Heredity (2012) 108, 87–95

www.nature.com/hdy

ORIGINAL ARTICLE

Molecular genetic and quantitative trait divergence associated with recent homoploid hybrid speciation: a study of *Senecio squalidus* (Asteraceae)

AC Brennan[1,2], D Barker[2], SJ Hiscock[3] and RJ Abbott[2]
[1]*Estación Biológica de Doñana CSIC, Seville, Spain;* [2]*School of Biology, University of St Andrews, Fife, UK and* [3]*School of Biological Sciences, University of Bristol, Bristol, UK*

Hybridization is increasingly seen as a trigger for rapid evolution and speciation. To quantify and qualify divergence associated with recent homoploid hybrid speciation, we compared quantitative trait (QT) and molecular genetic variation between the homoploid hybrid species *Senecio squalidus* and its parental species, *S. aethnensis* and *S. chrysanthemifolius*, and also their naturally occurring Sicilian hybrids. *S. squalidus* originated and became invasive in the United Kingdom following the introduction of hybrid plants from Mount Etna, Sicily, about 300 years ago. We recorded considerable molecular genetic differentiation between *S. squalidus* and its parents and their Sicilian hybrids in terms of both reduced genetic diversity and altered allele frequencies, potentially due to the genetic bottleneck associated with introduction to the United Kingdom. *S. squalidus* is also distinct from its parents and Sicilian hybrids for QTs, but less so than for molecular genetic markers. We suggest that this is due to resilience of polygenic QTs to changes in allele frequency or lack of selection for hybrid niche divergence in geographic isolation. While *S. squalidus* is intermediate or parental-like for most QTs, some trangressively distinct traits were observed, which might indicate emerging local adaptation in its invasive range. This study emphasizes the important contribution of founder events and geographic isolation to successful homoploid hybrid speciation.
Heredity (2012) **108,** 87–95; doi:10.1038/hdy.2011.46; published online 10 August 2011

Keywords: asteraceae; hybrid speciation; hybrid zones; hybridization; *Senecio*; speciation

Introduction

The importance of hybridization as a trigger for evolutionary change and speciation has long been appreciated in plants (Stebbins 1959; Grant, 1981; Rieseberg, 1997; Rieseberg *et al*., 2003; Arnold, 2004) and also more recently in animals (Nolte *et al*., 2005; Gompert *et al*., 2006; Mavárez *et al*., 2006; Mallet, 2007; Jiggins *et al*., 2008). Traditionally, hybridization associated with a multiplication of chromosome number, that is, allopolyploid speciation, has been considered to be the major mode of speciation resulting from hybridization for two reasons. First, it is a conceptually simple means of conferring balanced chromosome numbers that restore meiosis and fertility in new hybrid taxa while creating a powerful reproductive barrier between hybrids and their progenitors. Second, allopolyploid species can be recognized relatively easily on the basis of altered chromosome numbers and associated hybrid phenotypes (Grant, 1981; Rieseberg, 1997). However, there is growing recognition that hybridization is also frequently associated with speciation without changes in chromosome number, known as homoploid hybrid speciation. This recognition has been facilitated by the increasing applicability of molecular genetic techniques to non-model organisms (Rieseberg, 1997; Gross and Rieseberg, 2005; Mallet, 2007; Abbott *et al*., 2010; Nolte and Tautz, 2010).

There are additional challenges associated with homoploid hybrid speciation compared to allopolyploid speciation. Incompatibilities between parental genomes after hybridization leading to hybrid unfitness due to increased mortality and/or reduced fertility might occur in the absence of polyploidy, thus tending to restrict successful homoploid hybrid speciation events to more closely related parents compared to allopolyploid speciation (Chapman and Burke, 2007; Paun *et al*., 2009). In addition, lack of reproductive, ecological and spatial isolation between new hybrids and their progenitors frequently leads to ongoing introgression and genetic swamping by more frequent parents, preventing hybrid establishment and subsequent independent evolution (Buerkle *et al*., 2000; Sobel *et al*., 2009; Nolte and Tautz, 2010). All known examples of homoploid hybrid speciation have overcome one or, more usually, several of these barriers to speciation.

Frequently, reproductive isolation seems to have been achieved by large-scale chromosomal rearrangements between parental genomes, termed recombinational speciation, that maximize hybrid fitness and fertility but also incidentally impose reproductive isolation with progenitors (Grant, 1981; Rieseberg, 1997).

Correspondence: Dr AC Brennan, Estación Biológica de Doñana (EBD-CSIC), Avda. Américo Vespucio s/n, 41092 Sevilla, Spain.
E-mail: acebrennan@ebd.csic.es
Received 8 November 2010; revised 14 April 2011; accepted 8 May 2011; published online 10 August 2011

Recombinational speciation has been most comprehensively demonstrated and investigated in homoploid hybrid desert and salt marsh *Helianthus* species (Rieseberg, 1997). However, ecological differentiation also appears to be a highly important factor contributing to successful homoploid hybrid speciation (Rieseberg, 1997; Buerkle *et al*., 2000; Gross and Rieseberg, 2005; Abbott *et al*., 2010). Hybrid superiority in a distinctive niche from either parent is a selective force for continued hybrid persistence (Buerkle *et al*., 2000). Furthermore, niche differentiation is often implicitly associated with decreased gene flow between hybrids and parents for various reasons such as increased spatial separation, different phenology, resource use or mate recognition cues (Filchak *et al*., 2000; Gross and Rieseberg, 2005; Mavárez *et al*., 2006; Jiggins *et al*., 2008; Abbott *et al*., 2010; Nolte and Tautz, 2010). In general, these factors are rarely absolute barriers to gene flow, but in combination they may be sufficient to tip the balance in favour of hybrid persistence leading to eventual speciation.

Divergence towards speciation is frequently aided by population isolation (allopatry and parapatry) and/or founder effects associated with colonization (Turelli *et al*., 2001; Templeton, 2008; Sobel *et al*., 2009). While it remains controversial whether founder effects and isolation ever lead to speciation without additional divergent selection, both can contribute to kick-starting the process (Slatkin, 1996; Turelli *et al*., 2001; Templeton, 2008). Founder events during colonization involve population bottlenecks followed by population growth that cause considerable changes in genetic diversity relative to the source population as the rate of genetic drift changes from high to low, generating novel variance on which selection can act (Slatkin, 1996; Templeton, 2008). Geographic population isolation limits or prevents gene flow and so is in itself an important prezygotic reproductive barrier (Coyne and Orr, 2004; Lowry *et al*., 2008; Sobel *et al*., 2009). In addition, divergence in isolation of quantitative traits (QTs) not directly associated with reproductive isolation, such as differences in phenology (Hall and Willis, 2006; Lowry *et al*., 2008), or niche partitioning (Gompert *et al*., 2006; Hall and Willis, 2006) frequently contribute to prezygotic mating barriers that are effective at limiting gene flow. Finally, conditions of population isolation accelerate genetic drift and divergent selection contributing to reproductive isolation due to Dobzhanksy–Muller incompatible interactions between loci with alternately fixed alleles between populations (Turelli and Orr, 2000).

The subject of this investigation, *Senecio squalidus*, combines elements of both hybrid and founder speciation. The recent and well-documented origin of *S. squalidus* makes it a useful and fascinating case study of ecogeographic homoploid hybrid speciation 'in action' (Abbott *et al*., 2010). *S. squalidus* is a homoploid hybrid species that recently arose following human-mediated introduction of material from a hybrid zone between *S. aethnensis* and *S. chrysanthemifolius* on Mount Etna, Sicily, to the Oxford Botanic Garden in Britain (Harris, 2002; James and Abbott, 2005). *S. aethnensis* grows on the higher slopes of Mount Etna, above 2000 m on recent lava flows, whereas *S. chrysanthemifolius* grows below 1000 m on arable agricultural and waste land. All three species are short lived, outcrossing perennial herbs with generalist hoverfly pollinators and wind-dispersed fruits. Furthermore, all species are self-incompatible but highly interfertile without disruption of self-incompatible function in hybrids (Chapman *et al*., 2005; AC Brennan and SJ Hiscock, unpublished). Thus, on Mount Etna, gene flow between *S. aethnensis* and *S. chrysanthemifolius* has generated a stable hybrid zone that exists at intermediate altitudes around 1100–1900 m between the two species' ranges (James and Abbott, 2005; Brennan *et al*., 2009). However, strong selection against hybrids prevents long-term persistence of individual hybrid lineages and prevents independent evolution and hybrid speciation within the hybrid zone itself (Brennan *et al*., 2009). In contrast, since its origin, *S. squalidus* has become a successful invasive species throughout large parts of the British Isles occupying disturbed urban habitats distinct from its parents, such as railway lines, motorway verges and waste land (James and Abbott, 2005). Here we present a comprehensive investigation measuring and comparing molecular genetic and QT differentiation between *S. squalidus*, *S. aethnensis* and *S. chrysanthemifolius*, and hybrid zone individuals from Sicily. In doing so, we identify the extent and nature of the differences that make *S. squalidus* distinct from Sicilian *Senecio* as a starting point for assessment of the relative contributions of hybridization, selection and population history to these differences.

Materials and methods

Plants

Plants of *S. aethnensis*, *S. chrysanthemifolius* and their hybrids investigated in the present study were the same as those described in Brennan *et al*. (2009). They were derived from seeds collected from separately sampled wild plants growing on Mount Etna, Sicily, typically within a 500 m^2 area per sampled population. Sample individuals were classified as *S. aethnensis*, *S. chrysanthemifolius* or their hybrids on the basis of field observations, sample location and subsequent analysis of QTs as described in Brennan *et al*. (2009). This resulted in a sample size of 28 *S. aethnensis*, 56 *S. chrysanthemifolius* and 101 hybrid individuals. In addition, 23 *S. squalidus* individuals were sampled as seed from separate maternal plants from urban sites of 500 m^2 area in Edinburgh and Cardiff in the United Kingdom, and were cultivated along with the wild-sampled *Senecio* from Sicily in a fully randomized design in a glasshouse. Most of the individuals cultivated (208) were surveyed for both molecular genetic and QT variation as detailed below.

Molecular genetic measures

Genomic DNA was extracted from plants as described in Brennan *et al*. (2009). Variation at 22 molecular genetic loci was surveyed across all 208 individuals. Results for 13 of these molecular genetic loci, including 6 loci encoding allozymes (AAT, ACP, GDH, IDH, PGI and PGM) and 7 anonymous simple sequence repeat loci (SSRs: S4, S10, S15, S20, S26, V44 and V45), were previously reported in Brennan *et al*. (2009) for the Sicilian *Senecio*. New molecular markers were developed by analyzing expressed sequence tag (EST) data for *S. aethnensis*, *S. chrysathemifolius*, *S. squalidus* and *S. vulgaris* from the online *Senecio* database

(www.seneciodb.org/database.htm). Sequence analysis involved using CAP3 to produce contigs (Huang and Madan, 1999), EMBOSS etandem (Rice *et al*., 1999) to detect SSRs and Perl scripts to visualize ESTs aligned against contigs and highlight polymorphisms resulting in six novel SSR markers (EC1019, ES1, ES29, ES43, ES45 and ES87; Supplementary Table S1). A further three insertion–deletion (indel) polymorphic markers were developed from sequence data for single copy genes in *S. aethnensis* and *S. chrysanthemifolius* (A26, C19 and D11, Supplementary Table S1; Chapman *et al*., 2007; M Chapman, unpublished). All markers were fluorescently PCR amplified and capillary sequencer genotyped as described in Brennan *et al*. (2009).

Molecular genetic analysis

Rare alleles that were observed in a single individual were excluded from analysis. Significant deviations from Hardy–Weinberg equilibrium and the presence of possible null alleles (specific non-detected alleles) were tested across loci and populations using GenAlex v6.1 (Peakall and Smouse, 2006) and Microchecker v2.2.0.3 (Van Oosterhout *et al*., 2004). To compare molecular genetic diversity between species, allelic richness (R_s) and expected heterozygosity (H_e) were calculated for each marker class (SSRs, EST SSRs, indels and allozymes) and corrected for sample size differences by rarefaction and data bootstrapping, respectively, for unbiased molecular diversity comparisons using Microsatellite Analyser Software (MSA v4.05; Dieringer and Schlötterer, 2003). Molecular genetic differentiation was investigated by measuring patterns of allele sharing between species using GenAlex v6.1 (Peakall and Smouse, 2006) and by measuring paired-species differentiation values (F_{st} and F_{st}' corrected for different marker diversities) and their 95% confidence intervals following bootstrapping using MSA v4.05 (Dieringer and Schlötterer, 2003). The influence of marker type (SSRs, EST SSRs, indels or allozymes) on patterns of differentiation was tested by comparing divergence measured for each marker type separately. Genotype data were used in a principal coordinate analysis and also a discriminant coordinate analysis of an absolute genetic distance matrix to discriminate between *S. aethnensis*, *S. chrysanthemifolius*, *S. squalidus* and Sicilian hybrid samples using the ade4 package in R v2.9 software (Dray and Dufour 2007; R Development Core Team 2009). Percentage variance explained was calculated from eigenvalues. The effectiveness of the first and second principal coordinates and discriminant scores at explaining paired-species differences was assessed by one-way analysis of variance after Bonferroni correction for multiple tests (three species pairs multiplied by two principal coordinates) using R v2.9 (R Development Core Team 2009). In addition, the influence of high, medium and low marker diversity on discriminant coordinate analysis results was investigated by analyzing different marker types separately divided into SSRs, EST SSRs, and combined indels and allozymes, respectively. Multivariate analysis of variance of the first two discriminant scores was used to assess the effectiveness of each marker class at discriminating between *S. aethnensis*, *S. chrysanthemifolius*, *S. squalidus* and Sicilian hybrid samples using R v2.9 (R Development Core Team 2009).

QT measures

Twenty QTs were measured on all sample plants raised from seed in the glasshouse and included: days to germinate, days to development of first true leaf, days to flowering, height, branch number, capitulum number, pedicel length, ray display, pollen number, pollen fertility, floret number, seed length, pappus length, photosynthetic rate at 20 °C, photosynthesis at 7 °C, leaf auricle width, leaf dissection, stomata number, leaf chlorophyll and leaf anthocyanin content as described in Brennan *et al*. (2009). Square root or natural logarithm transformations of the QT data were used when necessary to improve homogeneity of variance as described in Brennan *et al*. (2009). Missing data (9.2%) were conservatively replaced with mean QT values.

QT analysis

Differences in paired-species QT means were assessed by one-way analysis of variance after Bonferroni correction for multiple tests (four species pairs multiplied by 20 QTs). Principal component analysis (PCA) and linear discriminant analysis (LDA) were performed using the ade4 package in R v2.9 software (Dray and Dufour 2007; R Development Core Team 2009). Percentage variance explained was calculated from standard deviations of eigenvalues and important traits were identified from principal component and discriminant function axis rotations. The effectiveness of the first and second QT principal components and discriminant scores at explaining paired-species differences was assessed by one-way analysis of variance after Bonferroni correction for multiple tests using R v2.9 (R Development Core Team 2009).

Results

Molecular genetic analysis

Genotypes at 22 molecular genetic marker loci were successfully recorded for 203 of the 208 sample individuals of *S. aethnensis*, *S. chrysanthemifolius*, their Sicilian hybrids, and *S. squalidus* (Supplementary Table S2). Overall, 26/294 (8.8%) of Hardy–Weinberg equilibrium tests indicated non-equilibrium genotype frequencies and 27/294 (9.2%) of permutation tests indicated null allele presence at a 5% significance level (Supplementary Table S3). This is a small excess over the 5% significant results expected by chance alone and indicate that sample populations are generally at or close to Hardy–Weinberg equilibrium with loci exhibiting few null alleles, satisfying the assumptions of later analyses. A comparison of different classes of markers indicated a pattern of increasing molecular diversity in terms of allelic richness and heterozygosity (R_s and H_e) moving from allozymes to indels to EST SSRs to anonymous SSRs (Table 1). A comparison of the different species exhibited increasing molecular genetic diversity in the order *S. squalidus*, *S. aethnensis*, *S. chrysanthemifolius* and Sicilian hybrid *Senecio* (Table 1). Analyses of molecular genetic differentiation (F_{st} and F_{st}' paired-species estimates) indicated that *S. squalidus* was more highly differentiated from *S. aethnensis* and *S. chrysanthemifolius* than from their Sicilian hybrids. This pattern held across all different marker types with different levels of diversity, but was strongest for SSR F_{st}' measures due to relatively

Table 1 Molecular genetic diversity indices of SSRs, indels and allozymes for 203 wild-sampled *Senecio aethnensis*, *S. chrysanthemifolius*, *S. squalidus* and Sicilian hybrid *Senecio* individuals

Marker type (number)	*Random SSRs (7)*		*EST SSRs (6)*		*Indels (3)*		*Allozymes (6)*		*All loci (22)*	
Mean common sample size	*21.00 (0.59)*		*16.33 (0.76)*		*16.67 (1.20)*		*20.92 (0.08)*		*19.11 (0.57)*	
Measure	R_s	H_e	R_s	H_e	R_s	H_e	R_s	H_e	R_s	H_e
S. aethnensis	6.10 (0.75)	0.75 (0.03)	2.83 (0.48)	0.29 (0.08)	2.05 (0.28)	0.21 (0.15)	1.96 (0.03)	0.23 (0.07)	3.53 (0.47)	0.41 (0.06)
S. chrysanthemifolius	6.17 (0.91)	0.70 (0.04)	2.43 (0.56)	0.40 (0.13)	2.63 (0.63)	0.48 (0.02)	1.49 (0.22)	0.06 (0.04)	3.39 (0.53)	0.41 (0.06)
Sicilian hybrid	7.29 (0.62)	0.76 (0.02)	3.22 (0.69)	0.38 (0.12)	2.62 (0.62)	0.54 (0.08)	2.07 (0.09)	0.26 (0.06)	4.12 (0.55)	0.49 (0.06)
S. squalidus	2.86 (0.26)	0.51 (0.05)	2.33 (0.21)	0.41 (0.07)	2.00 (0.58)	0.37 (0.19)	1.50 (0.22)	0.23 (0.11)	2.23 (0.17)	0.39 (0.05)

Abbreviations: EST, expressed sequence tag; H_e, expected heterozygosity; indels, insertion–deletions; R_s, allelic richness; SSRs, simple sequence repeats.
Molecular genetic diversity indices were R_s and H_e, both corrected for the minimum species sample size per locus. Standard errors are given in brackets after each measure.

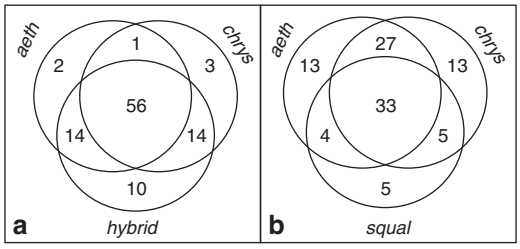

Figure 1 Patterns of allele sharing (**a**) between *Senecio aethnensis*, *S. chrysanthemifolius* and their hybrids from Sicily, and (**b**) between *S. aethnensis*, *S. chrysanthemifolius* and the hybrid species *S. squalidus* from the United Kingdom. Values within labelled circles and their overlaps are the percentage of alleles observed in each sharing class derived from 180 individuals and 111 alleles and from 108 individuals and 102 alleles for comparisons involving Sicilian hybrids and UK *S. squalidus*, respectively.

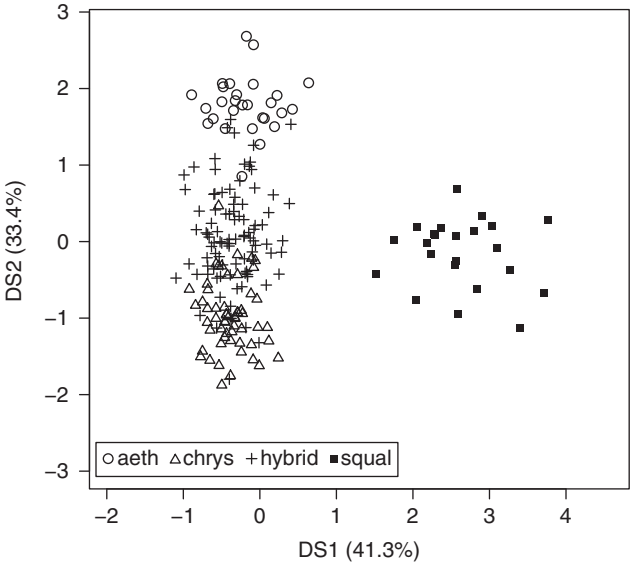

Figure 2 Plot of first and second discriminant coordinate scores resulting from an analysis of a between-individual genetic distance matrix for 22 molecular genetic markers surveyed across 203 wild-sampled *Senecio aethnensis*, *S. chrysanthemifolius*, *S. squalidus* and Sicilian hybrid individuals. Axes are labelled with the percentage variance explained by the first and second discriminant functions, respectively. The same scale has been used for both axes.

fewer shared SSR alleles between *S. squalidus* and Sicilian *Senecio* than for other less diverse marker types (Supplementary Figure S1). Patterns of between-species allele sharing indicated that, while *S. squalidus* was not distinctive in terms of many unique alleles, many more alleles were shared between Sicilian hybrids and both parents than between *S. squalidus* and both parents (Figure 1). Molecular genetic differentiation in paired-species allele frequencies as measured by F_{st} was significantly greater than zero and of a similar magnitude for comparisons between *S. aethnensis* and *S. chrysanthemifolius* and between each parent and *S. squalidus* meaning that *S. squalidus* is as distinct from each of its parent species as they are from each other (F_{st} (95% CI) = 0.21 (0.14–0.31), 0.23 (0.15–0.32), 0.21 (0.15–0.28); *aeth.* vs *chrys.*, *aeth.* vs *squal.*, *chrys.* vs *squal.*, respectively; Supplementary Figure S1). There was also significant molecular genetic differentiation between *S. squalidus* and the full range of Sicilian hybrids (F_{st} [95% CI] = 0.16 [0.11–0.22]; hybrid v. *squal*; Supplementary Figure S1) that further emphasizes the genetic distinctiveness of *S. squalidus* relative to its Sicilian progenitors. The discriminant coordinate analysis was highly effective at distinguishing distinct species clusters with the first and second discriminant scores (DS1 and DS2) explaining 41.3 and 33.4% of between-group variance, respectively (Figure 2). The principal coordinate analysis was also highly effective at distinguishing distinct species clusters with the first and second principal coordinates (PC1 and PC2) explaining 13.9 and 7.0% of total genetic distance variation, respectively (Supplementary Table S4 and Supplementary Figure S2). *S. squalidus* was highly distinct from *S. aethnensis*, *S. chrysanthemifolius* and Sicilian hybrids based on DS1 values (Figure 2, Table 2), while the parental species were distinguished by their DS2 values with Sicilian hybrid *Senecio* and *S. squalidus* having intermediate DS2 values (Figure 2, Table 2). This pattern of species differentiation was similar for subsets of high, medium and low diversity markers indicating that the molecular genetic distinctiveness of *S. squalidus* is not simply a feature of more highly discriminating high diversity markers (Supplementary Figure S3).

QT analysis

Phenotype data for 20 QTs for 208 individuals of *S. aethnensis*, *S. chrysanthemifolius*, their Sicilian hybrids,

Table 2 Comparison of quantitative traits and discriminant scores between *S. aethnensis*, *S. chrysanthemifolius*, their Sicilian hybrids and *S. squalidus*

Trait (units, transformation)	*Mean (standard error)*				F *value (DF)*				S. squalidus *phenotype*
	S. aeth.	S. chrys.	S. squal.	*Sicilian hybrid*	*aeth vs chrys*	*aeth vs squal*	*chrys vs squal*	*hybrid vs squal*	
Leaf anthocyanin concentration (mM per g, s)	0.61 (0.04)	0.73 (0.05)	0.63 (0.06)	0.68 (0.04)	1.21 (77)	0.05 (45)	0.63 (72)	0.16 (116)	No differences
Leaf chlorophyll concentration (mg per g, n)	0.50 (0.03)	0.61 (0.02)	0.71 (0.03)	0.56 (0.01)	9.89 (77)	**27.04 (44)**	8.30 (71)	**19.76 (112)**	chrys transgressive (ns)
Leaf stomata number (count, l)	18.41 (0.99)	27.50 (1.45)	27.32 (2.08)	24.05 (0.88)	**16.08 (79)**	**13.45 (47)**	0.00 (74)	1.82 (113)	chrys-like
Mean florets per capitulum (count, n)	100.49 (2.63)	78.37 (1.47)	82.19 (2.00)	91.31 (1.48)	**62.92 (82)**	**28.67 (49)**	2.10 (77)	7.98 (121)	chrys-like
Mean fruit length (mm, s)	3.36 (0.05)	2.20 (0.02)	2.21 (0.03)	2.70 (0.05)	**684.93 (82)**	**383.37 (45)**	0.05 (73)	**20.61 (110)**	chrys-like, distinct from hybrid
Mean pappus length (mm, s)	6.29 (0.11)	5.14 (0.05)	5.06 (0.13)	5.71 (0.06)	**117.46 (82)**	**54.89 (45)**	0.61 (73)	**21.50 (111)**	chrys-like, distinct from hybrid
Mean pollen number (per 3/40 florets, s)	106.60 (6.3)	66.00 (2.67)	61.73 (5.18)	85.61 (3.26)	**43.61 (78)**	**31.93 (45)**	0.77 (75)	11.08 (118)	chrys-like, distinct from hybrid (ns)
Mean poor pollen (proportion, n)	0.40 (0.03)	0.34 (0.02)	0.40 (0.03)	0.33 (0.01)	2.53 (71)	0.01 (40)	2.74 (73)	5.10 (115)	No differences
Photosynthetic rate at 20 °C (μM CO_2 per m per min, n)	−0.01 (<0.01)	−0.01 (<0.01)	−0.01 (<0.01)	−0.01 (<0.01)	0.72 (63)	3.01 (47)	0.73 (58)	2.20 (104)	No differences
Photosynthetic rate at 7 °C (μM CO_2 per m per min, n)	−0.01 (<0.01)	0.01 (<0.01)	<0.01 (<0.01)	<0.01 (<0.01)	8.50 (46)	0.21 (42)	3.12 (46)	0.14 (88)	No differences
Primary capitulum display area (mm^2, l)	2319.32 (125.18)	968.74 (32.69)	1349.30 (67.27)	1434.52 (59.64)	**184.86 (82)**	**46.45 (48)**	**25.99 (76)**	<0.01 (117)	Intermediate
Primary capitulum pedicel length (cm, s)	1.89 (0.11)	1.40 (0.08)	1.18 (0.07)	1.68 (0.06)	**13.70 (82)**	**28.21 (48)**	2.25 (76)	**14.93 (117)**	chrys transgressive (ns)
Primary inflorescence capitulum number (count, n)	7.64 (0.42)	12.09 (0.42)	8.95 (0.69)	10.07 (0.33)	**44.18 (82)**	2.91 (48)	**15.25 (76)**	2.12 (117)	aeth-like
Primary stem branch number (count, l)	7.96 (0.6)	13.52 (0.97)	12.14 (0.91)	10.53 (0.57)	11.27 (82)	11.35 (48)	0.03 (76)	2.45 (117)	chrys-like (ns)
Primary stem height (cm, n)	31.17 (1.21)	36.97 (1.00)	19.36 (1.26)	34.00 (0.76)	12.41 (82)	**45.07 (48)**	**98.53 (76)**	**73.24 (117)**	aeth transgressive
Primary stem midleaf auricle width (mm, l)	4.46 (0.52)	1.30 (0.17)	1.71 (0.39)	2.41 (0.17)	**81.44 (78)**	**29.41 (49)**	2.41 (73)	5.00 (118)	chrys-like
Primary stem midleaf dissection, perimeter to area ratio (per mm, l)	0.21 (0.02)	0.83 (0.04)	0.44 (0.03)	0.48 (0.03)	**288.56 (78)**	**55.95 (49)**	**54.1 (73)**	0.18 (118)	Intermediate
Time first true leaf to flowering (days, n)	74.38 (1.95)	71.89 (1.35)	70.53 (3.20)	74.45 (1.02)	1.11 (78)	1.17 (43)	0.21 (71)	2.22 (105)	No differences
Time germination to first true leaf (days, n)	10.77 (0.37)	11.93 (0.34)	13.20 (0.57)	11.35 (0.31)	4.40 (78)	**13.88 (44)**	3.78 (72)	6.73 (111)	No differences
Time sowing to germination (days, n)	14.00 (1.33)	15.96 (0.95)	13.70 (1.16)	14.65 (0.72)	1.41 (78)	0.03 (44)	1.73 (72)	0.33 (111)	No differences
Mol. Gen. DS1	−0.19 (0.07)	−0.39 (0.03)	2.62 (0.12)	−0.35 (0.03)	8.80 (83)	**449.91 (49)**	**1055.80 (78)**	**1141.80 (116)**	Transgressive (ns)
Mol. Gen. DS2	1.79 (0.07)	−0.97 (0.06)	−0.15 (0.09)	0.09 (0.07)	**789.91 (83)**	**298.57 (49)**	**53.93 (78)**	2.94 (116)	Intermediate
QT DS1	1.57 (0.07)	−0.84 (0.04)	−0.95 (0.11)	0.25 (0.08)	**985.27 (82)**	**402.63 (49)**	1.25 (77)	**51.72 (122)**	chrys-like
QT DS2	0.33 (0.11)	−0.59 (0.09)	1.76 (0.16)	−0.16 (0.08)	**37.03 (82)**	**59.56 (49)**	**179.22 (77)**	**110.67 (122)**	Transgressive

Abbreviations: ANOVA, analysis of variance; DS1 and DS2, first and second discriminant scores; l, natural logarithm; Mol. Gen, molecular genetic diversity; n, none; QT, quantitative trait; s, square root.

Transformations to improve homogeneity of variance were n, s or l. Transformed QT means were compared using one-way ANOVA. Bold text F ratios indicate significantly different trait means at a 5% significance level or less after Bonferroni correction for multiple tests. The *S. squalidus* phenotype column summarizes *S. squalidus* QTs relative to its parents, *S. aethnensis* and *S. chrysanthemifolius*, on the basis of mean trait values and the ANOVA. Traits were defined as transgressive if *S. squalidus* was significantly different from both parents and not intermediate. The (ns) postscript in this column indicates that *S. squalidus* trait definition was not significant at an overall 5% significance level.

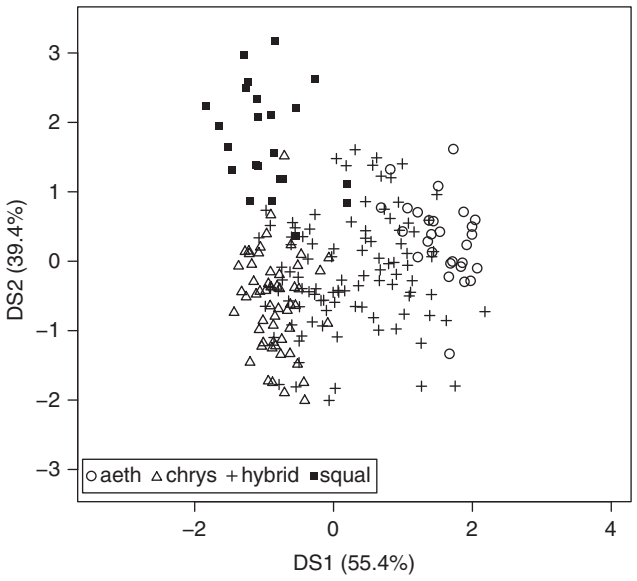

Figure 3 Plot of first and second discriminant component scores resulting from an analysis of 20 quantitative genetic traits measured on 208 wild-sampled *Senecio aethnensis*, *S. chrysanthemifolius*, *S. squalidus* and Sicilian hybrid individuals. Axes are labelled with the percentage variance explained by the first and second discriminant functions, respectively. The same scale has been used for both axes.

and *S. squalidus* are presented in Supplementary Table S5. The LDA of QT variation distinguished species based on the first and second discriminant scores that explained 55.4 and 39.4% of the between-group variation, respectively (Figure 3). The corresponding PCA of QT variation also distinguished species with PC1 and PC2 explaining 23.5 and 9.3% of the total variation, respectively (Supplementary Table S4, Supplementary Figure S4). The Sicilian parent species were distinguished by DS1 while their Sicilian hybrids exhibited intermediate values for DS1 (Figure 3, Table 2). The three traits that contributed most to the DS1 axis were fruit length, pappus length and stomata number. The LDA DS2 values showed that *S. squalidus* individuals were distinct from Sicilian *Senecio* (Figure 3, Table 2) due largely to differences in plant height, leaf dissection and ray display. Comparisons of paired-species means for each of the twenty QTs permitted assessment of the *S. squalidus* phenotype relative to its parents (Table 2). This showed that although there was no significant variation between the three species for seven of twenty QTs, *S. squalidus* was intermediate to its parents for leaf dissection and ray display, more similar to *S. chrysanthemifolius* for seven QTs and more similar to *S. aethnensis* for just one QT, capitulum number. We defined transgressive traits in *S. squalidus* as not intermediate and significantly (or close to significantly) different from both parents. According to this definition, *S. squalidus* was transgressively shorter for plant height, had shorter pedicel length and contained more concentrated leaf chlorophyll than either parent. Thus, while for many QTs, *S. squalidus* is phenotypically similar to *S. chrysanthemifolius*, for one QT it is more similar to *S. aethnensis*, while for others it is transgressively distinct from both parents. Differences between *S. squalidus* and Sicilian hybrids were identified for fruit length, pappus length and pollen number, as well as for the three trangressive traits described above (Table 2).

Discussion

Our detailed comparisons of molecular and QT diversity between the recently originated homoploid hybrid species, *S. squalidus*, and its progenitor species, *S. aethnensis* and *S. chrysanthemifolius*, have emphasized this hybrid species' distinctiveness from both of its parents and their Sicilian hybrids. Thus, *S. squalidus* supports the consensus that successful, established hybrid species are genetically and phenotypically distinct from progenitors (Rieseberg, 1997; Rieseberg *et al*., 2003; Arnold, 2004; Gross and Rieseberg, 2005; Nolte and Tautz, 2010). However, a particularly interesting finding to emerge from the present study was that molecular genetic divergence between *S. squalidus* and its parents is greater than QT divergence.

Molecular genetic divergence of *S. squalidus* from its parents

S. squalidus was as differentiated from both parents for molecular genetic diversity as the latter were from each other, in striking contrast with Sicilian hybrids that only exhibited intermediate discriminant scores between the parental species (Figure 2 and Supplementary Figure S2). Importantly, *S. squalidus* was also highly genetically differentiated from the full range of Sicilian hybrid genotypes and these patterns of differentiation were remarkably consistent across different marker types that exhibited different levels of allelic diversity (Figure 2, Supplementary Figures S1 and S2).

Investigation of the homoploid hybrid origin of *S. squalidus* is greatly facilitated by historical records documenting the timing and approximate population sizes of key events in its population history such as its introduction to the United Kingdom and its period of cultivation before exponential population growth (Harris, 2002; Abbott *et al*., 2009). Thus, it is unsurprising to find that unbiased comparisons of molecular genetic diversity indicate reduced diversity in *S. squalidus* relative to Sicilian *Senecio* in terms of allele richness and expected heterozygosity (R_s and H_e) across all investigated marker types that themselves exhibit a wide range of molecular genetic diversities (Table 1). However, the extensive molecular genetic differentiation observed between *S. squalidus* and its parental species and Sicilian hybrids (Figure 2) is not all simply due to reduced genetic diversity and therefore requires further explanation. One consequence of our background knowledge of the population history of *S. squalidus* is that we know that the extensive molecular genetic differentiation between *S. squalidus* relative to its parents and Sicilian hybrids must have accumulated in the 300 generations (assuming one generation per year) since introduction to the United Kingdom. Founder effects and genetic drift have probably had an important, if not predominant, role in generating and shaping genetic differentiation between *S. squalidus* and its progenitors during this period. The known population history has involved an extreme population bottleneck at the time of introduction of hybrid material from Mount Etna to the United Kingdom, followed by a long period of 90–150 years/

generations of small cultivated population size prior to invasiveness. Such demographic history is likely to have contributed to rapid genetic change through loss of diversity and drift (Slatkin, 1996). Subsequently, as *S. squalidus* experienced rapid population growth during its invasive phase, the remaining genetic diversity would have been preserved and augmented by mutation while preserving altered allele frequencies from the introduction phase (Slatkin, 1996; Templeton, 2008). This study identified that at least three alleles present in UK *S. squalidus* were absent from both its parents and Sicilian hybrids (Figure 1). In theory, selection generating further molecular genetic differentiation is enhanced in growing populations because relaxed drift prevents more new favourable mutations or allele interactions from being lost when initially very rare soon after appearance (Slatkin, 1996).

A feature of the allele sharing analysis was that 10% of alleles were observed in Sicilian hybrid samples that were not shared with either parent species (Figure 1). It is possible that this is an effect of limited sample sizes failing to identify some of these alleles if they were present at low frequency in parent species samples. It is also possible that the hybrid zone could be acting as a reservoir of molecular genetic diversity retaining alleles introgressed from both parent species that have later been lost due to drift in the parents. The *Senecio* hybrid zone on Mount Etna is both persistent and large, having persisted for more than 300 generations and estimated to extend between 1.5–3.0 km in width (Brennan *et al*., 2009). The pattern of alleles from both parents in *S. squalidus* is suggestive of the importance of the hybrid zone as the potential original source of introduced *S. squalidus*. We are currently using individual-based, forward simulations, under a plausible range of demographic population histories for *S. squalidus*, to assess the relative contribution of natural selection and neutral processes to its divergence from its parental species (Barker *et al*., submitted).

Quantitative trait divergence of *S. squalidus* from its parents

LDA revealed a detailed and complex picture of QT differentiation between the species (Table 2). For some traits, *S. squalidus* exhibited an intermediate phenotype between its two differentiated parents, as expected in the case of QTs controlled by multiple quantitative trait loci (QTLs) of individually small additive effect acting mainly in the same trait direction within parental species and in mainly opposite directions between them (Rieseberg *et al*., 1999; Lexer *et al*., 2003). Frequently, however, QTs exhibited a parental-like phenotype in *S. squalidus* that instead indicates dominant expression of the QTL alleles of one parent over the other, or a biased contribution in *S. squalidus* of QTL alleles from one parent only. The majority of parental-like QTs in *S. squalidus* were in the direction of *S. chrysanthemifolius* (seven versus one *S. aethnensis*-like QT, Table 2) placing *S. squalidus* close to *S. chrysanthemifolius* in terms of the first LDA discriminant function (Figure 3).

The pattern of QT divergence between *S. squalidus* and its progenitors resembles in some respects that reported for certain other homoploid hybrid species and their progenitors. For example, a glasshouse-based QT investigation of the hybrid species *Argyranthemum sundingii* and its progenitors, *A. broussonetii* and *A. frutescens*, observed a similar pattern of hybrid QT variation partially overlapping parental and synthetic hybrid QT variation, whereas for certain QTs the hybrid species resembled the *A. frutescens* parent in particular (Brochmann *et al*., 2000).

Three of the QTs we examined exhibited transgressive or more extreme QT means relative to both parent species (Table 2). *S. squalidus* was also distinct from Sicilian hybrids on the basis of six QTs, including the three transgressive traits (Table 2). These findings are in accordance with observations of other homoploid hybrid species, such as *H. anomalus* and *H. deserticola*, which showed that the hybrid species were distinct from early-generation hybrids and exhibit particular QTs that are transgressive relative to their parents *H. annuus* and *H. petiolaris* (Schwarzbach *et al*., 2001; Rosenthal *et al*., 2002, 2005). Transgressive QT expression can be spontaneously generated in new hybrids when QTLs with fixed but opposing directions of effects within species segregate in heterozygous hybrids and also by interaction of divergent gene expression mechanisms in hybrids (Rieseberg *et al*., 1999; Hegarty and Hiscock, 2008). Furthermore, in the aforementioned *Helianthus* homoploid hybrids, selection among parental QTL alleles in novel hybrid habitats has been implicated as a major cause of adaptive transgressive character expression (Rieseberg *et al*., 1999, 2003; Lexer *et al*., 2003; Gross and Rieseberg, 2005).

Comparison of amounts of molecular genetic and QT divergence between *S. squalidus* and its parent species

Direct comparisons between levels of QT differentiation (Q_{st}) and molecular genetic differentiation (F_{st} or G_{st}) are a complementary means of investigating the extent to which adaptation has contributed to quantitative differentiation (Merilä and Crnokrak, 2001; McKay and Latta, 2002). However, the population sampling approach of our study does not permit adequate partitioning of genetic and environmental contributions to quantitative variance, preventing Q_{st} estimates that, validly, can be directly compared with molecular genetic estimates. Therefore, we chose instead to focus on multivariate analysis of QTs that have been successfully applied in other hybrid studies (for example, Brochmann *et al*., 2000; Rosenthal *et al*., 2002, 2005).

Our analyses indicated that there was less QT divergence between *S. squalidus* and its progenitors relative to molecular genetic divergence. Thus, whereas *S. squalidus* showed some overlap with *S. chrysanthemifolius* and the Sicilian hybrids for QTs, there was no such overlap in terms of molecular genetic divergence (Table 2 and Figures 2 and 3). In contrast to randomly sampled molecular genetic markers representing neutral variation, QTs represent phenotypic variation that is potentially exposed to selection. Accordingly, greater QT differentiation than neutral molecular genetic differentiation is typically observed in studies investigating both forms of variation (Merilä and Crnokrak, 2001; McKay and Latta, 2002). Furthermore, divergent ecological selection is increasingly seen as a key feature required for hybrid species establishment and speciation more generally (Rieseberg, 1997; Buerkle *et al*., 2000; Sobel *et al*., 2009; Abbott *et al*., 2010; Nolte and Tautz, 2010), and invasion of

novel urban habitats by *S. squalidus* appears to conform to this expectation. Indeed, some local adaptation has been identified across the UK range of *S. squalidus* for flowering time and drought and temperature stress tolerance traits indicating that rapid QT evolution is possible for this species (Allan and Pannell, 2009). In theory, for QTs controlled by multiple QTLs of individual small effect, interactions between alleles at different loci permit evolution of a considerable range of QT variation with only small changes in allele frequency at individual loci (Latta, 1998; Le Corre and Kremer, 2003). However, our observation of greater molecular genetic differentiation than quantitative genetic differentiation contradicts these expectations and requires explanation.

Some of the reasons for this might be due to biases associated with estimating QT differentiation using the PCA approach. The PCA might be prone to underestimating overall QT differentiation due to variance of QTs that are not divergent between species. Thus the resulting principal components will consist of eigenvectors that seek to explain total variation that might not necessarily correspond to maximum species divergence. However, we found that the corresponding LDA analysis that determines the optimal discriminant function of QT loadings to distinguish defined species groups, and that is not susceptible to the potential weakness of the PCA, provided a similar picture of species differentiation to the PCA. A different possible explanation might therefore be that complex interactions between QTs and their underlying genetics, such as pleiotropic constraints, imposed by interactions between multiple QTLs or QTs could generate evolutionary stasis despite selection due to the complex multivariate nature of adaptation (Merilä *et al*., 2001; Walsh and Blows, 2009). Alternatively, divergent selection pressures might have been weak in this particular case of hybrid speciation because complete geographic isolation between the introduced hybrid and Sicilian progenitors would have removed selection pressures for ecological differentiation as a means of spatial and reproductive isolation that seems to be such an important feature of *in situ* homoploid hybrid speciation (Buerkle *et al*., 2000; Lexer *et al*., 2003; Gross and Rieseberg, 2005). That said, it is unlikely that no adaptation to ecological differences between the intermediate altitude slopes of Mount Etna and urban British habitats was required in the transition from Sicilian hybrid *Senecio* to *S. squalidus*, given that the QT analysis presented here finds them to be distinct (Figure 3 and Table 2) and that local adaptation may have subsequently occurred within the UK range of *S. squalidus* (Allan and Pannell, 2009). Finally, it is worth emphasizing that the observed QT differentiation between *S. squalidus* and its progenitors is very recent and is likely to represent an example of speciation-in-progress having been underway for only ~300 generations, in comparison with other new homoploid species examined where speciation has been underway for far longer (Gross and Rieseberg, 2005). Further studies of adaptation, particularly reciprocal transplant experiments, are required to better understand selection for QT differentiation between this new hybrid species and its parents.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank Marie Chadburn, David Forbes, Harry Hodge, Cameron Hunt and Ailan Wang for plant care and contributions to plant genotyping and phenotyping and to three anonymous reviewers for their constructive comments. The authors acknowledge the financial support of a NERC grant (NE/D014166/1) to RJA (Principal Investigator) and to DB and SJH (Co-Investigators), and a Research Councils UK Academic Fellowship to DB. ACB is currently supported by funding from FP7-REGPOT 2010-1, Grant No. 264125 EcoGenes.

References

Abbott RJ, Brennan AC, James JK, Forbes DF, Hegarty MJ, Hiscock SJ (2009). Recent hybrid origin and invasion of the British Isles by a self-incompatible species, Oxford ragwort (*Senecio squalidus* L., Asteraceae). *Biol Invasions* **11**: 1145–1158.

Abbott RJ, Hegarty MJ, Hiscock SJ, Brennan AC (2010). Homoploid hybrid speciation in action. *Taxon* **59**: 1375–1386.

Allan R, Pannell JR (2009). Rapid divergence in physiological and life-history traits between northern and southern populations of the British introduced neo-species, *Senecio squalidus* L. *Oikos* **118**: 1053–1061.

Arnold ML (2004). Transfer and origin of adaptations through natural hybridization: were Anderson and Stebbins right? *Plant Cell* **16**: 562–570.

Barker D, Brennan SJ, Hiscock RJ. Evolution during hybrid speciation and invasion; causes of genetic change in Senecio squalidus. *Bioinformatics* (submitted).

Brennan AC, Bridle JR, Wang AL, Hiscock SJ, Abbott RJ (2009). Adaptation and selection in the *Senecio* (Asteraceae) hybrid zone on Mount Etna, Sicily. *New Phytol* **183**: 702–717.

Brochmann C, Borgen L, Stabbetorp OE (2000). Multiple diploid hybrid speciation of the Canary Island endemic *Argyranthemum sundingii* (Asteraceae). *Plant Syst Evol* **220**: 77–92.

Buerkle CA, Morris RJ, Asmussen MA, Rieseberg LH (2000). The likelihood of homoploid hybrid speciation. *Heredity* **84**: 441–451.

Chapman MA, Burke JM (2007). Genetic divergence and hybrid speciation. *Evolution* **61**: 1773–1780.

Chapman MA, Chang J, Weisman D, Kesseli RV, Burke JM (2007). Universal markers for comparative mapping and phylogenetic analysis in the Asteraceae (Compositae). *Theor Appl Genet* **115**: 747–755.

Chapman MA, Forbes DG, Abbott RJ (2005). Pollen competition among two species of *Senecio* (Asteraceae) that form a hybrid zone on Mt Etna Sicily. *Am J Bot* **92**: 730–735.

Coyne JA, Orr HA (2004). *Speciation*. Sinauer Associates: Sunderland, MA.

Dieringer D, Schlötterer C (2003). Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol Ecol Notes* **3**: 167–169.

Dray S, Dufour AB (2007). The ade4 package: implementing the duality diagram for ecologists. *J Stat Software* **22**: 1–40.

Filchak KE, Roethele JB, Feder JL (2000). Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature* **407**: 739–742.

Gompert Z, Fordyce JA, Forister M, Shapiro AM, Nice CC (2006). Homoploid hybrid speciation in an extreme habitat. *Science* **314**: 1923–1925.

Grant V (1981). *Plant speciation*. Columbia University Press: New York.

Gross BL, Rieseberg LH (2005). The ecological genetics of homoploid hybrid speciation. *J Hered* **96**: 241–252.

Hall MC, Willis JH (2006). Divergent selection on flowering time contributes to local adaptation in *Mimulus guttatus* populations. *Evolution* **60**: 2466–2477.

Harris SA (2002). Introduction of Oxford ragwort, *Senecio squalidus* L. Asteraceae), to the United Kingdom. *Watsonia* **24**: 31–43.
Hegarty MJ, Hiscock SJ (2008). Genomic clues to the evolutionary success of polyploid plants. *Current Biol* **18**: 435–444.
Huang X, Madan A (1999). CAP3: a DNA sequence assembly program. *Genome Res* **9**: 868–877.
James JK, Abbott RJ (2005). Recent, allopatric, homoploid hybrid speciation: the origin of Oxford ragwort, *Senecio squalidus* (Asteraceae), in the British Isles from a hybrid zone on Mount Etna, Sicily. *Evolution* **59**: 2533–2547.
Jiggins CD, Salazar C, Linares M, Mavárez J (2008). Hybrid trait speciation and *Heliconius* butterflies. *Phil Trans R Soc Lond B* **363**: 3047–3054.
Latta RG (1998). Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. *Am Nat* **151**: 283–292.
Le Corre V, Kremer A (2003). Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics* **164**: 1205–1219.
Lexer C, Welch ME, Raymond O, Rieseberg LH (2003). The origin of ecological divergence in *Helianthus paradoxus* (Asteraceae): selection on transgressive characters in a novel hybrid habitat. *Evolution* **57**: 1989–2000.
Lowry DB, Modliszewski JL, Wright KM, Wu CA, Willis JH (2008). The strength and genetic basis of reproductive isolating barriers in flowering plants. *Phil Trans R Soc Lond B* **363**: 3009–3021.
Mallet J (2007). Hybrid speciation. *Nature* **446**: 279–283.
Mavárez J, Salazar CA, Bermingham E, Salcedo C, Jiggins CD, Linares M (2006). Speciation by hybridization in *Heliconius* butterflies. *Nature* **441**: 868–871.
McKay JK, Latta RG (2002). Adaptive population divergence: markers, QTL and traits. *Trends Ecol Evol* **17**: 285–291.
Merilä J, Crnokrak P (2001). Comparison of genetic differentiation at marker loci and quantitative traits. *J Evol Biol* **14**: 892–903.
Merilä J, Sheldon BC, Kruuk LEB (2001). Explaining stasis: microevolutionary studies in natural populations. *Genetica* **112**: 199–222.
Nolte AW, Freyhof J, Stemshorn KC, Tautz D (2005). An invasive lineage of sculpins, *Cottus* sp (Pisces, Teleostei) in the Rhine with new habitat adaptations has originated from hybridization between old phylogeographic groups. *Proc R Soc B* **272**: 2379–2387.
Nolte AW, Tautz D (2010). Understanding the onset of hybrid speciation. *Trends Genet* **26**: 54–58.
Paun O, Forest F, Fay MF, Chase MW (2009). Hybrid speciation in angiosperms: parental divergence drives ploidy. *New Phytol* **182**: 507–518.
Peakall R, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* **6**: 288–295.
R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna.
Rice P, Longden I, Bleasby A (1999). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**: 276–277.
Rieseberg LH (1997). Hybrid origins of plant species. *Ann Rev Ecol Syst* **28**: 359–389.
Rieseberg LH, Archer MA, Wayne RK (1999). Transgressive segregation, adaptation and speciation. *Heredity* **83**: 363–372.
Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T *et al.* (2003). Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**: 1211–1216.
Rosenthal D, Schwarzbach AE, Donovan LA, Raymond O, Rieseberg LH (2002). Phenotypic differentiation between three ancient hybrid taxa and their parental species. *Int J Plant Sci* **163**: 387–398.
Rosenthal DM, Rieseberg LH, Donovan LA (2005). Re-creating ancient hybrid species' complex phenotypes from early-generation synthetic hybrids: three examples using wild sunflowers. *Am Nat* **166**: 26–41.
Schwarzbach AE, Donovan LA, Rieseberg LH (2001). Transgressive character expression in a hybrid sunflower species. *Am J Bot* **88**: 270–277.
Slatkin M (1996). In defense of founder-flush theories of speciation. *Am Nat* **147**: 493–505.
Sobel JM, Chen GE, Watt LR, Schemske DW (2009). The biology of speciation. *Evolution* **64**: 295–315.
Stebbins GL (1959). The role of hybridization in evolution. *Proc Amer Phil Soc* **103**: 231–251.
Templeton A (2008). The reality and importance of founder speciation in evolution. *Bioessays* **30**: 470–479.
Turelli M, Barton NH, Orr JA (2001). Theory and speciation. *Trends Ecol Evol* **16**: 330–343.
Turelli M, Orr HA (2000). Dominance, epistasis and the genetics of postzygotic isolation. *Genetics* **154**: 1663–1679.
Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004). Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* **4**: 535–538.
Walsh B, Blows MW (2009). Abundant genetic variation plus strong selection = multivariate genetic constraints: a geometric view of adaptation. *Ann Rev Ecol Syst* **40**: 41–59.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)