

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

3

4 Running title: Interploidy hybridization among *Chrysanthemum*

5

6 Shuai Qi¹, Alex D. Twyford^{2,3}, Junyi Ding¹, James S. Borrell⁴, Yueping Ma⁵, Nian
7 Wang^{1,5,7*}

8

9 ¹State Forestry and Grassland Administration Key Laboratory of Silviculture in
10 Downstream Areas of the Yellow River, College of Forestry, Shandong Agricultural
11 University, Tai'an 271018, China.

12 ²Institute of Evolutionary Biology, Ashworth Laboratories, the University of
13 Edinburgh, Edinburgh EH9 3FL, UK.

14 ³Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, UK.

15 ⁴Royal Botanic Gardens Kew, Richmond, Surrey TW9 3AB, UK.

16 ⁵College of Life and Health Sciences, Northeastern University, Shenyang 110004,
17 China.

18 ⁶Mountain Tai Forest Ecosystem Research Station of State Forestry and Grassland
19 Administration, College of Forestry, Shandong Agricultural University, Tai'an
20 271018, China.

21 ⁷State Key Laboratory of Crop Biology, Shandong Agricultural University, Tai'an,
22 China.

23 *Author for correspondence:

24 Nian Wang; Email: nian.wang@sdau.edu.cn

25

26 **Abstract**

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

46

47 **Keywords** *Chrysanthemum*, hybrid, microsatellite marker, symmetrical introgression,
48 *trnL-trnF*

49

50

51 **1. Introduction**

52 Hybridization has played an important role in plant domestication and diversification
53 through human history (Arnold, 2014; Heslop-Harrison and Schwarzacher, 2007;
54 Cornille et al., 2014). Multiple important crops have been generated through
55 hybridization either between wild species or through introgression from crop wild
56 relatives into cultivated lineages. Major examples include modern strawberries
57 (*Fragaria ananassa*) (Bringhurst and Voth, 1984) and triploid bananas (Simmonds
58 and Shepherd, 1955; Heslop-Harrison and Schwarzacher, 2007), but also ornamental
59 species such as tree peonies (*Paeonia suffruticosa*) (Zhou et al., 2014), cherry (*Prunus*
60 *yedoensis*) (Baek et al., 2018) and dahlia (*Dahlia variabilis*) (Saar et al., 2003).

61 Understanding the frequency, phylogenetic distribution and propensity for
62 hybridization in wild populations could not only inform breeders as to the possible
63 range of interspecific hybrids, but could also reduce the laborious, time consuming
64 and frequently unsuccessful process of artificial crossing and *de novo* hybrid
65 generation (Lim et al., 2008; Kuligowska et al., 2016).

66 Hybridization occurs more easily between species of the same ploidy level than
67 differing ploidy levels. For example, hybridization between diploid and tetraploid
68 species is often limited as triploid hybrids are usually inviable and less fit, preventing
69 backcross formation (Wang et al., 2014; Zohren et al., 2016; Husband and Sabara,
70 2003). However, species of contrasting higher ploidy levels appear to have weaker
71 reproductive barriers and hybridize more easily than diploids and tetraploids
72 (Sonnleitner et al., 2013; Sutherland and Galloway, 2017). For example, within the
73 *Campanula rotundifolia* polyploid complex, postzygotic isolation was lower in
74 tetraploid–hexaploid species than in diploid–tetraploid crosses (Sutherland and
75 Galloway, 2017). To date only a small number of studies have investigated

76 hybridization between higher ploidy levels, and the prevalence of higher level
77 cross-ploidy hybridization across plant families remains unclear.

78 *Chrysanthemum* L. (Asteraceae) provides an excellent model for studying
79 hybridization between high ploidy levels, with ploidy ranging from diploid ($2n = 2x =$
80 18) to decaploid ($2n = 10x = 90$) (Ma et al., 2015; Zhou and Wang, 1997; Tahara,
81 1915; Li et al., 2013; Luo et al., 2017) and with different cytotypes within species
82 (Chen, 2012; Yan et al., 2019; Dowrick, 1952; Dowrick, 1953; Liu et al., 2012).

83 *Chrysanthemum* includes approximately 37 wild species, of which 17 occur in China
84 (Shih and Fu, 1983) where they have captured great public interest. Multiple wild
85 species have been crossed by humans to generate numerous cultivars, and they are
86 among the most famous Chinese flowers, with significant commercial and medicinal
87 value (Kim and Lee, 2005; Shahrajabian et al., 2019). The evolutionary history of
88 polyploidy within *Chrysanthemum* is currently unknown, though some polyploids are
89 thought to be allopolyploid in origin (Liu et al., 2012; Chen et al., 1996) and subject
90 to multiple historical polyploidization events (Yang et al., 2006).

91 Chrysanthemums were first cultivated in China ~1600 years ago, and were later
92 introduced to Japan and Europe (Chen, 2012; Shih et al., 2011; Chen, 1985). Modern
93 cultivated chrysanthemums are mainly hexaploids and hybridization and subsequent
94 artificial selection are thought to give rise to numerous cultivars (Chen, 1985; Dai et
95 al., 2002). The ancestry of modern chrysanthemums remains elusive, but several wild
96 species are thought to be involved, including *C. indicum* (4x), *C. vestitum* (6x), *C.*
97 *lavandulifolium* (2x), *C. nankingense* (2x) and *C. zawadskii* (2x) (Chen, 1985; Dai
98 and Chen, 1997; Fukai, 2003; Ma et al., 2016; Ma et al., 2020). *Chrysanthemum*
99 *indicum* and *C. vestitum* are key species in the origin and evolution of cultivated
100 chrysanthemums (Dai et al., 2002; Dai et al., 1998) based on two lines of evidence.

101 First, ancient literature documents multiple uses for *C. indicum* in central China,
102 which is consistent with the geographic origin of modern cultivars (see Chen, 2012).
103 Second, artificial hybridization between *C. indicum* and *C. vestitum* can generate
104 hybrids resembling the prototype of modern chrysanthemums (Chen, 2012).

105 In this study, we investigate whether hybridization occurs between
106 *Chrysanthemum* species with different ploidy levels, focusing on *C. indicum* and *C.*
107 *vestitum* as well as varieties of those species. We ask: (1) Does interploidy
108 hybridization naturally occur between tetraploid and hexaploid *Chrysanthemum*
109 species, which are likely involved in origin of modern cultivated chrysanthemums? (2)
110 If hybridization occurs, is there evidence for a greater propensity at high ploidy levels?
111 (3) Do some modern chrysanthemum cultivars share chloroplast haplotypes with the
112 three wild *Chrysanthemum* taxa? To this end, we genotyped 317 samples at 13
113 microsatellite markers and sequenced chloroplast *trnL-trnF* for a subset of 103
114 samples. In addition, we extracted *trnL-trnF* from the chloroplast genomes of 28 taxa,
115 representing 15 wild species of *Chrysanthemum*, 12 cultivars and one sample of
116 *Ajania varifolium*. We analyze these data in a population genetic and phylogeography
117 context, and use them to investigate a case where traditionally reproductive isolation
118 caused by differences in ploidy levels would be expected to be strong.

119 **2. Material and Methods**

120 **2.1 Study species**

121 *Chrysanthemum indicum* is tetraploid with a wide distribution across China,
122 though narrowly distributed diploid and hexaploid cytotypes have also been reported
123 (Liu et al., 2012). *Chrysanthemum vestitum* is a hexaploid distributed across Hubei
124 and Henan provinces in central China (Zhao and Chen, 1999) though the variety *C.*
125 *vestitum* var. *latifolium* is hexaploid with a restricted distribution in the Dabie

126 Mountains in Anhui province.

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

136 **2.2 Sampling across hybridizing populations**

137 To understand the extent and distribution of hybridization, we surveyed sympatric
138 populations of the three focal taxa: *C. indicum*, *C. vestitum*, and *C. vestitum* var.
139 *latifolium*. We identified and sampled five populations where *C. indicum* and *C.*
140 *vestitum* co-occur and two where *C. indicum* and *C. vestitum* var. *latifolium* co-occur
141 (Table S1). In addition, we collected *C. indicum* from two allopatric populations, TA
142 and ZP (Table S1). Samples were collected at random within populations, ensuring at
143 least ten meter spacing between individuals. Healthy and pest free leaf tissue was
144 collected and stored in silica gel. A total of 317 samples were collected including
145 between 14 and 74 from each of the seven hybridizing populations, five from TA and
146 three from ZP (Table S1). A Global Position System (GPS, Unistrong) was used to
147 record the coordinates of each population. Sampling locations are illustrated in Fig. 2.

148 **2.3 Microsatellite genotyping**

149 Genomic DNA was isolated from dried leaves of all individuals, following a modified
150 2x CTAB (cetyltrimethylammoniumbromide) protocol (Wang et al., 2013). The

151 quality of isolated genomic DNA was assessed on 1.0 % agarose gels, and then
152 diluted to a concentration of 10-20 ng/ul for genotyping and sequencing. Thirteen
153 microsatellite loci were used for genotyping (Zhang et al., 2014; Jo et al., 2015). The
154 5' terminus of the forward primer was labeled with FAM, HEX or TAM fluorescent
155 probes. Each microsatellite locus was amplified individually prior to being combined
156 into four multiplexes (Table S2). The PCR protocol follows Hu et al. (2019).

157 **2.4 Population genetic analysis**

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

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

173 In addition, to identify the most likely K value across populations we included all
174 populations in a combined STRUCTURE analysis, testing K values from 1 to 10. The
175 number of genetic clusters was estimated using the Evanno test (Evanno et al., 2005)

176 in the program Structure Harvester 0.6.94 (Earl and vonHoldt, 2012). Ten replicates
177 of the STRUCTURE analysis were performed with 1,000,000 iterations and a burn-in
178 of 100,000 for each run. The admixture model, with an assumption of correlated allele
179 frequencies, was used. Individuals were assigned to clusters based on the highest
180 membership coefficient averaged over the ten independent runs. Replicate runs were
181 grouped based on a symmetrical similarity coefficient of >0.9 using the Greedy
182 algorithm in CLUMPP 1.1.2 (Jakobsson and Rosenberg, 2007) and visualized in
183 DISTRUCT 1.1 (Rosenberg, 2004). In populations where *C. indicum* and *C. vestitum*
184 or *C. indicum* and *C. vestitum* var. *latifolium* co-occur, we estimated Q scores in
185 STRUCTURE with 95% confidence intervals to define pure *C. indicum*, pure *C.*
186 *vestitum*, pure *C. vestitum* var. *latifolium* or putative hybrids. We distinguish
187 individuals with confidence intervals overlapping 1 as pure *C. indicum*, with 0 as pure
188 *C. vestitum* and those remaining as putative hybrids, in populations where the two
189 species co-occur. Similarly, in populations where *C. indicum* and *C. vestitum* var.
190 *latifolium* co-occur, we distinguish individuals with confidence intervals overlapping
191 with 1 as pure *C. indicum*, with 0 as pure *C. vestitum* var. *latifolium* and those
192 remaining as hybrids.

193 We compared the average allele number per individual between *C. indicum*, *C.*
194 *vestitum*, *C. vestitum* var. *latifolium* and their hybrids at each microsatellite locus,
195 using the Kruskal–Wallis test in the R package agricolae v1.3-3 (de Mendiburu, 2020).
196 We would expect that the average allele number is higher for *C. vestitum* and *C.*
197 *vestitum* var. *latifolium* than *C. indicum*. In addition, we tested if introgression was
198 symmetric between *C. indicum* and *C. vestitum* or between *C. indicum* and *C.*
199 *vestitum* var. *latifolium* in each hybridizing population, using the function ‘wilcox.test’
200 in R v4.0.3 (R Core Team, 2020). The putative hybrids identified as above excluded

201 from such comparisons.

202 **2.6 *trnL-trnF* sequencing and analysis**

203 In order to detect the potential maternal parents of the hybrids, and to estimate the
204 haplotype diversity of these taxa, we amplified *trnL-trnF* for a subset of 103 samples,
205 including between five and 34 individuals from each population. Reactions were
206 performed in 20 ul volumes containing 13 ul ddH₂O, 5 ul 2×Taq PCR mix
207 (TIANGEN, China), 0.5 ul each primer (*trnL* and *trnF* (Taberlet et al., 1991)) and 1 ul
208 DNA template. PCR products were outsourced for purification and sequencing, to
209 Qingdao, China. Sequences were manually edited and aligned using BioEdit v7.2.5
210 (Hall, 1999). The R package pegas v0.14 (Paradis, 2010) was used to construct
211 haplotype networks, using default settings, with gaps treated as a fifth state. The total
212 number of sites, polymorphic sites, parsimony informative sites, and nucleotide
213 diversity were computed using DnaSPv6.12.03 (Rozas et al., 2017). All sequences
214 obtained in this study were submitted to GenBank with accession number MZ032043
215 - MZ032145.

216 To aid in a broader phylogenetic analysis, we extracted the *trnL-trnF* region
217 from the available whole chloroplast genomes for 15 of the 17 wild species of
218 *Chrysanthemum* occurring in China, 12 cultivars and *Ajania varifolium*. A
219 phylogenetic tree was estimated using the maximum-likelihood method (ML) in
220 RAxML v. 8.1.16 (Stamatakis, 2006). *Ajania varifolium* was selected as the outgroup.
221 A rapid bootstrap analysis with 100 bootstrap replicates and 10 tree searches was
222 performed under the GTR + GAMMA nucleotide substitution model. The consensus
223 tree generated from the bootstrap replicates was visualized in FigTree v.1.3.1
224 (Rambaut and Drummond, 2009).

225

226 **3. Results**

227 **3.1 Hybridization across ploidy levels inferred from microsatellites**

228 Genetic diversity estimates were similar among the three taxa and hybrids, with allelic

229 richness ranging from 3.74 to 4.26 and gene diversity from 0.75 to 0.84 (Table S3).

230 On average, the number of alleles scored in *C. vestitum* and *C. vestitum* var. *latifolium*

231 was significantly higher than *C. indicum* at eight and seven loci, respectively ($P <$

232 0.05); this was expected for a hexaploid possessing more chromosome copies than a

233 tetraploid species (Fig. S1).

234 Principal Coordinate (PCO) analysis based on Bruvo's genetic distances among

235 all samples revealed three clusters, with coordinates 1 and 2 explaining 13.8% and 7.8%

236 of the total variation, respectively (Fig. 2). Coordinate 1 separated *C. indicum* from *C.*

237 *vestitum* and *C. vestitum* var. *latifolium* and coordinate 2 separated *C. vestitum* from *C.*

238 *vestitum* var. *latifolium* (Fig. 2). Most hybrids identified by the STRUCTURE

239 analysis fell between *C. indicum* and *C. vestitum* in the PCO plot (Fig. 2).

240 The combined STRUCTURE analysis across all samples identified $K = 3$ as the

241 optimal K value (Fig. S2a), with the three clusters corresponding to *C. indicum*, *C.*

242 *vestitum* and *C. vestitum* var. *latifolium* (Fig. 3a). At $K = 2$, *C. indicum* formed one

243 cluster and *C. vestitum* and *C. vestitum* var. *latifolium* formed another cluster (Fig.

244 S2b), supporting the close relationship between these two taxa and the inference that

245 *C. vestitum* var. *latifolium* is a subspecies of *C. vestitum*. Interestingly, there are three

246 *C. vestitum* individuals possessing a considerable level of introgression from *C.*

247 *vestitum* var. *latifolium* and one *C. indicum* individual in population TZ with

248 substantial introgression from *C. vestitum* (Fig. 3a).

249 A total of 20 hybrids were detected when STRUCTURE was performed for each

250 hybridizing population separately, of which 19 were hybrids between *C. indicum* and

251 *C. vestitum* and one was a hybrid between *C. indicum* and *C. vestitum* var. *latifolium*.
252 Fourteen hybrids between *C. indicum* and *C. vestitum* were from population XG, two
253 from each of NX and PH and one from WFS. The only hybrid between *C. indicum*
254 and *C. vestitum* var. *latifolium* was from population TZ. No hybrids were detected
255 from population LJS where *C. indicum* and *C. vestitum* co-occur and from BMJ where
256 *C. indicum* and *C. vestitum* var. *latifolium* co-occur (Fig. 3b).

257 Introgression occurred symmetrically between *C. indicum* and *C. vestitum* and
258 between *C. indicum* and *C. vestitum* var. *latifolium* in all hybridizing populations
259 except WFS where only three *C. indicum* were collected (Fig. 3c). Introgression was
260 limited in populations TZ, BMJ and LJS but was extensive in population XG (Fig.
261 3c).

262 **3.2 Plastid diversity and directionality of hybrid formation**

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

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

280 **4. Discussion**

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

291 **4.1 Tetraploid-hexaploid hybridization and symmetrical introgression**

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

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

307 Nineteen out of the 20 hybrids were between *C. indicum* and *C. vestitum* and one
308 was between *C. indicum* and *C. vestitum* var. *latifolium*. This difference may reflect
309 different levels of reproductive isolation. Higher fruit set in artificial *C. indicum*-*C.*
310 *vestitum* crosses than artificial *C. indicum*-*C. vestitum* var. *latifolium* crosses, partially
311 supports this hypothesis (Zhou, 2009). However, environmental selection against
312 hybrids between *C. indicum* and *C. vestitum* var. *latifolium* may also account for its
313 rarity. *Chrysanthemum indicum* and *C. vestitum* occupy similar habitats and are
314 usually intermixed in sympatric populations (Shuai Qi, personal observations). This
315 may enhance their opportunity for hybridization, while their similar habitat
316 preferences may reduce the chance of ecological selection on the hybrids. In contrast,
317 *C. vestitum* var. *latifolium* and *C. indicum* are adapted to different conditions, and
318 hybrids may fail to survive due to the breakdown of suites of co-adapted genes.
319 Within some diploid-tetraploid systems, introgression is more common from diploids
320 to tetraploids (Zohren et al., 2016; Wang et al., 2020). However, we observe
321 symmetrical introgression between *C. indicum* and *C. vestitum* and between *C.*
322 *indicum* and *C. vestitum* var. *latifolium*, indicating that hybrids can backcross with
323 both parents. This is in line with recent studies showing that pentaploids can mediate
324 gene flow between species with different ploidy levels (Peskoller et al., 2021). We
325 couldn't distinguish clearly between male and female parental taxa because *C.*

326 *indicum* and *C. vestitum* share some plastid haplotypes. However, *C. indicum* can
327 serve as maternal parent as one hybrid shared a haplotype with *C. indicum* whereas
328 this haplotype is absent from all *C. vestitum* and *C. vestitum* var. *latifolium* samples.
329 In addition, one hybrid had a unique haplotype (H16), which is possibly from
330 unsampled *C. indicum* and *C. vestitum* or introgressed from other *Chrysanthemum*
331 species.

332 **4.2 Variable hybridizing among populations**

333 The number of hybrids between *C. indicum* and *C. vestitum* varied substantially
334 among the hybridizing populations (Fig. 3b). Fourteen out of 19 hybrids are in
335 population XG, and no hybrids are found in population LJS. These populations are
336 approximately 50 km apart, and differential reproductive isolation seems unlikely to
337 account for such differences. However, we note that population LJS is closer to
338 human habitation and human activities may have an impact on the persistence of
339 hybrids.

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

347 However, under future climate change, *C. indicum* may move to higher altitudes
348 and come into closer contact with *C. vestitum* var. *latifolium*, producing more hybrids,
349 as seen in population TZ. This has implications for conserving *C. vestitum* var.
350 *latifolium*. *Chrysanthemum indicum* is widespread and abundant whereas *C. vestitum*

351 var. *latifolium* is restricted to the Dabie Mountains. Hybridization between abundant
352 *C. indicum* and rare *C. vestitum* var. *latifolium* may be predicted to drive the rare
353 species to extinction through genetic or demographic swamping (Todesco et al.,
354 2016).

355 **4.3 The presence of orphan hybrids**

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

362 We detected a few individuals showing considerable admixture between *C.*
363 *vestitum* var. *latifolium* and *C. vestitum* in populations WFS and NX where *C.*
364 *vestitum* var. *latifolium* is not known to occur. We also detected one *C. indicum*
365 individual showing considerable admixture from *C. vestitum* in population TS in
366 Shandong province. This indicates the presence of hybrids in the absence of one or
367 both parental species, which has been demonstrated in some plant species, such as
368 oaks (Dodd and Afzal-Rafii, 2004) and pines (Lanner and Phillips, 1992). A plausible
369 explanation is the existence of undetected *C. vestitum* var. *latifolium* near populations
370 WFS and NX, or within travelling distance of pollinators or seed dispersal.

371 **4.4 Implications for the origins of cultivated chrysanthemums**

372 Our results based on an analysis of plastid *trnL-trnF* showed a monophyletic clade
373 composed of *C. lavandulifolium*, *C. chonetii*, *C. zawadskii* and five haplotypes of *C.*
374 *indicum*, *C. vestitum*, *C. vestitum* var. *latifolium* and the hybrid between *C. indicum*
375 and *C. vestitum* (Fig. 5). This indicates that either of these species potentially acted as

376 the maternal parent of cultivated chrysanthemums. This is consistent with previous
377 research implicating their involvement (Dai et al., 1998; Dai et al., 2005; Fukai, 2003).
378 However, the specific maternal parental progenitor of modern cultivated
379 chrysanthemums remains elusive. The modern cultivated chrysanthemums were
380 placed in a monophyletic clade that was sister to *C. lavandulifolium*. However, the
381 chloroplast genome of *C. lavandulifolium* has some unique mutations compared with
382 cultivated chrysanthemums. This leads to the hypothesis that the maternal progenitor
383 of modern cultivated chrysanthemums has gone extinct (Ma et al., 2020). However,
384 this requires further evaluation as only one or two whole chloroplast genomes were
385 included for each wild *Chrysanthemum* species and cultivar. This means the diversity
386 of chloroplast genomes was not sufficiently represented. Given the high haplotype
387 diversity of the three taxa and the fact that these haplotypes did not form a
388 monophyletic clade (Fig. 5), we predict that the ultimate maternal progenitor of
389 modern cultivated chrysanthemums may potentially be any of the wild
390 *Chrysanthemum* species, and there may be multiple maternal progenitors for
391 chrysanthemum cultivars.

392 **Acknowledgements**

393 This work was funded by the National Natural Science Foundation of China
394 (31770230, 31600295 and 31872710), Shandong Provincial Natural Science
395 Foundation, China (ZR2018PC022) and Funds of Shandong “Double Tops” Program
396 (SYL2017XTTD13).

397

398 **References**

399 **Ainouche ML, Baumel A, Salmon A, Yannic G. 2003.** Hybridization, polyploidy
400 and speciation in *Spartina* (Poaceae). *New Phytologist* **161**: 165–172.

401 **Arnold ML. 2014.** Natural hybridization and the evolution of domesticated, pest and
402 disease organisms. *Molecular Ecology* **13**: 97–1007.

403 **Baek S, Choi K, Kim GB, Yu HJ, Cho A, Jang H, Kim C, Kim HJ, Chang KS,**
404 **Kim JH, Mun JH. 2018.** Draft genome sequence of wild *Prunus yedoensis*
405 reveals massive inter-specific hybridization between sympatric flowering
406 cherries. *Genome Biology* **19**: 127.

407 **Bringhurst RS, Voth V. 1984.** Breeding octoploid strawberries. *Iowa State Journal of*
408 *Research* **58**: 371–382.

409 **Bruvo R, Michiels NK, D'Souza TG, Schulenburg H. 2004.** A simple method for
410 the calculation of microsatellite genotype distances irrespective of ploidy level.
411 *Molecular Ecology* **13**: 2101–2106.

412 **Chen F, Chen P, Li H. 1996.** Genome analysis and their phylogenetic relationships of
413 several wild species of *Dendranthema* in China. *Acta Horticulturae Sinica* **23**:
414 67–72.

415 **Chen JY. 1985.** Studies on the origin of Chinese florist's chrysanthemum. *Acta*
416 *Horticulturae* **167**: 349–361.

417 **Chen JY. 2012.** *The Origin of Garden Chrysanthemum*. Hefei: Anhui Science &
418 Technology Publishing House.

419 **Clark LV, Jasieniuk M. 2011.** POLYSAT: an R package for polyploid microsatellite
420 analysis. *Molecular Ecology Resources* **11**: 562–566.

421 **Cornille A, Giraud T, Smulders MJM, Roldán-Ruiz I, Gladieux P. 2014.** The
422 domestication and evolutionary ecology of apples. *Trends in Genetics* **30**: 57–
423 65.

424 **Dai SL, Chen JY. 1997.** Cladistic study on some *Chrysanthemum* spp. in China.
425 *Journal of Wuhan Botanical Research* **15**: 27–34.

426 **Dai SL, Chen JY, Li WB. 1998.** Application of RAPD analysis in the study on the
427 origin of Chinese cultivated chrysanthemum. *Acta Botanica Sinica* **40**: 1053–
428 1059.

429 **Dai SL, Wang WK, Huang JP. 2002.** Advance of researches on phylogeny of
430 *Dendranthema* and origin of chrysanthemum. *Acta Scent Nat Univ Pekinensis*
431 **24**: 230–234.

432 **Dai SL, Wang WK, Xu YX. 2005.** Phylogenetic relationship of *Dendranthema* (DC.)
433 Des Moul. revealed by fluorescent in situ hybridization. *Journal of Integrative
434 Plant Biology* **7**: 783–791.

435 **de Mendiburu F. 2020.** agricolae: Statistical procedures for agricultural research. R
436 package version 1.3 - 3. Available from
437 <https://CRAN.r-project.org/package=agricolae> [accessed 30 March 2021].

438 **Dodd RS, Afzal-Rafii Z. 2004.** Selection and dispersal in a multispecies oak hybrid
439 zone. *Evolution* **58**: 261–269.

440 **Dowrick GJ. 1952.** The chromosome of chrysanthemum (I). *Heredity* **6**: 365–375.

441 **Dowrick GJ. 1953.** The chromosomes of chrysanthemum (II) garden varieties.
442 *Heredity* **7**: 59–72.

443 **Earl DA, vonHoldt BM. 2012.** STRUCTURE HARVESTER: a website and program
444 for visualizing STRUCTURE output and implementing the Evanno method.
445 *Conservation Genetics Resources* **4**: 359–361.

446 **Evanno G, Regnaut S, Goudet J. 2005.** Detecting the number of clusters of
447 individuals using the software STRUCTURE: a simulation study. *Molecular
448 Ecology* **14**: 2611–2620.

449 **Fukai S. 2003.** Dendranthema species as chrysanthemum genetic resources. *Acta
450 Horticulturae* **620**: 223–230.

451 **Goudet J. 1995.** FSTAT (version 1.2): a computer program to calculate F-statistics.
452 *Journal of Heredity* **86**: 485–486.

453 **Groh JS, Percy DM, Björk CR, Cronk QCB. 2019.** On the origin of orphan hybrids
454 between *Aquilegia formosa* and *Aquilegia flavescens*. *AoB Plants* **11**: plv071.

455 **Hülber K, Sonnleitner M, Suda J, Krejčíková J, Schönswitter P, Schneeweiss
456 GM, Winkler M. 2015.** Ecological differentiation, lack of hybrids involving
457 diploids, and asymmetric gene flow between polyploids in narrow contact
458 zones of *Senecio carniolicus* (syn. *Jacobaea carniolica*, Asteraceae). *Ecology
459 and Evolution* **5**: 1224–1234.

460 **Hall TA. 1999.** BioEdit: a user-friendly biological sequence alignment editor and
461 analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**:
462 95–98.

463 **Heslop-Harrison JS, Schwarzacher T. 2007.** Domestication, genomics, and the
464 future for banana. *Annals of Botany* **100**: 1073–1084.

465 **Hu YN, Zhao L, Buggs RJA, Zhang XM, Li J, Wang N. 2019.** Population structure
466 of *Betula albosinensis* and *Betula platyphylla*: evidence for hybridization and
467 a cryptic lineage. *Annals of Botany* **123**: 1179–1189.

468 **Husband BC, Sabara HA. 2003.** Reproductive isolation between autotetraploids and
469 their diploid progenitors in fireweed, *Chamerion angustifolium* (Onagraceae).
470 *New Phytologist* **161**: 703–713.

471 **Huxel GR. 1999.** Rapid displacement of native species by invasive species: effects of
472 hybridization. *Biological Conservation* **89**: 143–152.

473 **Jakobsson M, Rosenberg NA. 2007.** CLUMPP: a cluster matching and permutation
474 program for dealing with label switching and multimodality in analysis of
475 population structure. *Bioinformatics* **23**: 1801–1806.

476 **James JK, Abbott RJ. 2005.** Recent, allopatric, homoploid hybrid speciation: the
477 origin of *Senecio squalidus* (Asteraceae) in the British Isles from a hybrid
478 zone on Mount Etna, Sicily. *Evolution* **59**: 2533–2547.

479 **Jo KM, Jo Y, Chu H, Lian S, Cho WK. 2015.** Development of EST-derived SSR
480 markers using next-generation sequencing to reveal the genetic diversity of 50
481 chrysanthemum cultivars. *Biochemical Systematics and Ecology* **60**: 37–45.

482 **Kim HJ, Lee YS. 2005.** Identification of new dicaffeoylquinic acids from
483 *Chrysanthemum morifolium* and their antioxidant activities. *Planta Medica* **71**:
484 871–876.

485 **Kuligowska K, Lütken H, Müller R. 2016.** Towards development of new
486 ornamental plants: status and progress in wide hybridization. *Planta* **244**: 1–

487 17.

488 **Lanner RM, Phillips AMI. 1992.** Natural hybridization and introgression of pinyon
489 pines in northwestern Arizona. *International Journal of Plant Sciences* **153**:
490 250–257.

491 **Levin DA, Francisco-Ortega J, Jansen RK. 1996.** Hybridization and the extinction
492 of rare plant species. *Conservation Biology* **10**: 10–16.

493 **Li J, Wan Q, Abbott RJ, Rao GY. 2013.** Geographical distribution of cytotypes in
494 the *Chrysanthemum indicum* complex as evidenced by ploidy level and
495 genome - size variation. *Journal of Systematics and Evolution* **51**: 196–204.

496 **Lim KB, Barba-Gonzalez R, Zhou S, Ramanna MS, van Tuyl JM. 2008.**
497 Interspecific hybridization in Lily (*Lilium*): Taxonomy and commercial
498 aspects of using species hybrids in breeding. In: da Silva JAT, ed. *Floriculture,
499 ornamental and plant biotechnology*. Japan: Global Science Books.

500 **Liu PL, Wan Q, Guo YP, Yang J, Rao GY. 2012.** Phylogeny of the genus
501 *Chrysanthemum* L.: evidence from single-copy nuclear gene and chloroplast
502 DNA sequences. *PLoS ONE* **7**: e48970.

503 **Luo C, Chen D, Cheng X, Zhao H, Huang C. 2017.** Genome size estimations in
504 *Chrysanthemum* and correlations with molecular phylogenies. *Genetic
505 Resources and Crop Evolution* **64**: 1451–1463.

506 **Ma Y, Wei J, Yu Z, B. Q, Dai S. 2015.** Characterization of ploidy levels in
507 *Chrysanthemum* L. by flow cytometry. *Journal of Forestry Research* **26**: 771–
508 775.

509 **Ma Y, Zhao L, Zhang W, Zhang Y, Xing X, Duan X, Hu J, Harris AJ, Liu P, Dai**
510 **S, Wen J. 2020.** Origins of cultivars of *Chrysanthemum* - evidence from the
511 chloroplast genome and nuclear LFY gene. *Journal of Systematics and*
512 *Evolution* **58**: 925–944.

513 **Ma YP, Chen MM, Wei JX, Zhao L, Liu PL, Dai SL, Wen J. 2016.** Origin of
514 *Chrysanthemum* cultivars — Evidence from nuclear low-copy *LFY* gene
515 sequences. *Biochemical Systematics and Ecology* **65**: 129–136.

516 **Marques I, Feliner GN, Draper Munt D, Martins-Louçao MA, Aguilar JF. 2010.**
517 Unraveling cryptic reticulate relationships and the origin of orphan hybrid
518 disjunct populations in *Narcissus*. *Evolution* **64**: 2353–2368.

519 **Moraes AP, Chinaglia M, Clarisse Palma - Silva C, Pinheiro F. 2013.** Interploidy
520 hybridization in sympatric zones: the formation of *Epidendrum fulgens* × *E.*
521 *puniceoluteum* hybrids (Epidendroideae, Orchidaceae). *Ecology and Evolution*
522 **3**: 3824 – 3837.

523 **Nakata M, Hong D, Qiu J, Uchiyama H, Tanaka R, Chen X. 1992.** Cytogenetic
524 studies on wild *Chrysanthemum* sensu lato in China. II. A natural hybrid
525 between *Dendranthema indicum* (2n = 36) and *D. vestitum* (2n = 54) from
526 Hubei Province. *Journal of Japan Botany* **67**: 92–100.

527 **Paradis E. 2010.** pegas: an R package for population genetics with an integrated–
528 modular approach. *Bioinformatics* **26**: 419–420.

529 **Peskoller A, Silbernagl L, Hülber K, Sonnleitner M, Schönswetter P. 2021.** Do
530 pentaploid hybrids mediate gene flow between tetraploid *Senecio disjunctus*

531 and hexaploid *S. carniolicus* s. str. (*S. carniolicus* aggregate, Asteraceae)?

532 *Alpine Botany*.

533 **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

534 levels in the Neotropical orchids *Epidendrum fulgens* and *E. puniceoluteum*

535 (Orchidaceae). *Molecular Ecology* **19**: 3981–3994.

536 **Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure

537 using multilocus genotype data. *Genetics* **155**: 945–959.

538 **R Core Team. 2020.** R: a language and environment for statistical computing.

539 Version 4.0.3. Vienna, Austria: R Foundation for Statistical Computing.

540 **Rambaut A, Drummond AJ. 2009.** FigTree version 1.3.1 [computer program].

541 <http://tree.bio.ed.ac.uk>.

542 **Rosenberg NA. 2004.** DISTRUCT: a program for the graphical display of population

543 structure. *Molecular Ecology Notes* **4**: 137–138.

544 **Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017.** DnaSP 6: DNA sequence

545 polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34**:

546 3299–3302.

547 **Saar DE, Polans NO, Sørensen PD. 2003.** A phylogenetic analysis of the genus

548 *Dahlia* (Asteraceae) based on internal and external transcribed spacer regions

549 of nuclear ribosomal DNA. *Systematic Botany* **28**: 627–639.

550 **Shahrajabian MH, Sun W, Zandi P, Cheng Q. 2019.** A review of chrysanthemum,

553 the eastern queen in traditional chinese medicine with healing power in
554 modern pharmaceutical sciences. *Applied Ecology and Environmental*
555 *Research* **17**: 13355–13369.

556 **Shih C, Christopher JH, Michael GG. 2011.** *Chrysanthemum*. In: Wu ZY, Raven
557 PH, Hong DY, eds. *Flora of China*. Beijing: Science Press; St. Louis: Missouri
558 Botanical Garden Press.

559 **Shih C, Fu GX. 1983.** *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

560 **Simmonds NW, Shepherd K. 1955.** The taxonomy and origins of the cultivated
561 bananas. *Botanical Journal of the Linnean Society* **55**: 302–312.

562 **Sonnleitner M, Weis B, Flatscher R, Escobar Garcia P, Suda J, Krejčíková J,**
563 **Schneeweiss GM, Winkler M, Schönswetter P, Hülber K. 2013.** Parental
564 ploidy strongly affects offspring fitness in heteroploid crosses among three
565 cytotypes of autopolyploid *Jacobsaea carniolica* (Asteraceae). *PloS ONE* **8**:
566 e78959.

567 **Stamatakis A. 2006.** RAxML-VI-HPC: maximum likelihood-based phylogenetic
568 analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–
569 2690.

570 **Sutherland BL, Galloway LF. 2017.** Postzygotic isolation varies by ploidy level
571 within a polyploid complex. *New Phytologist* **213**: 404–412.

572 **Taberlet P, Gielly L, Pautou G, Bouvet J. 1991.** Universal primers for amplification
573 of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**:
574 1105–1109.

575 **Tahara M. 1915.** Cytological studies in *Chrysanthemum* (A preliminary note). *Bot*
576 *Mag Tokyo* **29**: 48–50.

577 **Todesco M, Pascual MA, Owens GL, Ostevik KL, Moyers BT, Hübner S,**
578 **Heredia SM, Hahn MA, Caseys C, Bock DG, Rieseberg LH. 2016.**
579 Hybridization and extinction. *Evolutionary Applications* **9**: 892–908.

580 **Wang J, Dong S, Yang L, Harris A, Schneider H, Kang M. 2020.** Allopolyploid
581 speciation accompanied by gene flow in a tree fern. *Molecular Biology and*
582 *Evolution* **37**: 2487–2502.

583 **Wang N, Borrell JS, Bodles WJA, Kuttapitiya A, Nichols RA, Buggs RJA. 2014.**
584 Molecular footprints of the Holocene retreat of dwarf birch in Britain.
585 *Molecular Ecology* **23**: 2771–2782.

586 **Wang N, Thomson M, Bodles WJA, Crawford RMM, Hunt HV, Featherstone**
587 **AW, Pellicer J, Buggs RJA. 2013.** Genome sequence of dwarf birch (*Betula*
588 *nana*) and cross-species RAD markers. *Molecular Ecology* **22**: 3098–3111.

589 **Yan W, Jung JA, Lim K, Cabahug RAM, Hwang Y. 2019.** Cytogenetic studies of
590 *chrysanthemum*: a review. *Flower Research Journal* **27**: 242–253.

591 **Yang W, Glover BJ, Rao GY, Yang J. 2006.** Molecular evidence for multiple
592 polyploidization and lineage recombination in the *Chrysanthemum indicum*
593 polyploid complex (Asteraceae). *New Phytologist* **171**: 875–886.

594 **Zhang Y, Dai SL, Hong Y, Song XB. 2014.** Application of genomic SSR locus
595 polymorphisms on the identification and classification of *Chrysanthemum*
596 cultivars in China. *PloS ONE* **9**: e104856.

597 **Zhao HE, Chen JY. 1999.** Investigation and studies on some native species of
598 *Dendranthema*. *Chinese Landscape Architecture* **3**: 60–61.

599 **Zhou J. 2009.** *Studies on the problem of origin of Chinese garden chrysanthemum*,
600 Beijing Forestry University, Beijing.

601 **Zhou SJ, Wang JW. 1997.** The cytological study on ten species of *Dendranthema*.
602 *Journal of Wuhan Botanical Research* **15**: 289–292.

603 **Zhou SL, Zou XH, Zhou ZQ, Liu J, Xu C, Yu J, Wang Q, Zhang DM, Wang XQ,**
604 **Ge S, Sang T, Pan KY, Hong DY. 2014.** Multiple species of wild tree peonies
605 gave rise to the 'king of flowers', *Paeonia suffruticosa* Andrews. *Proceedings*
606 *of the Royal Society B: Biological Sciences* **281**: 1687–1694.

607 **Zohren J, Wang N, Kardailsky I, Borrell JS, Joecker A, Nichols RA, Buggs RJA.**
608 **2016.** Unidirectional diploid-tetraploid introgression among British birch trees
609 with shifting ranges shown by restriction site-associated markers. *Molecular*
610 *Ecology* **25**: 2413–2426.

611

612

613 **Tables**

614 **Table S1** Detailed sampling information for natural populations of *C. indicum*, *C.*

615 *vestitum* and *C. vestitum* var. *latifolium*.

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

617 **Table S3** Genetic diversity of 14 populations of *C. indicum*, *C. vestitum* and *C.*

618 *vestitum* var. *latifolium* based on microsatellite markers.

619 **Table S4** The chloroplast haplotypes and their variable sites based on *trnL-trnF*

620 sequences.

621

622 **Figure legends**

623 **Fig. 1** Diversity present in wild *Chrysanthemum* species included in this study.

624 Individual photographs and leaf morphology for *C. indicum* (a), *C. vestitum* (b) and *C.*
625 *vestitum* var. *latifolium* (c).

626 **Fig. 2** Principal coordinate analysis of *Chrysanthemum* samples based on 13
627 microsatellite markers. Hybrids were defined based on Q scores in a STRUCTURE
628 analysis (see text).

629 **Fig. 3** Hybridization among *C. indicum*, *C. vestitum* and *C. vestitum* var. *latifolium*. (a)
630 Hybridization across natural populations in a combined analysis of *C. indicum*, *C.*
631 *vestitum* and *C. vestitum* var. *latifolium* with K = 3. (b) Admixture in sympatric
632 population of *C. indicum* and *C. vestitum* and each sympatric population of *C.*
633 *indicum* and *C. vestitum* var. *latifolium*. Barplots on the left or the right of
634 STRUCTURE plot represent the number of individuals having different levels of
635 genetic admixture. Populations XG, NX and PH were analyzed together due to their
636 short geographic distance. Allopatric populations TA and ZP were served as a control.
637 Blue, green and orange represent *C. indicum*, *C. vestitum* and *C. vestitum* var.
638 *latifolium*, respectively. (c) Admixture value for each population among *C. indicum*, *C.*
639 *vestitum* and *C. vestitum* var. *latifolium*. Hybrids were excluded for such comparisons.

640 **Fig. 4** Haplotype network based on plastid *trnL-trnF* sequences in natural
641 *Chrysanthemum* populations. (a) The geographic distribution of haplotypes. Number
642 in brackets indicates the number of samples used for sequencing; (b) Haplotype
643 network graph. Each haplotype is represented by a circle with size proportional to the
644 number of individuals. Color within the circle represents species sharing the
645 haplotype.

646 **Fig. 5** The phylogenetic relationships of *Chrysanthemum* species inferred from the

647 chloroplast *trnL-trnF* region. Branch support values above 50% are depicted.

648 Haplotypes identified in the present study are indicated in brackets. Species names

649 before the identified haplotypes represent species sharing the haplotype. Colored in

650 blue, green, orange and purple represent *C. indicum*, *C. vestitum*, *C. vestitum* var.

651 *latifolium* and hybrids, respectively.

652 **Fig. S1** Allele number per individual at each of the 13 microsatellite loci among *C.*

653 *indicum*, *C. vestitum*, *C. vestitum* var. *latifolium* and hybrids. The difference in the

654 number of allele was assessed using Kruskal–Wallis test. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P <$

655 0.001.

656 **Fig. S2** The output of Structure Harvester showing that $K = 3$ is the optimal value (a)

657 and STRUCTURE results at $K = 2$ and 3 (b).

658

659

(a)



C. indicum

(b)



C. vestitum

(c)



C. vestitum var. *latifolium*

2 cm











