Natural hybridization between two butterfly bushes in Tibet: dominance of F1 hybrids promotes strong reproductive isolation

Rongli Liao
Kunming Institute of Botany Chinese Academy of Sciences https://orcid.org/0000-0002-6695-2739
Weibang Sun (wbsun@mail.kib.ac.cn)
Kunming Institute of Botany, Chinese Academy of Sciences
Yongpeng Ma
Kunming Institute of Botany Chinese Academy of Sciences

Research article

Keywords: Butterfly bushes, Buddleja, Hybridization, F1-dominated hybrids, Reproductive isolation

Posted Date: January 21st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-42245/v3

Version of Record: A version of this preprint was published on March 10th, 2021. See the published version at https://doi.org/10.1186/s12870-021-02909-7.

Abstract

Background: It has been recognized that certain amount of habitat disturbance is a prerequisite for occurrence of natural hybridization, yet we are currently still not aware of any studies exploring hybridization and reproductive barriers to those plants preferably occupying disturbed habitats. *Buddleja* plants (also called butterfly bush) generally grow in disturbed habitat, and several species with hybrid origin only on basis of morphology evidence have been proposed.

Results: In the present study, we test the natural hybridization origin hypothesis of *B. × wardii* in two sympatric populations of three taxa including *B. × wardii* and its parents (*B. alternifolia* and *B. crispa*) plus 4 referenced parental populations, using four nuclear genes and three chloroplast intergenic spacers, as well as with 10 morphological characters. Our results suggest that at both sites *B. × wardii* was likely to be hybrids between *B. alternifolia* and *B. crispa*, and moreover, most of the hybrids examined were confirmed to be F_1s. This was further supported by morphology as no transgressive characters were detected. *B. crispa* was found to be the maternal parent in Bahe (BH) population from the cpDNA. While in the Taji (TJ) population was difficult to distinguish the hybridization direction due to the shared haplotypes of cpDNA between *B. alternifolia* and *B. crispa*, we still predicted the similar unidirectional hybridization pattern due to results from cross-specific pollination treatments which supported the “SI x SC rule”.

Conclusions: Hybrids mainly consisting of F_1s can successfully impede gene flow and thus maintain species boundaries of parental species in its typical distribution of *Buddleja*, i.e. disturbed habitats.

Background

Natural hybridization in plants is a common phenomenon, and is thought to play an important role in plant diversity and speciation [1-4]. Certain amount of habitat disturbance is a prerequisite for occurrence of natural hybridization [5-6], yet we are currently still not aware of any studies exploring hybridization and reproductive barriers to those plants preferably occupying disturbed and/or newly alternated habitats. In the long term, hybrid speciation via homoploidy or polyploidy can be predicted, though the former is still assumed to be very rare in plant [7]. In the short term, many empirical studies have focused on onset of natural hybridization and examined issues like reproductive barriers [8-9], backcrossing and introgression which sometimes involved in transference of adaptation [10-12] or acceleration of extinction by genetic swamp [13-15]. A previously assumed to be rare but recently proved to be common phenomenon is hybrids mainly consisting of F_1s [16-20], due to intrinsic incompatibilities include F_1 hybrid sterility and/or inviability [4,16,21], and extrinsic selection with other genotypes of hybrids being outcompeted due to strong habitat selection [16-17,22-23].

Unlike later generation hybrids, that generally exhibit wide morphological variance because of genetic segregation, F_1 hybrids from particular parental species tend to have similar morphologies due to the complete combination of parental genomes [18]. Because of this, F_1s have been often inaccurately described as new species, especially by taxonomists concerned chiefly with morphology. An example of this is *Rhododendron agastum* from Yunnan, China, which has long be treated as a good species [24]. Recently, however, its hybrid origin has been confirmed, and population studies at the type locality suggest that most hybrids are F_1s [16].

The genus *Buddleja* L. (Scrophulariaceae), comprises more than 100 species and is widely distributed throughout tropical, subtropical and temperate regions of the Americas, Africa and Asia [25-26]. The Sino-Himalayan region is a center of diversity for this genus, with over 75% of the Asian *Buddleja* species occurring in this area [27-28]. Moreover, most *Buddleja* species prefer to grow in disturbed habitats (e.g. sides of roads and on riverbanks), which is typical of pioneer plants [29-31], and some species (e.g. *B. davidii*) can become invasive if introduced to new environments [32-34]. Due to substantial overlaps in distribution and flowering periods, as well as shared pollinators, interspecific hybridization is assumed to be relatively common in the genus [25-26,29]. However, to date, only a single case of natural hybridization has been confirmed using molecular data [24].

Buddleja wardii C. Marquand was originally described as a new species from Tibet by C. Marquand in 1929 [35]. Leeuwenberg [29] subsequently considered *B. wardii* to be a hybrid between the sympatric species *B. crispa* and *B. alternifolia* based on the morphological characteristics of the type specimen and Ludlow c.s. 4098 (BM, E, K), as well as these two species are the only one otherwise represented in the area. An important trait when determining *B. × wardii* plants is that the leaves are sometimes alternate and sometimes opposite on the same plant, even on the same stem (Figure 1).

The aim of this study was to test the hybrid origin hypothesis of *B. ×wardii* in two sites in Tibet. Specifically, we used three chloroplast DNA region and four nuclear genes applied in former studies on *Buddleja* (nrETS, gapC2, PPR24, PPR123) [19,36] to answer the following questions: 1) Are the *B. wardii* plants in fact hybrids between *B. crispa* and *B. alternifolia* at these two sites? 2) If yes, are there any differences between the genetic patterns at these two hybrid zones and 3) does the genetic structure of hybrid zones such as these provide clues to the mechanism of reproductive isolation between parental species?

Results

Morphological analysis

Of the three leaf characters, leaf length (L), leaf width (W) and ratio of leaf length to width(L/W), putative hybrid individuals consistently had two morphological characters, leaf length and leaf width, intermediate in value between *B. alternifolia* and *B. crispa* (Table 1). The L/W were significantly larger in *B. alternifolia* than the other two taxa (Table 1). Of the seven floral characters, corolla tube width (TW) and anther height (AH) of *B. × wardii* were intermediate between the values of the two assumed parental species, whereas herkogamy (HE) did not differ significantly between the three taxa (Table 1). The remaining four floral characters, corolla tube length (TL), corolla lobe length (CLL), corolla lobe width (CLW) and style length (SL) all showed a similar pattern, with the characters in *B. alternifolia* having significantly lower values than those measured in *B. crispa* or *B. × wardii* (Table 1).

The two putative parental species are morphologically clearly distinct. In the PCA of 10 morphological characteristics, 52.17% and 12.98% of the variation in total was explained by the top two principal components, respectively. The two-dimensional scatter diagram based on PC1 and PC2 showed clearly the separation of *B. alternifolia* and *B. crispa*. Individuals of *B. × wardii* fell between the two parent species, with a slight overlap with *B. alternifolia* and a large overlap with *B. crispa*. Apart from the character HE, there is little difference in the correlation coefficients between the other nine traits (0.29-0.38) (Figure 2a).

Petal color reflectance

Difference in reflectance spectrum of corolla was showed among there three taxa. Both *B. crispa* and *B. × wardii* were observed at 485 nm with obvious peak in the reflectance spectrum, with extremely low variation. However, there was no obvious peak in the reflectance spectrum in *B. alternifolia* (Figure 2b).

Pollination treatments

These two interspecific cross-pollination treatments with *B. crispa* or *B. alternifolia* as the mother species showed significant differences in fruit set, mean number of seeds per fruit, seed set, and mean number of vigor seeds per fruit. When *B. crispa* is the maternal plant, higher fruit set (64.71% vs 6.45%), more seed numbers per capsule (43.42±16.01 vs 6.5±0.71), higher seed set (68.79% vs 0) and more vigor seeds per fruit (26.90 vs 0) were coincidentally examined when comparing these parameters to cross-specific pollination when *B. alternifolia* as maternal plant (Supplemental Table S1). Additionally, comparatively large amount of seeds was produced when *B. crispa* was attributing to geitonogamy pollination indicating self-compatible breeding system (seed set: 48.68%) whereas no seeds were obtained when *B. alternifolia* was attributing to hand self-pollination.

Sequence analyses of the four nuclear genes

NrETS: The total length of the nrETS region alignment was 380 bp in all individuals, including 30 nucleotide substitutions and one 1-bp insertion/deletion (for variation sites, Supplemental Table S2). These variable sites generated 40 haplotypes, among them five, six and thirteen haplotypes belong to *B. alternifolia*, *B. crispa* and *B. × wardii* in BH population, whereas six, three and seven haplotypes in TJ populations, respectively. For the four reference populations, samples of JZA, KDA, BSC and KMC had two, three, three and four specific haplotypes, respectively. Two major clusters were identified by haplotype network analysis with six nucleotide mutations. One cluster composed mainly haplotypes of *B. alternifolia* and *B. × wardii*, comprising five haplotypes of BHA, one haplotype of BHC, seven haplotypes of BHW, six haplotypes of TJA, one haplotype of TJC, three haplotypes of TJW and all haplotypes of *B. alternifolia* from allopatric populations. The other cluster was mainly consisting of haplotypes of *B. crispa* and *B. × wardii,* including five haplotypes of BHC, six haplotypes of BHW, two haplotypes of TJC, four haplotypes of TJW and all haplotypes of *B. crispa* from allopatric populations (Figure 3 a).

Two haplotypes from *B. crispa* nested with the *B. alternifolia* cluster, and was found in BHCR2, BHCR7 and TJCR13, which show different haplotypes from both clusters. Except for two individuals, BHWI9 and TJWI1, all *B. × wardii* individuals with two haplotypes clustered into two divergent clusters and shared the mainly haplotypes of both *B. alternifolia* and *B. crispa* from the BH and TJ. The remaining two individuals had only one haplotype shared with *B. crispa* and were homozygous at this locus (Figure 3a).

GapC2: The total length of the gapC2 region alignment was 606 bp for all individuals, including 67 nucleotide substitutions, one 2-bp and one 1-bp insertion/deletion (for variation sites, Supplemental Table S2). A total of 41 haplotypes were observed from these loci among which one, nine and nine haplotypes belong to *B. alternifolia*, *B. crispa* and *B. × wardii* in BH population, whereas three, six and ten haplotypes in TJ populations, respectively. For the four reference populations, samples of JZA, KDA, BSC and KMC had one, five, eleven and four specific haplotypes, respectively. The haplotype network fall into three major clusters with one cluster consisted of all haplotypes of KDA, the second cluster contained one haplotype of BHA, one haplotype of BHC, two haplotypes of BHW, two haplotypes of TJA, three haplotypes of TJW, and the only one haplotype of JZA that share with the main

haplotype of TJA. And the third cluster contained eight haplotypes of BHC, seven haplotypes of BHW, six haplotypes of TJC, seven haplotypes of TJW and all haplotypes of *B. crispa* from allopatric populations.

For *B. crispa* in both populations of TJ and BH, two individuals (BHCR2 and BHCR7) had a haplotype found in the second cluster but all other haplotypes were found in the third cluster. For *B. alternifolia*, there was an exception of two individuals, TJAL6 and TJAL12, which had a haplotype found in the third cluster, but all the other haplotypes were found in the second cluster. All *B. × wardii* individuals but one (BHWI9) showed two divergent haplotypes originating from both the second and third clusters and shared the mainly haplotypes of both *B. alternifolia* and *B. crispa* from the BH and TJ. The remaining individual BHWI9 was homozygous for a haplotype from the third cluster (Figure 3b).

PPR24: The total length of the PPR24 region alignment was 647 bp for all individuals, including 70 nucleotide substitutions (for variation sites, Supplemental Table S2). This locus has the maximum haplotypes of 53. *B. alternifolia*, *B. crispa* and *B. × wardii* in BH populations had four, ten and 14 haplotypes, whereas there were one, 11 and 15 haplotypes in TJ population, respectively. For the four reference populations, samples of JZA, KDA, BSC and KMC had one, one, seven and seven specific haplotypes, respectively. In the haplotypes network analysis, two major clusters were identified with 16 nucleotide substitutions. One cluster contained four haplotypes of BHA, one haplotype of BHC, seven haplotypes of BHW, the only haplotype of TJA, eight haplotypes of TJW and the unique haplotype for each of the two referenced populations of *B. alternifolia*. The other cluster contained nine haplotypes of BHC, seven haplotypes of BHW, all haplotypes of TJC, seven haplotypes of TJW, and all haplotypes of *B. crispa* from allopatric populations. One *B. crispa* individual (BHCR7) and all *B. × wardii* individuals showed two divergent haplotypes originating from both clusters (Figure 3c).

PPR123: After sequence alignment, the PPR123 region was 735 bp in length for all samples including 59 nucleotide substitutions (for variation sites, Supplemental Table S2). A total of 45 haplotypes were identified for this gene, among which two, four and eight belong to *B. alternifolia*, *B. crispa* and *B. × wardii* in BH population, whereas two, ten and ten haplotypes in TJ population, respectively. For the four reference populations, samples of JZA, KDA, BSC and KMC had two, four, eleven and two specific haplotypes, respectively. As shown in the haplotypes network, all haplotypes were clustered into four groups with seven, six or 13 variations between them. The first clade comprised two haplotypes of BHA, one haplotype of BHC, five haplotypes of BHW, two haplotypes of TJA, one haplotype of TJA and all haplotypes of *B. alternifolia* from allopatric populations. The second clade comprised three haplotypes of BHC, three haplotypes of BHW, five haplotypes of TJC and seven haplotypes of TJW. The third clade comprised five haplotypes of TJC and two haplotypes of TJW. The remaining clade comprised all haplotypes of *B. crispa* from allopatric populations (Figure 3d).

All haplotypes from *B. alternifolia* individuals fell into the first cluster. For the *B. crispa* individuals in BHC and TJC, only one from BHC nested with the first clade derived from BHCR7, ten individuals (TJCR3/5/6/7/9/10/11/12/13/15) had two divergent haplotypes, falling into the second and third clusters, four individuals (TJCR1/4/8/16) had haplotypes nested in the second clade and two individuals (TJCR2/14) had haplotypes nested in the third clade. Of individuals of *B. × wardii*, all but BHWI9 had two different haplotypes, one clustered in the first cluster and share the major haplotypes with *B. alternifolia*, the other clustered in either the second or the third cluster and share the major haplotypes with *B. crispa*. The remaining BHWI9 had two haplotypes found in the second clade (Figure 3d).

Sequence analyses for the combined chloroplast regions

The combined aligned length of the cpDNA fragment alignment (including rpl16, trnD-trnT, trnS-trnfM) was 2014 bp, and contained 38 nucleotide substitutions, one 1-bp, one 2-bp and one 6-bp insertion/deletion (for variation sites, Supplemental Table S2). A total of 15 haplotypes were found in all samples, among which three, three and one were specific to the reference population of KMC, BSC and KDA, respectively. Haplotype network analysis indicated that these three reference populations differed from other populations by at least seven DNA base variants. All individuals in TJ had only two haplotypes, of them one was derived from six individuals of TJA (TJAL2/3/4/5/8/11) and shared with the only haplotype from the reference population JZA, and the remaining 42 individuals shared another haplotype consistent with the haplotype only in individual BHAL13. Each taxon had three haplotypes in BH population, among which two, one and one were specific to *B. alternifolia*, *B. crispa* and *B. × wardii*, respectively. Most of the individuals in BHW (75%) shared haplotypes with *B. crispa* (BHC), while they did not share any haplotypes with *B. alternifolia*. The remaining four *B. × wardii* individuals (25%) had a unique haplotype with one mutational step away from the predominant haplotype of *B. crispa* (Figure 3e).

NewHybrids analysis

Analysis of the four studied nuclear genes using NewHybrids among the three taxa showed that individuals morphologically identified as *B. × wardii* in TJ and most individuals in BH were F_1 hybrids between *B. alternifolia* and *B. crispa*. At BH, 31 of 33 individuals with putative parental morphology were assigned to be the pure parental species with high posterior probabilities (>98.8%). All individuals of BHW but two (BHWI2 and BHWI9) were identified as F_1s with >96.8% posterior probability. Of these two, BHWI2 was identified as F_1 with a 77.8% posterior probability, whereas BHWI9 was identified as *B. crispa* with a 99.7% posterior probability (Figure 4a).

At TJ population, only TJCR13 was identified as *B. crispa* with low probability of 84.0%, and all the rest of the putative parental individuals had a high posterior probability of >99.0% to be recognized as pure parents. The posterior probabilities of all putative hybrids being presumed to be F_1s were >97.4% (Figure 4a). Therefore, in both populations, three individuals were identified as backcross to *B. crispa* with a probability of less than 16%, which was far lower than their probability of being *B. crispa* or F_1s, so the hybridization was not beyond the F_1s.

Population structure analysis

Structure analysis of a total of 213 variation loci for all individuals showed that the highest value of $\triangle K$ was obtained when K=2, suggesting that all 152 individuals were clustered into two types of genetic clusters. When K=2, for the four reference allopatric populations, all individuals of KMC and BSC formed a pure cluster ($q \geq 0.999$), while all individuals of KDA and JZA formed another pure cluster ($q \geq 0.998$).

As showed in the Figure 4b, in BH population, all morphologically determined to be *B. alternifolia* clustered into one cluster with their high proportion containing one genetic component ($q \geq 0.997$). Except for two individuals, BHCR2 and BHCR7, all individuals identified as *B. crispa* formed another cluster with the proportion containing another genetic component ≥0.998. While BHCR2 and BHCR7 contained the same genetic component of BHC were 0.755 and 0.510, respectively. All *B. × wardii* individuals showed approximately equal proportions from both clusters ($q=0.494 \pm 0.084$ for BHA), expect for BHWI9, which had a proportion of 0.820 genetic component from BHC cluster.

Similar to the BH population, most of the *B. alternifolia* (TJA) and *B. crispa* (TJC) in the TJ population were segregated into a high proportion of different genetic components ($q \geq 0.988$ for TJA; $q \geq 0.999$ for TJC). Of the

remaining three individuals, the probability that TJAL6 and TJAL12 contained low levels of genetic components from TJA was 0.959 and 0.931, respectively, while TJCR13 possessed genetic components from TJC was 0.830. All individuals of *B. × wardii* (TJW) possessed close proportions of genetic components from both TJA and TJC (q=0.523±0.033 for TJA).

Discussion

B. × wardii displayed intermediate morphological characters in the four quantitative characters assessed, whereas the remaining five characters were similar to those of *B. crispa* but different from those of *B. alternifolia.* No transparent traits were observed, indicating very early generation hybrids with limited genetic segregation. Chromatogram additivity of *B. × wardii* at each of the differentiated nuclear genes of the parental species (Table 2), as well as estimation of genotypes using NewHybrids further confirmed that most *B. × wardii* individuals should be considered to be F_1s. Thus evidence from both morphology and molecular markers allowed us to reject the hypothesis that *B. × wardii* had undergone sufficient genetic recombination, generally assumed to be vital for establishing a diploid hybrid species [37]. Alternatively, F_1s can successfully impede gene flow and thus maintain species boundaries of *B. alternifolia* and *B. crispa* in sympatric areas. Intriguingly, it was noted that both parental species frequently occurred in distributed habitats and thus the empirical hypothesis of intermediate habitat which was deemed to be vital for hybrid establishment owing to the assumption of lower fitness of hybrids in both native habitats of parental species was rejected [5].

The present hybrid populations were found at very high elevation (both hybrid zones > 3400m) in Tibet, and most traditionally considered prerequisites were fulfilled in *B. × wardii* [1]. Sympatric distribution, overlapping flowering periods (*B. alternifolia* ranging from April to June, while *B. crispa* from March to August) [38-39] and shared pollinators (bees, bumblebees and butterflies) [40] (personal observation) all facilitate natural hybridization between the two parental species. Another factor favoring natural hybridization between *B. alternifolia* and *B. crispa* was that they are both diploid with 2n = 38 [27], which may facilitate hybridization [41-42]. In addition, *Buddleja* in Asia is a young clade that began diversifying approximately 10 Ma during the uplift of the Himalayas, and reproductive isolation between them is still incomplete [27,43].

Notably, we never found isolated populations *B. × wardii,* and in both populations, it grows together with both putative parents, *B. × wardii* can not likely to be self sustainable [44-45]. Taking together, *B. × wardii* would never be considered as a hybrid species. We therefore recommended that the literature in the future should avoid the name *B. wardii*.

Although F_1 dominant hybrid zones were confirmed in both study areas, heritable patterns of cpDNA were different. In most angiosperms, chloroplast DNA is maternally inherited [46]. The sequencing of cpDNA in the BH population showed that *B. crispa* and *B. alternifolia* each have their own specific chloroplast DNA haplotypes, and most *B. × wardii* individuals have the same haplotypes as *B. crispa*. The unique haplotypes seen in the remaining four *B. × wardii* individuals (BHWI8/11/15/16) may due to unsampled polymorphism of the parents. Several hypotheses have already been suggested to interpret asymmetric heritage of cpDNA in hybrids, including differences of breeding systems, flowering time, pollinator behavior and the local abundance of parental taxa in contact zone [8,47-49]. For the present study, we have no evidence of pollinator behavior data to relate the asymmetry hybridization. It has been hypothesized that self-compatible species might be more common as

maternal parent than the self-incompatible species [16], which was called “SI x SC rule”. Previous studies have confirmed self-compatibility in *B. crispa* from the substantial seed set attributed to geitonogamous pollination, similar to that following out-crossing when flowers [40], whereas no seeds were obtained in self-pollination treatments in *B. alternifolia*. Additionally, the peak flowering time of *B. crispa* is earlier than *B. alternifolia* with only very limited overlap. Based on our observation of flowering time towards the two parental species that had been introduced into Kunming Botanical Garden from Tibet, *B. crispa* was nearly finished flowering when *B. alternifolia* started flowering. Species with earlier flowering time are more likely to have more pollen from heterospecific onto the stigmas near the end of flowering due to the lack of homozygous pollen, and thus are more likely to be the maternal parent for hybrids [16,50], assuming the shared pollinator (s) did not have visitation preference to one parent.

Additionally, in the TJ hybrid zone, the maternal parent could not be identified due to the cpDNA results (only 6 individuals of *B. alternifolia* had a haplotype different from other individuals, with all other individuals sharing the same haplotype). Due to the fact that almost all hybrids were F_1s and only a very limited number with low probabilities backcrossed to the parents (NewHybrids: *B. alternifolia*: 0; *B. crispa*: 1; Structure: *B. alternifolia*: 2; *B. crispa*: 1), the explanation that repeated backcrossing contributed to the pattern of cpDNA in this hybrid zone was rejected. However, it is possible that other haplotypes of cpDNA in *B. crispa* existed historically but have disappeared, as we found that > 500 mature *B. alternifolia* and < 100 *B. crispa* individuals were currently distributed in TJ. This may also be due to the slow evolution of chloroplast gene [51-52]which makes it impossible to identify the maternal parent of hybrids in some species pair with close relatives. However as was discussed above, *B. crispa* could be still maternal plant of the hybrids in TJ due to the “SI x SC rule”.

Conclusions

We investigated patterns of hybridization in two butterfly bush species in two areas in Tibet, both with elevations above 3400 m. Both morphological and molecular analyses supported the hypothesis that the putative hybrid plants are F_1s, without evidence of transparental traits or genetic recombination. This system is further effectively promoting nearly complete reproductive isolation between the parental plants. The self-compatible breeding system, earlier peak flowering and small number of *B. crispa* individuals suggest that *B. crispa* is the maternal parent, which is supported by the cpDNA analysis of hybrids. Overall, the present study provides mechanismal insight into the maintenance of reproductive isolation, in particular for sympatrically growing pioneer plants in disturbed habitats, which have to date been largely ignored in natural hybridization studies.

Methods

Species and plant material in this study

Both *B. crispa* and *B. alternifolia* are vigorous deciduous shrubs or small trees to 2-4 m high. *B. crispa* is a widespread species distributed in the hot/warm-dry valleys, growing on forest edges, in shrubs, on exposed cliffs, and in rocks crevices at elevations of 1400-4300 m, across the Himalaya-Hengduan area [26,53]. *B. alternifolia* is distributed in northwest of China throughout Tibet to Loess Plateau, where it is naturally found growing along river banks or dried up streams in thickets at an altitude of 1500-4000m [26,54]. Ecologically, *B. crispa* and *B. alternifolia* are highly susceptible to habitat disturbance [29,53-54]. The two species occupy similar habitats and often occur sympatrically where their altitudinal ranges overlap [38,53-54]. Both are diploid with a

chromosome number of 2n=38 [27,29,55-56], and start flowering in the spring (*B. crispa*: March to August; *B. alternifolia*: April to June) [38-39].

All material for morphological characters and molecular analysis were field-collected. In Lhasa and Nyingchi, Tibet, China, individuals of intermediate morphologies between *B. alternifolia* and *B. crispa* were found co-occurring with sympatric populations of the two species along two branches of Brahmaputra river: the Ni-yang River and the Lhasa River (Figure 5). In this study, we sampled 17, 15 and 16 individuals of *B. alternifolia*, *B. crispa* and *B. × wardii* from Bahe town (BH) in Nyingchi, and 15, 16 and 17 individuals of *B. alternifolia*, *B. crispa* and *B. × wardii* from Taji county (TJ) in Lhasa for molecular analysis. In addition, four allopatric populations of *B. alternifolia* and *B. crispa* were sampled as pure parental populations for reference: *B. alternifolia* from Jiangzi county in Rikaze, Tibet, and Kangding county in Tibetan Autonomous Prefecture of Garzê, Sichuan; *B. crispa* from Sishan mountain in Kunming, Yunnan, and Basu county in Qamdo, Tibet. Sampling information is shown in Table 3. The Flora of China (FOC) was used for specific identification [38]. Voucher specimens were deposited at the herbarium of Kunming Institute of Botany, Chinese Academy of Sciences (Supplemental Table S3). In both sampled populations, many (more than 500) *B. alternifolia* individuals were found. However, at both BH and TJ, population size estimates of both *B. crispa* and the putative hybrid were fewer than 100 plants per population.

Measurements and data analysis of morphological traits

Thirty adult plants were randomly selected per taxon, and a fresh mature leave on the middle of current-growth long vegetative branchlets were sampled without pests, diseases, or other obvious damage in June. Three leaf morphological characters were measured for each of these sampled leaves. Seven morphological flower characters were measured from 30 randomly selected healthy flowers during the flowering period in April from 30 plants per taxon (one flower per individual). These 10 characteristics were measured using digital calipers to the nearest 0.01mm as follows: 1) leaf length (L); 2) leaf width (W); 3) ratio of leaf length to leaf width (L/W); 4) corolla tube length (TL); 5) corolla tube width (TW, diameter of the top of corolla tube); 6) corolla lobe length (CLL); 7) corolla lobe width (CLW); 8) anther height (AH, from the top of stamen to the start of ovary); 9) style length (SL, from the top of stigma to the base of ovary); 10) herkogamy (HE, separation between tip of style and base of stamen) (Supplemental Table S4).

A trimmed means was used to calculate the means for every 30 individuals of each taxon [57]. One-way ANOVA was used to analyze these traits among the three taxa in SPSS Statistics 16.0 [58], and the significance of differences between the means was determined using the standard F statistic. Where the data did not satisfy the criterion of homogeneity of variance, a Welch statistic was employed. And post hoc Tamhane's test was performed in pairwise comparison. The data were then subjected to two-dimensional Procrustes fitting in PAST (PAlaeontology STatistical) software ver. 3.26 [59] to standardize landmark coordinates, followed by a shape principal components analysis (PCA) to perform multivariate analyses of the measured morphological characters (leaf and floral traits) [60].

Petal color analysis

To assess patterns of light reflection in these three taxa at different wavelengths, a USB2000+ miniature fiber optic spectrometer with a DH2000 deuterium-halogen light source (Ocean Optics, Dunedin, FL, USA) was performed for spectral measurements of the corolla color. Measurements were taken in increments of 0.45 and

0.55 nm over the range of 250 nm to 750 nm [61]. We choose 30 petals from 30 individuals per taxon (one flower per individual) and took one measurement per petal (Supplemental Table S5).

Artificial hybridization and seed germination

B. alternifolia from Tibet were successfully introduced in the Kunming Botanical Garden (KBG) and flowered in the spring of 2020. Artificial hybridization experiments were carried out on *B. alternifolia* and *B. crispa* growing in KBG as parents for three interspecific cross-pollination treatments (Supplemental Table S1). For all treatments, more than ten inflorescences with 1-3 single flowers and a total of 24, 31and 34 single flowers were randomly selected, and then hand-pollinated after artificial emasculation and bagged. Capsules were harvested just before they cracked and the number of fruits was counted. The number of seeds per fruit was counted after the capsules cracked naturally.

The collected seeds of the two interspecific pollination treatments were evenly sown in sterilized disposable Petri dished of 9 cm diameter, lined with three pieces of filter paper and moistened with distilled water. A maximum of 50 seeds were sown per Petri dish and seeds were germinated at 20 ℃ for 12 h photoperiod in a plant growth chambers (MTI-202; Tokyo Rikakikai Co., Ltd., Japan) [62].

DNA extraction, PCR amplification and DNA sequencing

Total DNA was extracted from approximately 50 mg dried leaves using the modified cetyl trimethyl ammonium (CTAB) method [63], and then stored at -20°C before further analyses. Three low-copy nuclear genes (gapC2, PPR24, PPR123), the nrETS region, and three plastid regions (rpl16, trnD-trnT, trnS-trnfM) that had been successfully PCR amplified in *Buddleja* were selected for sequencing by previous publications [19,36,64]. NrETS, gapC2 and PPR are part of the nuclear ribosomal external transcribed spacer, the gapC gene family and the pentatricopeptide repeat gene family, respectively. Sequences of all the primers used are listed in Supplemental Table S6.

PCR was conducted using 2×Taq PCR Master-mix (Tiangen) or the KOD-FX DNA polymerase (for PPR24). The PCR conditions were set as follows: initial denaturation at 94°C for 4 min; then 32 cycles of 94°C for 30s, corresponding annealing temperature 52 to 54°C for 40s, and 72°C for 50s; and an extension step at 72°C for 10 min. After testing by running them on a 1.2% agarose gel, the PCR products were then sent for DNA sequencing on an ABI 3730 DNA analyzer (Applied Biosystems). Sequences have been deposited in GenBank with accession numbers MT7333350-MT733514 and XXXXXXX. In addition, sequence data for the ETS and gapC2 genes of 12 individuals from KMC were obtained from the paper published by Liao et al. [19].

Data analysis

The program SeqMan^{TM} in DNASTAR was used for alignment, assembling, and comparison of DNA sequences [4,19]. PHASE in DNASP ver. 5.10.01 was used to infer the haplotypes, and to calculate the number of and haplotypes [65-67]. For the few haplotypes estimates with uncertain phases, we choose the haplotype with the highest probability for analysis. Network ver. 5.0.1.1 was used to construct the haplotype network of each gene [60]. NewHybrids ver. 1.1 was used for the genotype class speculation of each individual: both parental species, F_1s and F_2s, backcross to each parental species [68] (Supplemental Table S7).

Genomic admixture proportions were determined using the program of Structure ver. 2.3.4 with the default settings [69-70]. Analyses were run with numbers of distinct clusters (K) varying from 1 to 15, with eight iterations performed for each K and a burn-in of 100,000 and a MCMC of 100,000 iterations (Supplemental Table S8). Structure Harvester web ver. 0.6.94 was used to obtain the optimal K of distinct groups [71-72] (Supplemental Figure S1). The membership coefficients at each of the suggested numbers of clusters for each individual were estimated across the 8 independent runs and graph of the population structuring were generated using Microsoft Excel 2016.

Abbreviations

ANOVA: Analysis of Variance; DNA: DeoxyriboNucleic Acid; cpDNA: chloroplast DNA; PCR: Polymerase Chain Reaction; MCMC: Markov Chain Monte Carlo.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The data sets supporting the results of the present study are included within this article and its additional files.

Competing interests

The authors declare no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (NSFC) (Grant No. 31770240), the Youth Innovation Promotion Association, Chinese Academy of Sciences (Grant No. 2018428) and the Reserve Talents for Academic and Technical Leaders of Middle-aged and Young People in Yunnan Province (Grant No. 2018HB066). The founder only provided the financial support and did not involve in the design of the experiment, collection, analysis and interpretation of data, decision to publish, or the preparation of the manuscript.

Authors' contributions

YM and RL collected the plant materials and wrote the manuscript; RL performed the experiments and analyzed the data; YM and WS designed the experiment and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We thank Ms. Lidan Tao for her help in the field investigations and Dr. Jane Marczewski for improving English. We thank the editor and anonymous reviewers for their comments on the manuscript.

References

1. Arnold ML. Natural hybridization and evolution. Demand: Oxford University Press. 1997.
2. Rieseberg LH. The role of hybridization in evolution-old wine in new skins. Am J Bot. 1995;82(7):944–953.
3. Soltis PS, Soltis DE. The role of hybridization in plant speciation. Annu Rev Plant Biol. 2009;60:561–588.
4. Zhang RS, Liu T, Wu W, Li YQ, Chao LF, Huang LS, Shi SH, Zhou RC. Molecular evidence for natural hybridization in the mangrove fern genus *Acrostichum*. BMC Plant Biol. 2013;13:74.
5. Anderson E. Hybridization of the habitat. Evolution. 1948;2:1–9.
6. Grant PR, Grant BR. Speciation undone. Nature. 2014;507(7491):178–179.
7. Schumer M, Rosenthal GG, Andolfatto P. How common is homoploid hybrid speciation? Evolution. 2014;68(6):1553–1560.
8. Zhou RC, Gong X, Boufford D, Wu CI, Shi SH. Testing a hypothesis of unidirectional hybridization in plants: observations on *Sonneratia*, *Bruguiera* and *Ligularia*. BMC Evol Biol. 2008;8:149.
9. Wei YK, Huang YB, Li GB. Reproductive isolation in sympatric *Salvia* species sharing a sole pollinator. Biodiversity Science, 2017;25(6):608–614. (In Chinese)
10. Arnold ML. (1993) *Iris nelsonii*: origin and genetic composition of a homoploid hybrid species. Am J Bot. 1993;80(5):577–583.
11. Yan LJ, Gao LM, Li DZ. Molecular evidence for natural hybridization between *Rhododendron spiciferum* and *R. spinuliferum* (Ericaceae). J Syst Evol. 2013;51(4):426–434.
12. Zhang W, Dasmahapatra KK, Mallet J, Moreira GR, Kronforst MR. Genome-wide introgression among distantly related *Heliconius* butterfly species. Genome Biol. 2016;17(1): 25.
13. Tagane S, Hiramatsu M, Okubo H. Hybridization and asymmetric introgression between *Rhododendron eriocarpum* and *R. indicum* on Yakushima Island, Southwest Japan. J Plant Res. 2008;121(4):387–395.
14. Todesco M, Pascual MA, Owens GL, Ostevik KL, Moyers BT, Hübner S, Heredia SM, Hahn MA, Caseys C, Bock DG, Rieseberg LH. Hybridization and extinction. Evol Appl. 2016;9(7):892–908.
15. Wu RZ, Zou PS, Tan GW, Hu ZY, Wang YQ, Ning ZL, Wu W, Liu Y, He SY, Zhou RC. Molecular identification of natural hybridization between *Melastoma malabathricum* and *Melastoama beccarianum* in Sarawak, Malaysia. Ecol Evol. 2019;9(10):5766–5776.
16. Zha HG, Milne RI, Sun H. Asymmetric hybridization in *Rhododendron agastum*: A hybrid taxon comprising mainly F_1s in Yunnan, China. Ann Bot. 2010;105:89–100.
17. Milne RI, Terzioglu S, Abbott R. A hybrid zone dominated by fertile F_1s: Maintenance of species barriers in *Rhododendron*. Mol Ecol. 2003;12(10):2719–2729.
18. Milne RI, Abbott RJ. Reproductive isolation among two interfertile *Rhododendron* species: low frequency of post-F_1 hybrid genotypes in alpine hybrid zones. Mol Ecol. 2008;17(4):1108–1121.
19. Liao RL, Ma YP, Gong WC, Chen G, Sun WB, Zhou RC, Marczewski T. Natural hybridization and asymmetric introgression at the distribution margin of two *Buddleja* species with a large overlap. BMC Plant Biol. 2015;15:146.

20. Zhang NN, Ma YP, Folk RA, Yu JJ, Pan YZ, Gong X. Maintenance of species boundaries in three sympatric *Ligularia* (Senecioneae, Asteraceae) species. J Integr Plant Biol. 2018;60(10):986–999.
21. Zhang NN, Yu JJ, Wang YH, Gong X. Molecular evidence for asymmetric hybridization in three closely related sympatric species. AoB Plants. 2018;10:ply011.
22. Kyhos DW, Clark C, Thompson WC. The hybrid nature of *Encelia-laciniata* (Compositae, Heliantheae) and control of population composition by post-dispersal selection. Syst Bot. 1981;6(4):399–411.
23. Nagano Y, Hirao AS, Itino T. Genetic structure of a hybrid zone between two violets, *Viola rossii* Hemsl. and *V. bissetii* Maxim.: Dominance of F_1 individuals in a narrow contact range. Plant Spec Biol. 2015;30(3):237–243.
24. Fang MY, Fang RZ, He MY, Hu LZ, Yang HB, Chamberlain DF. *Rhododendron*. In: Wu ZY, Raven PH, editors. Flora of China, Vol. 14. Beijing and St Louis: Science Press and Missouri Botanical Garden, 2005. p. 260–455.
25. Norman EM. Buddlejaceae. New York: The New York Botanical Garden Press. 2000.
26. Stuart DD. Buddlejas. Portland: Timber Press. 2006.
27. Chen G, Sun WB, Sun H. Ploidy variation in *Buddleja* (Buddlejaceae) in the sino-himalayan region and its biogeographical implications. Bot J Linn Soc. 2007;154(3):305–312.
28. Ge J, Cai L, Bi GQ, Chen G, Sun WB. Characterization of the complete chloroplast genomes of *Buddleja colvilei* and *B. sessilifolia*: Implications for the taxonomy of *Buddleja*. Molecules. 2018;23(6):1248.
29. Leeuwenberg AJM. The Loganiaceae of Aafrice . *Buddleja* . Revision of the African and Asiatic species. Wageninge: Mededelingen Landbouwhogeschool. 1979.
30. Gong WC, Chen G, Liu CQ, Dunn BL, Sun WB. Comparison of floral scent between and within *Buddleja fallowiana* and *Buddleja officinalis* (Scrophulariaceae). Biochem Syst Ecol. 2014;55:322–328.
31. Raja S, Ramya I. A review on ethnopharmacology, phytochemistry and pharmacology of *Buddleja asiatica*. Int J Pharm Sci Res. 2016;7(12):4697–4709.
32. Sheppard AW, Shaw RH, Sforza R. Top 20 environmental weeds for classical biological control in Europe: A review of opportunities, regulations and other barriers to adoption. Weed Res. 2006;46(2):93–117.
33. Tallent-Halsell NG, Watt MS. 2009. The invasive *Buddleja davidii* (butterfly bush). Bot Rev. 2009;75(3):292–325.
34. Wittig R. Frequency of *Buddleja davidii* Franch. (Buddlejaceae) in Germany along ecological gradients. Flora. 2012;207:133–140.
35. Marquand CVB. The botanical collection made by Captain F. Kingdon Ward in the eastern Himalaya and Tibet in 1924–25. Bot J Linn Soc. 1929;48(321):149–229.
36. Chau JH, O'Leary N, Sun WB, Olmstead RG. Phylogenetic relationships in tribe Buddlejeae (Scrophulariaceae) based on multiple nuclear and plastid markers. Bot J Linn Soc. 2017;184(2):137–166.
37. Ma YP, Zhang CQ, Zhang JL, Yang JB. Natural hybridization between *Rhododendron delavayi* and *R. cyanocarpum* (Ericaceae), from morphological, molecular and reproductive evidence. J Integr Plant Biol. 2010;52(9):844–851.
38. Li PT, Leeuwenberg AJM. Loganiaceae. In: Wu ZY, Raven PH, editors, Flora of China, Vol. 15. Beijing: Science Press. 1996. p. 320–338.
39. Wang HY, Zhang XS, Li XQ, Zhao HH, Shi XB. Rapid propagation of *Buddleja alternifolia*. Guizhou Agri Sci. 2010;38(6):30–31.

40. Gong WC, Chen G, Vereecken NJ, Dunn BL, Ma YP, Sun WB. Floral scent composition predicts bee pollination system in five butterfly bush (*Buddleja*, Scrophulariaceae) Species. Plant Biol. 2014;17:245–255.
41. Kou YX, Xiao K, Lai XR, Wang YJ, Zhang ZY. Natural hybridization between *Torreya jackii* and *T. grandis* (Taxaceae) in southeast China. J Syst Evol. 2017;55(1):25–33.
42. Ning H, Pan YZ, Gong X. Molecular evidence for natural hybridization between *Ligularia nelumbifolia* and *Cremanthodium stenoglossum* (Asteraceae, Senecioneae). Botany. 2019;97(1):53–69.
43. Chau JH. Systematics of *Buddleja* (Scrophulariaceae): Phylogenetic relationships, historical biogeography, and phylogenomics. Washington: University of Washington. 2017.
44. Abbott RJ. Sex, Sunflowers, and Speciation. Science. 2003;301(5637):1189–1190.
45. Feliner GN, Alvarez I, Fuertes-Aguilar J, Heuertz M, Marques I, Moharrek F, Pineiro R, Riina R, Rossello JA, Soltis PS, Villa-Machio I. Is homoploid hybrid speciation that rare? An empiricist's view. Heredity. 2017;118(6):513–516.
46. Morgensen HL. The hows and whys of cytoplasmic inheritance in seed plants. Am J Bot. 1996:83(3), 383–404.
47. Ma YP, Milne R, Zhang CQ, Yang JB. Unusual patterns of hybridization involving a narrow endemic *Rhododendron* Species (Ericaceae) in Yunnan, China. Am J Bot. 2010;97(10): 1749–57.
48. Ma YP, Xie WJ, Tian XL, Sun WB, Wu ZK, Milne R. Unidirectional hybridization and reproductive barriers between two heterostylous primrose species in north-west Yunnan, China. Ann Bot. 2014;113:753–761.
49. Tong ZY, Huang SQ. Pre- and post-pollination interaction between six co-flowering *Pedicularis* species via heterospecific pollen transfer. New Phytol. 2016;211(4):1452–1461.
50. Zhang JJ, Montgomery BR, Huang SQ. Evidence for asymmetrical hybridization despite pre- and post-pollination reproductive barriers between two Silene species. AoB Plants. 2016;8:plw032.
51. Wolfe KH, Li WH, Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci. 1987;84(24): 9054–9058.
52. Drouin G, Daoud H, Xia J. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. Mol Phylogenet Evol. 2008;49(3): 827–831.
53. Yue LL, Chen G, Sun WB, Sun H. Phylogeography of *Buddleja crispa* (Buddlejaceae) and its correlation with drainage system evolution in southwestern China. Am J Bot. 2012;99(10):1726–1735.
54. Wang JH, Zhang JG, Xu Y, Wu HW, Sun GP. Study on biology characteristics of drought enduring shrub *Buddleja alternidolia*. Chinese Wild Plant Resour. 2007;26(6):23–25.
55. Moore RJ. Cytotaxonomic studies in the Loganiaceae Ⅱ. Chromosome numbers and phylogeny in the Loganaceae. Am J Bot. 1947;34(10):527–538.
56. Moore R. Polyploidy, phylogeny, and photoperiodism in Old World *Buddleja*. Evolution. 1961;15(3): 272–280.
57. Marczewski T, Ma YP, Zhang XM, Sun WB, Marczewski AJ. Why is population information crucial for taxonomy? A case study involving a hybrid swarm and related varieties. AoB Plants. 2016;8:plw070.
58. SPSS. SPSS Statistics 16.0 SPSS Inc., Chicago. 2007.
59. Hammer Ø, Harper DAT, Ryan PD. Past: Paleontological statistics software package for education and data analysis. Palaeontologia Electronica. 2001;4(1):9.
60. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 1999;16(1):37–48.

61. Ma YP, Wu ZK, Dong K, Sun WB, Marczewski T. Pollination biology of *Rhododendron cyanocarpum* (Ericaceae): An alpine species endemic to NW Yunnan, China. J Syst Evol. 2015;53(1):63–71.
62. Sun WB, Kong FC, Mickael LW. Effect of light and temperature on seed germination of *Buddleja crispa*. Plant Physiol Commun. 2002;38(6):557–558.
63. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bulletin. 1987;19(1):11–15.
64. Borg AJ, McDade LA, Schönenberger J. Molecular phylogenetics and morphological evolution of Thunbergioideae (Acanthaceae). Taxon. 2008;57(3):811–822.
65. Librado P, Rozas J. Dnasp v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25(11):1451–1452.
66. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet. 2001;68(4):978–989.
67. Stephens M, Scheet P. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet. 2005;76(3):449–462.
68. Anderson EC and Thompson EA. A model-based method for identifying species hybrids using multilocus genetic data. Genetics. 2002;160(3):1217–1229.
69. Hubisz MJ, Falush D, Stephens M, Pritchard JK. Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour. 2009;9(5):1322–1332.
70. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155(2):945–959.
71. Earl DA, Vonholdt BM. Structure harvester: A website and program for visualizing structure output and implementing the evanno method. Conserv Genet Resour. 2012;4(2):359–361.
72. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol Ecol. 2005;14(8):2611–2620.

Tables

Table 1 Morphological traits used to distinguish between *B. alternifolia, B. crispa* and *B. × wardii.*

Characters	*B. alternifolia*	*B. × wardii*	*B. crispa*	F	Welch	P value
L (mm)	16.00±3.15[a]	27.31±10.02[b]	64.65±25.47[c]		62.971	<0.001
W (mm)	3.62±0.66[a]	14.00±3.26[b]	36.61±11.72[c]		235.112	<0.001
L/W	4.52±0.99[a]	1.94±0.47[b]	1.76±0.31[b]		97.410	<0.001
TL (mm)	7.10±0.80[a]	9.77±1.04[b]	10.25±1.51[b]		79.584	0.017
TW (mm)	0.93±0.17[a]	1.09±0.15[b]	1.20±0.18[c]	17.584		0.641
CLL (mm)	1.43±0.27[a]	2.29±0.37[b]	2.17±0.40[b]	49.695		0.073
CLW (mm)	1.60±0.20[a]	2.41±0.35[b]	2.45±0.46[b]		80.282	0.004
AH (mm)	5.22±0.76[a]	5.97±0.93[b]	6.66±0.96[c]	18.266		0.539
SL (mm)	2.89±0.47[a]	4.14±0.91[b]	4.41±1.05[b]		46.621	0.001
HE (mm)	1.08±0.67	0.75±0.72	0.95±0.87	1.261		0.588

1 L: leaf length, W: leaf width, L/W: ratio of leaf length to leaf width, TL: corolla tube length, TW: corolla tube width, CLL: corolla length, CLW: corolla lobe width, AH: anther height, SL: style length, HE: herkogamy.

Table 2 Haplotypes and genotypes of *B. × wardii* between *B. alternifolia* and *B. crispa* at four nuclear genes and the combined chloroplast regions (cpDNA). The code of each haplotype correspond to those in Figure 3.

duals	ETS-B	GapC2	PPR24	PPR123	cpDNA
/I1	Ba1(Bw1)	Ba1(Bw1)	Bw6	Ba1(Bw1, Bc1)	Bc3(Bw2)
	Bc3(Bw8)	Bc4(Bw5)	Bc2(Bw9)	Bc3(Bw7)	
/I2	Bw2	Ta1(Tw1, Bw2, Bc1, Ja1)	Bw2	Bw5	Bc1(Bw1)
	Bc3(Bw8)	Bc6(Bw6, Tw2, Tc1, Kc5)	Bw13	Bc3(Bw7)	
/I3	Bw5	Ta1(Tw1, Bw2, Bc1, Ja1)	Bw6	Ba1(Bw1, Bc1)	Bc1(Bw1)
	Bc5(Bw10)	Bc9(Bw8)	Bc5(Bw10)	Bc2(Bw6)	
/I4	Bw5	Ta1(Tw1, Bw2, Bc1, Ja1)	Ba2(Bw4)	Ba1(Bw1, Bc1)	Bc1(Bw1)
	Bc3(Bw8)	Bc6(Bw6, Tw2, Tc1, Kc5)	Bc5(Bw10)	Bc3(Bw7)	
/I5	Ba1(Bw1)	Ba1(Bw1)	Bw6	Ba1(Bw1, Bc1)	Bc3(Bw2)
	Bc3(Bw8)	Bc4(Bw5)	Bc6(Bw11)	Bc3(Bw7)	
/I6	Ba3(Bw3)	Ba1(Bw1)	Bw5	Ba1(Bw1, Bc1)	Bc3(Bw2)
	Bc3(Bw8)	Bc6(Bw6, Tw2, Tc1, Kc5)	Bw8	Bw8	
/I7	Ba1(Bw1)	Ta1(Tw1, Bw2, Bc1, Ja1)	Bw6	Ba1(Bw1, Bc1)	Bc1(Bw1)
	Bw11	Bc6(Bw6, Tw2, Tc1, Kc5)	Bc2(Bw9)	Bc2(Bw6)	
/I8	Ba1(Bw1)	Ba1(Bw1)	Bw2	Ba2(Bw4)	Bw3
	Bw13	Bw7	Bc5(Bw10)	Bc3(Bw7)	
/I9	Bc3(Bw8)	Bw4	Bw3	Bc3(Bw7)	Bc3(Bw2)
	Bc3(Bw8)	Bw4	Bc2(Bw9)	Bw8	
/I10	Ba1(Bw1)	Ta1(Tw1, Bw2, Bc1, Ja1)	Bw2	Ba1(Bw1, Bc1)	Bc1(Bw1)
	Bw12	Bc3(Bw3)	Bw8	Bc2(Bw6)	
/I11	Ba1(Bw1)	Ta1(Tw1, Bw2, Bc1, Ja1)	Ba4(Bw7)	Ba2(Bw4)	Bw3
	Bc3(Bw8)	Bc6(Bw6, Tw2, Tc1, Kc5)	Bc5(Bw10)	Bc3(Bw7)	
/I12	Ta1(Bw4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Ta1(Tw2, Ba1, Bw2, Ja1)	Bw2	Bc1(Bw1)
	Bc3(Bw8)	Bc6(Bw6, Tw2, Tc1, Kc5)	Bc7(Bw12)	Bc3(Bw7)	
/I13	Ta2(Tw1, Tc1, Ba5, Bw7)	Ta1(Tw1, Bw2, Bc1, Ja1)	Ta1(Tw2, Ba1, Bw2, Ja1)	Ba1(Bw1, Bc1)	Bc1(Bw1)

	Bc3(Bw8)	Bc9(Bw8)	Bc9(Bw14)	Bc3(Bw7)	
7I14	Ta2(Tw1, Tc1, Ba5, Bw7)	Ba1(Bw1)	Ta1(Tw2,Ba1, Bw2, Ja1)	Bw3	Bc1(Bw1)
	Bc3(Bw8)	Bw9	Bc2(Bw9)	Bc3(Bw7)	
7I15	Bw6	Ta1(Tw1, Bw2, Bc1, Ja1)	Ta1(Tw2, Ba1, Bw2, Ja1)	Ba1(Bw1, Bc1)	Bw3
	Bc3(Bw8)	Bc6(Bw6, Tw2, Tc1, Kc5)	Bc5(Bw10)	Bc3(Bw7)	
7I16	Ba1(Bw1)	Ta1(Tw1, Bw2, Bc1, Ja1)	Ta1(Tw2, Ba1, Bw2, Ja1)	Ba1(Bw1, Bc1)	Bw3
	Bw9	Bc6(Bw6, Tw2, Tc1, Kc5)	Bc2(Bw9)	Bc2(Bw6)	
1	Tw3	Ta2(Tw3)	Bc1(Tw1)	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Tw3	Tc3(Tw5)	Tc2(Tw11)	Tc7(Tw5)	
2	Tc2(Tw2, Bc4)	Ta2(Tw3)	Bc1(Tw1)	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Tc3(Tw5)	Tc1(Tw9)	Tw4	
3	Tc2(Tw2, Bc4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Bc1(Tw1)	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Tw7	Tc2(Tw11)	Tw8	
4	Tc2(Tw2, Bc4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Tw8	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Tw6	Tc2(Tw11)	Tw7	
5	Tc2(Tw2, Bc4)	Ta2(Tw3)	Bc1(Tw1)	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Tc4(Tw8, Ta3)	Tc8(Tw15)	Tw10	
6	Tc2(Tw2, Bc4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Bc1(Tw1)	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Tc4(Tw8, Ta3)	Tc1(Tw9)	Tc4(Tw3)	
7	Tc2(Tw2, Bc4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Tw7	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Tc6(Tw10)	Tc8(Tw15)	Tc7(Tw5)	
8	Tc2(Tw2, Bc4)	Tw4	Tw7	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Bc6(Bw6, Tw2, Tc1, Kc5)	Tc5(Tw14)	Tc7(Tw5)	
9	Tc2(Tw2, Bc4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Bc1(Tw1)	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)

	Ta2(Tw1, Tc1, Ba5, Bw7)	Tc4(Tw8, Ta3)	Tw13	Tc7(Tw5)	
10	Tc2(Tw2, Bc4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Tw4	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Tc3(Tw5)	Tc1(Tw9)	Tc4(Tw3)	
11	Tc2(Tw2, Bc4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Tw4	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Tw6	Bc6(Bw6, Tw2, Tc1, Kc5)	Tc1(Tw9)	Tc4(Tw3)	
12	Tw5	Ta1(Tw1, Bw2, Bc1, Ja1)	Tw4	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Tc5(Tw9)	Tc1(Tw9)	Tc2(Tw2)	
13	Tc2(Tw2, Bc4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Tw6	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Tw6	Bc6(Bw6, Tw2, Tc1, Kc5)	Tc8(Tw15)	Tc2(Tw2)	
15	Tc2(Tw2, Bc4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Ta1(Tw2, Ba1, Bw2, Ja1)	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta6(Tw7)	Tc4(Tw8, Ta3)	Tc8(Tw15)	Tc7(Tw5)	
16	Tc2(Tw2, Bc4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Tw3	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Tc4(Tw8, Ta3)	Tw12	Tc7(Tw5)	
17	Tc2(Tw2, Bc4)	Ta1(Tw1, Bw2, Bc1, Ja1)	Ta1(Tw2, Ba1, Bw2, Ja1)	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Tc4(Tw8, Ta3)	Tw10	Tw9	
18	Tw4	Ta1(Tw1, Bw2, Bc1, Ja1)	Tw5	Ta1(Tw1)	Ba3(Ta1, Tc1, Tw1)
	Ta2(Tw1, Tc1, Ba5, Bw7)	Bc6(Bw6, Tw2, Tc1, Kc5)	Tc8(Tw15)	Tc8(Tw6)	

Note. BHWI1-BHWI16 and TJWI1-TJWI16 in the first column are 33 individuals of the *B. × wardii* from BH and TJ population, respectively. “Ba”, “Bc” and “Bw” refer to *B. alternifolia*, *B. crispa*, and *B. × wardii* from BH population, respectively; while “Ta”, “Tc” and “Tw” refer to *B. alternifolia*, *B. crispa*, and *B. × wardii* from TJ population, respectively. In addition, “Kc” indicates *B. crispa* from KM population and “Ja” indicates *B. alternifolia* from JZ population.

Table 3 Details of the sampling of *B. alternifolia*, *B. crispa* and *B. × wardii* in this study. The collection codes of the individual samples are given in brackets.

Taxon	Sampling location	Number of individuals
B. alternifolia	Bahe, Nyingchi, Tibet, China (BHA)	17 (BHAL1-17)
	Taji, Lhasa, Tibet, China (TJA)	15 (TJAL1-15)
	Jiangzhi, Rikaze, Tibet, China (JZA)	15 (JZAL1-15)
	Kangding, Tibetan Autonomous Prefecture of Garzê, Sichuan, China (KDA)	14 (KDAL1-10,12-15)
B. crispa	Bahe, Nyingchi, Tibet, China (BHC)	15 (BHCR1-15)
	Taji, Lhasa, Tibet, China (TJC)	16 (TJCR1-16)
	Basu, Qamdo, Tibet, China (BSC)	15 (BSCR1-15)
	Xishan, Kunming, Yunnan, China (KMC)	12 (KMCR1-12)
B. × wardii	Bahe, Nyingchi, Tibet, China (BHW)	16 (BHWI1-16)
	Taji, Lhasa, Tibet, China (TJW)	17 (TJWI1-13,15-18)

Figures

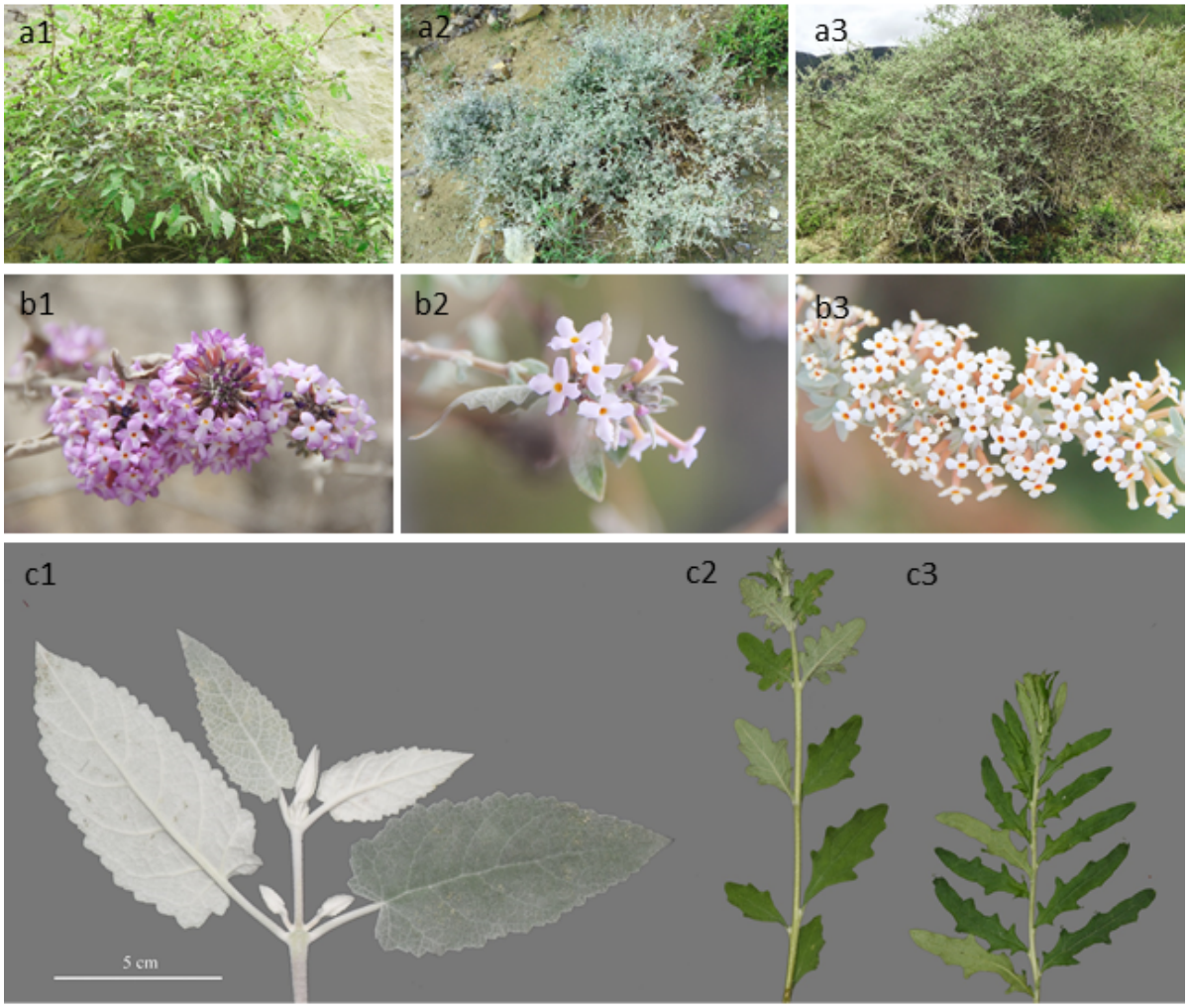

Figure 1

Morphological details of B. crispa (a1, b1, c1), B. × wardii (a2, b2, c2) and B. alternifolia (a3, b3, c3).

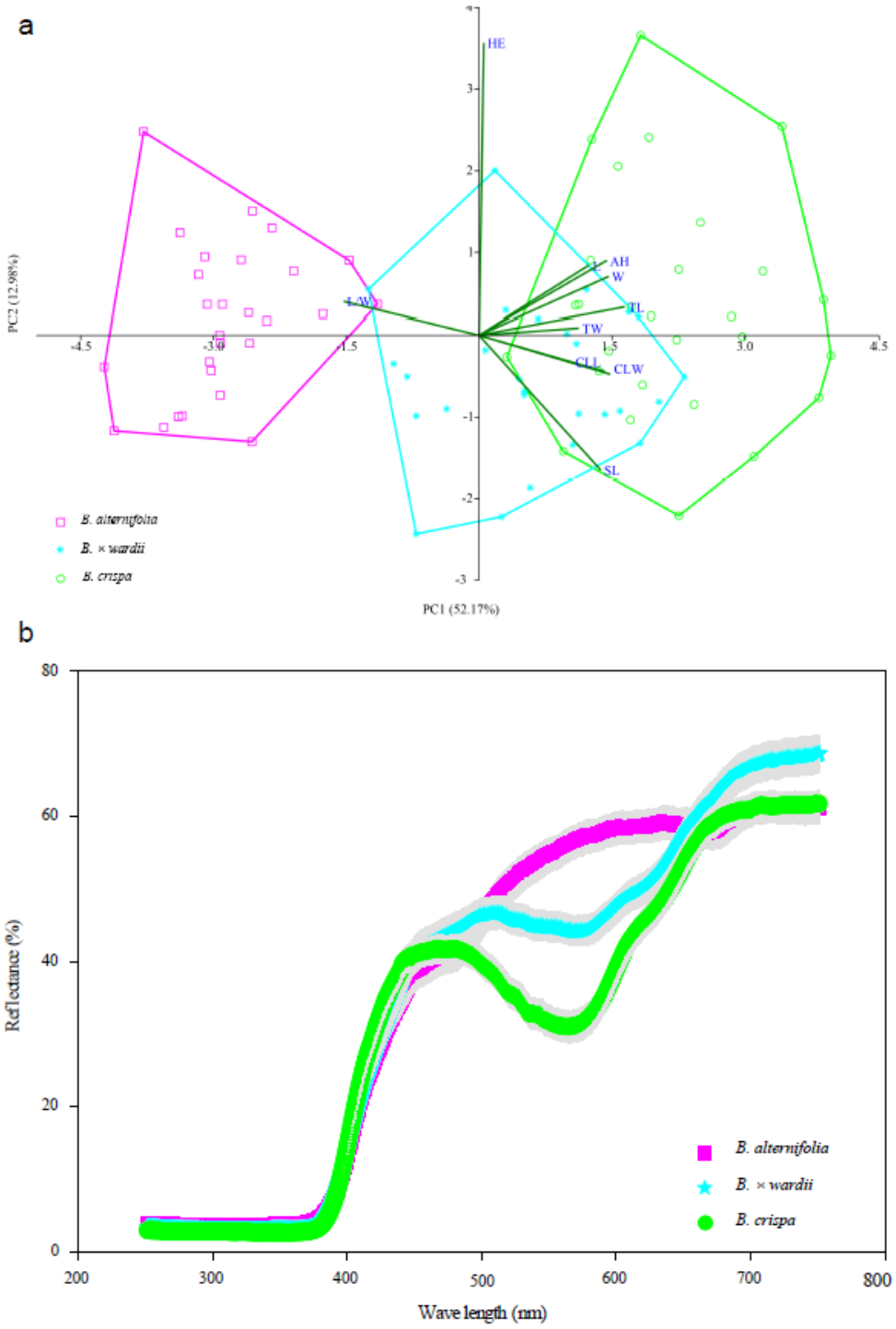

Figure 2

Two-dimensional scatter diagram of the first and second components from the PCA using 10 morphological characteristics (a), and petal reflectance spectra (b) in B. alternifolia, B. crispa, and B. × wardii. Pink square, green circle and calamine blue star represent B. alternifolia, B. crispa and B. × wardii, respectively.

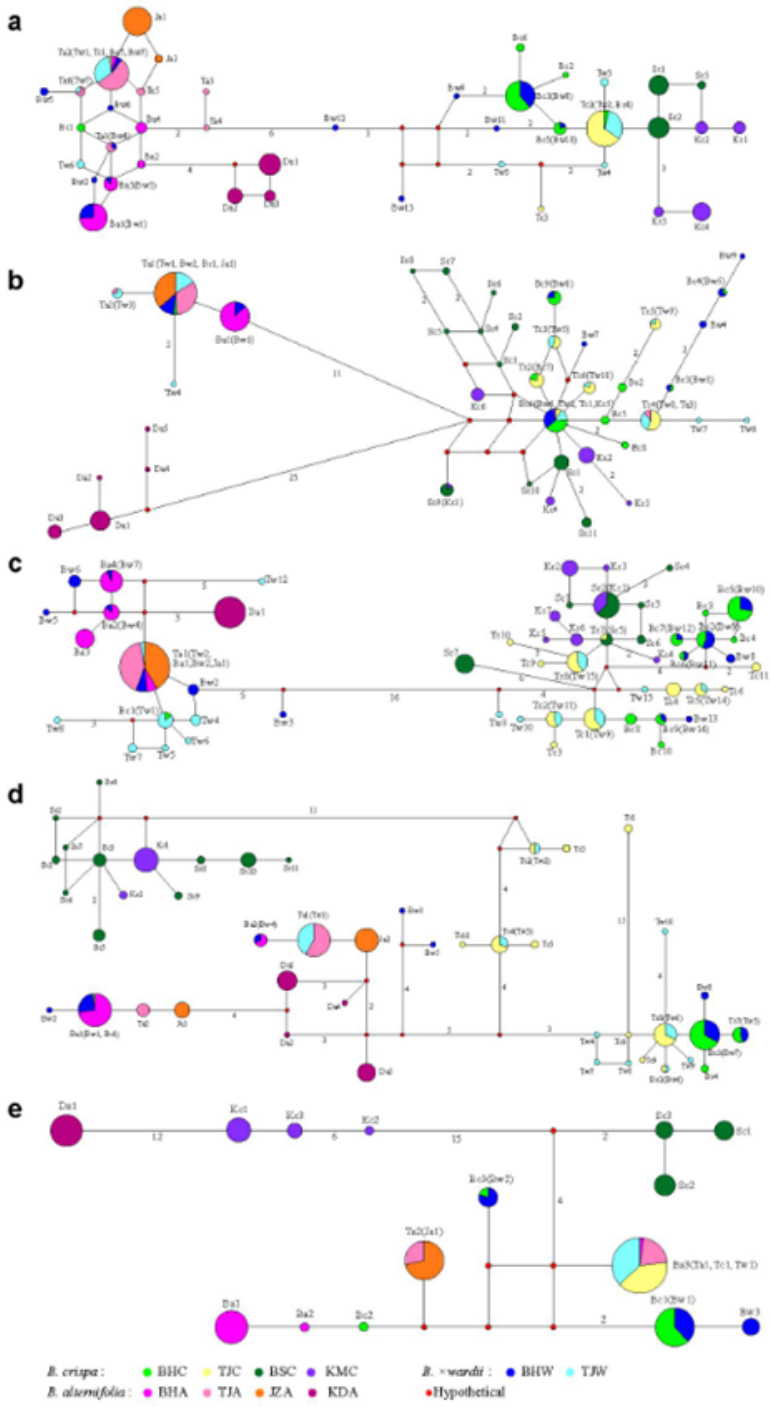

Figure 3

Haplotype network for ETS (a), gapC2 (b), PPR24 (c), PPR123 (d), cpDNA (e). Haplotypes of each taxon are denoted using the first letter of its population and species name (“Ba”, “Bc” and “Bw” refer to B. alternifolia, B. crispa, and B. × wardii from BH population, respectively; while “Ta”, “Tc” and “Tw” refer to B. alternifolia, B. crispa, and B. × wardii from TJ population, respectively; “Da”, “Ja”, “Sc” and “Kc” refer to B. alternifolia from KD and JZ, and B. crispa from BS and KM, respectively. The number of mutations separating two adjacent haplotypes is showed by the number shown on the connecting lines, the number is omitted for those with only one mutational step, and node size is proportional to the frequency of each haplotype. Color circles represent haplotypes of different species as follows: green, yellow, dark green and blue-purple represent B. crispa; pink, peach, orange and

magenta represent B. alternifolia; blue and calamine blue represent B. × wardii. Small red circles represent hypothetical or unsampled haplotypes.

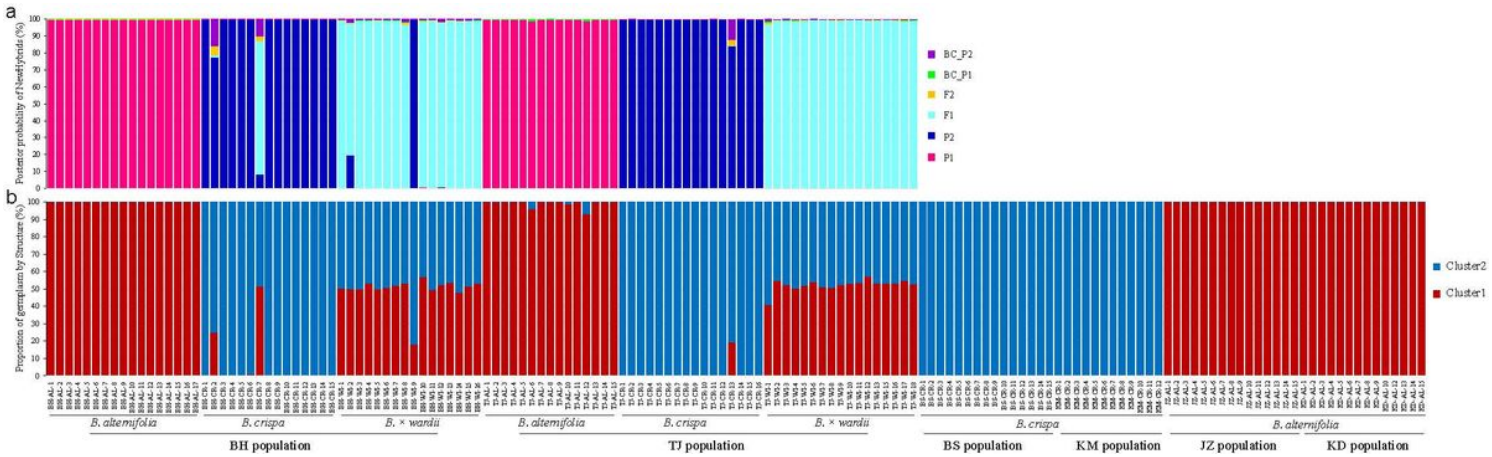

Figure 4

Genotype category assignment by NewHybrids (a) and Structure (b) for B. alternifolia, B. crispa, and B. × wardii. For the label of each accession, “BH”, “TJ”, “BS”, “KM”, “JZ” and “KD” refer to individuals collected in populations of BH, TJ, BS, KM, JZ and KD, respectively. “AL”, “WI” and “CR” refer to B. alternifolia, B. crispa, and B. × wardii, respectively. Bars in different color in a represent different genotype class: pink represent Parent □; Blue represent Parent □; calamine blue represent F1 hybrid; orange, red and purple represent F2 hybrid, back cross to Parent □ and back cross to Parent □, respectively. Bars in different color in b represent different clusters: blue and red represent Cluster 1 and Cluster 2, respectively.

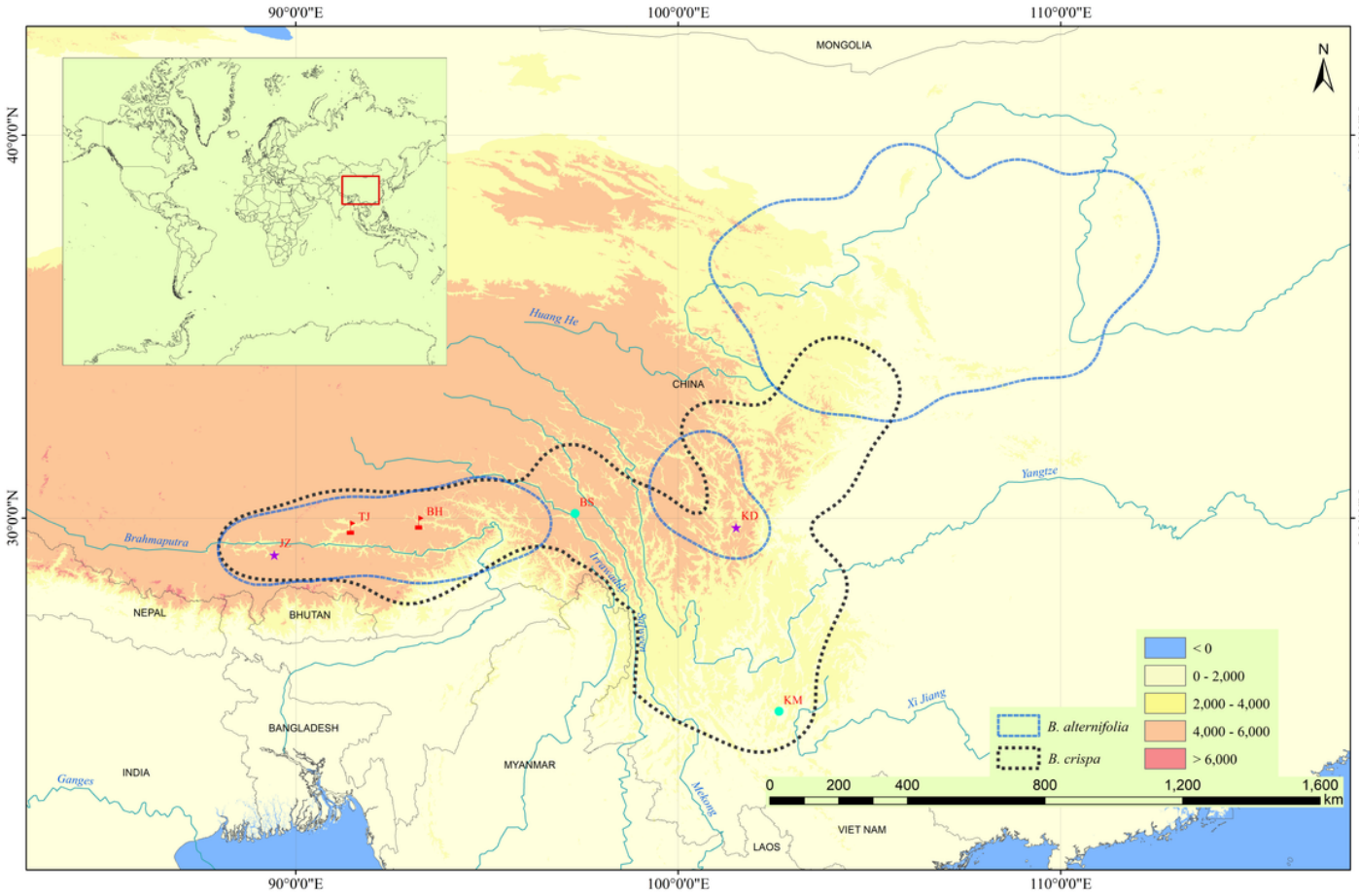

Figure 5

Known geographical distributions of B. alternifolia and B. crispa in China. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS4.xlsx
- TableS5.xlsx
- TableS6.xlsx
- TableS7.xlsx
- TableS8.xlsx

- renameda92ea.png
- FigureS1deltaK.pdf
- FigureS2.pdf