Rev Ecol Syst* 2003, **34**:339–364.
19. Barton N, Hewitt G: *7 Hybrid Zones and Speciation*, Evolution and speciation: essays in honor of MJD White; 1981:109.
20. Nosil P, Schluter D: **The genes underlying the process of speciation.** *Trends Ecol Evol* 2011, **26**(4):160–167.
21. Harrison R: **Hybrid zones: windows on evolutionary process.** *Oxf Surv Evol Biol* 1990, **7**:69–128.
22. Thibert-Plante X, Hendry AP: **When can ecological speciation be detected with neutral loci?** *Mol Ecol* 2010, **19**(11):2301–2314.
23. Feder JL, Flaxman SM, Egan SP, Nosil P: **Hybridisation and the build-up of genomic divergence during speciation.** *J Evol Biol* 2013, **26**:261–266.
24. Niemiller ML, Fitzpatrick BM, Miller BT: **Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: Gyrinophilus) inferred from gene genealogies.** *Mol Ecol* 2008, **17**(9):2258–2275.
25. Lexer C, Kremer A, Petit RJ: **COMMENT: shared alleles in sympatric oaks: recurrent gene flow is a more parsimonious explanation than ancestral polymorphism.** *Mol Ecol* 2006, **15**(7):2007–2012.
26. Rosenblum EB, Harmon LJ: **"Same same but different": replicated ecological speciation at white sands.** *Evolution* 2011, **65**(4):946–960.
27. Papadopoulos AST, Baker WJ, Crayn D, Butlin RK, Kynast RG, Hutton I, Savolainen V: **Speciation with gene flow on Lord Howe Island.** *Proc Natl Acad Sci* 2011, **108**(32):13188–13193.
28. Bergsten J: **A review of long-branch attraction.** *Cladistics* 2005, **21**(2):163–193.
29. Savolainen V, Anstett M-C, Lexer C, Hutton I, Clarkson JJ, Norup MV, Powell MP, Springate D, Salamin N, Baker WJ: **Sympatric speciation in palms on an oceanic island.** *Nature* 2006, **441**(7090):210–213.
30. Osborne OG, Batstone TE, Hiscock SJ, Filatov DA: **Rapid speciation with gene flow following the formation of Mt. Etna.** *Genome Biol Evol* 2013, **5**(9):1704–1715.
31. Willett CS, Ford MJ, Harrison RG: **Inferences about the origin of a field cricket hybrid zone from a mitochondrial DNA phylogeny.** *Heredity* 1997, **79**(5):484–494.
32. Key KHL: **The concept of sympatric speciation.** *Syst Biol* 1968, **17**(1):14–22.
33. Vazquez P, Cooper SJB, Gosalez J, Hewitt GM: **Nuclear DNA introgression across a Pyrenean hybrid zone between parapatric subspecies of the grasshopper *Chorthippus parallelus*.** *Heredity* 1994, **73**(4):436–443.
34. Maroja LS, Andr  s JA, Walters JR, Harrison RG: **Multiple barriers to gene exchange in a field cricket hybrid zone.** *Biol J Linn Soc* 2009, **97**(2):390–402.
35. Kawakami T, Butlin RK, Adams M, Paull DJ, Cooper SJB: **Genetic analysis of a chromosomal hybrid zone in the Australian morabine grasshoppers (*Vandiemena*, *Viatica* species group).** *Evolution* 2009, **63**(1):139–152.
36. Bigelow RS: *The Grasshoppers of New Zealand*. Christchurch, New Zealand: University of Canterbury; 1967.
37. McGlone MS: **Plant biogeography and the late Cenozoic history of New Zealand.** *N Z J Bot* 1985, **23**:723–749.
38. McGlone MS, Duncan RP, Heenan PB: **Endemism, species selection and the origin and distribution of the vascular plant flora of New Zealand.** *J Biogeogr* 2001, **28**(2):199–216.
39. Wardle P: **Distribution of native forest in the upper Clutha district, Otago, New Zealand.** *N Z J Bot* 2001, **39**(3):435–446.
40. Jamieson CD: *Distribution and Abundance of *Sigaus Childi* Jamieson (Orthoptera: Acrididae), a Central Otago Endemic Grasshopper.* Science for Conservation 110. Wellington, New Zealand: Department of Conservation; 1999.
41. Trewick SA: **Identity of an endangered grasshopper (Acrididae: *Brachaspis*): taxonomy, molecules and conservation.** *Conserv Genet* 2001, **2**(3):233–243.
42. Trewick SA, Morris S: *Diversity and Taxonomic Status of Some New Zealand Grasshoppers. DOC Research & Development Series #290*. Wellington, New Zealand: Science & Technical Publishing, Department of Conservation; 2008.
43. Trewick SA: **DNA Barcoding is not enough: mismatch of taxonomy and genealogy in New Zealand grasshoppers (Orthoptera: Acrididae).** *Cladistics* 2008, **24**(2):240–254.
44. Pritchard JK, Stephens M, Donnelly P: **Inference of population structure using multilocus genotype data.** *Genetics* 2000, **155**(2):945–959.
45. Ganley AR, Kobayashi T: **Highly efficient concerted evolution in the ribosomal DNA repeats: total rDNA repeat variation revealed by whole-genome shotgun sequence data.** *Genome Res* 2007, **17**(2):184–191.
46. Applebaum S, Heifetz Y: **Density-dependent physiological phase in insects.** *Annu Rev Entomol* 1999, **44**(1):317–341.
47. Garnier BJ: *The Climate of New Zealand, a Geographic Survey*. London: E. Arnold; 1958.
48. Hubbard JCE, Wilson JB: **A survey of the lowland vegetation of the Upper Clutha District of Otago, New Zealand.** *N Z J Bot* 1988, **26**(1):21–35.
49. van Oppen MJH, McDonald BJ, Willis B, Miller DJ: **The evolutionary history of the coral Genus *Acropora* (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence?** *Mol Biol Evol* 2001, **18**(7):1315–1329.
50. Jason DW, Zink RM, Blackwell-Rago RC, Nelson DA: **Anomalous variation in mitochondrial genomes of white-crowned (*Zonotrichia leucophrys*) and golden-crowned (*Z. atricapilla*) sparrows: pseudogenes, hybridization, or incomplete lineage sorting?** *Auk* 2001, **118**(1):231–236.
51. Coyne JA, Orr HA: *Speciation*. Sunderland, MA: Sinauer Associates; 2004.
52. Feder JL, Egan SP, Nosil P: **The genomics of speciation-with-gene-flow.** *Trends Genet* 2012, **28**(7):342–350.
53. Wang RL, Wakeley J, Hey J: **Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives.** *Genetics* 1997, **147**(3):1091–1106.
54. Peccoud J, Ollivier A, Plantegenest M, Simon J-C: **A continuum of genetic divergence from sympatric host races to species in the pea aphid complex.** *Proc Natl Acad Sci* 2009, **106**(18):7495–7500.
55. Cui R, Schumer M, Kruesi K, Walter R, Andolfatto P, Rosenthal GG: **Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes.** *Evolution* 2013, **67**(8):2166–2179.
56. Jasinska AK, Wachowiak W, Muchewicz E, Boratynska K, Montserrat JM, Boratynski A: **Cryptic hybrids between *Pinus uncinata* and *P. sylvestris*.** *Bot J Linn Soc* 2010, **163**(4):473–485.
57. Forsman A, Karlsson M, Wennersten L, Johansson J, Karpesten E: **Rapid evolution of fire melanism in replicated populations of pygmy grasshoppers.** *Evolution* 2011, **65**(9):2530–2540.
58. Orr MR, Smith TB: **Ecology and speciation.** *Trends Ecol Evol* 1998, **13**(12):502–506.
59. Palumbi SR: *The Evolution Explosion: how Humans Cause Rapid Evolutionary Change*. New York: Norton; 2002.
60. Stockwell CA, Hendry AP, Kinnison MT: **Contemporary evolution meets conservation biology.** *Trends Ecol Evol* 2003, **18**(2):94–101.
61. McSaveney MJ, Stirling MW: *Central Otago: Basin and Range Country, Landforms of New Zealand*; 1992.
62. McWethy DB, Whitlock C, Wilmshurst JM, McGlone MS, Li X: **Rapid deforestation of South Island, New Zealand, by early Polynesian fires.** *The Holocene* 2009, **19**(6):883–897.
63. Biern   N, Gagnaire P, David P: **The geography of introgression in a patchy environment and the thorn in the side of ecological speciation.** *Current Zoology* 2013, **59**(1):72–86.
64. Morris S: *Distribution and Taxonomic Status of New Zealand Endangered Grasshoppers (Orthoptera: Acrididae). Internal Department of Conservation Report*. Wellington, New Zealand: Department of Conservation; 2003.
65. Morris S: **Two new species of *Sigaus* from Fiordland, New Zealand (Orthoptera: Acrididae).** *New Zealand Entomologist* 2003, **26**:65–74.
66. Morris S: *Identification Guide to Grasshoppers (Orthoptera: Acrididae) in Central Otago and Mackenzie Country*. Wellington, New Zealand: Department of Conservation; 2002.
67. Minitab: **Minitab 15 (Statistical Software)**. State College, PA: 2007. Minitab Inc: www.minitab.com.
68. Abramoff MD, Magalhaes PJ, Ram SJ: **Image processing with ImageJ.** *Biophoton Int* 2004, **11**(7):36–42.
69. Klingenberg CP: **MorphoJ: an integrated software package for geometric morphometrics.** *Mol Ecol Resour* 2011, **11**(2):353–357.
70. Sunnucks P, Hales DF: **Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of the genus *Sitobion* (Hemiptera: Aphididae).** *Mol Biol Evol* 1996, **13**(3):510–524.
71. Trewick SA, Morgan-Richards M: **After the deluge: mitochondrial DNA indicates Miocene radiation and Pliocene adaptation of tree and giant weta (Orthoptera: Anostomatidae).** *J Biogeogr* 2005, **32**(2):295–309.
72. Simon C, Frati F, Beckenbach A, Crespi B, Liu UH, Flook P: *Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain Reaction Primers*, Volume 87. Lanham, MD: Entomological Society of America; 1994.

73. Rambaut A: **Se-Al: Sequence Alignment Editor. v2.0a11.** Oxford, UK: University of Oxford; 2002. Available from <http://evolve.200.ox.ac.uk/>.
74. **Geneious version v5.3.4 created by Biomatters.** Available from <http://www.geneious.com>.
75. Bandelt H-J, Forster P, Röhl A: **Median-joining networks for inferring intraspecific phylogenies.** *Mol Biol Evol* 1999, **16**:37–48.
76. Mantel N: **The detection of disease clustering and a generalized regression approach.** *Cancer Res* 1967, **27**:209–220.
77. Jensen JL, Bohonak AJ, Kelley ST: **Isolation by Distance, web Service.** *BMC Genetics* 6: **13.** v.3.16. In; 2005. <http://ibdws.sdsu.edu>.
78. Excoffier L, Laval G, Schneider S: **Arlequin ver. 3.0: an integrated software package for population genetics data analysis.** *Evol Bioinformatics Online* 2005, **1**:47–50.
79. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P: **Micro-checker: software for identifying and correcting genotyping errors in microsatellite data.** *Mol Ecol Notes* 2004, **4**(3):535–538.
80. Carlsson J: **Effects of microsatellite null alleles on assignment testing.** *J Hered* 2008, **99**(6):616–623.
81. Evanno G, Regnaut S, Goudet J: **Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study.** *Mol Ecol* 2005, **14**:2611–2620.
82. Rosenberg NA: **Distruct: a program for the graphical display of population structure.** *Mol Ecol Notes* 2004, **4**(1):137–138.
83. Jakobsson M, Rosenberg NA: **CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure.** *Bioinformatics* 2007, **23**(14):1801–1806.
84. White TJ, Bruns T, Lee S, Taylor J: *Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for Phylogenetics.* San Diego: Academic Press; 1990.
85. Zerbino D, Birney E: **Velvet: algorithms for de novo short read assembly using de Bruijn graphs.** *Genome Res* 2008, **18**(5):821–829.
86. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nat Meth* 2012, **9**(4):357–359.
87. Milne I, Bayer M, Cardle L, Shaw P, Stephen G, Wright F, Marshall D: **Tablet-next generation sequence assembly visualization.** *Bioinformatics* 2010, **26**(3):401–402.
88. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE: **Double digest RADseq: an inexpensive method for De novo SNP discovery and genotyping in model and Non-model species.** *PLoS ONE* 2012, **7**(5):e37135.
89. Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH: **Stacks: building and genotyping loci De novo from short-read sequences.** *G3 (Bethesda)* 2011, **1**(3):171–182.
90. Fischer MC, Foll M, Excoffier L, Heckel G: **Enhanced AFLP genome scans detect local adaptation in high-altitude populations of a small rodent (*Microtus arvalis*).** *Mol Ecol* 2011, **20**(7):1450–1462.
91. Meirmans PG, Van Tienderen PH: **Genotype and genodive: two programs for the analysis of genetic diversity of asexual organisms.** *Mol Ecol Notes* 2004, **4**(4):792–794.
92. Beerli P: **Comparison of Bayesian and maximum-likelihood inference of population genetic parameters.** *Bioinformatics* 2006, **22**(3):341–345.
93. Beerli P: *How to use MIGRATE or why are Markov Chain Monte Carlo Programs Difficult to use?* New York: Cambridge University Press; 2009.
94. Foll M, Fischer MC, Heckel G, Excoffier L: **Estimating population structure from AFLP amplification intensity.** *Mol Ecol* 2010, **19**(21):4638–4647.
95. Foll M, Gaggiotti O: **A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective.** *Genetics* 2008, **180**(2):977–993.
96. R Core Team: **R: A Language and Environment for Statistical Computing.** Vienna, Austria: R foundation for Statistical Computing; 2013. Retrieved from <http://www.R-project.org>.
97. Johnson M, Zaretskaya I, Raytselis Y, Merezuk Y, McGinnis S, Madden TL: **NCBI BLAST: a better web interface.** *Nucleic Acids Res* 2008, **36**(suppl 2):W5–W9.

doi:10.1186/s12862-014-0216-x

Cite this article as: Dowle et al.: Morphological differentiation despite gene flow in an endangered grasshopper. *BMC Evolutionary Biology* 2014 **14**:216.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit

