

Maintenance of species boundaries in three sympatric *Ligularia* (Senecioneae, Asteraceae) species

Ningning Zhang^{1,2}, Yongpeng Ma¹, Ryan A. Folk³, Jiaojun Yu¹, Yuezhi Pan¹ and Xun Gong^{1*}

1. Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, the Chinese Academy of Sciences, Kunming 650201, China

2. University of the Chinese Academy of Sciences, Beijing 100049, China

3. Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA

*Correspondence: Gong Xun (gongxun@mail.kib.ac.cn)

doi: 10.1111/jipb.12674

Abstract The key process in speciation concerns the formation and maintenance of reproductive isolating barriers between diverging lineages. Although species boundaries are frequently investigated between two species across many taxa, reproductive isolating barriers among multiple species (>2) that would represent the most common phenomenon in nature, remain to be clarified. Here, we use double digest restriction-site associated DNA (ddRAD) sequencing to examine patterns of hybridization at a sympatric site where three *Ligularia* species grow together and verify whether those patterns contribute to the maintenance of boundaries among species. The results based on the RAD SNP datasets indicated hybridization *Ligularia cyathiceps* × *L. duciformis* and *L. duciformis* × *L. yunnanensis* were both restricted to

F₁s plus a few first-generation backcrosses and no gene introgression were identified, giving rise to strong reproductive isolation among hybridizing species. Moreover, hybrid swarm simulation, using HYBRIDLAB, indicated the RAD SNP datasets had sufficient discriminatory power for accurate hybrid detection. We conclude that parental species show strong reproductive isolation and they still maintain species boundaries, which may be the key mechanism to maintain species diversity of *Ligularia* in the eastern Qinghai-Tibetan Plateau and adjacent areas. Moreover, this study highlights the effectiveness of RAD sequencing in hybridization studies.

Edited by: Meng-Xiang Sun, Wuhan University, China

Received Apr. 10, 2018; **Accepted** May 31, 2018; **Online on** Jun. 7, 2018

INTRODUCTION

A key process in speciation concerns the origin and maintenance of reproductive isolating barriers between species (Coyne 1994; Levin 2000; Wu 2001). Given a general scenario that interfertile species grow in sympatry and interspecific hybridization occurs, reproductive isolation could still be maintained by some processes that occur within hybrid zones. For instance, in the early stages of hybridization, hybrids may be only restricted to the F₁ generation with no production of later-generation hybrids, thus maintaining species boundaries (Surget-Groba and Kay 2013; Zhang et al. 2013). Both intrinsic incompatibilities and extrinsic selection could contribute to the removal of all other

hybrid types (Milne et al. 2003; Milne and Abbott 2008; Pinheiro et al. 2016). Two common phenomena with regard to intrinsic incompatibilities are hybrid inviability and hybrid sterility, both of which can be triggered by Dobzhansky-Muller incompatibilities or chromosomal rearrangements (Dobzhansky 1936; Muller 1939; Stebbins 1950, 1958; White 1978; Rieseberg et al. 1999). Alternatively, in an F₁ dominant hybrid zone, fertile F₁s can outcompete all other hybrid types due to strong extrinsic selection (Milne et al. 2003; Milne and Abbott 2008; Zha et al. 2009). In addition, species in sympatry usually interact with each other and form complicated relationships, such as the network evolutionary history for some taxa where many reticulate evolutionary events occur (Stebbins 1950; Arnold 1997;

Otto and Whitton 2000; Linder and Rieseberg 2004). Although species boundaries are frequently investigated between two species across many taxa, reproductive isolating barriers among multiple sympatric species (>2) that would represent the most common phenomenon in nature, remain to be clarified.

Ligularia Cassini, belonging to the Subtribe Tussilagininae (Senecioneae, Asteraceae), consists of some 140 species, which are primarily distributed in Asia, with the exception of two species present in Europe (Liu and Illarionova 2011). More than 100 *Ligularia* species are distributed in the eastern Qinghai-Tibetan Plateau (QTP) region and adjacent areas, among which more than 60 are endemic species (Liu 1989, 2004). The Hengduan Mountains region, which is located at the southeastern edge of the QTP and is considered among the global biodiversity hotspots (Myers et al. 2000), is recognized as a main center of evolution and diversification of *Ligularia* (Liu et al. 1994).

Species of *Ligularia* are mostly diploid (Liu 2004) and occupy a great variety of habitats, such as alpine meadow, swamp and forest edge, at altitudes ranging from 1,000 to 5,000 m (Gong 2005). *Ligularia* as currently circumscribed is not a monophyletic group; it comprises the monophyletic *Ligularia-Cremanthodium-Parasenecio* (L-C-P) complex, together with other related genera in the QTP (Liu et al. 2006; Ren 2012). Allopatric speciation during the uplifts of the QTP and interspecific hybridization, caused by secondary sympatry, may have given rise to the high species diversity observed in the L-C-P complex (Liu et al. 2006). In *Ligularia*, many species occur together with overlapping flowering periods; natural hybridization has been frequently observed (e.g. Pan et al. 2008; Yu et al. 2011, 2014a, 2014b). To understand the mechanisms by which the remarkable diversity of *Ligularia* species in these areas arose, it is necessary to examine how species barriers are maintained within this group.

Double digest restriction-site associated DNA (ddRAD) sequencing is a next-generation sequencing approach that can generate massive genome-wide single nucleotide polymorphism (SNP) datasets, without any prior genomic information, thus making it widely applied to evolutionary studies, such as population genetics and phylogenetic research (Cromie et al. 2013; Eaton and Ree 2013; Reitzel et al. 2013; Hipp et al. 2014). Recently, RAD sequencing was applied in research on speciation and genetic introgression

(Pujolar et al. 2014; Capblancq et al. 2015; Zohren et al. 2016). Next-generation RAD sequencing can be used to analyze genome-wide patterns of variation in non-model organisms; these methods have improved hybrid identification and enabled examination of the genetic and evolutionary consequences of species hybridization (Pujolar et al. 2014). However, most studies concentrate on animals, with a few examples in plant taxa to date (e.g. Stölting et al. 2013; Christe et al. 2016; Zohren et al. 2016).

In this study, we used genome-wide SNP data obtained from ddRAD sequencing to examine the hybridization patterns in three *Ligularia* species (*L. cyathiceps* Handel-Mazzetti, *L. duciformis* (C. Winkler) Handel-Mazzetti and *L. yunnanensis* (Franchet) C. C. Chang), in the sympatric area, to verify whether those patterns contribute to the maintenance of boundaries among species. There are two hypotheses where rampant gene introgression occurs and blurs the species boundaries among sympatric *Ligularia* species, or hybridization is limited to early generation hybrids where species boundaries can be maintained.

RESULTS

ddRAD sequencing

Sequencing of the ddRAD libraries of 68 individuals generated 1.6 billion first (left) paired-end reads, in total (we only used the first reads for analysis), after quality filtering in *process_radtags*. On average, 23.6 million reads were processed per individual for these 68 individuals, ranging from 16.8 million (H26) to 72.7 million (D17) (Table S1). All reads were trimmed to a length of 140 bp, resulting in 3.3 Gb data retained, on average, per individual. The mean quality score of retained reads was 41, indicating high quality base calling. The average GC content was 37.2%, without apparent differences among investigated individuals. Retained reads were assembled into an average of 940,143 putative loci or RAD tags, per individual, with the average coverage depth per tag of $13.5 \times$ (Table S2). A catalog of 4,137,372 putative loci was constructed by *cstacks* and an average of 934,863 putative loci, per individual, was matched to the catalog in *sstacks*. Using *populations* in STACKS, a total of 2,150 SNPs was identified in dataset A, where the data matrix was 81.6% complete. There were 156 and 1,447 SNPs retained in dataset B and C, with 89.1% and

86.8% complete data, respectively. Moreover, the missing data content, per individual, was less than 20% for all the individuals in these three datasets.

Genetic structure and genomic differentiation

STRUCTURE analysis, based on 2,150 SNPs, showed that when $K = 2$, the posterior log probability of the data, $L(K)$, achieved the largest increase (Figure 1A) and ΔK value was the maximum ($\Delta K = 3,520.2$; Figure 1B). Therefore, the optimal K value was 2 and the 68 investigated individuals were divided into two genetic clusters. At $K = 2$ (Figure 1C), *L. cyathiceps*, *L. duciformis* and their hybrids (morphotypes A and B) comprised one pure group, whereas *L. yunnanensis* clustered into

another pure group. By contrast, most hybrid individuals of morphotypes C and D possessed mixed genetic components from two groups, with the exception of three individuals (H12, S1, and S4) that consisted of alleles inferred entirely to be from *L. yunnanensis*.

The genetic differentiation coefficient F_{ST} value between *L. cyathiceps* and *L. duciformis* ($F_{ST} = 0.682$; Table 1) was the lowest compared to differentiation between *L. yunnanensis* and *L. cyathiceps*/*L. duciformis* ($F_{ST} = 0.819/0.768$; Table 1). It is possible that this relatively low genetic differentiation, observed between *L. cyathiceps* and *L. duciformis*, resulted in their treatment as a single genetic cluster in STRUCTURE when compared to *L. yunnanensis*.

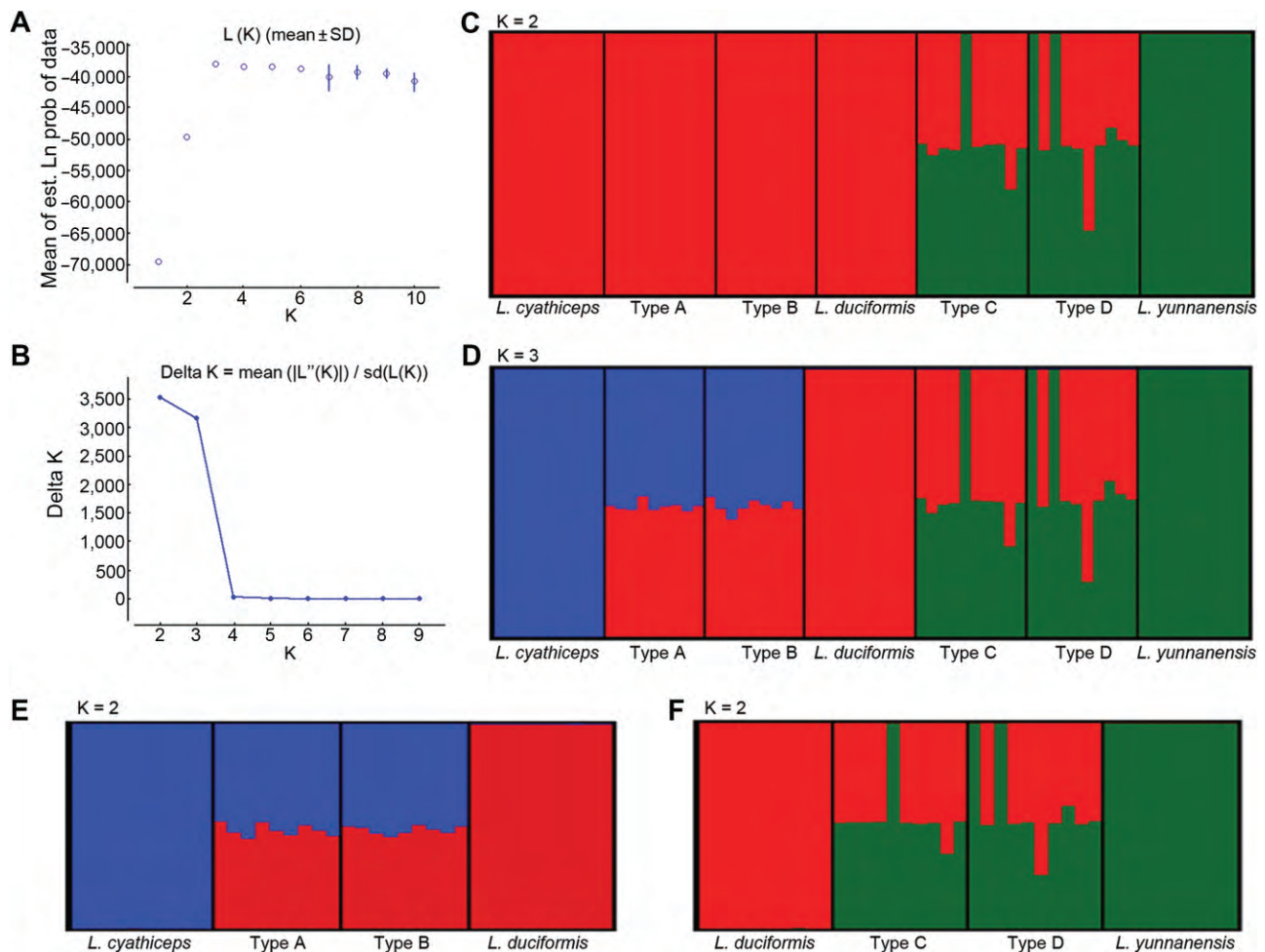


Figure 1. STRUCTURE plots for different datasets

(A–D) 2,150 SNPs, 68 individuals. (A) Mean log probability of data $L(K)$ ($\pm SD$) over 10 runs for each K value. (B) Rate of change in the log probability of data between successive K values. (C) Genetic admixture of 68 individuals when $K = 2$. (D) Genetic admixture of 68 individuals when $K = 3$. (E) Genetic admixture of 38 individuals in hybridization group *L. cyathiceps* × *L. duciformis* at 156 SNPs. (F) Genetic admixture of 40 individuals in hybridization group *L. duciformis* × *L. yunnanensis* at 1,447 SNPs.

Table 1. Pairwise F_{ST} values for three *Ligularia* species and four hybrid morphotypes based on 2,150 SNPs

Taxon	<i>L. cyathiceps</i>	morphotype A	morphotype B	<i>L. duciformis</i>	morphotype C	morphotype D	<i>L. yunnanensis</i>
<i>L. cyathiceps</i>		<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
morphotype A	0.232		0.288	<0.001	<0.001	<0.001	<0.001
morphotype B	0.239	0.000		<0.001	<0.001	<0.001	<0.001
<i>L. duciformis</i>	0.682	0.278	0.287		<0.001	<0.001	<0.001
morphotype C	0.618	0.354	0.352	0.327		0.667	<0.001
morphotype D	0.630	0.379	0.379	0.377	0.000	0.269	<0.001
<i>L. yunnanensis</i>	0.819	0.688	0.694	0.768	0.303	0.269	<0.001

Corresponding P -values are above the diagonal. Significant comparisons are shown in bold ($\alpha = 0.05$).

To investigate this possibility, we increased the K value to three. When this was done, the posterior log probability of the data ($L(K)$) still increased substantially and ΔK value, while suboptimal, was much higher than that obtained for $K \geq 4$, and the increase of $L(K)$ was minor for values greater than $K = 3$ (Figure 1A, B). Taking into account both model fit statistics and *a priori* taxonomic assignments, $K = 3$ appears to better describe genetic clustering patterns in the present analysis, a conclusion that is also supported by the results of principle coordinate analysis (Figure 2) and phylogenetic analysis (Figure 3).

At $K = 3$ (Figure 1D), parental species *L. cyathiceps*, *L. duciformis* and *L. yunnanensis* corresponded to three pure groups, implying no gene introgression occurred among these three species. Hybrid morphotypes A and B showed admixed genetic composition, possessing nearly equal proportion of alleles from *L. cyathiceps* and *L. duciformis* ($q_i = 0.488 \pm 0.016$ for morphotype A; $q_i = 0.488 \pm 0.025$ for morphotype B). Hybrid morphotypes C and D showed similar genetic composition with that inferred using $K = 2$, where most of the individuals had alleles from *L. duciformis* and *L. yunnanensis* in nearly equal proportion ($q_i = 0.501 \pm 0.016$ for morphotype C; $q_i = 0.509 \pm 0.016$ for morphotype D). However, three individuals (H12, S1, and S4) of morphotypes C and D were inferred purely to have alleles from *L. yunnanensis*. In addition, two individuals (H26 and S8) possessed alleles in a higher proportion from *L. duciformis* ($q_i = 0.662$ and 0.795 , respectively), and one individual (S10) had alleles in a higher proportion from *L. yunnanensis* ($q_i = 0.584$). Results obtained in the separate STRUCTURE analyses for the two hybridization groups, were similar to those inferred in the combined analysis of all investigated individuals (Figure 1E, F).

Principle coordinate analysis (PCoA), based on 2,150 SNPs, identified five well-differentiated clusters (Figure 2), corresponding to three parental species *L. cyathiceps*, *L. duciformis* and *L. yunnanensis*, hybrids generated from *L. cyathiceps* \times *L. duciformis* (morphotypes A and B), and hybrids produced from *L. duciformis* \times *L. yunnanensis* (morphotypes C and D). There was relatively higher differentiation among *L. yunnanensis* individuals along the PCoA1 (14.24% of the variation), compared to *L. cyathiceps* and *L. duciformis*. Individuals of morphotypes A and B had very low differentiation and clustered into one group intermediate between *L. cyathiceps* and *L. duciformis*

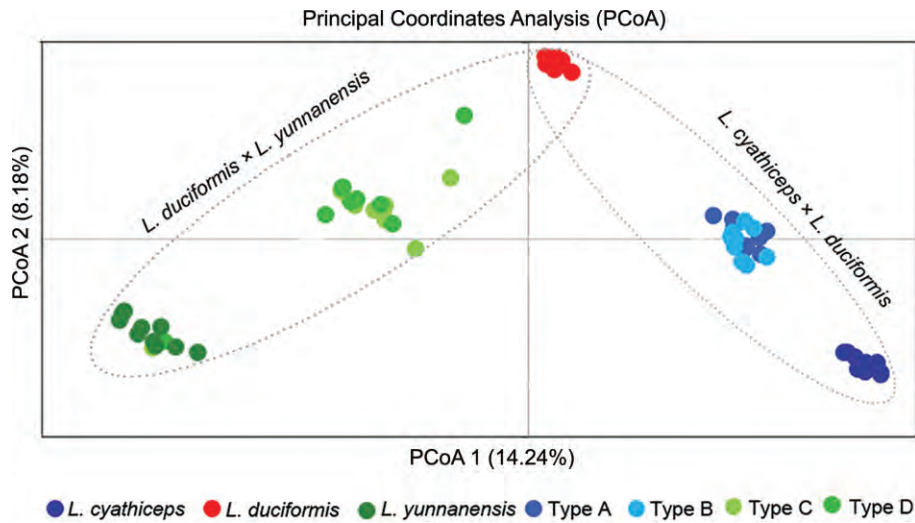


Figure 2. Principal coordinates analysis for 68 individuals based on 2,150 SNPs

The first principal component explains 14.24% of the variance and separates two hybridization groups. The second principle component explains 8.18% of the variance and separates hybrids from parental species in two hybridization groups.

clusters. In concordance with these results, the F_{ST} value between morphotypes A and B indicated no genetic divergence between them ($F_{ST} = 0.000$; Table 1). Among individuals of morphotypes C and D, three individuals (H12, S1, and S4) clustered with *L. yunnanensis*, consistent with STRUCTURE results. The remaining individuals of morphotypes C and D formed a cluster intermediate between *L. duciformis* and *L. yunnanensis* and showed no difference among individuals ($F_{ST} = 0.000$; Table 1).

The best tree from a full maximum-likelihood search showed that the 68 investigated individuals formed three clades with high bootstrap support (Figure 3). Individuals of three parental species, *L. cyathiceps*, *L. duciformis* and *L. yunnanensis*, were recovered as clades; as with other analyses, three individuals from morphotypes C and D (H12, S1, and S4) clustered with *L. yunnanensis*. Moreover, for individuals S8 and H26, inferred to possess more alleles from *L. duciformis* in STRUCTURE analyses, S8 showed a closer relationship

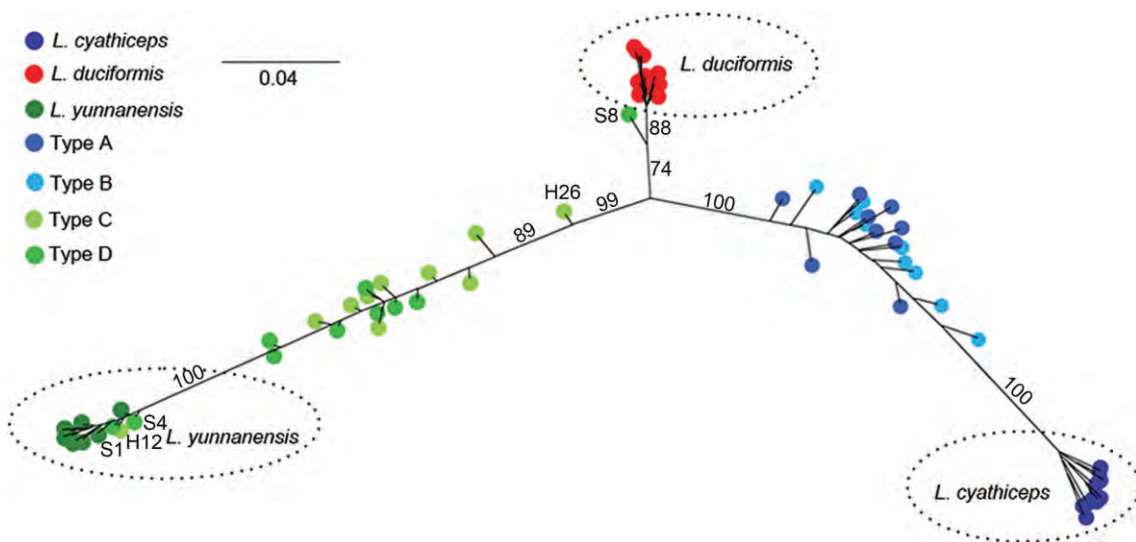


Figure 3. Maximum-likelihood tree based on 2,150 SNPs for 68 individuals using RAxML

Circle colors indicate the three species and four hybrid morphotypes. Numbers above branches are bootstrap values, based on 1000 replicates. Numbers of some individuals mentioned in the text, such as H12, H26, S1, S4, and S8, are shown.

with *L. duciformis* whereas H26 was not recovered as closely related to other individuals of hybrid morphotypes C and D.

Hybrid identification

Posterior probabilities for individuals of two hybridization groups falling into six different hybrid categories were computed, separately, using NewHybrids software. For the hybridization group *L. cyathiceps* × *L. duciformis* (144 SNPs), NewHybrids analysis indicated that individuals of *L. cyathiceps* and *L. duciformis* were assigned to one parent, respectively (Figure 4A), consistent with morphological identification. Hybrid individuals of morphotypes A and B were all robustly assigned to the F₁ class ($P = 1.000$). For the *L. duciformis* × *L. yunnanensis* hybridization group, individuals of *L. duciformis* and *L. yunnanensis* were assigned to one parent, respectively (Figure 4B), in accordance with morphological identification. For hybrids of morphotypes C and D, most individuals (14 of 20) were assigned to F₁ generation. Individuals H26 and S8 were assigned as backcross hybrids to *L. duciformis*, and individual S10 was assigned to backcross hybrid to *L. yunnanensis*, consistent with STRUCTURE results. Three individuals

(H12, S1, and S4) were assigned to the parent *L. yunnanensis*, also consistent with the results of STRUCTURE analysis, PCoA and phylogenetic analysis. All these individuals were assigned with robust probabilities ($P = 1.000$).

Furthermore, the genotype assignment to 12 hybrid categories after three generations of hybridization for two hybridization groups was in accordance with the assigning results above (Figure S1), indicating that hybrid individuals are restricted to F₁ generation and some first-generation backcrosses and there are no later-generation hybrids. The absence of introgression between parental species was also supported by the vast fixed differences between them. There are 82.2% of alleles present in *L. cyathiceps* individuals but absent in *L. duciformis* individuals, and vice versa, whereas 91.7% of species-specific alleles are present in *L. duciformis* and *L. yunnanensis* individuals.

Hybrid swarm simulation

Using STRUCTURE, the reassignment of simulated individuals for two hybridization groups *L. cyathiceps* × *L. duciformis* (Figure S2A) and *L. duciformis* ×

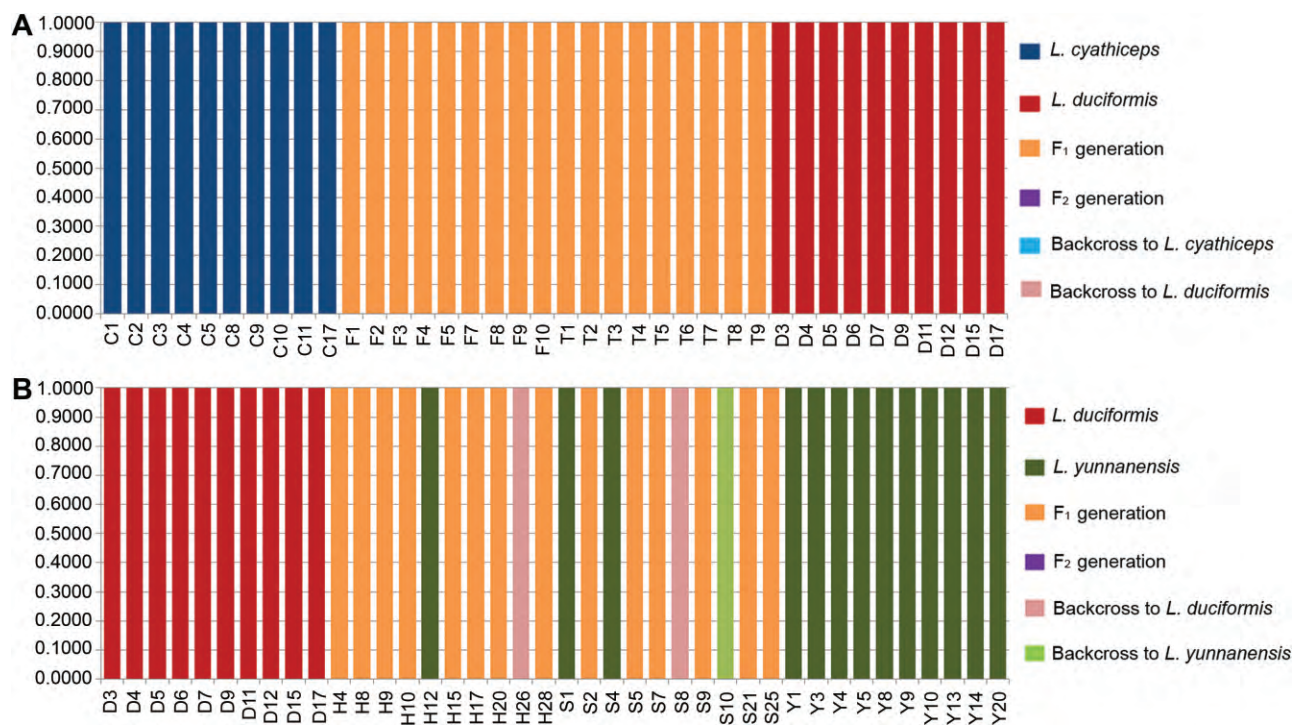


Figure 4. Genotype class assignment by using NewHybrids program

(A) Assignment of 38 individuals in hybridization group *L. cyathiceps* × *L. duciformis* using 144 SNPs. (B) Assignment of 40 individuals in hybridization group *L. duciformis* × *L. yunnanensis* using 349 SNPs.

L. yunnanensis (Figure S2B) both obtained similar results to those observed in the simulation of Pujolar et al. (2014). Compared to 156 SNPs used in the simulation for *L. cyathiceps* × *L. duciformis*, simulation for *L. duciformis* × *L. yunnanensis* with 1,447 SNPs showed greater accuracy in average admixture proportion values. For example, pure *L. duciformis* and pure *L. yunnanensis* showed admixture proportion values of 1.0000 and 0.0000, respectively. Although F₁ and F₂ hybrids, and second-generation backcrosses, such as BLD × F₁, BLY × LD and BLD × LY, BLY × F₁, were indistinguishable with approximately equivalent admixture proportion values, most hybrid categories could be distinguished in STRUCTURE, indicating our marker loci showed sufficient power to discriminate hybrids.

Using NewHybrids, we classified data for simulated individuals to six and 12 categories after two and three generations of hybridization for two hybridization groups. For *L. cyathiceps* × *L. duciformis* and *L. duciformis* × *L. yunnanensis*, all 300 simulated individuals of six categories for each hybridization group were assigned correctly (Table 2; Figure S3). For the case of 600 simulated individuals in 12 hybrid categories after three generations of hybridization, simulated hybrid swarms of *L. cyathiceps* × *L. duciformis* showed decreased assignment accuracy; 470 individuals (78.3%) were assigned correctly, 95 individuals (15.8%) were not assigned to a single category and 35 individuals (5.8%) were assigned to an incorrect category (Table 2; Figure S4A). For the simulated swarms of *L. duciformis* × *L. yunnanensis*, the larger SNP dataset exhibited greater assignment accuracy. In the case of 600 simulated individuals, 565 individuals (94.2%) were assigned to the correct category, 20 individuals (3.3%) were unassigned and 15 individuals (2.5%) were assigned incorrectly (Table 2; Figure S4B). In general, our genetic markers showed sufficient power to accurately assign early generation hybrids.

DISCUSSION

SNPs derived from RAD sequencing for detecting hybridization

The STRUCTURE analysis, PCoA, and ML tree based on genome-wide SNPs confirmed the occurrence of natural hybridization between the species pairs *L. cyathiceps* × *L. duciformis* and *L. duciformis* × *L. yunnanensis*. Simulation studies indicated the SNP dataset we collected had sufficient discriminatory power for accurate hybrid detection, consistent with previous researches demonstrating that SNP markers from RAD sequencing are effective markers for hybridization studies (Pujolar et al. 2014; Capblancq et al. 2015; Bernal et al. 2017). The hybrid identification efficiency of two Bayesian methods implemented in STRUCTURE and NewHybrids also depends on the number and the level of genetic divergence of loci (Vähä and Primmer 2006), as indicated in our results by the increase in accuracy of simulated hybrid assignment in STRUCTURE and NewHybrids for the larger SNP dataset recovered in the species pair *L. duciformis*/*L. yunnanensis* as compared to *L. cyathiceps*/*L. duciformis*.

Although the number and genetic characteristics of genetic markers may have an impact on the accuracy of hybrid identification (Zohren et al. 2016), most studies on natural hybridization involving the comparison of RAD SNPs with alternative genetic markers show congruent results with higher resolution using RAD data (Beacham et al. 2005; Bradbury et al. 2015). This study highlights the effectiveness of RAD sequencing in hybridization studies and the genome-wide SNPs providing a valuable tool for future genomic studies of *Ligularia*.

Outcomes of natural hybridization and maintenance of species boundaries

RAD sequencing data indicated that the species *L. cyathiceps*, *L. duciformis* and *L. yunnanensis* formed

Table 2. NewHybrids assignment for simulated hybrids of two hybridization groups *L. cyathiceps* × *L. duciformis* and *L. duciformis* × *L. yunnanensis*

Hybrid group	Number of SNPs used	Number of hybrid categories used	Assigned to correct category	Unassigned	Incorrect category
<i>L. cyathiceps</i> × <i>L. duciformis</i>	144	6	300 (100%)	0	0
<i>L. cyathiceps</i> × <i>L. duciformis</i>	144	12	470 (78.3%)	95 (15.8%)	35 (5.8%)
<i>L. duciformis</i> × <i>L. yunnanensis</i>	349	6	300 (100%)	0	0
<i>L. duciformis</i> × <i>L. yunnanensis</i>	349	12	565 (94.2%)	20 (3.3%)	15 (2.5%)

two hybridization groups at a sympatric site, where *L. duciformis* hybridized with *L. cyathiceps* and *L. yunnanensis*, respectively. Hybrids of *L. cyathiceps* and *L. duciformis* were limited to F₁ generation and no introgression was found between two species, suggesting strong reproductive isolation between *L. cyathiceps* and *L. duciformis*. The sterility of F₁s may be the reason for no occurrence of later-generation hybrids. Hybrid sterility seems to be one important reason for the reproductive isolation of hybridizing species in *Ligularia* and it is not rare in other cases of hybridization in *Ligularia*. For example, *Ligularia* × *maoniushanensis*, a natural hybrid between *Ligularia paradoxa* Handel-Mazzetti and *L. duciformis*, shows abnormal meiotic behavior and inviability of seeds; ISSR markers indicate all observed hybrids are F₁ generation, implying the sterility of F₁s (Pan et al. 2008). In addition, hybrids of *Ligularia nelumbifolia* (Bureau & Franchet) Handel-Mazzetti and *Ligularia subspicata* (Bureau & Franchet) Handel-Mazzetti are mostly composed of F₁s and their seed germination rate is extremely low (Yu 2010). The theoretical expectation of F₁s is 25% when parental individuals practice a selfing rate of 0.5 and outcross randomly with respect to taxon, thus the observation of few F₁ hybrids is evidence for the existence of strong reproductive isolation (Twyford et al. 2015). This is also the case in the hybridization between *L. cyathiceps* and *L. duciformis*, where hybrids of morphotypes A and B were rare when compared to abundant parental individuals.

Although some hybrids between *L. duciformis* and *L. yunnanensis* were backcrosses, the F₁ generation still comprised the majority of hybrids and no gene introgression was observed between parental species, suggesting two primarily reproductively isolated species. Nevertheless, three individuals (H12, S1, and S4), identified as hybrids using morphological data, were identified to be *L. yunnanensis* individuals according to the results of STRUCTURE analysis, ML tree, PCoA and NewHybrids analyses. The hypothesis that three individuals are generated by repetitive backcrosses with *L. yunnanensis* can likely be excluded, since in a genome-wide dataset it is unlikely that we would obtain loci deriving from only one of the parental species for these three individuals. Take dataset C (1,447 SNPs, *L. duciformis* × *L. yunnanensis*) for example, the 1,447 genome-wide SNP loci of individuals H12, S1 and S4 (if they are backcross individuals) would not be totally

consistent with those of *L. yunnanensis* without any polymorphisms. This unexpected result may stem from incorrect morphological identification, which may also imply that some hybrids and parental individuals have indistinguishable morphological traits. Additionally, *L. yunnanensis* prefers shady and humid environments, and they occupy habitats under the trees at the study site. However, some individuals (*L. yunnanensis* or hybrids) are exposed to sunlight, due to constant tree felling, and show unique morphological characteristics, such as purple stems and inflorescences, thus making it difficult to distinguish *L. yunnanensis* from hybrids.

Hybridization and introgression may occur between isolated species with strong sterility barriers, given that the accumulation of sufficient F₁s may bypass the sterility bottleneck resulting in the formation of backcrossed individuals (Yatabe et al. 2007; Twyford et al. 2015). Moreover, human disturbance and the resultant establishment of intermediate “hybridized habitats” can promote opportunities for hybridization (Anderson and Stebbins 1954; Bleeker and Hurka 2001; Lamont et al. 2003; Ma et al. 2014) and affect the presence of hybrid classes by favoring segregating hybrid derivatives (Abbott 1992; Rieseberg and Carney 1998; Milne et al. 2003). Hence, backcross individuals may have been generated by hybridization between *L. duciformis* and *L. yunnanensis* under circumstances of habitat disturbance.

This study demonstrates that natural hybridization between sympatric *Ligularia* species in the eastern QTP and adjacent areas is mostly limited to F₁ generation, without the occurrence of gene introgression. Although species boundaries can be maintained through reproductive isolation with the occurrence of introgressive hybridization (Christe et al. 2016; Xie et al. 2017), no gene flow between hybridizing species with F₁s contributes to the maintenance of species integrity. Moreover, even under the circumstances that human disturbance may promote the occurrence of hybridization, these three *Ligularia* species are still reproductively isolated and maintain species boundaries. Hence, the presence of reproductive isolation may be a key mechanism to maintain species diversity of *Ligularia* in the eastern QTP and adjacent areas in the face of opportunities for hybridization. In addition, natural hybridization among multiple species (>2) may represent the most common phenomenon in the wild; studies on reproductive isolating mechanisms in these scenarios are significant

to understanding the maintenance of species boundaries among sympatric species.

MATERIALS AND METHODS

Study system

We selected three species, *L. cyathiceps*, *L. duciformis* and *Ligularia yunnanensis*, as the study taxa. *L. cyathiceps*, belonging to series *Ligularia* in section *Ligularia*, possesses racemose inflorescence and ray florets in the margin of the capitulum (Liu 1989). The other two species, *L. duciformis* and *L. yunnanensis*, belonging to series *Retuase* in section *Corymbosae*, have capitula in corymb comprised of tubular florets (Liu 1989). *L. cyathiceps* and *L. yunnanensis* are endemic to the eastern QTP and adjacent areas, while *L. duciformis* is distributed more widely in China (Liu and Illarionova 2011). According to specimen records in the Chinese Virtual Herbarium (CVH, <http://www.cvh.ac.cn/>) and our field investigation, only one sympatric area was identified for the three species in Northwest Yunnan (Tianchi, Zhongdian, 27.617087, 99.747011, 3,901 m; Table S3), a site that has experienced human disturbance, including grazing, tree felling, and tourism.

For the three *Ligularia* species, molecular evidence from nuclear loci has demonstrated the natural occurrence of two hybrid combinations, *L. cyathiceps* × *L. duciformis* and *L. duciformis* × *L. yunnanensis* (NN Zhang, unpubl. data). Morphologically intermediate individuals, here designated as morphotypes A and B, are produced by *L. cyathiceps* × *L. duciformis*; the primary morphological difference between morphotypes A and B is the presence of ray florets in the margin of capitula in morphotype A. Morphologically intermediate individuals, here designated as morphotypes C and D, are from *L. duciformis* × *L. yunnanensis*; morphotype C is morphologically distinct from morphotype D in having inflorescences covered with a short brown indumentum. In the study site, *L. cyathiceps* and *L. duciformis* grow abundantly in sunny and disturbed habitats, such as hillsides and roadsides. In contrast, *L. yunnanensis* prefers shady and humid environments, and occupies relatively undisturbed habitats. Moreover, with the exception that the population of *L. yunnanensis* plants is small, the individuals of the other two parental species (*L. cyathiceps* and *L. duciformis*) greatly outnumber putative hybrids.

ddRAD sequencing

A total of 68 individuals (10 *L. cyathiceps* [C1-10], 10 *L. duciformis* [D1-10], 10 *L. yunnanensis* [Y1-10], nine morphotype A [F1-9], nine morphotype B [T1-9], 10 morphotype C [H1-10], and 10 morphotype D [S1-10]) were used for ddRAD sequencing. The voucher specimens were deposited in the herbarium of the Kunming Institute of Botany, Chinese Academy of Sciences (KUN). High-quality genomic DNA was extracted from silica-dried leaves, with a modified CTAB method (Doyle 1991). Library construction followed the protocol of Peterson et al. (2012); ddRAD sequencing was carried out by Biozeron Technologies Corporation (Shanghai, China). Genomic DNA of each investigated individual was digested with two restriction enzymes: *EcoRI* (cut site 5'-GAATTC-3') and *NlaIII* (cut site 5'-CATG-3'); fragments were size-selected to a range of 400–600 bp. Library sequencing was performed on an Illumina HiSeq 4000 System (San Diego, CA, USA) with 150 bp paired-end reads. Libraries were pooled to a target of 5 Gb raw data per individual (approximately equal to the genome size of *Ligularia fischeri* (Ledebour) Turczaninow; JJ Yu, unpubl. data).

The *de novo* pipeline in STACKS v. 1.42 (Catchen et al. 2011, 2013) was implemented to identify SNPs. Raw reads were de-multiplexed and quality-filtered using *process_radtags*. During the process, all reads were trimmed to 140 bp to remove low quality bases at the end of sequences and reads with ambiguous or low-quality bases (below a Phred score of Q10) were discarded. Read quality and GC-content were assessed using FastQC v. 0.11.4 (Andrews 2015). For each sample, *ustacks* with a maximum likelihood framework was used to merge short reads into loci. The deleveraging algorithm (d) and removal algorithm (r) were implemented to remove highly repetitive stacks and over-merged tags. The minimum depth of coverage required to create a stack (m) was set to five, and the maximum nucleotide mismatch allowed between stacks (M) was two. The catalog of loci was built using *cstacks* with the mismatch number allowed between sample tags (n) of five. Finally, the loci of each sample were matched against the catalog to determine alleles using *sstacks*.

Using *populations* in the STACKS pipeline, we treated three parental species (*L. cyathiceps*, *L. duciformis*, and *L. yunnanensis*) and four hybrid morphotypes (A, B, C, and D) as seven different populations. To filter loci, we set parameters in three

different permutations described as follows, resulting in three final datasets. First, for all the 68 investigated individuals, we required for locus retention that a locus be present in each of the seven populations ($p = 7$) and in at least 50% of the individuals for each of these populations ($r = 0.5$), with minimum coverage (m) of five; hereafter this dataset is referred to as dataset A. Due to software limitations present in NewHybrids when using a large number of SNPs, we filtered loci and reduced SNPs using more stringent parameters ($r = 0.7$, $p = 4$, $m = 10$, $MAF \geq 0.3$) following Bell et al. (2015), producing two further datasets intended specifically for the STRUCTURE and NewHybrids analyses. For the 38 individuals in the *L. cyathiceps* × *L. duciformis* hybridization group (10 *L. cyathiceps*, 10 *L. duciformis*, nine morphotype A, nine morphotype B), we retained only those loci that occurred in at least 70% of the individuals ($r = 0.7$) in each of the four populations ($p = 4$), with coverage of $10\times$ or greater ($m = 10$) and a minor allele frequency (MAF) of 0.3 or greater; this dataset is hereafter referred to as dataset B. The same parameters were set for the 40 individuals in *L. duciformis* × *L. yunnanensis* hybridization group (10 *L. duciformis*, 10 *L. yunnanensis*, 10 morphotype C, 10 morphotype D); this dataset is hereafter referred to as dataset C. Given these filtering conditions, appropriate genotype outputs were generated for these three datasets by implementing the ‘-write_random_snp’ option (i.e., producing one randomly selected SNP from each locus). Datasets A, B, and C consisted of 2,150, 156 and 1,447 SNP loci, respectively. Further format conversion for downstream analyses was performed using the program PGDSpider v. 2.1.0.3 (Lischer and Excoffier 2012).

ddRAD-seq data analysis

Genetic admixture of the 68 individuals was assessed using the Bayesian approach implemented in STRUCTURE v. 2.3.4 (Pritchard et al. 2000) using 2,150 SNPs. The analysis was run with an initial burn-in of 100,000 replicates followed by 100,000 replicates without population priors, using the admixture model and correlated allele frequencies. These parameters were repeated in 10 independent analyses for each K (number of genetic clusters) from 1 to 10. The most likely number of genetic clusters (K) was chosen using the delta K method as well as visual inspection of the change in the $\ln P(D)$ of each model in STRUCTURE HARVESTER (Earl 2012). CLUMPP v. 1.1.2 (Jakobsson and Rosenberg 2007)

was used to average admixture proportions of 10 replicate runs for each K value, and admixture plots were created from this using DISTRUCT v. 1.1 (Rosenberg 2004). In addition, admixture analyses with datasets B and C were conducted with the same parameter settings in STRUCTURE.

A principle coordinate analysis was also conducted for the 68 individual dataset using GenAlex v. 6.5 (Peakall and Smouse 2012) with 2,150 SNPs. The genetic distance matrix was calculated treating the data as codominant and using the covariance-standardized method. Pairwise F_{ST} values for three parental species and four morphotypes of hybrids based on 2,150 SNPs were calculated using Arlequin v. 3.5 (Schneider et al. 2000; Excoffier and Lischer 2010).

To assess phylogenetic relationships among three *Ligularia* species and investigated hybrids, a phylogenetic tree was constructed with 2,150 SNPs using maximum-likelihood (ML), which was implemented in RAxML-VI-HPC (Stamatakis 2006) and its graphical interface raxmlGUI v. 1.3 (Silvestro and Michalak 2012). We used the GTR- Γ nucleotide substitution model, assessing branch support with 1,000 bootstrap replicates (rapid bootstrap algorithm).

NewHybrids v. 1.1 (Anderson and Thompson 2002) was used to compute the posterior probability of individual membership among six genotypes (P_1 , P_2 , F_1 , F_2 , BC to P_1 , BC to P_2). This program uses dominant markers that are not fixed in either parent species, and assumes loci are not under selection. Therefore, BAYESCAN v. 2.1 (Foll and Gaggiotti 2008) was used to detect loci under selection in 2,150 SNPs and it was implemented with default priors and parameter settings (39 SNP loci were inferred to be under selection). We removed SNP loci under selection in datasets B and C, leaving 144 SNPs to be used for the NewHybrids calculation for the *L. cyathiceps* × *L. duciformis* hybridization group. Because dataset C was still too large after removing loci under selection for NewHybrids, 349 SNPs were selected randomly among those remaining for the *L. duciformis* × *L. yunnanensis* hybridization group. NewHybrids was run using 100,000 burn-in iterations and 100,000 subsequent Markov Chain Monte Carlo iterations with default priors for allele frequencies and mixing proportions. Moreover, the same assignment was obtained for individuals in the *L. duciformis* × *L. yunnanensis* hybridization group based on different randomly selected SNP loci. To verify if there are some

later-generation hybrids, individuals in hybridization groups *L. cyathiceps* × *L. duciformis* and *L. duciformis* × *L. yunnanensis* were assigned to 12 hybrid categories after three generations of hybridization using NewHybrids with the same parameters.

Hybrid swarm simulation

The discriminatory power of markers used in STRUCTURE and NewHybrids for admixture analysis and hybrid class assignment was assessed by simulating SNP data from hypothetical hybrid swarms. The two species pair datasets used in STRUCTURE and NewHybrids were used for separate simulations using HYBRIDLAB v. 1.0 (Nielsen et al. 2006), a program for generating simulated hybrids from population samples. First, to simulate STRUCTURE analyses, HYBRIDLAB was used to generate 50 genotypes for each of 12 categories after three generations of simulated hybridization for both hybridization groups (*L. cyathiceps* × *L. duciformis*, 156 SNPs; *L. duciformis* × *L. yunnanensis*, 1,447 SNPs) for a total of 600 simulated genotypes per hybridization group. Then simulated hybrid data for the two hybridization groups were analyzed in STRUCTURE separately with $K=2$, using the admixture model and no prior population information. Finally, we followed the approach of Pujolar et al. (2014) to count the average admixture proportion values and corresponding standard deviation values for each category and plotted figures accordingly. Second, we repeated these analyses using the reduced datasets generated for hybrid classification. Both hybridization groups (*L. cyathiceps* × *L. duciformis*, 144 SNPs; *L. duciformis* × *L. yunnanensis*, 349 SNPs) used in NewHybrids analysis were used to generate 50 genotypes of each of six and 12 categories per hybridization group after two and three generations of hybridization, respectively. For the resultant simulated datasets, individuals were reassigned to the most probable category using NewHybrids.

Take *L. cyathiceps* × *L. duciformis* for example, the 12 categories after three generations of hybridization (following Pujolar et al. 2014) are: (i) *L. cyathiceps* (LC), (ii) *L. duciformis* (LD), (iii) F_1 hybrid (F_1), (iv) F_2 hybrid (F_2), (v) first-generation backcross *L. cyathiceps* × F_1 hybrid (BLC), (vi) first-generation backcross *L. duciformis* × F_1 hybrid (BLD), (vii) second-generation backcross between first-generation backcross (*L. cyathiceps* × F_1 hybrid) and *L. cyathiceps* (BLC × LC),

(viii) second-generation backcross between first-generation backcross (*L. cyathiceps* × F_1 hybrid) and *L. duciformis* (BLC × LD), (ix) second-generation backcross between first-generation backcross (*L. cyathiceps* × F_1 hybrid) and F_1 hybrid (BLC × F_1), (x) second-generation backcross between first-generation backcross (*L. duciformis* × F_1 hybrid) and *L. cyathiceps* (BLD × LC), (xi) second-generation backcross between first-generation backcross (*L. duciformis* × F_1 hybrid) and *L. duciformis* (BLD × LD), (xii) second-generation backcross between first-generation backcross (*L. duciformis* × F_1 hybrid) and F_1 hybrid (BLD × F_1).

ACKNOWLEDGEMENTS

We thank Guoqian Yang and Yujuan Zhao for their help with data analysis. This research was supported by the National Natural Science Foundation of China (No. 31470336 to X.G. and 31600178 to J.Y.).

AUTHOR CONTRIBUTIONS

X.G. designed the research; X.G. and J.Y. collected the plant materials; N.Z. performed the experiments, analyzed the data and wrote the manuscript; Y.M., R.A.F., Y.P., J.Y. and X.G. revised the manuscript.

REFERENCES

- Arnold ML (1997) *Natural Hybridization and Evolution*. Oxford University Press, New York
- Abbott RJ (1992) Plant invasions, interspecific hybridization and the evolution of new plant taxa. **Trends Ecol Evol** 7: 401–405
- Anderson E, Stebbins JG (1954) Hybridization as an evolutionary stimulus. **Evolution** 8: 378–388
- Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. **Genetics** 160: 1217–1229
- Andrews S (2015) Fast QC: A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (last accessed 9 August 2017)
- Beacham TD, Hay DE, Le KD (2005) Population structure and stock identification of eulachon (*Thaleichthys pacificus*), an anadromous smelt, in the Pacific Northwest. **Mar Biotechnol** 7: 363–372

- Bell RC, Drewes RC, Zamudio KR (2015) Reed frog diversification in the Gulf of Guinea: Overseas dispersal, the progression rule, and in situ speciation. **Evolution** 69: 904–915
- Bernal MA, Gaither MR, Simison WB, Rocha LA (2017) Introgression and selection shaped the evolutionary history of sympatric sister-species of coral reef fishes (genus: *Haemulon*). **Mol Ecol** 26: 639–652
- Bleeker W, Hurka H (2001) Introgressive hybridization in *Rorippa* (Brassicaceae): Gene flow and its consequences in natural and anthropogenic habitats. **Mol Ecol** 10: 2013–2022
- Bradbury IR, Hamilton LC, Dempson B, Robertson MJ, Bourret V, Bernatchez L, Verspoor E (2015) Transatlantic secondary contact in Atlantic Salmon, comparing microsatellites, a single nucleotide polymorphism array and restriction-site associated DNA sequencing for the resolution of complex spatial structure. **Mol Ecol** 24: 5130–5144
- Capblancq T, Després L, Rioux D, Mavárez J (2015) Hybridization promotes speciation in *Coenonympha* butterflies. **Mol Ecol** 24: 6209–6222
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: An analysis tool set for population genomics. **Mol Ecol** 22: 3124–3140
- Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: Building and genotyping loci *de novo* from short-read sequences. **G3-Genes Genom Genet** 1: 171–182
- Christe C, Stölting KN, Bresadola L, Fussi B, Heinze B, Wegmann D, Lexer C (2016) Selection against recombinant hybrids maintains reproductive isolation in hybridizing *Populus* species despite F1 fertility and recurrent gene flow. **Mol Ecol** 25: 2482–2498
- Coyne JA (1994) Ernst Mayr and the origin of species. **Evolution** 48: 19–30
- Cromie GA, Hyma KE, Ludlow CL, Garmendia-Torres C, Gilbert TL, May P, Huang AA, Dudley AM, Fay JC (2013) Genomic sequence diversity and population structure of *Saccharomyces cerevisiae* assessed by RAD-seq. **G3-Genes Genom Genet** 3: 2163–2171
- Dobzhansky TH (1936) Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. **Genetics** 21: 113
- Doyle J (1991) DNA protocols for plants-CTAB total DNA isolation. In: Hewitt GM, ed. *Molecular Techniques in Taxonomy*. Springer, Berlin. pp. 283–293
- Earl DA (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. **Conserv Genet Resour** 4: 359–361
- Eaton DA, Ree RH (2013) Inferring phylogeny and introgression using RADseq data: An example from flowering plants (*Pedicularis*: Orobanchaceae). **Syst Biol** 62: 689–706
- Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. **Mol Ecol Resour** 10: 564–567
- Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. **Genetics** 180: 977–993
- Gong X (2005) *Study on a Natural Hybrid of Ligularia and Artificial Hybridization of Magnoliaceae*. Sun Yat-sen University, Guangzhou
- Hipp AL, Eaton DA, Cavender-Bares J, Fitzek E, Nipper R, Manos PS (2014) A framework phylogeny of the American oak clade based on sequenced RAD data. **PLoS ONE** 9: e93975
- Jakobsson M, Rosenberg NA (2007) CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. **Bioinformatics** 23: 1801–1806
- Lamont BB, He T, Enright NJ, Krauss SL, Miller BP (2003) Anthropogenic disturbance promotes hybridization between *Banksia* species by altering their biology. **J Evol Biol** 16: 551–557
- Levin D (2000) *The Origin, Expansion and Demise of Plant Species*. Oxford University Press, New York
- Linder CR, Rieseberg LH (2004) Reconstructing patterns of reticulate evolution in plants. **Am J Bot** 91: 1700–1708
- Lischer H, Excoffier L (2012) PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. **Bioinformatics** 28: 298–299
- Liu JQ, Wang YJ, Wang AL, Hideaki O, Abbott RJ (2006) Radiation and diversification within the *Ligularia-Cremnathodium-Parasenecio* complex (Asteraceae) triggered by uplift of the Qinghai-Tibetan Plateau. **Mol Phylogen Evol** 38: 31–49
- Liu JQ (2004) Uniformity of karyotypes in *Ligularia* (Asteraceae: Senecioneae), a highly diversified genus of the eastern Qinghai-Tibet Plateau highlands and adjacent areas. **Bot J Linn Soc** 144: 329–342
- Liu SW (1989) *Flora Reipublicae Popularis Sinicae*. Science Press, Beijing
- Liu SW, Deng DS, Liu JQ (1994) Origin, evolution and distribution of *Ligularia* Cass.(Compositae). **Acta Phytotaxon Sin** 32: 514–524
- Liu SW, Illarionova I (2011) *Flora of China*. Science Press & St. Louis: Missouri Botanical Garden Press, Beijing. pp. 371–544
- Ma YP, Xie WJ, Tian XL, Sun WB, Wu ZK, Milne R (2014) Unidirectional hybridization and reproductive barriers between two heterostylous primrose species in north-west Yunnan, China. **Ann Bot** 113: 763–775
- Milne RI, Abbott RJ (2008) Reproductive isolation among two interfertile *Rhododendron* species: Low frequency of post-F1 hybrid genotypes in alpine hybrid zones. **Mol Ecol** 17: 1108–1121
- Milne RI, Terzioglu S, Abbott RJ (2003) A hybrid zone dominated by fertile F1s: Maintenance of species barriers in *Rhododendron*. **Mol Ecol** 12: 2719–2729

- Muller HJ (1939) Reversibility in evolution considered from the standpoint of genetics. **Biol Rev** 14: 261–280
- Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. **Nature** 403: 853–858
- Nielsen EE, Bach LA, Kotlicki P (2006) HYBRIDLAB (version 1.0): A program for generating simulated hybrids from population samples. **Mol Ecol Notes** 6: 971–973
- Otto SP, Whitton J (2000) Polyploid incidence and evolution. **Annu Rev Genet** 24: 401–437
- Pan YZ, Shi SH, Gong X, Kuroda C (2008) A natural hybrid between *Ligularia paradoxa* and *L. duciformis* (Asteraceae, Senecioneae) from Yunnan, China. **Ann Mo Bot Gard** 95: 487–494
- Peakall R, Smouse PE (2012) GenAEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. **Bioinformatics** 28: 2537–2539
- Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. **PLoS ONE** 7: e37135
- Pinheiro F, Cozzolino S, Cafasso D, Cardoso-Gustavson P, Suzuki RM, Palma-Silva C (2016) Strong but permeable barriers to gene exchange between sister species of *Epidendrum*. **Am J Bot** 103: 1472–1482
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. **Genetics** 155: 945–959
- Pujolar JM, Jacobsen MW, Als TD, Frydenberg J, Magnussen E, Jónsson B, Jiang X, Cheng L, Bekkevold D, Maes GE, Bernatchez L, Hansen MM (2014) Assessing patterns of hybridization between North Atlantic eels using diagnostic single-nucleotide polymorphisms. **Heredity** 112: 627–637
- Reitzel AM, Herrera S, Layden MJ, Martindale MQ, Shank TM (2013) Going where traditional markers have not gone before: Utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. **Mol Ecol** 22: 2953–2970
- Ren C (2012) *Cytotaxonomy, Molecular Systematics, and Biogeography of the Ligularia-Cremanthodium-Parasenecio Complex (Asteraceae-Senecioneae)*. South China Botanical Garden, Chinese Academy of Sciences, Guangzhou
- Rieseberg LH, Carney SE (1998) Plant hybridization. **New Phytol** 140: 599–624
- Rieseberg LH, Whitton J, Gardner K (1999) Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. **Genetics** 152: 713–727
- Rosenberg NA (2004) DISTRUCT: A program for the graphical display of population structure. **Mol Ecol Notes** 4: 137–138
- Schneider S, Roessli D, Excoffier L (2000) Arlequin: A software for population genetics data analysis. **User Manual Ver** 2: 2496–2497
- Silvestro D, Michalak I (2012) raxmlGUI: A graphical front-end for RAxML. **Org Divers Evol** 12: 335–337
- Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. **Bioinformatics** 22: 2688–2690
- Stebbins GL (1950) *Variation and Evolution in Plants*. Columbia University Press, New York
- Stebbins GL (1958) The inviability, weakness, and sterility of interspecific hybrids. **Adv Genet** 9: 147–215
- Stölting KN, Nipper R, Lindtke D, Caseys C, Waeber S, Castiglione S, Lexer C (2013) Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. **Mol Ecol** 22: 842–855
- Surget-Groba Y, Kay KM (2013) Restricted gene flow within and between rapidly diverging Neotropical plant species. **Mol Ecol** 22: 4931–4942
- Twyford AD, Kidner CA, Ennos RA (2015) Maintenance of species boundaries in a Neotropical radiation of *Begonia*. **Mol Ecol** 24: 4982–4993
- Vähä JP, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. **Mol Ecol** 15: 63–72
- White MJD (1978) *Modes of Speciation*. CA: W. H. Freeman, San Francisco
- Wu CI (2001) The genic view of the process of speciation. **J Evol Biol** 14: 851–865
- Xie YP, Zhu XF, Ma YP, Zhao JL, Li L, Li QJ (2017) Natural hybridization and reproductive isolation between two *Primula* species. **J Integr Plant Biol** 59: 526–530
- Yatabe Y, Kane NC, Scotti-Saintagne C, Rieseberg LH (2007) Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. **Genetics** 175: 1883–1893
- Yu JJ (2010) *Hybrid Speciation of Ligularia Cass. -Natural Hybridization between L. subspicata and L. nelumbifolia in Yunnan, China*. Kunming Institute of Botany, Chinese Academy of Sciences, Kunming
- Yu JJ, Kuroda C, Gong X (2014a) Natural hybridization and introgression between *Ligularia cymbulifera* and *L. tongolensis* (Asteraceae, Senecioneae) in four different locations. **PLoS ONE** 9: e115167
- Yu JJ, Li P, Pan YZ, Gong X (2014b) Natural Hybrids between *Ligularia vellerea* and *L. subspicata* (Asteraceae: Senecioneae). **Plant Divers Resour** 36: 219–226
- Yu JJ, Kuroda C, Gong X (2011) Natural hybridization and introgression in sympatric *Ligularia* species (Asteraceae, Senecioneae). **J Syst Evol** 49: 438–448
- Zha HG, Milne RI, Sun H (2009) Asymmetric hybridization in *Rhododendron agastum*: A hybrid taxon comprising mainly F1s in Yunnan, China. **Ann Bot** 105: 89–100
- Zhang RS, Liu T, Wu W, Li YQ, Chao LF, Huang LS, Huang YL, Shi SH, Zhou RC (2013) Molecular evidence for natural hybridization in the mangrove fern genus *Acrostichum*. **BMC Plant Biol** 13: 74
- Zohren J, Wang N, Kardailsky I, Borrell JS, Joecker A, Nichols RA, Buggs RJ (2016) Unidirectional diploid-tetraploid

introgression among British birch trees with shifting ranges shown by restriction site-associated markers. *Mol Ecol* 25: 2413–2426

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article: <http://onlinelibrary.wiley.com/doi/10.1111/jipb.12674/supinfo>

Figure S1. Genotype class assignment to twelve hybrid categories after three generations of hybridization by using NewHybrids program

(A) Assignment of 38 individuals in hybridization group *L. cyathiceps* × *L. duciformis* using 144 SNPs. (B) Assignment of 40 individuals in hybridization group *L. duciformis* × *L. yunnanensis* using 349 SNPs.

Figure S2. Admixture analysis of 600 simulated individuals for each hybridization group in STRUCTURE. Twelve hybrid categories after three generations of hybridization were simulated. Fifty individuals were simulated for each category. Bottom numbers are the average admixture proportion values and corresponding standard deviation values for each category.

(A) *L. cyathiceps* × *L. duciformis*. (B) *L. duciformis* × *L. yunnanensis*.

Figure S3. NewHybrids analysis of 300 simulated individuals for each hybridization group

Six hybrid categories after two generations of hybridization were simulated and fifty individuals were simulated for each category. The symbols for each simulated taxa are listed below. (A) *L. cyathiceps* × *L. duciformis*. (B) *L. duciformis* × *L. yunnanensis*.

Figure S4. NewHybrids analysis of 600 simulated individuals for each hybridization group

Twelve hybrid categories after three generations of hybridization were simulated and fifty individuals were simulated for each category. The symbols for each simulated taxa are listed below. (A) *L. cyathiceps* × *L. duciformis*. (B) *L. duciformis* × *L. yunnanensis*.

Table S1. Statistics describing cleaned reads for each investigated individual after quality filtering using *process_radtags*

Table S2. Summary of loci recovered for each investigated sample in *ustacks* and number of matched loci in *sstacks*

Table S3. Geographic distribution of *L. cyathiceps*, *L. duciformis* and *L. yunnanensis* in China