THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Natural interploidy hybridization among the key taxa involved in the origin of horticultural chrysanthemums

Citation for published version:

Qi, S, Twyford, AD, Ding, J, Borrell, JS, Wang, LZ, Ma, Y-P & Wang, N 2021, 'Natural interploidy hybridization among the key taxa involved in the origin of horticultural chrysanthemums', *Journal of Systematics and Evolution*. <https://doi.org/10.1111/jse.12810>

Digital Object Identifier (DOI):

[10.1111/jse.12810](https://doi.org/10.1111/jse.12810)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Journal of Systematics and Evolution

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Research Article

Natural interploidy hybridization among the key taxa involved in the origin of horticultural chrysanthemums

Shuai Qi^{1,2}, Alex D. Twyford^{3,4}, Jun-Yi Ding^{1,2}, James S. Borrell⁵, Long-Zhen Wang^{1,2}, Yue-Ping Ma⁶ , and Nian Wang^{1,2,7*} ¹State Forestry and Grassland Administration Key Laboratory of Silviculture in Downstream Areas of the Yellow River, College of Forestry, Shandong Agricultural University, Tai'an 271018, Shandong, China²Mountain Tai Forest Ecosystem Research Station of State Forestry and Grassland Administration, College of Forestry, Shandong Agricultural University, Tai'an 271018, Shandong, China³Institute of Evolutionary Biology, Ashworth Laboratories, University of Edinburgh, Edinburgh EH9 3FL, UK⁴Royal Botanic Garden Edinburgh, Edinburgh EH3 5LR, UK⁵Royal Botanic Gardens Kew, Richmond TW9 3AB, UK⁶College of Life and Health Sciences, Northeastern University, Shenyang 110004, China⁷State Key Laboratory of Crop Biology, Shandong Agricultural University, Tai'an 271018, Shandong, China

*Author for correspondence. E-mail: nian.wang@sdau.edu.cn

Received 26 May 2021; Accepted 30 July 2021; Article first published online 18 August 2021

Abstract Understanding hybridization and introgression between natural plant populations can give important insights into the origins of cultivated species. Recent studies suggest differences in ploidy might not create such strong reproductive barriers as once thought, and thus studies into cultivated origins should examine all co-occurring taxa, including those with contrasting ploidy levels. Here, we characterized hybridization between *Chrysanthemum indicum* L., *Chrysanthemum vestitum* (Hemsley) Ling and *Chrysanthemum vestitum* var. *latifolium* (Zhou & Chen), the most important wild species involved in the origins of cultivated chrysanthemums. We analyzed the population structure of 317 *Chrysanthemum* accessions based on 13 microsatellite markers and sequenced chloroplast *trnL-trnF* for a subset of 103 *Chrysanthemum* accessions. We identified three distinct genetic clusters, corresponding to the three taxa. We detected 20 hybrids between species of different ploidy levels, of which 19 were between *C. indicum* (4x) and *C. vestitum* (6x) and one was between *C. indicum* and *C. vestitum* var. *latifolium* (6x). Fourteen hybrids between *C. indicum* and *C. vestitum* were from one of the five study sites. *Chrysanthemum vestitum* and *C. vestitum* var. *latifolium* share only one chloroplast haplotype. The substantially different number of hybrids between hybridizing species was likely due to different levels of reproductive isolation coupled with environmental selection against hybrids. In addition, human activities could play a role in the different patterns of hybridization among populations.

Key words: *Chrysanthemum*, hybrid, microsatellite marker, symmetrical introgression, *trnL-trnF*.

1 Introduction

Hybridization has played an important role in plant domestication and diversification through human history (Heslop-Harrison & Schwarzacher, 2007; Arnold, 2014; Cornille et al., 2014). Multiple important crops have been generated through hybridization either between wild species or through introgression from crop wild relatives into cultivated lineages. Major examples include modern strawberries (*Fragaria ananassa* Duch.) (Bringhurst & Voth, 1984) and triploid bananas (*Musa* ssp. L.) (Simmonds & Shepherd, 1955; Heslop-Harrison & Schwarzacher, 2007), but also ornamental species such as tree peonies (*Paeonia*

suffruticosa Andrews) (Zhou et al., 2014), cherry (*Prunus yedoensis* Matsumura) (Baek et al., 2018), and dahlia (*Dahlia variabilis* L.) (Saar et al., 2003). Understanding the frequency, phylogenetic distribution, and propensity for hybridization in wild populations could not only inform breeders as to the possible range of interspecific hybrids, but could also reduce the laborious, time-consuming, and frequently unsuccessful process of artificial crossing and *de novo* hybrid generation (Lim et al., 2008; Kuligowska et al., 2016).

Hybridization occurs more easily between species of the same ploidy level than differing ploidy levels. For example, hybridization between diploid and tetraploid

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

species is often limited as triploid hybrids are usually inviable and less fit, preventing backcross formation (Husband & Sabara, 2003; Wang et al., 2014; Zohren et al., 2016). However, species of contrasting higher ploidy levels appear to have weaker reproductive barriers and hybridize more easily than diploids and tetraploids (Sonnleitner et al., 2013; Sutherland & Galloway, 2017). For example, within the *Campanula rotundifolia* L. polyploid complex, postzygotic isolation was lower in tetraploid–hexaploid species than in diploid–tetraploid crosses (Sutherland & Galloway, 2017). To date, however, only a small number of studies have investigated hybridization between higher ploidy levels, and the prevalence of higher level cross-ploidy hybridization across plant families remains unclear.

Chrysanthemum L. (Asteraceae) provides an excellent model for studying hybridization between high ploidy levels, with ploidy ranging from diploid ($2n = 2x = 18$) to decaploid ($2n = 10x = 90$) (Tahara, 1915; Zhou & Wang, 1997; Li et al., 2013; Ma et al., 2015; Luo et al., 2017) and with different cytotypes within species (Dowrick, 1952, 1953; Chen, 2012; Liu et al., 2012; Yan et al., 2019). *Chrysanthemum* includes approximately 37 wild species, of which 17 occur in China (Shih & Fu, 1983) where they have captured great public interest. Multiple wild species have been crossed by humans to generate numerous cultivars, and they are among the most famous Chinese flowers, with significant commercial and medicinal value (Kim & Lee, 2005; Shahrajabian et al., 2019). The evolutionary history of polyploidy within *Chrysanthemum* is currently unknown, although some polyploids are thought to be allopolyploid in origin (Chen et al., 1996; Liu et al., 2012) and subject to multiple historical polyploidization events (Yang et al., 2006).

Chrysanthemums were first cultivated in China approximately 1600 years ago, and were later introduced to Japan and Europe (Chen, 1985, 2012; Shih et al., 2011). Modern cultivated chrysanthemums are mainly hexaploids and hybridization and subsequent artificial selection are thought to give rise to numerous cultivars (Chen, 1985; Dai et al., 2002). The ancestry of modern chrysanthemums remains elusive, but several wild species are thought to be involved, including *Chrysanthemum indicum* (4x), *Chrysanthemum vestitum* (6x), *Chrysanthemum lavandulifolium* (Fischer ex Trautvetter) Kitamura (2x), *Chrysanthemum nankingense* (Handel-Mazzetti) Cui (2x), and *Chrysanthemum zawadskii* (Herbert) Tzvelev (2x) (Chen, 1985; Dai & Chen, 1997; Fukai, 2003; Ma et al., 2016, 2020). *Chrysanthemum indicum* and *C. vestitum* are key species in the origin and evolution of cultivated chrysanthemums (Dai et al., 1998, 2002) based on two lines of evidence. First, ancient literature documents multiple uses for *C. indicum* in central China, which is consistent with the geographic origin of modern cultivars (see Chen, 2012). Second, artificial hybridization between *C. indicum* and *C. vestitum* can generate hybrids resembling the prototype of modern chrysanthemums (Chen, 2012).

In this study, we investigate whether natural hybridization occurs between *Chrysanthemum* species with different ploidy levels, focusing on *C. indicum* and *C. vestitum* as well as varieties of those species. We ask: (i)

Does interploidy hybridization naturally occur between tetraploid and hexaploid *Chrysanthemum* species, which are likely involved in origin of modern cultivated chrysanthemums? (ii) If hybridization occurs, is there evidence for a greater propensity at high ploidy levels? (iii) Do some modern chrysanthemum cultivars share chloroplast haplotypes with the three wild *Chrysanthemum* taxa, consistent with their involvement in their origins? To this end, we genotyped 317 samples at 13 microsatellite markers and sequenced chloroplast *trnL-trnF* for a subset of 103 samples. In addition, we extracted *trnL-trnF* from the chloroplast genomes of 28 taxa, representing 15 wild species of *Chrysanthemum*, 12 cultivars, and one sample of *Ajania variifolia* (Chang) Tzvelev. We analyze these data in a population genetic and phylogeography context, and use them to investigate a case where, traditionally, reproductive isolation caused by differences in ploidy levels would be expected to be strong.

2 Material and Methods

2.1 Study species

Chrysanthemum indicum is tetraploid with a wide distribution across China, although narrowly distributed diploid and hexaploid cytotypes have also been reported (Liu et al., 2012). *Chrysanthemum vestitum* is a hexaploid distributed across Hubei and Henan provinces in central China (Zhao & Chen, 1999); the variety *C. vestitum* var. *latifolium* is hexaploid with a restricted distribution in the Dabie Mountains in Anhui province.

The three taxa are outcrossing perennials (Chen, 2012) and are likely to be pollinated by bees (Wang N, pers. obs., 2020). Morphologically, *C. indicum* has yellow florets and smooth leaves with deep serrations. Both *C. vestitum* and *C. vestitum* var. *latifolium* have white florets and pubescent leaves and stems (Fig. 1). Putative hybrids either with an intermediate morphology or ploidy level between *C. indicum* and *C. vestitum* have been found in localities where the species co-occur (Nakata et al., 1992; Zhao & Chen, 1999). These hybrids exhibit continuous phenotypes between *C. indicum* and *C. vestitum* (Zhao & Chen, 1999), indicating the existence of putative hybrid swarms.

2.2 Sampling across hybridizing populations

To understand the extent and distribution of hybridization, we surveyed sympatric populations of the three focal taxa: *C. indicum*, *C. vestitum*, and *C. vestitum* var. *latifolium*. We identified and sampled five populations where *C. indicum* and *C. vestitum* co-occur and two where *C. indicum* and *C. vestitum* var. *latifolium* co-occur (Table S1). In addition, we collected *C. indicum* from two allopatric populations, TA and ZP (Table S1). Samples were collected at random within populations, ensuring spacing of at least 10 m between individuals. Healthy and pest-free leaf tissue was collected and stored in silica gel. A total of 317 samples were collected including between 14 and 74 from each of the seven hybridizing populations, five from TA, and three from ZP (Table S1). A global position system (UniStrong) was used to record the coordinates of each population. Sampling locations are illustrated in Fig. 2.

Fig. 1. Diversity present in wild *Chrysanthemum* species included in this study. **A–C**, Individual photographs and leaf morphology for *Chrysanthemum indicum* (**A**), *Chrysanthemum vestitum* (**B**), and *Chrysanthemum vestitum* var. *latifolium* (**C**).

Fig. 2. Principal coordinate analysis of *Chrysanthemum* samples based on 13 microsatellite markers. Hybrids were defined based on Q scores in a STRUCTURE analysis.

2.3 Microsatellite genotyping

Genomic DNA was isolated from dried leaves of all individuals, following a modified 2x CTAB protocol (Wang et al., 2013). The quality of isolated genomic DNA was

assessed on 1.0% agarose gels, and then diluted to a concentration of 10–20 ng/μl for genotyping and sequencing. Thirteen microsatellite loci were used for genotyping (Zhang et al., 2014; Jo et al., 2015). The 5' terminus of the forward primer was labeled with FAM, HEX, or TAM fluorescent probes. Each microsatellite locus was amplified individually prior to being combined into four multiplexes (Table S2). The polymerase chain reaction (PCR) protocol follows Hu et al. (2019).

2.4 Population genetic analysis

It is often difficult to assign microsatellite genotypes for mixed ploidy species, as the frequency of different alleles can be difficult to quantify. In hexaploids, each microsatellite locus would be expected to have up to six alleles per individual. We chose to score each allele separately using the software GeneMarker 2.4.0 (SoftGenetics), and checked each genotype manually. We then calculated allele richness for each population using FSTAT 2.9.4 (Goudet, 1995) and undertook principal coordinate (PCO) analysis in POLYSAT 1.7-4 (Clark & Jasieniuk, 2011), based on Bruvo's pairwise genetic distances (Bruvo et al., 2004).

We undertook STRUCTURE analysis for each hybridizing population separately using STRUCTURE 2.3.4 (Pritchard et al., 2000) with ploidy specified as 6n. We combined XG, NX, and PH into one hybridizing population as the three localities are separated by only a few kilometers. We set the number of genetic clusters (K) to 2 when analyzing the genetic structure of each hybridizing population as only two parental species are present. The allopatric *C. indicum* populations (TA and ZP) were used as a reference population.

In addition, to identify the most likely K value across populations we included all populations in a combined STRUCTURE analysis, testing K values from 1 to 10. The number of genetic clusters was estimated using the Evanno test (Evanno et al., 2005) in the program STRUCTURE HARVESTER 0.6.94 (Earl & vonHoldt, 2012). Ten replicates of the STRUCTURE analysis were carried out with 1 000 000 iterations and a burn-in of 100 000 for each run. The admixture model, with an assumption of correlated allele frequencies, was used. Individuals were assigned to clusters based on the highest membership coefficient averaged over the 10 independent runs. Replicate runs were grouped based on a symmetrical similarity coefficient of >0.9 using the Greedy algorithm in CLUMPP 1.1.2 (Jakobsson & Rosenberg, 2007) and visualized in DISTRUCT 1.1 (Rosenberg, 2004). In populations where *C. indicum* and *C. vestitum* or *C. indicum* and *C. vestitum* var. *latifolium* co-occur, we estimated Q scores in STRUCTURE with 95% confidence intervals to define pure *C. indicum*, pure *C. vestitum*, pure *C. vestitum* var. *latifolium*, or putative hybrids. We distinguish individuals with confidence intervals overlapping 1 as pure *C. indicum*, with 0 as pure *C. vestitum*, and those remaining as putative hybrids, in populations where the two species co-occur. Similarly, in populations where *C. indicum* and *C. vestitum* var. *latifolium* co-occur, we distinguish individuals with confidence intervals overlapping with 1 as pure *C. indicum*, with 0 as pure *C. vestitum* var. *latifolium*, and those remaining as hybrids.

We compared the average allele number per individual between *C. indicum*, *C. vestitum*, *C. vestitum* var. *latifolium*,

Fig. 3. Hybridization among *Chrysanthemum indicum*, *Chrysanthemum vestitum*, and *Chrysanthemum vestitum* var. *latifolium*. **A**, Hybridization across natural populations in a combined analysis of *C. indicum*, *C. vestitum*, and *C. vestitum* var. *latifolium* with $K = 3$. **B**, Admixture in sympatric population of *C. indicum* and *C. vestitum* and each sympatric population of *C. indicum* and *C. vestitum* var. *latifolium*. Barplots on the left or the right of the STRUCTURE plot represent the number of individuals having different levels of genetic admixture. Populations XG, NX, and PH were analyzed together due to their close proximity. Allopatric populations TA and ZP served as controls. Blue, green, and orange represent *C. indicum*, *C. vestitum*, and *C. vestitum* var. *latifolium*, respectively. **C**, Admixture value for each population among *C. indicum*, *C. vestitum*, and *C. vestitum* var. *latifolium*. Hybrids were excluded for such comparisons. The map in the top left corner of (B) was supplied from the State Bureau of Surveying and Mapping (<http://bzdt.ch.mnr.gov.cn/>).

and their hybrids at each microsatellite locus, using the Kruskal–Wallis test in the R package agricolae version 1.3-3 (de Mendiburu, 2020). We would expect that the average allele number is higher for *C. vestitum* and *C. vestitum* var. *latifolium* than for *C. indicum*. In addition, we tested whether introgression was symmetric between *C. indicum* and *C. vestitum* or between *C. indicum* and *C. vestitum* var. *latifolium* in each hybridizing population, using the function wilcox.test in R version 4.0.3 (R Core Team, 2020). The putative hybrids identified as above were excluded from such comparisons.

2.5 trnL-trnF sequencing and analysis

In order to detect the potential maternal parents of the hybrids, and to estimate the haplotype diversity of these taxa, we amplified trnL-trnF for a subset of 103 samples, including between five and 34 individuals from each population. Reactions were carried out in 20 μ l volumes

containing 13 μ l ddH₂O, 5 μ l 2 $\times$ Taq PCR mix (Tiangen), 0.5 μ l each primer (trnL and trnF; Taberlet et al., 1991), and 1 μ l DNA template. The PCR products were outsourced for purification and sequencing to Qingdao. Sequences were manually edited and aligned using BioEdit version 7.2.5 (Hall, 1999). The R package pegas version 0.14 (Paradis, 2010) was used to construct haplotype networks, using default settings, with gaps treated as a fifth state. The total number of sites, polymorphic sites, parsimony informative sites, and nucleotide diversity were computed using DnaSP version 6.12.03 (Rozas et al., 2017). All sequences obtained in this study were submitted to GenBank with accession numbers MZ032043–MZ032145.

To aid in a broader phylogenetic analysis, we extracted the trnL-trnF region from the available whole chloroplast genomes for 15 wild taxa of *Chrysanthemum* occurring in China, 12 cultivars, and *Ajania variifolia*. A phylogenetic tree was estimated using the maximum likelihood method in

RAxML version 8.1.16 (Stamatakis, 2006). *Ajania variifolia* was selected as the outgroup. A rapid bootstrap analysis with 100 bootstrap replicates and 10 tree searches was carried out under the GTR + GAMMA nucleotide substitution model. The consensus tree generated from the bootstrap replicates was visualized in FigTree version 1.3.1 (Rambaut & Drummond, 2009).

3 Results

3.1 Hybridization across ploidy levels inferred from microsatellites

Genetic diversity estimates were similar among the three taxa and hybrids, with allelic richness ranging from 3.74 to 4.26 and gene diversity from 0.75 to 0.84 (Table S3). On average, the number of alleles scored in *Chrysanthemum vestitum* and *C. vestitum* var. *latifolium* was significantly higher than *Chrysanthemum indicum* at eight and seven loci, respectively ($P < 0.05$); this was expected for a hexaploid possessing more chromosome copies than a tetraploid species (Fig. S1).

The PCO analysis based on Bruvo's genetic distances among all samples revealed three clusters, with coordinates 1 and 2 explaining 13.8% and 7.8% of the total variation, respectively (Fig. 2). Coordinate 1 separated *C. indicum* from *C. vestitum* and *C. vestitum* var. *latifolium* and coordinate 2 separated *C. vestitum* from *C. vestitum* var. *latifolium* (Fig. 2). Most hybrids identified by the STRUCTURE analysis fell between *C. indicum* and *C. vestitum* in the PCO plot (Fig. 2).

The combined STRUCTURE analysis across all samples identified $K = 3$ as the optimal K value (Fig. S2A), with the three clusters corresponding to *C. indicum*, *C. vestitum*, and *C. vestitum* var. *latifolium* (Fig. 3A). At $K = 2$, *C. indicum* formed one cluster and *C. vestitum* and *C. vestitum* var. *latifolium* formed another cluster (Fig. S2B), supporting the close relationship between these two taxa and the inference that *C. vestitum* var. *latifolium* is a subspecies of *C. vestitum*. Interestingly, there are three *C. vestitum* individuals possessing a considerable level of introgression from *C. vestitum* var. *latifolium* and one *C. indicum* individual in population TZ with substantial introgression from *C. vestitum* (Fig. 3A).

A total of 20 hybrids were detected when STRUCTURE was used for each hybridizing population separately, of which 19 were hybrids between *C. indicum* and *C. vestitum* and one was a hybrid between *C. indicum* and *C. vestitum* var. *latifolium*. Fourteen hybrids between *C. indicum* and *C. vestitum* were from population XG, two from each of NX and PH and one from WFS. The only hybrid between *C. indicum* and *C. vestitum* var. *latifolium* was from population TZ. No hybrids were detected from population LJS where *C. indicum* and *C. vestitum* co-occur and from BMJ where *C. indicum* and *C. vestitum* var. *latifolium* co-occur (Fig. 3B).

Introgression occurred symmetrically between *C. indicum* and *C. vestitum* and between *C. indicum* and *C. vestitum* var. *latifolium* in all hybridizing populations except WFS, where only three *C. indicum* were collected (Fig. 3C). Introgression was limited in populations TZ, BMJ, and LJS but was extensive in population XG (Fig. 3C).

3.2 Plastid diversity and directionality of hybrid formation

A total of 17 chloroplast haplotypes were detected across samples, with 12, nine, and four in *C. indicum*, *C. vestitum*, and *C. vestitum* var. *latifolium*, respectively (Fig. 4; Table S4). Both *C. indicum* and *C. vestitum* harbored three private haplotypes and *C. vestitum* var. *latifolium* harbored one (haplotype H3). *Chrysanthemum indicum* shared six haplotypes with *C. vestitum* and three haplotypes with *C. vestitum* var. *latifolium*. By contrast, *C. vestitum* only shared the most common haplotype, H6, with *C. vestitum* var. *latifolium*. Frequencies of particular haplotypes varied considerably between species as some haplotypes found in *C. indicum* or *C. vestitum* were absent in *C. vestitum* var. *latifolium* (Fig. 4). Three haplotypes (H6, H16, and H17) were found in hybrids, with haplotype H6 shared among the three taxa, and haplotype H16 found to be private to the hybrid (Fig. 4). One hybrid shared haplotype H17 with *C. indicum*, whereas this haplotype was absent from all *C. vestitum* and *C. vestitum* var. *latifolium* samples (Fig. 4).

All modern *Chrysanthemum morifolium* Ramat. cultivars formed a clade with full support, and this clade was nested in a large monophyletic clade including *Chrysanthemum dichrum* Shih, *Chrysanthemum zawadskii*, and *Chrysanthemum chanetii* (Léveillé) Shih (Fig. 5). Five haplotypes (1, 3, 6, 16, and 17) from the three taxa and hybrid were also nested in this monophyletic clade (Fig. 5).

4 Discussion

In this study, we provide genetic evidence of natural interploidy hybridization between two pairs of taxa involved in the formation of modern *Chrysanthemum* horticultural hybrids. We detected more hybrids between *Chrysanthemum indicum* and *Chrysanthemum vestitum* than between *C. indicum* and *C. vestitum* var. *latifolium*, possibly due to different levels of reproductive isolation. In addition, we show that *C. vestitum* var. *latifolium* formed a genetic cluster distinct from *C. vestitum*, and as such deserves its varietal status. Here, we discuss the importance of hybridization between different ploidy levels in natural *Chrysanthemum* populations, before considering the dynamics of different hybrid swarms. We finish with the wider implications of our findings for understanding the origin of horticultural chrysanthemum hybrids.

4.1 Tetraploid–hexaploid hybridization and symmetrical introgression

Interploidy hybridization is more common in genera containing many polyploids, and where species readily co-occur and hybridize, such as *Betula* (Zohren et al., 2016; Hu et al., 2019) and *Spartina* (Ainouche et al., 2003). Diploid–tetraploid hybridization produces mostly sterile triploids; however, where hybrids are fertile, introgression usually occurs preferentially from diploids to tetraploids (Pinheiro et al., 2010; Moraes et al., 2013). In contrast, hybridization between tetraploids and hexaploids has been proposed to be easier, as pentaploid hybrids are formed frequently (Hülber et al., 2015) and are more fertile than triploids (Sutherland & Galloway, 2017). Consistent with this, we detected 20 hybrids among *C. indicum*, *C. vestitum*, and

Fig. 4. Haplotype network based on plastid *trnL-trnF* sequences in natural *Chrysanthemum* populations. **A**, Geographic distribution of haplotypes. Numbers in brackets indicate the number of samples used for sequencing. **B**, Haplotype network graph. Each haplotype is represented by a circle with size proportional to the number of individuals. Color within the circle represents species sharing the haplotype. The map in the top left corner of (A) was supplied from the State Bureau of Surveying and Mapping (<http://bzdt.ch.mnr.gov.cn/>).

C. vestitum var. *latifolium*, indicating incomplete reproductive isolation. The number of hybrids is likely to be underestimated in our study, because of the stringent confidence threshold applied in the *STRUCTURE* analysis. Three individuals with admixture between 26.2% and 31.4% from population XG and two individuals with admixture between 34.5% and 48.7% were not supported by Q scores with 95% confidence intervals, but could prove to be hybrids such as later generation backcrosses.

Nineteen of the 20 hybrids were between *C. indicum* and *C. vestitum* and one was between *C. indicum* and *C. vestitum* var. *latifolium*. This difference might reflect different levels of reproductive isolation. Higher fruit set in artificial *C. indicum*–*C. vestitum* crosses than in artificial *C. indicum*–*C. vestitum* var. *latifolium* crosses partially supports this hypothesis (Zhou, 2009). However, environmental selection against hybrids between *C. indicum* and *C. vestitum* var. *latifolium* could also account for its rarity. *Chrysanthemum indicum* and *C. vestitum* occupy similar habitats and are usually intermixed in sympatric populations (Qi S, pers. obs., 2019). This could enhance their opportunity for hybridization, while their similar habitat preferences might reduce the chance of ecological selection on the hybrids. In contrast, *C. indicum* and *C. vestitum* var. *latifolium* are adapted to different conditions, and hybrids might

fail to survive due to the breakdown of suites of coadapted genes.

Within some diploid–tetraploid systems, introgression is more common from diploids to tetraploids (Zohren et al., 2016; Wang et al., 2020). However, we observe symmetrical introgression between *C. indicum* and *C. vestitum* and between *C. indicum* and *C. vestitum* var. *latifolium*, indicating that hybrids can backcross with both parents. This is in line with recent studies showing that pentaploids can mediate gene flow between species with different ploidy levels (Peskoller et al., 2021). We could not distinguish clearly between male and female parental taxa because *C. indicum* and *C. vestitum* share some plastid haplotypes. However, *C. indicum* can serve as maternal parent as one hybrid shared a haplotype with *C. indicum*, whereas this haplotype is absent from all *C. vestitum* and *C. vestitum* var. *latifolium* samples. In addition, one hybrid had a unique haplotype (H16), which is possibly from unsampled *C. indicum* and *C. vestitum* or introgressed from other *Chrysanthemum* species.

4.2 Variable hybridizing among populations

The number of hybrids between *C. indicum* and *C. vestitum* varied substantially among the hybridizing populations (Fig. 3B). Fourteen of 19 hybrids were in population XG,

widespread and abundant, whereas *C. vestitum* var. *latifolium* is restricted to the Dabie Mountains. Hybridization between abundant *C. indicum* and rare *C. vestitum* var. *latifolium* could be predicted to drive the rare species to extinction through genetic or demographic swamping (Todesco et al., 2016).

Unexpectedly, one hybrid between *C. indicum* and *C. vestitum* var. *latifolium* is from population TZ and none are from BMJ. Moreover, the extent of genetic admixture in TZ seems to be higher than in BMJ (Fig. S1). In TZ, *C. indicum* and *C. vestitum* var. *latifolium* grow closely together, meaning there are enhanced opportunities for hybridization, and likely relaxed selection against hybrids. However, in BMJ, *C. indicum* and *C. vestitum* var. *latifolium* are segregated by altitude; this could limit the survival of hybrids due to environmental selection.

Hybrids usually occur in sympatry with their parental species, but sometimes they occur separately, due to (often human-mediated) long-distance dispersal of hybrids, or natural colonization of sterile hybrid taxa (James & Abbott, 2005). Alternatively, parental species might die out due to genetic swamping or competitive exclusion (Levin et al., 1996; Huxel, 1999). Regardless of the mechanism, these result in orphan hybrids (Marques et al., 2010; Groh et al., 2019).

We detected a few individuals showing considerable admixture between *C. vestitum* var. *latifolium* and *C. vestitum* in populations WFS and XG, where *C. vestitum* var. *latifolium* is not known to occur. We also detected one *C. indicum* individual showing considerable admixture from *C. vestitum* in population TA in Shandong province.

This indicates the presence of hybrids in the absence of one or both parental species, which has been reported in some plant species, such as oaks (Dodd & Afzal-Rafii, 2004) and pines (Lanner & Phillips, 1992). A plausible explanation is the existence of undetected *C. vestitum* var. *latifolium* near populations WFS and XG, or within travelling distance of pollinators, or seed dispersal.

4.4 Implications for the origins of cultivated chrysanthemums

Our results based on an analysis of plastid *trnL-trnF* showed a monophyletic clade composed of *C. lavandulifolium*, *C. chanetii*, *C. zawadskii*, and five haplotypes of *C. indicum*, *C. vestitum*, *C. vestitum* var. *latifolium*, and the hybrid between *C. indicum* and *C. vestitum* (Fig. 5). This indicates that either of these species potentially acted as the maternal parent of cultivated chrysanthemums. This is consistent with previous research implicating their involvement (Dai et al., 1998, 2005; Fukai, 2003). However, the specific maternal parental progenitor of modern cultivated chrysanthemums remains elusive. The modern cultivated chrysanthemums were placed in a monophyletic clade that was sister to *C. lavandulifolium*. However, the chloroplast genome of *C. lavandulifolium* has some unique mutations compared with cultivated chrysanthemums. This leads to the hypothesis that the maternal progenitor of modern cultivated chrysanthemums has gone extinct (Ma et al., 2020). However, this requires further evaluation as only one or two whole chloroplast genomes were included for each wild *Chrysanthemum* species and cultivar. This means the diversity of chloroplast genomes was not sufficiently represented. Given the high haplotype diversity of the three taxa and the fact that these haplotypes did not form a monophyletic clade (Fig. 5), we predict that the ultimate maternal progenitor of modern cultivated chrysanthemums could potentially be any of the wild *Chrysanthemum* species, and there could be multiple maternal progenitors for chrysanthemum cultivars.

Acknowledgements

This work was funded by the National Natural Science Foundation of China (31770230, 31600295, and 31872710), Natural Science Foundation of Shandong Province, China (ZR2018PC022), and Funds of Shandong “Double Tops” Program (SYL2017XTTD13).

References

- Ainouche ML, Baumel A, Salmon A, Yannic G. 2003. Hybridization, polyploidy and speciation in *Spartina* (Poaceae). *New Phytologist* 161: 165–172.
- Arnold ML. 2014. Natural hybridization and the evolution of domesticated, pest and disease organisms. *Molecular Ecology* 13: 97–1007.
- Baek S, Choi K, Kim GB, Yu HJ, Cho A, Jang H, Kim C, Kim HJ, Chang KS, Kim JH, Mun JH. 2018. Draft genome sequence of wild *Prunus yedoensis* reveals massive inter-specific hybridization between sympatric flowering cherries. *Genome Biology* 19: 127.
- Bringham RS, Voth V. 1984. Breeding octoploid strawberries. *Iowa State Journal of Research* 58: 371–382.
- Bruvo R, Michiels NK, D'Souza TG, Schulenburg H. 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology* 13: 2101–2106.
- Chen F, Chen P, Li H. 1996. Genome analysis and their phylogenetic relationships of several wild species of *Dendranthema* in China. *Acta Horticulturae Sinica* 23: 67–72.
- Chen JY. 1985. Studies on the origin of Chinese florist's chrysanthemum. *Acta Horticulturae* 167: 349–361.
- Chen JY. 2012. *The origin of garden chrysanthemum*. Hefei: Anhui Science & Technology Publishing House.
- Clark LV, Jasieniuk M. 2011. POLYSAT: An R package for polyploid microsatellite analysis. *Molecular Ecology Resources* 11: 562–566.
- Cornille A, Giraud T, Smulders MJM, Roldán-Ruiz I, Gladieux P. 2014. The domestication and evolutionary ecology of apples. *Trends in Genetics* 30: 57–65.
- Dai SL, Chen JY. 1997. Cladistic study on some *Chrysanthemum* spp. in China. *Journal of Wuhan Botanical Research* 15: 27–34.
- Dai SL, Chen JY, Li WB. 1998. Application of RAPD analysis in the study on the origin of Chinese cultivated chrysanthemum. *Acta Botanica Sinica* 40: 1053–1059.
- Dai SL, Wang WK, Huang JP. 2002. Advance of researches on phylogeny of *Dendranthema* and origin of chrysanthemum. *Acta Scientiarum Naturalium Universitatis Pekinensis* 24: 230–234.
- Dai SL, Wang WK, Xu YX. 2005. Phylogenetic relationship of *Dendranthema* (DC.) Des Moul. revealed by fluorescent in situ hybridization. *Journal of Integrative Plant Biology* 7: 783–791.
- de Mendiburu F. 2020. agricolae: Statistical procedures for agricultural research. R package version 1.3-3 [online]. Available from <https://CRAN.r-project.org/package=agricolae> [accessed 30 March 2021].
- Dodd RS, Afzal-Rafii Z. 2004. Selection and dispersal in a multispecies oak hybrid zone. *Evolution* 58: 261–269.
- Dowrick GJ. 1952. The chromosomes of *Chrysanthemum* I: The species. *Heredity* 6: 365–375.
- Dowrick GJ. 1953. The chromosomes of *Chrysanthemum* II: Garden varieties. *Heredity* 7: 59–72.
- Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.
- Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14: 2611–2620.
- Fukai S. 2003. *Dendranthema* species as chrysanthemum genetic resources. *Acta Horticulturae* 620: 223–230.
- Goudet J. 1995. FSTAT (version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* 86: 485–486.
- Groh JS, Percy DM, Björk CR, Cronk QCB. 2019. On the origin of orphan hybrids between *Aquilegia formosa* and *Aquilegia flavescens*. *AoB Plants* 11: ply071.
- Hülber K, Sonnleitner M, Suda J, Krejčíková J, Schönschetter P, Schneeweiss GM, Winkler M. 2015. Ecological differentiation, lack of hybrids involving diploids, and asymmetric gene flow between polyploids in narrow contact zones of *Senecio carniolicus* (syn. *Jacobaea carniolica*, Asteraceae). *Ecology and Evolution* 5: 1224–1234.
- Hall TA. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.
- Heslop-Harrison JS, Schwarzacher T. 2007. Domestication, genomics, and the future for banana. *Annals of Botany* 100: 1073–1084.

- Hu YN, Zhao L, Buggs RJA, Zhang XM, Li J, Wang N. 2019. Population structure of *Betula albosinensis* and *Betula platyphylla*: Evidence for hybridization and a cryptic lineage. *Annals of Botany* 123: 1179–1189.
- Husband BC, Sabara HA. 2003. Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, *Chamerion angustifolium* (Onagraceae). *New Phytologist* 161: 703–713.
- Huxel GR. 1999. Rapid displacement of native species by invasive species: Effects of hybridization. *Biological Conservation* 89: 143–152.
- 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.
- James JK, Abbott RJ. 2005. Recent, allopatric, homoploid hybrid speciation: The origin of *Senecio squalidus* (Asteraceae) in the British Isles from a hybrid zone on Mount Etna, Sicily. *Evolution* 59: 2533–2547.
- Jo KM, Jo Y, Chu H, Lian S, Cho WK. 2015. Development of EST-derived SSR markers using next-generation sequencing to reveal the genetic diversity of 50 chrysanthemum cultivars. *Biochemical Systematics and Ecology* 60: 37–45.
- Kim HJ, Lee YS. 2005. Identification of new dicaffeoylquinic acids from *Chrysanthemum morifolium* and their antioxidant activities. *Planta Medica* 71: 871–876.
- Kuligowska K, Lütken H, Müller R. 2016. Towards development of new ornamental plants: Status and progress in wide hybridization. *Planta* 244: 1–17.
- Lanner RM, Phillips AMI. 1992. Natural hybridization and introgression of pinyon pines in northwestern Arizona. *International Journal of Plant Sciences* 153: 250–257.
- Levin DA, Francisco-Ortega J, Jansen RK. 1996. Hybridization and the extinction of rare plant species. *Conservation Biology* 10: 10–16.
- Li J, Wan Q, Abbott RJ, Rao GY. 2013. Geographical distribution of cytotypes in the *Chrysanthemum indicum* complex as evidenced by ploidy level and genome-size variation. *Journal of Systematics and Evolution* 51: 196–204.
- Lim KB, Barba-Gonzalez R, Zhou S, Ramanna MS, van Tuyl JM. 2008. Interspecific hybridization in Lily (*Lilium*): Taxonomy and commercial aspects of using species hybrids in breeding. In: da Silva JAT ed. *Floriculture, ornamental and plant biotechnology*. Japan: Global Science Books. 146–151.
- Liu PL, Wan Q, Guo YP, Yang J, Rao GY. 2012. Phylogeny of the genus *Chrysanthemum* L.: Evidence from single-copy nuclear gene and chloroplast DNA sequences. *PLoS One* 7: e48970.
- Luo C, Chen D, Cheng X, Zhao H, Huang C. 2017. Genome size estimations in *Chrysanthemum* and correlations with molecular phylogenies. *Genetic Resources and Crop Evolution* 64: 1451–1463.
- Ma Y, Wei J, Yu Z, Qin B, Dai S. 2015. Characterization of ploidy levels in *Chrysanthemum* L. by flow cytometry. *Journal of Forestry Research* 26: 771–775.
- Ma Y, Zhao L, Zhang W, Zhang Y, Xing X, Duan X, Hu J, Harris AJ, Liu P, Dai S, Wen J. 2020. Origins of cultivars of *Chrysanthemum* – Evidence from the chloroplast genome and nuclear *LFY* gene. *Journal of Systematics and Evolution* 58: 925–944.
- Ma YP, Chen MM, Wei JX, Zhao L, Liu PL, Dai SL, Wen J. 2016. Origin of *Chrysanthemum* cultivars – Evidence from nuclear low-copy *LFY* gene sequences. *Biochemical Systematics and Ecology* 65: 129–136.
- Marques I, Feliner GN, Draper Munt D, Martins-Loução MA, Aguilar JF. 2010. Unraveling cryptic reticulate relationships and the origin of orphan hybrid disjunct populations in *Narcissus*. *Evolution* 64: 2353–2368.
- Moraes AP, Chinaglia M, Clarisse Palma-Silva C, Pinheiro F. 2013. Interoecological hybridization in sympatric zones: The formation of *Epidendrum fulgens* × *E. puniceolutes* hybrids (Epidendroideae, Orchidaceae). *Ecology and Evolution* 3: 3824–3837.
- Nakata M, Hong D, Qiu J, Uchiyama H, Tanaka R, Chen X. 1992. Cytogenetic studies on wild *Chrysanthemum sensu lato* in China. II. A natural hybrid between *Dendranthema indicum* (2n = 36) and *D. vestitum* (2n = 54) from Hubei Province. *Journal of Japan Botany* 67: 92–100.
- Paradis E. 2010. pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* 26: 419–420.
- Peskoller A, Silbernagl L, Hülber K, Sonnleitner M, Schönschwetter P. 2021. Do pentaploid hybrids mediate gene flow between tetraploid *Senecio disjunctus* and hexaploid *S. carniolicus* s. str. (*S. carniolicus* aggregate, Asteraceae)? *Alpine Botany*. <https://doi.org/10.1007/s00035-021-00254-x>
- Pinheiro F, de Barros F, Palma-Silva C, Meyer D, Fay MF, Suzuki RM, Lexer C, Cozzolino S. 2010. Hybridization and introgression across different ploidy levels in the Neotropical orchids *Epidendrum fulgens* and *E. puniceolutes* (Orchidaceae). *Molecular Ecology* 19: 3981–3994.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- R Core Team. 2020. R: A language and environment for statistical computing Version 4.0.3. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rambaut A, Drummond AJ. 2009. FigTree version 1.3.1 [computer program] [online]. Available from <http://tree.bio.ed.ac.uk> [accessed 29 March 2021].
- Rosenberg NA. 2004. DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes* 4: 137–138.
- Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* 34: 3299–3302.
- Saar DE, Polans NO, Sørensen PD. 2003. A phylogenetic analysis of the genus *Dahlia* (Asteraceae) based on internal and external transcribed spacer regions of nuclear ribosomal DNA. *Systematic Botany* 28: 627–639.
- Shahrajabian MH, Sun W, Zandi P, Cheng Q. 2019. A review of *Chrysanthemum*, the eastern queen in traditional Chinese medicine with healing power in modern pharmaceutical sciences. *Applied Ecology and Environmental Research* 17: 13355–13369.
- Shih C, Christopher JH, Michael GG. 2011. *Chrysanthemum*. In: Wu ZY, Raven PH, Hong DY eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press. 20: 669–676.
- Shih C, Fu GX. 1983. *Flora Reipublicae Popularis Sinicae*. 76. Beijing: Science Press.
- Simmonds NW, Shepherd K. 1955. The taxonomy and origins of the cultivated bananas. *Botanical Journal of the Linnean Society* 55: 302–312.
- Sonnleitner M, Weis B, Flatscher R, Escobar Garcia P, Suda J, Krejčíková J, Schneeweiss GM, Winkler M, Schönschwetter P, Hülber K. 2013. Parental ploidy strongly affects offspring fitness in heteroploid crosses among three cytotypes of autopolyploid *Jacobaea carniolica* (Asteraceae). *PLoS One* 8: e78959.
- Stamatakis A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.

- Sutherland BL, Galloway LF. 2017. Postzygotic isolation varies by ploidy level within a polyploid complex. *New Phytologist* 213: 404–412.
- Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.
- Tahara M. 1915. Cytological studies in *Chrysanthemum* (a preliminary note). *Botanical Magazine (Tokyo)* 29: 48–50.
- Todesco M, Pascual MA, Owens GL, Ostevik KL, Moyers BT, Hübner S, Heredia SM, Hahn MA, Caseys C, Bock DG, Rieseberg LH. 2016. Hybridization and extinction. *Evolutionary Applications* 9: 892–908.
- Wang J, Dong S, Yang L, Harris A, Schneider H, Kang M. 2020. Allopolyploid speciation accompanied by gene flow in a tree fern. *Molecular Biology and Evolution* 37: 2487–2502.
- Wang N, Borrell JS, Bodles WJA, Kuttapitiya A, Nichols RA, Buggs RJA. 2014. Molecular footprints of the Holocene retreat of dwarf birch in Britain. *Molecular Ecology* 23: 2771–2782.
- Wang N, Thomson M, Bodles WJA, Crawford RMM, Hunt HV, Featherstone AW, Pellicer J, Buggs RJA. 2013. Genome sequence of dwarf birch (*Betula nana*) and cross-species RAD markers. *Molecular Ecology* 22: 3098–3111.
- Yan W, Jung JA, Lim K, Cabahug RAM, Hwang Y. 2019. Cytogenetic studies of chrysanthemum: A review. *Flower Research Journal* 27: 242–253.
- Yang W, Glover BJ, Rao GY, Yang J. 2006. Molecular evidence for multiple polyploidization and lineage recombination in the *Chrysanthemum indicum* polyploid complex (Asteraceae). *New Phytologist* 171: 875–886.
- Zhang Y, Dai SL, Hong Y, Song XB. 2014. Application of genomic SSR locus polymorphisms on the identification and classification of *Chrysanthemum* cultivars in China. *PLoS One* 9: e104856.
- Zhao HE, Chen JY. 1999. Investigation and studies on some native species of *Dendranthema*. *Chinese Landscape Architecture* 3: 60–61.
- Zhou J. 2009. *Studies on the problem of origin of Chinese garden chrysanthemum*. Beijing: Beijing Forestry University.
- Zhou SJ, Wang JW. 1997. The cytological study on ten species of *Dendranthema*. *Journal of Wuhan Botanical Research* 15: 289–292.
- Zhou SL, Zou XH, Zhou ZQ, Liu J, Xu C, Yu J, Wang Q, Zhang DM, Wang XQ, Ge S, Sang T, Pan KY, Hong DY. 2014. Multiple species of wild tree peonies gave rise to the ‘king of flowers’, *Paeonia suffruticosa* Andrews. *Proceedings of the Royal Society B: Biological Sciences* 281: 1687–1694.
- Zohren J, Wang N, Kardailsky I, Borrell JS, Joecker A, Nichols RA, Buggs RJA. 2016. Unidirectional diploid-tetraploid introgression among British birch trees with shifting ranges shown by restriction site-associated markers. *Molecular Ecology* 25: 2413–2426.

Supplementary Material

The following supplementary material is available online for this article at <http://onlinelibrary.wiley.com/doi/10.1111/jse.12810/supinfo>:

Fig. S1. Allele number per individual at each of the 13 microsatellite loci among *Chrysanthemum indicum*, *Chrysanthemum vestitum*, *Chrysanthemum vestitum* var. *latifolium*, and hybrids. The difference in the number of alleles was assessed using the Kruskal–Wallis test. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Fig. S2. Output of STRUCTURE HARVESTER showing that $K = 3$ is the optimal value (A) and STRUCTURE results at $K = 2$ and 3 (B).

Table S1. Detailed sampling information of *Chrysanthemum indicum*, *Chrysanthemum vestitum*, and *Chrysanthemum vestitum* var. *latifolium*.

Table S2. Details of microsatellite markers used in this study.

Table S3. Genetic diversity of 14 populations of *Chrysanthemum indicum*, *Chrysanthemum vestitum*, and *Chrysanthemum vestitum* var. *latifolium* based on microsatellites.

Table S4. Chloroplast haplotypes and their variable sites based on trnL-trnF sequences.