Received: 13 November 2015 | Revised: 7 March 2017 | Accepted: 3 April 2017

DOI: 10.1111/mec.14150

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Genomic clustering of adaptive loci during parallel evolution of an Australian wildflower

Federico Roda[1,2] | Greg M. Walter[1] | Rick Nipper[3] | Daniel Ortiz-Barrientos[1]

[1]School of Biological Sciences, The University of Queensland, St. Lucia, QLD, Australia

[2]Harvard University, Boston, MA, USA

[3]Floragenex Inc, Eugene, OR, USA

Correspondence
Federico Roda, Harvard University, Boston, MA, USA.
Email: federicoroda@fas.harvard.edu

Funding information
This work was funded by research grants DP120104559 and DP0986172 from the Australian Research Council to DOB.

Abstract

The build-up of the phenotypic differences that distinguish species has long intrigued biologists. These differences are often inherited as stable polymorphisms that allow the cosegregation of adaptive variation within species, and facilitate the differentiation of complex phenotypes between species. It has been suggested that the clustering of adaptive loci could facilitate this process, but evidence is still scarce. Here, we used QTL analysis to study the genetic basis of phenotypic differentiation between coastal populations of the Australian wildflower *Senecio lautus*. We found that a genomic region consistently governs variation in several of the traits that distinguish these contrasting forms. Additionally, some of the taxon-specific traits controlled by this QTL cluster have evolved repeatedly during the adaptation to the same habitats, suggesting that it could mediate divergence between locally adapted forms. This cluster contains footprints of divergent natural selection across the range of *S. lautus*, which suggests that it could have been instrumental for the rapid diversification of this species.

KEYWORDS
adaptation, ecotypes, parallel speciation, QTL mapping, supergenes, transplant experiment

1 | INTRODUCTION

The many traits that distinguish species often segregate as stable sets of taxon-specific traits, suggesting the existence of mechanisms that prevent the generation of maladaptive intermediate phenotypes (Fisher, 1930). One of these mechanisms could be the clustering of genes governing multiple traits (Schwander, Libbrecht, & Keller, 2014; Thompson & Jiggins, 2014; Yeaman, 2013). Arrays of tightly linked loci often govern fundamental processes like sex determination, self-incompatibility and immunity (Schwander et al., 2014; Thompson & Jiggins, 2014; Yeaman, 2013) but also can promote creating divergence between interbreeding taxa (Kirkpatrick, 2010; Kirkpatrick & Barrett, 2015; Kirkpatrick & Barton, 2006). The genes causing characteristic features of species or ecotypes are often grouped in chromosomal regions where recombination rate is low (Schwander et al., 2014; Thompson & Jiggins, 2014; Yeaman, 2013). For instance, quantitative trait loci (QTLs) for traits mediating divergence between annual and perennial forms of monkeyflowers map to an inversion, suggesting that this chromosomal difference prevented the dissolution of allelic combinations that drive local adaptation (Lee, Fishman, Kelly, & Willis, 2016b; Lowry & Willis, 2010; Twyford & Friedman, 2015).

The role of gene clusters during species formation remains controversial due to the scarcity of evidence documenting their build-up during population divergence (Kirkpatrick & Barrett, 2015; Thompson & Jiggins, 2014), particularly during speciation in the face of gene flow. Theory suggests that the interplay between natural selection and maladapted gene flow can create gene clusters through mechanisms that suppress recombination rate between adaptive loci, or favour chromosomal rearrangements that bring together previously unlinked genes (see Ortiz-Barrientos, Engelstädter, & Rieseberg, 2016 for a review). However, it remains largely unexplored whether the same gene clusters evolve repeatedly across populations exposed to the similar environments.

By integrating results from population genomics and QTL analyses of adaptive traits across multiple population pairs that evolved

repeatedly and independently, one can test the role of specific genomic regions in evolution (Anderson, Willis, & Mitchell-Olds, 2011; Barrett & Hoekstra, 2011; Colautti, Lee, & Mitchell-Olds, 2012; Mitchell-Olds, Feder, & Wray, 2008; Mitchell-Olds, Willis, & Goldstein, 2007; Savolainen, Lascoux, & Merila, 2013; Seehausen et al., 2014). Such systems can reveal whether the genetic basis of ecotypic or species differences is concentrated in few genomic regions, and whether or not independently evolved taxon pairs have produced similar clusters. Cases of parallel evolution are particularly good systems to conduct this type of study as they constitute naturally replicated experiments (Seehausen et al., 2014). Here, we examined the genetic basis of phenotypic convergence in an Australian wildflower to test the importance of concentrated genomic architectures during species divergence in parapatric populations.

The variable groundsel *Senecio lautus* (Asteraceae) is a suitable model to study the parallel build-up of phenotypic differences as it consists of numerous recently diverged ecotypes (the most recent common ancestor of the clade lived 200,000–500,000 years, Melo, 2014) with strikingly different morphologies and life histories (Ali, 1964, 1969; Radford & Cousens, 2006); especially, interesting are populations growing in adjacent sand dunes and rocky headlands along the coast as they appear to have evolved similar phenotypes repeatedly and independently in the presence of gene flow: nearby populations are more genetically similar independently of the environment that they inhabit (Roda et al., 2013a, 2013b). Plants on rocky headlands (hereafter H population) are short-lived perennials, and their growth habit is prostrate. In contrast, individuals found in sand dunes (hereafter D population) are annual and erect (Melo, 2014; Melo, Grealy, Brittain, Walter, & Ortiz-Barrientos, 2014; Radford & Cousens, 2000, 2006; Walter et al., 2016). These differences are maintained when plants are grown in a common environment, indicating that they are genetically based (Melo, 2014; Melo et al., 2014; Radford & Cousens, 2000, 2006). H and D populations from the same phylogenetic clade are reproductively compatible (Melo, 2014; Ornduff, 1964; Richards & Ortiz-Barrientos, 2016) but survive preferentially in their respective habitats (Melo et al., 2014; Radford & Cousens, 2000; Richards & Ortiz-Barrientos, 2016; Richards, Walter, McGuigan, & Ortiz-Barrientos, 2016; Walter et al., 2016). How was this suite of traits repeatedly assembled in a relatively short amount of evolutionary time? Here, we evaluated whether the clustering of loci governing traits has contributed to the repeated evolution of these complex phenotypes. We first investigated the genetic architecture of adaptation in a pair of parapatric populations by conducting quantitative trait loci (QTL) analysis of phenotypic variation. Then, we use bulk segregant analysis (BSA) of survivorship to test whether these clusters are important for environmental adaptation. Finally, we used previously published genomewide screens of differentiation (Roda et al., 2013a, 2013b) to determine whether these clusters show footprints of divergent natural selection across the *S. lautus* range. Overall, our results show that a genomic region governs multiple traits and presents strong genetic differentiation across parapatric populations, suggesting that QTL clusters may have driven the divergence of coastal *S. lautus* forms.

2 | MATERIALS AND METHODS

2.1 | QTL mapping of morphological variation

2.1.1 | Plant material

We used a backcross-like population to construct a linkage map (LM) of *S. lautus* and conduct QTL analyses of phenotypic traits. We created this population by crossing an individual from the dune population (D01) at Lennox Head (New South Whales) with an individual from the adjacent headland population (H01). A single F1 was then crossed with a full-sibling of the initial H parent (*S. lautus* is self-incompatible, and therefore, individual genotypes cannot be propagated) to create a "backcross-towards-H" (BCH) population composed of 139 individuals. Crosses were performed reciprocally. We decided to use a backcross-like population because the dune phenotype is dominant for some of the convergent traits that we are interested in mapping (i.e., F1s are similar to their dune parental, Table S1). To evaluate whether traits that distinguish parental families also differentiate their populations of origin, we also grew 11 individuals from the D family, 13 individuals from the H family and 31 F1s. Plants were grown in a controlled temperature room at 25°C with 12 hr light period using 8 × 8 × 8 cm pots. Plants were randomized every 3 days to remove effects of plant position in the growth chamber.

2.1.2 | Phenotyping

To identify traits that distinguish the dune and headland populations and detect QTLs, we measured 13 phenotypic traits related to plant architecture and life history in the parental families, F1s and BCH individuals (Table S1). These measurements were conducted during the first flowering season in a controlled temperature room (see Table S1). We compared the phenotypes of natural D and H populations as well as F1s and BCHs. We used the JMP9 software (SAS Institute Inc) to compare phenotypic means for each trait using *t* tests and to calculate pairwise correlations between traits.

2.1.3 | RAD library production, illumina sequencing and bioinformatics analysis

DNA was extracted from leaf samples of 139 backcross individuals and the two parental plants as described previously (Roda et al., 2013a). Individual samples were submitted for RAD sequencing at Floragenex (Eugene, OR) following the general protocol described in Reference (Chutimanitsakun et al., 2011; Slavov et al., 2014). After library construction, RAD libraries were submitted for 100-bp single-end Illumina sequencing on a HiSeq 2500 at the Oregon State University Center for Genome Research and Biocomputing (Corvallis, OR). The PstI enzyme was used for DNA digestion as it produces a relatively large number of reads, enriched in GC-rich regions, and is insensitive to DNA methylation (Davey & Blaxter, 2010; Davey et al., 2011). In previous studies (Roda et al., 2013a, 2013b), we found that 80% of RAD tags obtained with this protocol map to

genes. Therefore, we were more likely to sample genetic variation in functional significance (Hoban et al., 2016). Over half a billion sequence reads (total reads: 576,384,120; average reads per sample: 3,920,980; median Depth: 21.59×) were generated for the 139 backcrosses and the two parentals. Bioinformatics processing of the data was performed by Floragenex to align sequence reads, identify polymorphisms and produce genotype calls for each polymorphic site and individual using reported criteria for quality control and filtering (Chutimanitsakun et al., 2011).

Because the headland plants used to create the F1 and BCH populations were different, we used "outbreeder full-sib family" (CP) model (van Ooijen, 2006), rather than a backcross model for linkage QTL mapping. The CP crossing scheme is ideal for mapping in outbred and heterozygous individuals, allowing us to maximize the number of useful genetic markers. We then used a pipeline that utilizes genotype data from two parental plants to score for markers segregating appropriately in the mapping population. This pipeline produced genotypic data in the "outbreeder full-sib family" (CP) JOINMAP format (van Ooijen, 2006) for 139 individuals. Markers suitable for linkage mapping were of three types (Amores, Catchen, Ferrara, Fontenot, & Postlethwait, 2011): (i) homozygous in the H and heterozygous in the F1 (type nn × np, segregating 1:1, testcross), (ii) heterozygous in the H and homozygous in the F1 (type lm × ll, segregating 1:1, testcross) and (iii) heterozygous in both parents (type hk × hk, segregating 1:2:1, intercross). We obtained 7,366 segregating RAD markers genotyped in more than 50% of the progeny with coverage greater than 20×. Of these, 1,007 markers were sampled in 80% of the population at >25×.

2.1.4 | Genetic map construction

A LM was constructed with JOINMAP 4 (van Ooijen, 2006). Missing genotypes and systematic errors in RAD-seq data can generate problems for LM construction (Henning, Lee, Franchini, & Meyer, 2014). To avoid these problems, we used stringent thresholds for coverage, missing data and segregation distortion (SD) before importing markers into the LM BUILDING software (Henning et al., 2014). We selected markers (i) being scored at least in 70% of the 139 individuals (i.e., 98 individuals), (ii) having a sequencing coverage greater than 25× and (iii) showing segregation ratios that did not deviate from Mendelian proportions (Bonferroni-corrected *p*-value of the chi-square statistic lower than 0.05). Information on markers, genotypes and coverages are provided in Table S3 and Figure S2.

Markers that passed these filters were first assigned to linkage groups (LGs) using a logarithm of the odds (LOD) threshold of seven, and then ordered using the regression-mapping algorithm with a recombination frequency threshold of 0.5, a LOD threshold of 0.5 and a jump threshold value of 5.0. A ripple was performed after addition of each locus. Map distances in centimorgans (cM) were calculated from recombination frequencies using Kosambi's mapping function (Kosambi, 1944). The map2 output containing 1,166 markers (RAD haplotypes) distributed across 20 LGs and 2,235.8 cM (Table S3) was used for the rest of the analyses.

We used the strongest cross-link function (SCL) from JOINMAP (van Ooijen, 2006) to locate ungrouped loci into map intervals. This function allows assigning markers to a constructed map. We assigned ungrouped markers to the map position showing the highest SCL score with a LOD threshold of six. This allowed localizing 4,038 extra SNPs in the original map. The objective of this analysis is to localize markers that were not included in the LM but were genotyped in bulk segregant analyses, and in analyses of population differentiation. It is important to note that the location of these markers in the map is less precise than for the markers used in map construction, and therefore, markers mapped with the SCL function were not used in linkage QTL analyses.

Finally, we positioned contigs from the draft genome of *S. lautus* into our LM. This shotgun Illumina assembly is composed of short contigs (N50 = 1,160 bp) but is enriched in gene-containing sequences, and therefore, most RAD tags can be mapped to it (Roda et al., 2013a, 2013b). We used *bwa-0.7.7* (Li & Durbin, 2009) and *samtools-0.1.19* (Li et al., 2009) as described in Ref. (Roda et al., 2013a) to map the RAD markers included in the LM to the genome draft. In this way, we were able to position 2,882 genomic contigs to the corresponding position in the LM.

2.1.5 | QTL analyses

We identified genomic regions governing phenotypic traits by conducting QTL mapping. QTL analysis was conducted with the software MAPQTL6 (van Ooijen, 2009). Genotypes and phenotypes from the 139 individuals used to build the LM were used in QTL analysis.

A two-way pseudo-testcross approach (Grattapaglia & Sederoff, 1994; van Ooijen, 2009) was applied during QTL analyses in the combined parental map. The interval-mapping (IM) algorithm was first used to identify major QTL peaks for each trait with 1-cM increments and conducting 200 iterations and 1,000 permutations. LOD thresholds were calculated using 1,000 permutations and a genomewide significance level of 5%. The automatic cofactor selection routine (ACS) was then used to select cofactors among the markers contained in LGs with significant IM peaks. Finally, a multiple QTL-mapping (MQM) analysis was ran using cofactors selected in parental maps, 1-cM increments, 200 iterations and 1,000 permutations.

Running multitrait QTL analysis will be advantageous to explore the contributions of pleiotropy and genetic clustering to phenotypic divergence. Unfortunately, software running such analyses (i.e., QTL cartographer) is currently not compatible with the CP crossing scheme that we used in this initial work.

2.1.6 | QTL clustering

To determine whether observed levels of QTL clustering were higher than expected by chance, we created random distributions of QTLs using the R software (R Core Team 2014). For this, we assigned QTLs for each trait to the map position presenting the highest LOD. Then, we permuted 10,000 times a loop composed of three steps.

First, we randomized QTLs for each trait across map positions. Second, we summed QTLs across traits for each map interval. Finally, we calculated the maximum number of QTLs per map position and evaluated Moran's I, a measure of spatial autocorrelation of QTLs across traits and map positions. For this analysis, we eliminated redundant traits that were governed by the same QTL (i.e., traits that measure similar aspects of the phenotype: total height/vegetative height and total flowers/number inflorescences). We conducted the analyses by linkage group and across all the genome.

2.2 | Bulk segregant analysis of survivorship

To investigate the genetic basis of survivorship in the field, we conducted reciprocal transplants of recombinant lines and used bulk segregant analysis to compare the genomes of survivors in the sand dunes and rocky headlands.

2.2.1 | Transplants

We created three replicate F8 recombinant lines using a North Carolina 2 breeding design (Lynch & Walsh, 1998; Rockman & Kruglyak, 2008). For each independent replicate line, we randomly crossed 23 D parentals with 22 H parentals. We used these F1 populations to create three independent recombinant populations. We grew an average of 58 individuals per replicate population and conducted two random crosses per family per generation (three for the F5 generation to increase population size).

We transplanted parental (D01 and H01) and recombinant populations in November 2011. Seeds were planted into nine experimental blocks per environment with each block containing three replicates per family (total 267 individuals from 89 full-sibling families per block) and 40 parents (20 for each parental population), totalling 2,403 F8s and 360 parents per environment. Seedling survivorship was truncated once mortality plateaued 85 days after planting. These data were used to calculate survivorship for each family (proportion of survivors per family) across all blocks. We collected two surviving individuals per family for families located in the upper 20% tail of the distribution of survivorship at 85 days (18 families planted in the dunes; 20 families planted in the rocky headland). We also grew in a growth chamber at the University of Queensland 831 F8 recombinants from the same families transplanted to the field. We selected 20% of the individuals grown under these benign conditions to compare the genotypic changes caused by selection in natural environments to the allele frequencies segregating in this control population. The size of the tails (20%) was selected to make sure to have a sufficient number of samples per pool.

2.2.2 | Bulk segregant analysis

We pooled DNAs from individuals belonging to the same replicate population F8 and environment (Table S7). DNA pools were created as described in Ref. (Roda et al., 2013a). Briefly, DNAs from individual plants were extracted, quantified and pooled in equimolar amounts. Libraries of restriction site associated DNA (RAD) tags (Baird et al., 2008) were synthesized at Floragenex following the methodology described in reference (Roda et al., 2013a) but increasing sequencing depth and coverage (50× depth and 2 M read/sample). To find SNPs associated with survivorship, we conducted pairwise comparisons between pools of F8 survivors from different environments (Futschik & Schlötterer, 2010; Gompert et al., 2013; Kover & Mott, 2012; Magwene, Willis, & Kelly, 2011; Michelmore, Paran, & Kesseli, 1991; Soria-Carrasco et al., 2014) using the POPOOLATION2 software version 1201 (Kofler, Pandey, & Schlötterer, 2011). Candidates were defined as loci fitting *simultaneously* three criteria:

First, we expected candidates for survivorship to have different allelic frequencies between pools of plants surviving in different environments. Therefore, we selected loci displaying significantly different allelic frequencies between pools of survivors from dune and headland environments (corrected *p*-value lower than 5% in a Fisher's exact test, FET).

Second, genes that respond to selection are expected to have different allelic frequencies before and after selection (Gompert et al., 2013; Soria-Carrasco et al., 2014). For this reason, we selected loci having significantly different allelic frequencies between pools of survivors and pools of plants grown under benign conditions (corrected *p*-value lower than 5% in a FET).

Finally, we expect D alleles to be more common in pools of plants surviving in the sand dunes while H alleles should be more common in the rocky headland environment. Therefore, we searched genomic regions enriched in local alleles. We defined D and H alleles as the more frequent allele in the dune and headland natural populations. These local alleles were defined at loci showing significantly different allele frequencies (corrected *p*-value lower than 5% in a Fisher's exact test, FET) in the natural populations from Lennox Head (Roda et al., 2013a, 2013b). The frequencies of these local alleles were then used to define "SNP indexes" (Takagi et al., 2013): the SNP index has a value of one if the local allele is fixed in the pool of survivors (i.e., when the D allele is fixed in the pool of dune survivors or the H allele is fixed in pool of headland survivors) and has a value of 0 when the alternative allele is fixed. Selection candidates where defined as having SNP index values located in the 5% uppermost tail for the genomewide distribution of SNP indexes in the pool.

SNPs fitting simultaneously these three criteria were declared candidates for survivorship. We identified map positions with a high density of these significant SNPs. For this, we used the resampling function of R to create 10,000 random distributions of significant SNPs across map positions. In each permutation, significant SNPs were randomly localized across the genome and the density of significant SNPs (significant SNPs/Total SNPs) across all 1-cM intervals was calculated. Intervals with densities greater than the upper 95% confidence interval of the distribution of maximum densities across the 10,000 permutations were declared significant. We used 1-cM intervals because this interval size is sufficiently small to contain

SNPs under strong LD but sufficiently big to contain a relatively large number of SNPs. Because the density of significant SNPs can be strongly affected by 1-cM intervals that had a very low number of total SNPs, we excluded from our analyses all intervals containing less than four SNPs. All the test distributions were calculated for each interval independently to account for differences in SNP density across intervals. This analysis identifies loci presenting parallel genotypic differences between natural and recombinant populations inhabiting different environments. Even though rare alleles can mediate adaptation, we are interested in finding alleles that were highly differentiated in natural and recombinant populations. We acknowledge that by doing so, we might lose minor alleles contributing to adaptation only in recombinant populations. However, our method reduces false positives while identifying candidate genes that are more likely to be under selection in both natural populations and recombinants.

2.3 | Footprints of selection in natural populations

2.3.1 | Differentiation

We evaluated whether QTLs showed high differentiation in natural populations using pool-seq analyses of 13 parapatric pairs distributed across the range of *S. lautus*. For this, we analysed previously reported genomic comparisons of eight parapatric populations (Roda et al., 2013a, 2013b) and genotyped four new parapatric pairs (see Table S6). For each comparison between two parapatric populations, we defined F_{st} outliers using the same methodology employed in reference (Roda et al., 2013a, 2013b), which involves searching for loci that present extreme (5% upper tail of genome wide F_{st} distribution) and significant (p-value <.05 in a FET) allele frequency differences between parapatric pairs. We calculated the density of these F_{st} outliers (i.e., number of outliers/number of polymorphic sites) across map positions and used 10,000 randomizations to define regions with extreme density of F_{st} outliers using the same resampling procedure described in the previous section.

2.3.2 | Allelic frequencies

We explored allelic variation at the QTLs cluster located at LG10 (see Results) to test the possibility that divergence between geographically widespread D and H populations relies on similar allelic variants at this locus. We used JMP9 (SAS Institute Inc) to conduct a principal component analysis of allelic frequencies at SNPs contained in (i) a 10-cM region centred at the QTL cluster from LG10 and (ii) in map markers not presenting QTLs, following the same procedure as in Ref. (Roda et al., 2013a, 2013b). Because the region containing the QTL cluster and displaying genetic differentiation across natural populations is relatively big (Figures 2 and 3), we conducted this analysis using a relatively broad region around the QTL cluster. However, we obtained similar results using 1-, 2- and 5-cM windows around the QTL cluster (data not shown). We expect that a clustering of populations by environment at the QTL cluster to be indicative of the spread of the same adaptive variants across the *S. lautus* range.

3 | RESULTS

3.1 | Phenotypic variation

As reported in previous studies (Ali, 1964; Radford & Cousens, 2000, 2006), we found that D and H plants showed significantly different phenotypes for traits related to flowering and plant architecture (Figure 1). Individuals from D families had larger flowers, were taller (total height, vegetative height, diameter of the main stem) and displayed more erect stems (angle of the main stem and angle of secondary branches). In contrast, H individuals were heavily branched (number of tertiary branches and number of leaves per branch), prostrate and produced succulent leaves. Tables show statistical results for all trait comparisons.

Most traits present continuous distributions in the recombinant populations, but for some traits that distinguish D and H ecotypes the mean phenotype of F1s does not differ from the phenotypic value of one of the parentals (i.e., for vegetative height, angle main stem, diameter main stem at 2 cm above the ground, number tertiary branches, succulence aerial, Table S1) suggesting some degree of genetic dominance.

We calculated pairwise correlations between phenotypic traits (Table S5) as strong correlations could reflect a common genetic basis or genetic linkage (Friedman, Twyford, Willis, & Blackman, 2015). As expected, pairwise trait correlations were stronger in parental families than in recombinants, but several traits maintained significantly strong correlations in mapping and recombinant populations (Table S5).

3.2 | Reciprocal transplants

Local adaptation can lead to speciation if organisms survive preferentially in their native environments. We tested this prediction by conducting reciprocal transplants of natural populations and recombinant families. As expected, survivorship was greatest in local individuals compared to nonlocals and recombinants (Figure 1). Specifically, D families had higher survivorship in the sand dunes ($t = 3.00$, $p = .0048$, $N = 39$ individuals), whereas H families performed better in the adjacent headland environment ($t = -4.85$, $p < .0001$, $N = 39$).

3.3 | Linkage map

We used RAD-seq to build the first linkage map of *Senecio lautus*. Given that *S. lautus* is self-incompatible and highly heterozygous, we constructed a linkage map and ran QTL analyses using the CP design of JOINMAP (van Ooijen, 2006, 2009, 2011). The map consisted of 1,166 markers (Table S3). In agreement with reported chromosome counts from karyotyping studies (Lawrence, 1980; Ornduff, 1960, 1964; Radford, Liu, & Michael, 1995), the markers assembled into 20 linkage groups (LGs). The average LG size was 110.8 ± 4.4 cM and ranged from 67.7 cM in LG5 to 147.6 in LG4. As the haploid size of *S. lautus* genome is about 1.39 GB (Liu, 2015), each cM approximately corresponds to 0.6 Mb.

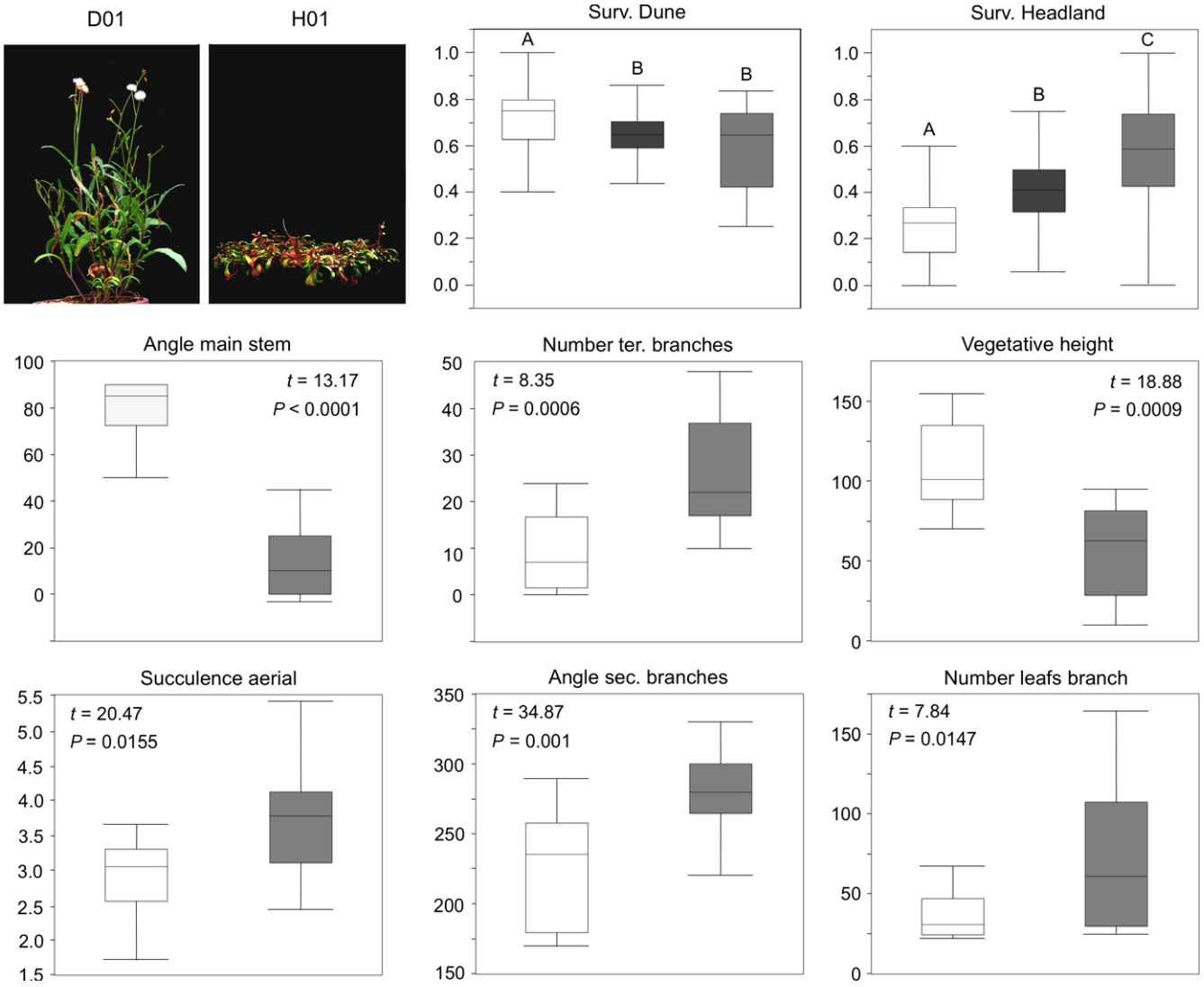

FIGURE 1 Phenotypic differences between parental families. Plants from the dune (D01) and headland (H01) populations are shown in the upper left. Box-plots show survivorship in reciprocal transplants (top row) as well as variation for six morphological traits. D01 individuals in white, H01 in light grey and F1s in dark grey. Results from *t* tests comparing phenotypic means are provided [Colour figure can be viewed at wileyonlinelibrary.com]

3.4 | QTL analysis

We identified QTLs for traits involved in plant architecture and flowering including many of the traits that distinguish D and H populations (Figure 1, Table S1). QTL effects are in the expected direction, where plants homozygous for the H alleles are more similar to the H parental than heterozygous individuals (Figure S3). These QTLs explained little to moderate proportions of the variance observed in the traits (between 3% and 43%, see Table S1).

We detected a positive genomic autocorrelation in QTL positions across traits (Table S2). Specifically, for LG10 Moran's I (Gompert et al., 2013; Moran, 1950) was significantly greater than expected under the null hypothesis of no autocorrelation ($I = 0.053$; $p < .001$). LG10 contained QTLs for eight traits (Figure 2, Table S1), including a map position (cM 22) with QTLs for five traits (length secondary branches, total height, vegetative height, total number of flowers/buds and number of inflorescences). This pattern, where a loci harbour variants associated with seemingly distinct traits is known as cross-phenotype association (CPA). CPAs can result from pleiotropy, where a single gene governs multiple phenotypic traits, or from the clustering of several causal genes (Solovieff, Cotsapas, Lee, Purcell, & Smoller, 2013). The size and SNP density of LG10 are not unusual relative to the rest of the LM (Table S2), indicating that the pattern of QTL clustering at this LG does not result from a larger number of markers.

3.5 | Bulk segregant analysis of survivorship

We tested for the presence of loci involved in local adaptation in QTL clusters by transplanting recombinant populations to natural environments and comparing the genomes of survivors from the different habitats (see methods). SNPs governing survivorship are

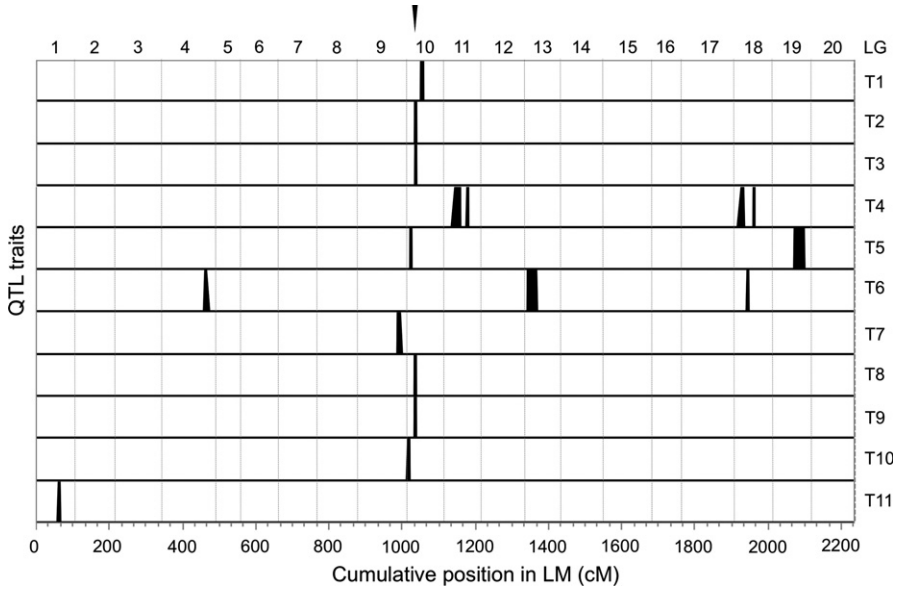

FIGURE 2 Loci governing morphological variation are clustered. We show the position of quantitative trait loci (QTLs) for 11 phenotypic traits. In the *x*-axis, we represent the cumulative position (in cM) across the 20 linkage groups (indicated in the top and delimited by dashed lines). Every horizontal panel represents a trait (T1-T11, not all traits are shown) and black vertical lines show QTLs for these traits. The triangle in the top indicates the map position presenting QTLs for multiple traits. T1 = Max number of capitula/inflorescence; T2 = Total flowers/buds; T3 = Number inflorescences; T4 = Internode length secondary stem; T5 = Aerial succulence; T6 = Diameter main stem at 2 cm above the ground; T7 = Average number leafs per branch; T8 = Total height; T9 = Vegetative height; T10 = Angle main stem; T11 = Number tertiary branches

expected to (i) present different frequencies between recombinant survivors from different environments and (ii) show similar allelic frequencies in recombinant survivors and the local natural populations (Anderson, Lee, Rushworth, Colautti, & Mitchell-Olds, 2013; Gompert et al., 2013; Soria-Carrasco et al., 2014). We found that the CPA at LG 10 also contains a significantly high density ($p < .05$) of survivorship-associated SNPs (Figure 3a).

3.6 | Genomic footprints of selection at QTLs

Divergent natural selection can create genetic differentiation at adaptive loci (Nielsen, 2005; Nielsen et al., 2005; Nosil, Funk, & Ortiz-Barrientos, 2009). To examine whether genomic regions containing QTLs have diverged between *S. lautus* populations adapted to different environments, we used differentiation indexes (F_{st}) from parapatric populations distributed across the range of the species (Roda et al., 2013a, 2013b). We found a significant excess of F_{st} outliers in parapatric populations in multiple QTLs positions from LG 10 (see methods and Figure 3b) relative to other positions in the genome. Notably, the interval presenting a CPA in LG10 is highly differentiated in multiple parapatric populations from the north and south of Australia (Figure 3b). This raises the possibility of recurrent selection on an adaptive haplotype present as standing genetic variation.

We tested this hypothesis by analysing natural genetic variation at these two positions (Colosimo et al., 2005) using a principal components analysis (PCA) of QTL clusters and the rest of the genome. Based in previous results, we predict that neutral genomic regions will produce a clustering of the populations by geographic region (Roda et al., 2013a, 2013b). A clustering of populations by environment at adaptive loci could be indicative of selection on the same genetic variant exchanged through gene flow or present as standing genetic variation. This analysis revealed a clustering of *S. lautus* populations by geography rather than environment at the QTL cluster and in the rest of the genome (Figure 4). We obtained the same results using windows of different sizes around the QTL cluster (1, 2, 5 cM; not shown).

4 | DISCUSSION

4.1 | Strong phenotypic divergence between locally adapted populations

We searched for the genetic basis of convergent traits in *S. lautus* to understand how parapatric populations maintain complex phenotypic differences despite the lack of geographic barriers to gene flow between them. Several of the traits that differed between the parental families used to construct the linkage map also distinguished dune and headland forms (Radford & Cousens, 2000, 2006; Walter et al., 2016): compared to the D, the H plants are prostrate and branched and have numerous smaller flowers. Most of these morphological differences have been reported in previous studies of coastal *Senecio lautus* ecotypes using data from multiple D and H populations (Radford & Cousens, 2006; Walter et al., 2016). This points to the existence of a strong correlation between habitat and morphological variation in the system. A prostrate architecture seems to be important to survive in rocky headlands as many plant

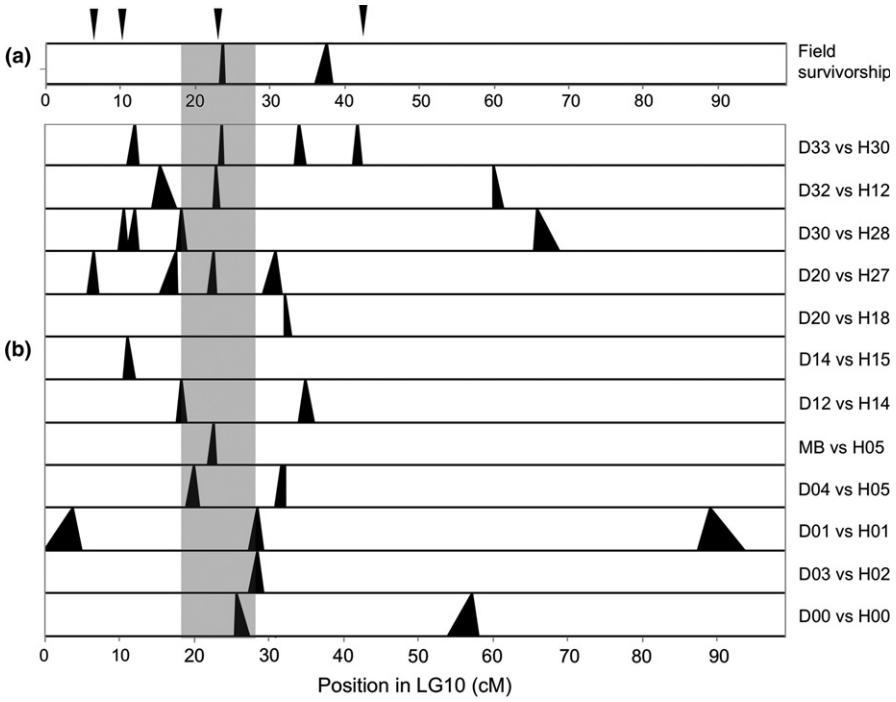

FIGURE 3 A cluster of QTLs diverged repeatedly during parallel evolution of coastal ecotypes: Results from bulked segregant analysis (BSA) to test the clustering of adaptive loci at linkage group 10 (LG10). In the *x*-axis, we present the position along LG10 (in cM). Map intervals displaying divergent allelic frequencies between pools of plants are shown as black triangles, where the width of the triangle represents the size of the interval with significant divergence. QTLs from linkage mapping are indicated by triangles in the top. A 10-cM window around the position containing a CPA is shaded. (a) Association to survivorship in reciprocal transplants. (b) Genomic differentiation between parapatric populations. Each panel represents a comparison between two populations

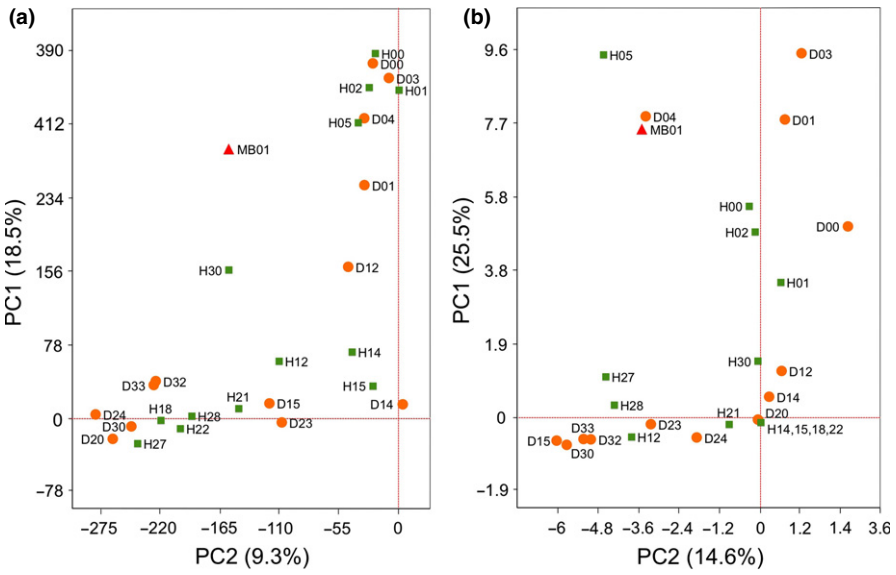

FIGURE 4 Allelic variation at QTLs from LG10. Principal components analysis of allelic frequencies at map positions without QTLs (a) and in a 10-cM region centred at the QTL cluster from LG10 (b). Population names are coded as presented in Table S6. Dune populations are shown as circles, headland populations shown as squares and bird rockery population shown as a triangle [Colour figure can be viewed at wileyonlinelibrary.com]

species found in these environments are prostrate (Adam, Stricker, Wiecek, & Anderson, 1989; Auld & Morrison, 1992), possibly as an adaptation to stress imposed by wind exposure (Schweingruber, Börner, & Schulze, 2013). Alternatively, an erect architecture could help prevent sand covering in the dunes. In fact, the transition between erect and prostrate growth habit is common during the divergence of annual and perennial forms of other plant species (Crawford, Chapman, & Smith, 1995; Fishman, Stathos, Beardsley, Williams, & Hill, 2013; Lowry, Rockwood, & Willis, 2008; McNaughton, 1984; Twyford & Friedman, 2015). The frequency of this transition reflects the important role of plant architecture in mediating plant adaptation but likely also reflects trade-offs in the determination of plant morphology.

Reciprocal transplant experiments showed that plants survive preferentially in their native environments (Figure 1), confirming that local adaptation is under genetic control. These results are consistent with data from *S. lautus* transplants conducted at other locations (Melo et al., 2014; Radford & Cousens, 2006; Richards & Ortiz-Barrientos, 2016; Richards et al., 2016; Walter et al., 2016) and indicate that although there are no intrinsic hybrid incompatibilities between parapatric populations from the same phylogenetic clade (Melo, 2014), immigrants and hybrid genotypes suffer from extrinsic selection, a quintessential signature of local adaptation (Blanquart, Kaltz, Nuismer, & Gandon, 2013).

The segregation of phenotypic variation in hybrid populations reflects the genetic build-up of morphological differences between

species (Grant, 1981). Most of the traits that distinguish *S. lautus* populations have continuous distributions in natural and recombinant populations (Figure 1, Table S1). However, some of the divergent traits showed dominance as revealed by F1 progenies that did not differ phenotypically from one of the parents (Table S1). Altogether, these results suggest that the genetic basis of adaptive traits is relatively complex, and some phenotypes might have been favoured by selection because of their "visibility" across individuals resulting from to their genetic dominance (Haldane, 1927).

Organisms consist of interrelated phenotypes that function jointly to define individual fitness. Therefore, the response to selection is controlled by genetic covariation (Blows, 2007; Blows & Hoffmann, 2005; Lande & Arnold, 1983), where natural selection creates correlations between traits that are important for local adaptation (Kirkpatrick, 2009; Schwander et al., 2014; Thompson & Jiggins, 2014; Walsh & Blows, 2009). Thus, it is expected that adaptive traits tend to display correlations in hybrids. For instance, in hybrid zones between hawkmoth-pollinated and hummingbird-pollinated forms of scarlet gilia (*Ipomopsis aggregata*), there are strong correlations between the traits responsible for the two pollination syndromes (Milano, Kenney, & Juenger, 2016). These correlations can be due to pleiotropy, physical genetic linkage or LD resulting from selection on correlated traits (Sinervo & Svensson, 2002). We found strong phenotypic correlations between putatively adaptive traits (traits that differed between dune and headland natural populations), in F1s and in backcross hybrids (Table S5), which suggest the existence of genetic correlations between putatively adaptive traits in the system. We inspected the possible causes of these correlations by searching for QTLs governing these traits.

4.2 | Clustering of adaptive QTLs

The number, effect size and distribution of loci responsible for the evolution of adaptive phenotypes determine the path followed by adaptive divergence (Orr, 1995, 2000; Rockman, 2012). Recent studies have found that oligogenic architectures underlie putatively adaptive traits (Anderson & Mitchell-Olds, 2010; Anderson et al., 2011; Ferris, Barnett, Blackman, & Willis, 2016; Mitchell-Olds, 2013; Mitchell-Olds & Schmitt, 2006; Whitney et al., 2015), but it is unclear whether this result reflects biased sampling arising from lack of experimental power (Rockman, 2012). For instance, we detected only one linkage QTL for most of the traits analysed (Figure 2), which explain relatively low proportions of the phenotypic variation (3%–26%, Table S1). An explanation is that the rest of the variation is governed by loci of small effect that our linkage mapping failed to detect (Hoban et al., 2016; Rockman, 2012). Nonetheless, our results are consistent with recent QTL studies of local adaptation in plants, where only a moderate proportion of the genetic determinants of phenotypic variation is captured (Friedman et al., 2015; Lowry et al., 2015; Whitney et al., 2015).

How are these genetic and phenotypic differences maintained in the face of gene flow? In many cases, the traits that distinguish species are regulated by "supergenes" (Ferris et al., 2016; Schwander et al., 2014; Yeaman & Whitlock, 2011), clusters of tightly linked loci that segregate as stable polymorphisms within or between natural populations. Supergenes provide integrated control of complex adaptive phenotypes segregating within species, preventing allele combinations leading to nonoptimal phenotypes (e.g., Kirkpatrick & Barrett, 2015; Schwander et al., 2014; Thompson & Jiggins, 2014). In a large number of organisms, chromosomal inversions and other modifiers of recombination promote adaptive gene clustering by preventing the dissolution of allelic combinations put together by natural selection (Ortiz-Barrientos et al., 2016; Yeaman, 2013; Yeaman & Whitlock, 2011). The antagonism between selection and recombination can also be solved by pleiotropic "magic genes" controlling reproductive isolation and adaptation simultaneously (Bomblies, 2010; Servedio, Doorn, Kopp, Frame & Nosil, 2011; Thibert-Plante & Gavrilets 2013). For instance, genes responsible for hybrid incompatibilities also mediate environmental adaptation mechanisms, like immune defence in plants (Bomblies, 2010; Bomblies et al., 2007; Chae et al., 2014).

QTLs for morphological variation in *S. lautus* displayed a nonrandom distribution in the genome, and multiple traits were often associated to the same map positions (Figures 2 and 3). In fact, LG 10 showed significant associations to traits involved in reproductive development, fitness and convergent traits (Figures 2 and 3). Convergent traits are likely adaptive (Arendt & Reznick, 2008; Losos, 2011; Stern, 2013) and flowering differences can underlie local adaptation and lead to reproductive isolation (Bischoff, Raguso, Jürgens, & Campbell, 2015; Hopkins, 2013; Waser, 1998). Therefore, the colocalization of alleles controlling plant morphology could facilitate the speciation process in *S. lautus*.

This colocalization of QTLs could arise from pleiotropy, fortuitous linkage, from the spread of allelic modifiers that locally suppress recombination or from cotransmission imposed by genetic structures in the genome, such as chromosomal inversions (Ortiz-Barrientos et al., 2016; Yeaman, 2013; Yeaman & Whitlock, 2011). Further experiments are required to distinguish these nonexclusive options and to disentangle the genetic mechanisms that promote the formation of these concentrated genetic architectures, and possible supergenes. This will require fine mapping of the mutations responsible for phenotypic transitions and testing their effects through transgenesis or gene editing. Additionally, a good quality genome and a refined linkage map are necessary to identify genomic rearrangements in QTL-containing regions. Although the existence of genomic rearrangements could complicate fine-mapping adaptive genes, new developments in transformation and gene editing can make the functional screening of relatively large numbers of candidates contained in these regions feasible.

4.3 | Adaptive trade-offs

Natural selection can favour different genotypes across environments. This process known as local adaptation could lead to genetic divergence and speciation if adaptation to one environment reduces the chances of surviving in other habitats (Blanquart et al., 2013; Kawecki

& Ebert, 2004; Levene, 1953). Some theories predict that genomic regions causing fitness trade-offs, where alleles that are beneficial in one environment are detrimental in others, would be reticent to gene flow (Anderson et al., 2013, 2011; Mitchell-Olds, 2013) and could drive genomic divergence between locally adapted populations (Via, 2012). Genetic differentiation at adaptive loci can also result from the effect of selective sweeps (Nielsen, 2005; Przeworski, 2002), where natural selection drives the fixation of adaptive alleles, as well as many nonadaptive processes (Cruickshank & Hahn, 2014).

Interestingly, several QTL positions of the *S. lautus* genome show extreme differentiation (enrichment of F_{st} outliers, Figure 3b) between parapatric *S. lautus* populations. More importantly, the two genomic positions on LG 10 governing convergent traits (height and angle of the main stem) are highly differentiated in multiple (but not all) parapatric populations, which suggest that they could have been repeatedly recruited by natural selection. Although we cannot determine whether differentiation at adaptive QTLs between these two parapatric populations is due to a local reduction of gene flow in this genomic area or to the fixation of favourable alleles in regions of reduced recombination and in the absence of gene flow (Cruickshank & Hahn, 2014), our results indicate that the same adaptive QTLs are important for genomic divergence between *S. lautus* populations.

Genetic trade-offs between environments can guarantee that adaptive alleles are not exchanged between habitats (Anderson et al., 2013; Mitchell-Olds, 2013; Mitchell-Olds et al., 2008). We found enrichment in loci displaying these trade-offs (i.e., D and H alleles fixed in survivors from the dune and headland environments, respectively) in the CPA position from LG 10. This suggests that this genomic region could maintain the alleles necessary to produce a locally adapted phenotype during the repeated adaptation of *Senecio lautus* to coastal environments.

Our results raise questions on the history of adaptive gene clusters during the parallel evolution of *S. lautus*: Did evolution recruit new mutations or standing genetic variation? What was the role of migration in the repeated recruitment of the same clusters? At what genomic and temporal scales does the clustering of adaptive genes occur? Are adaptive genes actively "recruited" from other parts of the genome? Are adaptive clusters exchanged between populations adapted to the same environment like in perennial mountain populations of the yellow monkeyflower (Lee, Fishman, Kelly, & Willis, 2016a; Lee et al., 2016b; Twyford & Friedman, 2015)? Did genetic clustering occur by chance in the absence of gene flow? The analysis of allelic frequencies at these clusters provides some insights into these questions. Previous genomic studies showed that *S. lautus* populations display a strong isolation by distance and are phylogenetically clustered according to their geographic origin rather than the environment that they inhabit, implying that *S. lautus* has adapted repeatedly to the same coastal environments (Roda et al., 2013a, 2013b). We found that populations are grouped according to their geographic location rather than their environment in a PCA of allelic frequencies at the QTL cluster from LG10 (Figure 4b). This is the same pattern obtained in "neutral" loci (Figure 4a), suggesting that the formation of this adaptive cluster probably involved the recruitment of different adaptive alleles from new mutations or standing genetic variation (Roda et al., 2013b). This interpretation has to be taken with caution because the genomic region containing this cluster is large and therefore expected to contain a great number of neutral loci that will distort phylogenetic signals. Additionally, the analysis of allelic frequencies tells us little about the assembly of this concentrated genetic architecture, because we ignore the identity and genomic location of the adaptive genes in the different *S. lautus* populations. We are currently working on the creation of assembled genomes and linkage maps from multiple population pairs in order to track the genetic changes responsible for the rapid evolution of these plants.

5 | CONCLUSIONS

We showed evidence suggesting that the clustering of adaptive alleles likely facilitated the fast and repeated adaptation of *S. lautus* to contrasting environments along the coast of Australia. A genomic region containing determinants for multiple adaptive traits seem to have been targeted by selection multiple times during the evolution of these plants. Interestingly, this divergent QTL cluster does not seem to have spread via gene flow across the range, like adaptive supergenes in Heliconius butterflies (Joron et al., 2011) or monkeyflowers (Lee et al., 2016a, 2016b), but seem to have been "assembled" in situ multiple times. Additionally, this QTL cluster explains a moderate proportion of phenotypic variation, supporting a scenario where some essential adaptive variants are repeatedly recruited by selection while other are locally selected in each population. Recent results (Walter et al., 2016) suggest that *S. lautus* populations are adapted to the specific location that they inhabit (i.e., have lower fitness when transplanted to the same environment in other locations), which suggest that some of these unique variants could help, fine tuning adaptation to specific conditions at each site.

ACKNOWLEDGEMENTS

We thank all members from the OB laboratory, D. Bernal, B. Brittain, J. Donohoe, M. James, H. Liu, M.C. Melo, A. Posada, T. Richards and M. Wilkinson for their help with field experiments and plant phenotyping. We thank Robin Hopkins, Sevan Suni and Shayla Salzman for their comments and corrections on the manuscript.

CONFLICT OF INTEREST

R.N. is an employee of Floragenex, an organization that offers commercial RAD sequencing services.

DATA ACCESSIBILITY

- Reads from BSA of natural populations and field survivorship were submitted to the Sequence Read Archives of the GenBank. Accession numbers are provided in Tables S6 and S7.

- Genotype calls and genetic differentiation statistics from BSA of parapatric populations and field survivorship were submitted to the Dryad Digital Repository (https://doi.org/10.5061/dryad.83f17).
- Custom scripts used for bulk segregant analysis are available upon request from the corresponding authors.

AUTHOR CONTRIBUTION

F.R. carried out the molecular laboratory, field and glasshouse work and analysed the data; G.W. carried out the field transplant experiment; R.N. sequenced the linkage map population and produced the genotypes; F.R. and D.O.B. conceived the study, designed the study and wrote the manuscript. D.O.B. supervised the study. All authors gave final approval for publication.

REFERENCES

Adam, P., Stricker, P., Wiecek, B., & Anderson, D. (1989). The vegetation of seacliffs and headlands in New South Wales, Australia. *Australian Journal of Ecology*, *14*, 515–545.

Ali, S. (1964). *Senecio lautus* complex in Australia. II. Cultural studies of populations. *Australian Journal of Botany*, *12*, 292–316.

Ali, S. (1969). *Senecio lautus* complex in Australia. V. Taxonomic interpretations. *Australian Journal of Botany*, *17*, 161–176.

Amores, A., Catchen, J., Ferrara, A., Fontenot, Q., & Postlethwait, J. H. (2011). Genome evolution and meiotic maps by massively parallel DNA sequencing: Spotted Gar, an outgroup for the teleost genome duplication. *Genetics*, *188*, 799–808.

Anderson, J. T., Lee, C.-R., Rushworth, C. A., Colautti, R. I., & Mitchell-Olds, T. (2013). Genetic trade-offs and conditional neutrality contribute to local adaptation. *Molecular Ecology*, *22*, 699–708.

Anderson, J. T., & Mitchell-Olds, T. (2010). Beyond QTL cloning. *PLoS Genetics*, *6*, e1001197.

Anderson, J. T., Willis, J. H., & Mitchell-Olds, T. (2011). Evolutionary genetics of plant adaptation. *Trends in Genetics*, *27*, 258–266.

Arendt, J., & Reznick, D. (2008). Convergence and parallelism reconsidered: What have we learned about the genetics of adaptation? *Trends in Ecology & Evolution*, *23*, 26–32.

Auld, T., & Morrison, D. (1992). Genetic determination of erect and prostrate growth habit in five shrubs from windswept headlands in the Sydney region. *Australian Journal of Botany*, *40*, 1–11.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., … Johnson, E. A. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, *3*, 1–7.

Barrett, R. D. H., & Hoekstra, H. E. (2011). Molecular spandrels: Tests of adaptation at the genetic level. *Nature Reviews Genetics*, *12*, 767–780.

Bischoff, M., Raguso, R. A., Jürgens, A., & Campbell, D. R. (2015). Context-dependent reproductive isolation mediated by floral scent and color. *Evolution*, *69*, 1–13.

Blanquart, F., Kaltz, O., Nuismer, S. L., & Gandon, S. (2013). A practical guide to measuring local adaptation. *Ecology Letters*, *16*, 1195–1205.

Blows, M. (2007). A tale of two matrices: Multivariate approaches in evolutionary biology. *Journal of Evolutionary Biology*, *20*, 1–8.

Blows, M. W., & Hoffmann, A. A. (2005). A reassessment of genetic limits to evolutionary change. *Ecology*, *86*, 1371–1384.

Bomblies, K. (2010). Doomed lovers: Mechanisms of isolation and incompatibility in plants. *Annual Review of Plant Biology*, *61*, 109–124.

Bomblies, K., Lempe, J., Epple, P., Warthmann, N., Lanz, C., Dangl, J. L., & Weigel, D. (2007). Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biology*, *5*, e236.

Chae, E., Bomblies, K., Kim, S. T., Karelina, D., Zaidem, M., Ossowski, S., … Lechner, S. (2014). Species-wide genetic incompatibility analysis identifies immune genes as hot spots of deleterious epistasis. *Cell*, *159*, 1341–1351.

Chutimanitsakun, Y., Nipper, R. W., Cuesta-Marcos, A., Cistué, L., Corey, A., Filichkina, T., … Hayes, P. M. (2011). Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. *BMC Genomics*, *12*, 4.

Colautti, R. I., Lee, C.-R., & Mitchell-Olds, T. (2012). Origin, fate, and architecture of ecologically relevant genetic variation. *Current Opinion in Plant Biology*, *15*, 199–204.

Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, … Kingsley, D. M. (2005). Widespread parallel evolution in Sticklebacks by repeated fixation of Ectodysplasin alleles. *Science*, *307*, 1928–1933.

Crawford, R., Chapman, H., & Smith, L. (1995). Adaptation to variation in growing season length in arctic populations of *Saxifraga oppositifolia* L. *Botanical Journal of Scotland*, *47*, 177–192.

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

Davey, J. W., & Blaxter, M. L. (2010). RAD MARKERSeq: Next-generation population genetics. *Briefings in Functional Genomics*, *9*, 416–423.

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

Ferris, K. G., Barnett, L. L., Blackman, B. K., & Willis, J. H. (2016). The genetic architecture of local adaptation and reproductive isolation in sympatry within the *Mimulus guttatus* species complex. *Molecular Ecology*, *26*, 208–224.

Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford: Oxford University Press.

Fishman, L., Stathos, A., Beardsley, P. M., Williams, C. F., & Hill, J. P. (2013). Chromosomal rearrangements and the genetics of reproductive barriers in Mimulus (monkey flowers). *Evolution*, *67*, 2547–2560.

Friedman, J., Twyford, A. D., Willis, J. H., & Blackman, B. K. (2015). The extent and genetic basis of phenotypic divergence in life history traits in *Mimulus guttatus*. *Molecular Ecology*, *24*, 111–122.

Futschik, A., & Schlötterer, C. (2010). The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics*, *186*, 207–218.

Gompert, Z., Comeault, A. A., Farkas, T. E., Feder, J. L., Parchman, T. L., Buerkle, C. A., & Nosil, P. (2013). Experimental evidence for ecological selection on genome variation in the wild. *Ecology Letters*, *17*, 369–379.

Grant, V. (1981). *Plant speciation*. New York: Columbia University Press. xii, 563p.-illus., maps, chrom. nos. En 2nd ed.

Grattapaglia, D., & Sederoff, R. (1994). Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: Mapping strategy and RAPD markers. *Genetics*, *137*, 1121–1137.

Haldane, J. B. S. (1927). A mathematical theory of natural and artificial selection, part V: Selection and mutation, vol. *23*, pp. 838–844.

Henning, F., Lee, H. J., Franchini, P., & Meyer, A. (2014). Genetic mapping of horizontal stripes in Lake Victoria cichlid fishes: Benefits and pitfalls of using RAD markers for dense linkage mapping. *Molecular Ecology*, *23*, 5224–5240.

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., … Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *The American Naturalist*, *188*, 379–397.

Hopkins, R. (2013). Reinforcement in plants. *New Phytologist*, *197*, 1095–1103.

Joron, M., Frezal, L., Jones, R. T., Chamberlain, N. L., Lee, S. F., Haag, C. R., ... Wilkinson, P. A. (2011). Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature*, *477*, 203–206.

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*, 1225–1241.

Kirkpatrick, M. (2009). Patterns of quantitative genetic variation in multiple dimensions. *Genetica*, *136*, 271–284.

Kirkpatrick, M. (2010). How and why chromosome inversions evolve. *PLoS Biology*, *8*, e1000501.

Kirkpatrick, M., & Barrett, B. (2015). Chromosome inversions, adaptive cassettes, and the evolution of species' ranges. *Molecular Ecology*, *24*, 2046–2055.

Kirkpatrick, M., & Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics*, *173*, 419–434.

Kofler, R., Pandey, R. V., & Schlötterer, C. (2011). PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, *27*, 3435–3436.

Kosambi, D. (1944). The estimation of map distance from recombination values. *Annals of Eugenics*, *12*, 172–175.

Kover, P. X., & Mott, R. (2012). Mapping the genetic basis of ecologically and evolutionarily relevant traits in *Arabidopsis thaliana*. *Current Opinion in Plant Biology*, *15*, 212–217.

Lande, R., & Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution*, *37*, 1210–1226.

Lawrence, M. (1980). Senecio L. (Asteraceae) in Australia: Chromosome numbers and the occurrence of polyploidy. *Australian Journal of Botany*, *28*, 151–165.

Lee, Y. W., Fishman, L., Kelly, J. K., & Willis, J. H. (2016a). Fitness variation Is generated by a segregating inversion in Yellow Monkeyflower (*Mimulus guttatus*). *Genetics*, *115*. https://doi.org/10.1534/genetics.115.183566

Lee, Y. W., Fishman, L., Kelly, J. K., & Willis, J. H. (2016b). A segregating inversion generates fitness variation in yellow monkeyflower (*Mimulus guttatus*). *Genetics*, *202*, 1473–1484.

Levene, H. (1953). Genetic equilibrium when more than one ecological niche is available. *The American Naturalist*, *87*, 331–333.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The Sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079.

Liu, H. (2015). Developing genomic resources for an emerging ecological model species *Senecio lautus*.

Losos, J. B. (2011). Convergence, adaptation, and constraint. *Evolution*, *65*, 1827–1840.

Lowry, D. B., Hernandez, K., Taylor, S. H., Meyer, E., Logan, T. L., Barry, K. W., ... Juenger, T. E. (2015). The genetics of divergence and reproductive isolation between ecotypes of *Panicum hallii*. *New Phytologist*, *205*, 402–414.

Lowry, D. B., Rockwood, R. C., & Willis, J. H. (2008). Ecological reproductive isolation of coast and inland races of *Mimulus guttatus*. *Evolution*, *62*, 2196–2214.

Lowry, D. B., & Willis, J. H. (2010). A Widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive Isolation. *PLoS Biology*, *8*, e1000500.

Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits*. Sunderland: Sinauer Associates Inc.

Magwene, P. M., Willis, J. H., & Kelly, J. K. (2011). The statistics of bulk segregant analysis using next generation sequencing. *PLoS Computational Biology*, *7*, e1002255.

McNaughton, S. (1984). Grazing lawns: Animals in herds, plant form, and coevolution. *American Naturalist*, *124*, 863–886.

Melo, M. C. (2014). *Ecological Speciation in Senecio lautus*, PhD, The University of Queensland.

Melo, M. C., Grealy, A., Brittain, B., Walter, G. M., & Ortiz-Barrientos, D. (2014). Strong extrinsic reproductive isolation between parapatric populations of an Australian groundsel. *New Phytologist*, *203*, 323–334.

Michelmore, R. W., Paran, I., & Kesseli, R. V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences*, *88*, 9828–9832.

Milano, E. R., Kenney, A. M., & Juenger, T. E. (2016). Adaptive differentiation in floral traits in the presence of high gene flow in scarlet gilia (*Ipomopsis aggregata*). *Molecular Ecology*, *25*, 5862–5875.

Mitchell-Olds, T. (2013). Selection on QTL and complex traits in complex environments. *Molecular Ecology*, *22*, 3427–3429.

Mitchell-Olds, T., Feder, M., & Wray, G. (2008). Evolutionary and ecological functional genomics. *Heredity*, *100*, 101–102.

Mitchell-Olds, T., & Schmitt, J. (2006). Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature*, *441*, 947–952.

Mitchell-Olds, T., Willis, J. H., & Goldstein, D. B. (2007). Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics*, *8*, 845–856.

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*, 17–23.

Nielsen, R. (2005). Molecular signatures of natural selection. *Annual Review of Genetics*, *39*, 197–218.

Nielsen, R., Williamson, S., Kim, Y., Hubisz, M. J., Clark, A. G., & Bustamante, C. (2005). Genomic scans for selective sweeps using SNP data. *Genome Research*, *15*, 1566–1575.

Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, *18*, 375–402.

van Ooijen, J. W. (2006). *JoinMap 4® Software for the calculation of genetic linkage maps in experimental populations*. Wageningen: Kyazma B.V.

van Ooijen, J. W. (2009). *MapQTL 6 Software for the mapping of quantitative trait loci in experimental populations of diploid species*. Wageningen: Kyazma B.V.

van Ooijen, J. W. (2011). Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. *Genetics Research*, *93*, 343–349.

Ornduff, R. (1960). An interpretation of the *Senecio lautus* complex in New Zealand. *Transactions of the Royal Society of New Zealand*, *88*, 1–24.

Ornduff, R. (1964). Evolutionary pathways of the *Senecio lautus* alliance in New Zealand and Australia. *Evolution*, *18*, 349–360.

Orr, H. A. (1995). The population genetics of speciation: The evolution of hybrid incompatibilities. *Genetics*, *139*, 1805–1813.

Orr, H. A. (2000). Adaptation and the cost of complexity. *Evolution*, *54*, 13–20.

Ortiz-Barrientos, D., Engelstädter, J., & Rieseberg, L. H. (2016). Recombination rate evolution and the origin of species. *Trends in Ecology & Evolution*, *31*, 226–236.

Przeworski, M. (2002). The signature of positive selection at randomly chosen loci. *Genetics*, *160*, 1179–1189.

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. 2013. ISBN 3-900051-07-0.

Radford, I. J., & Cousens, R. D. (2000). Invasiveness and comparative life-history traits of exotic and indigenous Senecio species in Australia. *Oecologia*, *125*, 531–542.

Radford, I. J., & Cousens, R. D. (2006). Morphological and genetic variation in the Senecio pinnatifolius complex: Are variants worthy of taxonomic recognition? *Australian Systematic Botany*, *17*, 29–48.

Radford, I., Liu, Q., & Michael, P. (1995). Chromosome counts for the Australian weed known as *Senecio madagascariensis* (Asteraceae). *Australian Systematic Botany*, *40*, 1029–1033.

Richards, T. J., & Ortiz-Barrientos, D. (2016). Immigrant inviability produces a strong barrier to gene flow between parapatric ecotypes of *Senecio lautus*. *Evolution*, *70*, 1239–1248.

Richards, T. J., Walter, G. M., McGuigan, K., & Ortiz-Barrientos, D. (2016). Divergent natural selection drives the evolution of reproductive isolation in an Australian wildflower. *Evolution*, *70*, 1993–2003.

Rockman, M. V. (2012). The QTN program and the alleles that matter for evolution: All that's gold does not glitter. *Evolution*, *66*, 1–17.

Rockman, M. V., & Kruglyak, L. (2008). Breeding designs for recombinant inbred advanced intercross lines. *Genetics*, *179*, 1069–1078.

Roda, F., Ambrose, L., Walter, G. M., Liu, H. L., Schaul, A., Lowe, A., … Ortiz-Barrientos, D. (2013a). Genomic evidence for the parallel evolution of coastal forms in the *Senecio lautus* complex. *Molecular Ecology*, *22*, 2941–2952.

Roda, F., Liu, H., Wilkinson, M. J., Walter, G. M., James, M. E., Bernal, D. M., … Ortiz-Barrientos, D. (2013b). Convergence and divergence during the adaptation to similar environments by an Australian groundsel. *Evolution*, *67*, 2515–2529.

Savolainen, O., Lascoux, M., & Merila, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, *14*, 807–820.

Schwander, T., Libbrecht, R., & Keller, L. (2014). Supergenes and complex phenotypes. *Current Biology*, *24*, R288–R294.

Schweingruber, F. H., Börner, A., & Schulze, E. D. (2013). Anatomical adaptations to environmental conditions. In *Atlas of stem anatomy in herbs, shrubs and trees*, Chapter 4 (Vol. *2*, pp. 355–370). Berlin: Springer Berlin Heidelberg.

Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., … Brelsford, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, *15*, 176–192.

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

Sinervo, B., & Svensson, E. (2002). Correlational selection and the evolution of genomic architecture. *Heredity*, *89*, 329–338.

Slavov, G. T., Nipper, R., Robson, P., Farrar, K., Allison, G. G., Bosch, M., … Jensen, E. (2014). Genome-wide association studies and prediction of 17 traits related to phenology, biomass and cell wall composition in the energy grass Miscanthus sinensis. *New Phytologist*, *201*, 1227–1239.

Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M., & Smoller, J. W. (2013). Pleiotropy in complex traits: Challenges and strategies. *Nature Reviews Genetics*, *14*, 483–495.

Soria-Carrasco, V., Gompert, Z., Comeault, A. A., Farkas, T. E., Parchman, T. L., Johnston, J. S., … Egan, S. P. (2014). Stick insect genomes reveal natural selection's role in parallel speciation. *Science*, *344*, 738–742.

Stern, D. L. (2013). The genetic causes of convergent evolution. *Nature Reviews Genetics*, *14*, 751–764.

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., … Innan, H. (2013). QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *The Plant Journal*, *74*, 174–183.

Thibert-Plante, X., & Gavrilets, S. (2013). Evolution of mate choice and the so-called magic traits in ecological speciation. *Ecology Letters*, *16*, 1004–1013.

Thompson, M., & Jiggins, C. (2014). Supergenes and their role in evolution. *Heredity*, *113*, 1–8.

Twyford, A. D., & Friedman, J. (2015). Adaptive divergence in the monkey flower Mimulus guttatus is maintained by a chromosomal inversion. *Evolution*, *69*, 1476–1486.

Via, S. (2012). Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 451–460.

Walsh, B., & Blows, M. W. (2009). Abundant genetic variation+ strong selection=multivariate genetic constraints: A geometric view of adaptation. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 41–59.

Walter, G. M., Wilkinson, M. J., James, M. E., Richards, T. J., Aguirre, J. D., & Ortiz-Barrientos, D. (2016). Diversification across a heterogeneous landscape. *Evolution*, *70*, 1979–1992.

Waser, N. M. (1998). Pollination, angiosperm speciation, and the nature of species boundaries. *Oikos*, *82*, 198–201.

Whitney, K. D., Broman, K. W., Kane, N. C., Hovick, S. M., Randell, R. A., & Rieseberg, L. H. (2015). Quantitative trait locus mapping identifies candidate alleles involved in adaptive introgression and range expansion in a wild sunflower. *Molecular Ecology*, *24*, 2194–2211.

Yeaman, S. (2013). Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proceedings of the National Academy of Sciences*, *110*, E1743–E1751.

Yeaman, S., & Whitlock, M. C. (2011). The genetic architecture of adaptation under migration-selection balance. *Evolution*, *65*, 1897–1911.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Roda F, Walter GM, Nipper R, Ortiz-Barrientos D. Genomic clustering of adaptive loci during parallel evolution of an Australian wildflower. *Mol Ecol*. 2017;26:3687–3699. https://doi.org/10.1111/mec.14150