Running title: Adaptation of *Aquilegia* species

Genetic and epigenetic signatures associated with the rapid radiation of *Aquilegia* species

Tianyuan Lu[1,2,3], Ming-Rui Li[1], Ning Ding[1], Zhen-Hui Wang[4], Li-Zhen Lan[1], Xiang Gao[5], and Lin-Feng Li[1,*]

[1]Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, School of Life Sciences, Fudan University, Shanghai 200438, China;

[2]McGill University and Genome Quebec Innovation Center, Montreal H3A 0G1, Quebec, Canada;

[3]Lady Davis Institute, SMBD JGH, Montreal H3A 1A3, Quebec, Canada;

[4]Faculty of Agronomy, Jilin Agricultural University, Changchun 130118, China;

[5]Key Laboratory of Molecular Epigenetics of the Ministry of Education, Northeast Normal University, Changchun 130024, China.

***Correspondence author:**

Lin-Feng Li (lilinfeng@fudan.edu.cn)

Abstract

Elucidating how the diverse evolutionary mechanisms interact to determine species diversification is crucial to understanding the evolution and persistence of biodiversity. The genus *Aquilegia* (columbine) is a model system to identify genetic and epigenetic signatures underpinning the rapid adaptive radiation. In this study, we surveyed the genomes and DNA methylomes of ten worldwide *Aquilegia* species to investigate whether specific genomic architectures were associated with rapid species diversification in the Asian, European and North American lineages. The resulting phylogenies and population structure inferences revealed clearly high genetic and DNA methylomic divergence among the three lineages. Genome-wide scanning demonstrated significantly higher positive and purifying selection pressures among the Asian species compared to the European and North American lineages. More importantly, candidate genes identified at the genetic and epigenetic levels are functionally related to diverse adaptation-related traits, such as stress tolerance, photosynthesis and cell reproduction. While a considerable proportion of the candidate genes share hotspots of intra- and inter-lineage divergence, genetic and epigenetic factors were found to act on complementary biological pathways. By assessing the interaction between genetic variations and epigenetic variability, we found that epigenetic variability is a partially independent factor that intertwines with genetic mechanism to facilitate the diversification of the *Aquilegia* species. Taken together, these findings suggest that specific genetic architectures have been involved in vital biological pathways in response to diverse environmental conditions and epigenetic modifications may play a complementary role in the process of adaptive speciation. Our study provides an exploratory overview of how the established genetic and epigenetic signatures are associated with the rapid species diversification of *Aquilegia* species.

Key words: Adaptive radiation; *Aquilegia*; Methylome; Speciation; Species diversification

Author Summary

Disentangling the genetic and epigenetic bases underpinning species diversification is crucial to understanding the evolution and persistence of biodiversity. The columbine (genus *Aquilegia*) is a model system to address biological mechanisms associated with rapid adaptive radiation. We surveyed the genomes and DNA methylomes of ten worldwide columbine species to investigate whether specific genetic and epigenetic architectures were involved in the diversification of Asian, European and North American columbine species. We pinpointed candidate genes featuring rapid adaptive radiation identified at either the genetic or the epigenetic levels. These candidate genes are functionally associated with diverse adaptation-related traits such as cell reproduction, plant growth, and stress tolerance. Such genetic and epigenetic signatures have potential contributed adaptabilities to the columbine species to cope with diverse environmental conditions. In addition, we also showed that epigenetic modifications could act as a complementary factor that intertwined with genetic mechanism to facilitate the diversification of the columbine species. In all, our study provides a genome-wide view of how the genetic and epigenetic factors are associated with the rapid species diversification of the columbine species.

Introduction

Adaptive radiation is the rapid diversification of a single ancestral species into a vast array of common descendants that inhabit different ecological niches or use a variety of resources, but differ in phenotypic traits required to exploit diverse environments [1-4]. Disentangling the evolutionary mechanisms underpinning adaptive radiation is fundamental to understanding the evolution and persistence of biodiversity (5, 6). This has been a key focus of many studies investigating different animal and plant lineages that diversified through adaptive radiation, including Hawaiian silversword, Caribbean anoles, Darwin's finches, and African cichlids (7-10). However, it remains under-investigated as to why some lineages could diversify rapidly but their close relatives or other sympatrically distributed lineages did not. In the past decades, accumulating evidence from diverse radiation lineages suggests that both the extrinsic environmental variables (e.g., resource availability) and intrinsic factors (e.g., genetic variants determine lineage-specific morphological variations) are the potential mechanisms in determining the rate and volume of species diversification [11]. Among the environmental triggers, ecological opportunity is considered as the primary mechanism that causes rapid adaptive radiation through the acquisition of key innovations, invasion of new environments and extinction of competitors [2, 12]. On the other hand, new species can also arise as a result of multiple evolutionary mechanisms (e.g. natural selection) acting on genetic factors (e.g., advantage alleles) that ultimately influence the phenotypic disparity among closely related species [13]. An illustration of rapid adaptive radiation is the African cichlid fishes, on which the extrinsic environmental factors (e.g., ecological specialization) and intrinsic mechanisms (e.g., adaptive introgression) acted together to provoke the repeated adaptive radiation in geographically isolated lakes [10, 11, 14, 15].

The genus *Aquilegia* L. (columbine) is a well-recognized model system to address the evolutionary mechanisms underlying rapid adaptive radiation [16, 17]. This genus includes approximately 70 recently diversified species that are widely distributed in the temperate areas of North America and Eurasia [18]. Phylogenetic and geographic inferences have illustrated two independent adaptive radiations of North American and European lineages from the distinct ancestral Asian species [16, 19]. For example, floral diversification of the North American *Aquilegia* species is highly correlated with the pollinator specialization [20-23]. In contrast, ecological adaptation and geographic isolation are considered as the major driving forces that promoted rapid radiation of the European species [16, 24]. In Asia, changes in pollinator and ecological habitat are both proposed to be the underpinning mechanisms that resulted in the diversification of more than 20 morphologically distinct species [25, 26]. These Asian *Aquilegia* species constitute four highly divergent lineages corresponding to respective geographic origins and

evolved relatively independent under distinct external conditions [25, 26]. This well-described evolutionary history renders the *Aquilegia* species an ideal system to address how the extrinsic and intrinsic factors shaped genome-wide genetic and epigenetic variation patterns.

In this study, we survey the genomes and DNA methylomes of 36 accessions from ten worldwide *Aquilegia* species to elucidate the genetic and epigenetic bases underlying rapid diversification of the Asian, European and North American *Aquilegia* species. Among the Asian species, four phylogenetically distinct species (*A. japonica*, *A. oxysepala, A. yabeana*, and *A. viridiflora*) are selected according to their geographic distributions and ecological habitats. *Aquilegia japonica* and *A. oxysepala* are sister species inhabiting alpine tundra and low altitude forest niches in northeastern China, respectively [26, 27]. Our previous studies have documented that natural selection during ecological specialization together with genetic drift under geographic isolation caused the rapid evolution of reproductive isolation between these two species [26, 28]. Here, we further investigate how diverse evolutionary driving forces shaped genetic and epigenetic architectures of the two species in the processes of speciation and adaptation. In addition, we also evaluated patterns of nucleotide variation and cytosine methylation in the *A. yabeana* and *A. viridiflora*. The former species shares highly similar morphological traits and ecological niches with the *A. oxysepala* but is allopatrically distributed in the northern China. In contrast, while the *A. viridiflora* is sympatrically distributed with *A. yabeana* and *A. oxysepala* in the northern and northeastern China, it often occupies rocky and sandy ecological niches. As a supplementary, we also examined nucleotide and cytosine methylation variation patterns of the North American and European lineages to investigate how the distinct evolutionary mechanisms shaped genomic and epigenomic architectures on a genome-wide scale. Our study will provide a genome-wide view of how the genetic and epigenetic architectures associated with the phenotypic divergence and ecological specialization of the *Aquilegia* species.

Results

Population structure and nucleotide variation pattern

Neighbor-joining (NJ) trees were reconstructed for the 36 *Aquilegia* accessions based on 689,123 homozygous SNPs. The established phylogenies showed that accessions of the ten species formed three distinct lineages corresponding to their geographic origins (**Figure 1a** and **Figure S1**). In brief, all the 22 accessions of the four East Asian species, *A. japonica*, *A. oxysepala*, *A. yabeana* and *A. viridiflora*, clustered as a monophyletic lineage, with the former two species and their hybrid forming a clade and the latter two species grouping as a sister clade. In contrast, the West Asian species *A. fragrans*

clustered with geographically adjoining European species. Within the North American lineage, phylogenetic relationships of the four species varied across the seven chromosomes. In particular, one European accession of the *A. alpina* var. *alba* fell into the North American lineage at the chromosomes 1 and 3. Likewise, the principal component analysis (PCA) and population structure inferences also revealed distinct genetic structure of the three phylogenetic lineages (**Figure 1b** and **c**). Consistent with the above phylogenies, one *A. alpina* var. *alba* accession shared the same ancestral genetic cluster with the North American lineage and the putative hybrid of the *A. oxysepala* and *A. japonica* possessed a mixed genomic constitution (**Figure 1b** and **c**).

To further gain genome-wide nucleotide variation pattern of the ten *Aquilegia* species, we calculated nucleotide diversity (π) and genetic divergence (F_{ST}) for each chromosome and 100-kb sliding window, respectively. Among the three phylogenetic lineages, the Asian *Aquilegia* species harbored relatively higher nucleotide diversity compared to the European and North American lineages across the seven chromosomes (**Figure S2**). By comparing the nucleotide diversity for each 100-kb sliding window, we observed a moderate correlation of genome-wide variation pattern among the three lineages (Spearman R = 0.40-0.52) and a high correlation between the *A. oxysepala* and *A. japonica* (Spearman R = 0.83) (**Figure S3**). In particular, 116 of 260 low genetic diversity genomic regions (LDGRs, the top 5% lowest π) were shared by at least two of the three lineages (**Figure S4**). Between the *A. oxysepala* and *A. japonica*, while we defined 148 LDGRs and 148 high divergence genomic regions (HDGRs, 5% highest F_{ST}), only seven candidate genomic regions showed overlap between the HDGRs and LDGRs (**Table S1**).

Identification of the genes indicating selection pressure and highly impactful genetic variations

Candidate genes associated with the adaptive divergence were determined according to the following three strategies. First, we considered the genes localized within the HDGRs and LDGRs more prone to be exposed to selection in the adaptive evolution process. Under this hypothesis, we identified 23 genes from the above seven candidate genomic regions between the *A. oxysepala* and *A. japonica* (**Table S1**). Genes within these genomic regions were functionally associated with meiotic nuclear division, adenine methyltransferase and basic cellar activities.

The first strategy mainly relied on genome-wide scanning for 100-kb non-overlapping sliding window. We next employed a functional annotation-based approach to identify highly impactful conservative clade-specific variations (CCVs; e.g., frameshift mutations, stop mutation and splicing alteration) from both the within and between lineage comparisons. Our results revealed that a considerable proportion (17.9-40.5%) of the CCVs were identified in the gene body regions (**Table S2**).

157 We then examined the potential impacts of these identified CCVs on gene functions. Between the *A.*
158 *oxysepala* and *A. japonica*, the CCV-carrying genes were enriched in several vital biological pathways
159 related to cell reproduction, including telomere maintenance, DNA repair, and DNA helicase activity
160 (**Figure 2** and **Table 1**). For example, two candidate genes (*Aqcoe6G160300* and *Aqcoe7G062500*) coding
161 for *XkIp2* (*TPX2*) were functionally correlated with spindle assembly during the mitotic process [29, 30].
162 Among the three phylogenetic lineages, the CCVs-harboring genes were also functionally involved in the
163 mitotic chromosome condensation, DNA ligase activity and aminopeptidase activity (**Figure 2** and **Table**
164 **1**). For instance, two CCV-containing genes (*Aqcoe2G276600* and *Aqcoe1G273400*) encoding DNA
165 mismatch repair proteins *MutS/MSH* and *MutS2* [31] carried one Asian-specific-to-American frameshift
166 variant.
167 Thirdly, we also performed genome-wide scanning to identify genes informative of selection
168 pressure from the within and between lineage comparisons. By estimating the pair-wise synonymous (d_S)
169 and non-synonymous (d_N) mutation rate, we found that species within the Asian lineage experienced
170 significantly higher positive ($d_N/d_S > 1$) and purifying ($d_N/d_S < 1$) selection pressures compared to the
171 European and North American lineages (Wilcoxon rank sum test, all Bonferroni-corrected *p* values <
172 1.5×10^{-16}) (**Figure S5**). Likewise, the European species showed significant purifying selection (Wilcoxon
173 rank sum test, Bonferroni-corrected *p* value = 7.8×10^{-8}) compared to the North American species.
174 Among the three lineages, all the comparisons showed significant purifying and positive selection
175 pressures (Wilcoxon rank sum test, all Bonferroni-corrected *p* values < 0.0018) except those that
176 indicated positive selection between the North American and the other two lineages (Wilcoxon rank
177 sum test, *p* value = 1).
178
179 **CG methylation patterns and differentially methylated genes**
180 In parallel with the above genomic analyses, we also investigated CG methylation pattern of the
181 representative *Aquilegia* species. Despite variance across the 36 *Aquilegia* accessions, the North
182 American, Asian and European species showed no distinguishable differences (t test, all Bonferroni-
183 corrected *p* values > 0.01) in overall percentage of methylated cytosines (**Figure 3a**). We then performed
184 PCA to examine CG-cytosine methylomic relatedness of all the *Aquilegia* accessions. The resulting
185 overall methylation pattern highly resembled the above genomic inferences, with the European and
186 American species forming two independent groups and the four Asian species forming three separated
187 clusters (**Figure S6**). We then assessed the CG methylation patterns for the European and North
188 American lineages as well as the three Asian species (*A. japonica*, *A. oxysepala* and *A. viridiflora*)

separately. Consistent with the described genomic features, heterogeneous pattern of the CG methylation was also observed among the seven chromosomes, with the chromosome 4 demonstrating obviously higher overall CG methylation divergence compared to the other six chromosomes (**Figure 3b**). We further quantified CG methylation deposited in the genic regions, putative *cis*-regulatory regions and CG island, respectively. In genic and regulatory regions, all three lineages shared similar modification patterns with apparent depletion of CG methylation around the transcription start site (TSS) and transcription end site (TES) (**Figure 3c**). However, the American lineage exhibited hyper-methylation (more than 10%) around the center of CG islands and a more drastic decrease throughout the CG island shores compared to the European and Asian species (**Figure 3d**).

To examine the biological impacts of CG methylation on the species diversification, differentially methylated regions (DMRs) and differentially methylated genes (DMGs) were identified from both the within and between thee three lineages, respectively (**Tables S3** and **S4**). Within the Asian lineage, 3,622 DMRs in 2,899 DMGs were identified between the *A. japonica* and *A. oxysepala*. Functional enrichment of these DMGs indicated that the two species may have different activities in photosynthesis-related pathways, including photosystem I, photosynthesis and chloroplast (**Figure 2**). For example, two photosynthesis-related genes, *PsaA*/*PsaB* and *CemA*, showed significantly differential methylation between the two species in the genic regions (**Figure S7a** and **b**). At the inter-lineage level, apparently more DMGs were identified between the North American and European species (6,087 genes) compared to those of between the two lineages and Asian species (3,308-5,003 genes) (**Table S3** and **S4**). DMGs characterized from the inter-lineage comparisons were mainly involved in the plant growth (e.g., response to auxin) and defense, response to biotic stimulus and wounding (**Figure 2**).

We then examined whether the candidate genes (CCV-carrying genes and DMGs) superimposed on the same signature of natural selection. We found while a considerable proportion of the candidate genes were shared for each of the genetic and epigenetic assessments (**Figure S8**), they showed a segregated distribution pattern across all comparisons (**Figure S9**). Likewise, the Gene Ontology (GO) enrichment analyses of the candidate gene identified from the genetic and epigenetic levels were enriched in functionally complementary pathways (**Figure 2**), suggesting co-existence of different underlying evolutionary mechanisms.

Association between epigenetic variability and genetic variations

Since both genetic variation and differential CG methylation seemed to have crucial and multifaceted influences on the adaptation of the ten *Aquilegia* species, we wondered whether differential epigenetic

221 modifications were dependent on genetic variability. Among the 588,659 CG sites examined, 224,222
222 (38.09%) carried a CG-loss variation. We then illustrated epigenetic variability for the variation-carrying
223 and non-variant CG sites, respectively. As shown in Figure 4, genetic-epigenetic associations of varying
224 magnitude were observed in both types of CG sites. The variation-carrying CG sites conveyed
225 information that highly resembled their genetic background. The overall methylation pattern was highly
226 conserved within the same species but exhibited obvious divergence across the ten *Aquilegia* species
227 (**Figure 4a**). In contrast, CG methylation divergence at the non-variant CG sites varied with higher
228 variability at both the intra- and inter-specific levels (**Figure 4b**). By examining the correlation of genetic
229 variability and cytosine methylation, we found that CG methylation divergence at variation-carrying CG-
230 site was largely attributable to the CG-loss variations (**Figure 4c**). In particular, 75% of the CG-loss
231 variations occurring at the most highly variable CG-methylated dinucleotides could explain at least 75%
232 of the total epigenetic variability *per se*. Nevertheless, there was still a considerable proportion of
233 epigenetic variability that could not be sufficiently explained by the variant-CG site (**Figure 4d**).
234 We also attempted to identify *cis*-driving mutations for each of the 1,229 DMRs between the *A.*
235 *japonica* and *A. oxysepala*. Our results revealed that only 568 out of the 1,229 (46.2%) DMRs were
236 significantly associated with at least one genetic variation inside or around a 500 base-pair (bp)
237 upstream/downstream genomic region, even under the least stringent *p* value threshold (5×10^{-5}),
238 indicating that the epigenetic changes were only partially dependent on *cis*-genetic driving mutations
239 (**Figure 4e**). Moreover, we observed significant, though weak, associations between differential CG
240 methylation and selection pressure. In most inter-lineage comparisons, DMGs were significantly more
241 prone to be under positive selection than non-DMGs (**Table 2**), suggesting that epigenetic modifications
242 could probably assist selection pressure in shaping genotypes. In contrast, DMGs were significantly less
243 prone to be under purifying selection (**Table 2**).
244
245 **Discussion**
246 **Genetic-driven mechanisms associated with the rapid diversification of *Aquilegia* species**
247 Elucidating evolutionary mechanisms underpinning species diversification is crucial to understanding the
248 evolution and persistence of biodiversity [2, 5, 6]. The genus *Aquilegia* provides an ideal system to
249 address how the diverse evolutionary mechanisms promoted rapid adaptive radiation [16, 17]. Although
250 various environmental conditions related to ecological opportunities, such as shifts in pollinator and
251 habitat, have been proposed to facilitate the evolution of reproductive isolation [22, 26], genetic basis
252 associated with the rapid diversification of *Aquilegia* species has still remained largely unclear. In this

253 study, we surveyed the genomes of ten worldwide *Aquilegia* species to address whether specific genetic
254 architectures have been involved in the rapid species diversification. Broadly consistent with previously
255 inferred phylogenies [16, 19, 26, 28], the ten *Aquilegia* species from Asia, Europe and North America
256 formed three phylogenetically independent lineages corresponding to their geographic origins. This
257 attribute offers the ten *Aquilegia* species a suitable system to identify genomic determinants associated
258 with the repeated adaptive speciation by extensive comparisons from different facets.

259 It has been proposed that if a genetic factor is the potential determinant promoting adaptive
260 speciation, one would expect to identify specific genetic architectures in the diversified lineages [9, 14,
261 32]. In Darwin's finches, for example, polyphyletic topology was observed as a general pattern in 14
262 morphologically distinct species, phenotypic diversity of the beak shape was mainly determined by
263 natural selection acting on the *ALX1* gene during the ecological specialization process [9]. A similar
264 phenomenon was observed in the East African cichlid fish where the radiating lineages are more
265 dynamic in terms of gene content and transcriptomic landscape compared to their non-radiating
266 relatives [14, 32]. In this study, while the Asian, European and North American species have clearly
267 diverged and evolved allopatrically in respective geographic regions, similar genome-wide nucleotide
268 variation pattern was observed among the three lineages, suggesting that a considerable of the genomic
269 variants might have evolved naturally during the diversification process. Nevertheless, while both the
270 North American and European lineages were descendants of the ancestral Asian species, shifts in
271 pollinator and ecological habitat are supposed to drive the independent adaptive radiations,
272 respectively [16, 17, 19]. This attribute suggest that specific genetic determinants are potentially related
273 to the repeated adaptive speciation of the *Aquilegia* species. As expected, our genome-wide scanning of
274 selection not only revealed higher positive and purifying selection pressures among the Asian species,
275 but also demonstrated significant differences in selection pressures in all inter-lineage comparisons
276 excepting these of experienced positive selection in North American and European lineages. More
277 importantly, the CCV-carrying genes identified from the three lineages are associated with cell
278 reproduction (e.g., telomere maintenance and mitotic chromosome condensation) and some other
279 functionally important pathways, indicating distinct selection pressures in specific lineages of these loss
280 function genes.

281 Similar genomic feature was also observed in the two ecological species *A. japonica* and *A.*
282 *oxysepala*. Our previous studies have demonstrated that natural selection and genetic drift together
283 resulted in the rapid evolution of reproductive isolation [26, 28]. Here, we further demonstrated that
284 candidate genes involved in the adaptive speciation are functionally enriched in the pathways related to

285 cell reproduction (e.g., telomere maintenance), stress tolerance (e.g., response to wounding) and basic
286 cellular activities. These findings from intra- and inter-lineage comparisons suggest that specific genetic
287 architectures might have played important roles in the repeated adaptive speciation of the *Aquilegia*
288 species. It should be noted that although a majority of the enriched pathways are specific to each
289 comparison, enrichment of cell reproduction-related pathways (e.g., telomere maintenance, DNA repair
290 and DNA helicase activity) is consistent across the intra- and inter-lineage comparisons. Taken together,
291 our study indicates that both the common and specific genetic determinants might have conferred high
292 adaptability to the *Aquilegia* species to cope with diverse environmental conditions.
293
294 **Evolutionary potential of cytosine methylation in the adaptation of *Aquilegia* species**
295 The role of epigenetic modification in the long-term evolutionary process has long been debated [33-35].
296 It has been proposed that epigenetic variations are frequently under the genetic control but which can
297 also alter rapidly as a result of environmental induction and stochastic epimutation [36, 37].
298 Nevertheless, it has also been recognized that some epigenetic variants can persist over generations and
299 be highly correlated with phenotypic diversity [35]. As illustrated in Arabidopsis, changes in cytosine
300 methylation can produce meiotically stable epialleles, which could eventually lead to phenotypic
301 diversity in the absence of genetic variations [38-40]. Here, we assessed whether the epigenetic
302 modifications are also associated with the adaptive speciation of the *Aquilegia* species. Consistent with
303 the genomic features detailed above, high divergence of cytosine methylation was observed across the
304 Asian, European and North American lineages. Nevertheless, differential cytosine methylation is not
305 only found across the seven chromosomes but also evident in the gene body of DMGs and CG island
306 region among the three lineages. Particularly, functional enrichment analyses identified significant
307 associations with adaptation-related traits, including plant growth, stress tolerance and basic cellular
308 activities. Between the *A. japonica* and *A. oxysepala*, for example, the identified DMGs showed
309 significant enrichment in categories related to diverse important phenotypic traits, such as
310 photosynthesis, embryo development and response to auxin. These epigenomic features indicate that
311 epigenetic factors may also play a role in response to diverse environmental conditions.
312 We noted that some candidate genes and enriched pathways shared hotspots of both genetic and
313 epigenetic disparities, especially those related to cell reproduction, plant growth and stress tolerance.
314 Many studies based on human and mouse have shown genetic variations can not only directly control
315 CG methylation patterns [41, 42], but also manipulate long-range CG methylation at specific sites to
316 further influence phenotypes, where CG methylation serves as a mediator [43, 44]. However, precedent

317 studies on explaining, in a time-span of evolutionary history, whether CG methylation functions
318 independent of genetic variations are few to our knowledge. To this end, we asked whether the
319 epigenetic mechanism works independently without certain genetic control. By analyzing the
320 associations between genetic and epigenetic variability, we conclude that while many CG-loss variations
321 can directly lead to depletion of CG methylation, a lot of DMRs are not manipulated by any *cis*-variations.
322 Since gene body CG methylation in plants generally stabilizes gene expression and is positively
323 correlated with gene expression [45-48], differential methylation in our study is indicative of the amount
324 of eventual gene products. Based on these attributes together with the informative associations
325 between differential methylation (DMGs) and positive selection (d_N/d_S), we propose that epigenetic
326 modification is a complementary mechanism facilitating phenotypic diversity of the *Aquilegia* species.
327

328 **Limitations and future directions**

329 Our study has important limitations. First of all, the small sample size in our study may introduce bias
330 and inflation of false positives, and we postulate that our findings should be interpreted carefully and
331 considered exploratory. When investigating the association between genetic divergence and
332 evolutionary events, it is impossible to deny the roles of other evolutionary forces. We acknowledge
333 that the lineage-specific allele frequencies are possibly consequences of genetic drift and that genetic
334 hitchhiking may lead to identification of candidate genes residing in neighboring genomic regions
335 representing the same driving force. Therefore, we claim that the candidate genes identified to be
336 associated with adaptive radiation do not necessarily point towards driving evolutionary mechanisms.
337 They may also be by-products of the long-term process of adaptive radiation. Besides, we only focused
338 on analysis of CG methylation as puzzles persist regarding the functional roles of CHG and CHH
339 methylation. While our sample size is not sufficient to yield enough statistical power even if such *trans*-
340 association tests were to be performed, our findings at least suggest that the epigenetic variability also
341 possesses autonomy to some extent.
342

343 **Methods and Materials**

344 **Sample collection, DNA extraction and sequencing**

345 In this study, a total of 36 accessions from ten worldwide *Aquilegia* species were collected (**Table S5**).
346 Among these samples, 18 accessions were collected to represent the two Asian species, *A. japonica* and
347 *A. oxysepala*, and their putative hybrid. In addition, four accessions were collected from the other two
348 Asian species *A. yabeana* and *A. viridiflora*. Furthermore, six and eight accessions were sampled from

349 the European and North American lineages, respectively. All the 36 accessions were grown in green
350 house under the same conditions (25°C/12 hours, 16°C/12 hours). Genomic DNA was extracted from
351 fresh mature leaves using TianGen plant genomic DNA kit. Whole genome resequencing and bisulfite
352 sequencing were performed on the extracted genomic DNA using the Illumina X-ten platform (Illumina,
353 California, USA). Short-insert (350 bp) DNA libraries of all accessions were constructed by NovoGene
354 (NovoGene, Tianjin, China). Genome assembly of an admixed species *A. coerulea* "Goldsmith" was
355 obtained from Phytozome v12.1 (https://phytozome.jgi.doe.gov) as the reference genome [17]. All data
356 generated from the study were submitted to EBI under the accession number PRJEB34182.
357
358 **Sequence assembly, population genetic structure and functional annotation**
359 Whole genome sequences of each accession were aligned against the reference genome using Burrows-
360 Wheeler Aligner (BWA) [49]. Raw assemblies were realigned using IndelRealigner provided in the
361 Genome Analysis Tool Kit [50]. Single nucleotide polymorphisms (SNPs) and insertions/deletions
362 (INDELs) were reported using SAMtools [51]. Only the high-quality variants (SNPs and INDELs) (read
363 depth > 3, mapping quality > 20 and missing allele < 1%) were retained for subsequent population
364 genomics analyses. Genomic context and predicted functional effects of the identified variants were
365 reported for each of the 36 samples separately. Intergenic variants occurring within 5 kb up- and down-
366 stream of TSS or TES were defined as regulatory variations. Functional annotation of each identified
367 variant was performed using SnpEff [52].
368 To infer the phylogenetic relationships of the ten *Aquilegia* species, NJ trees were reconstructed for
369 each chromosome and the whole genome dataset using MEGA 7 [53]. PCA was also carried out to
370 examine the genetic relatedness of the 36 *Aquilegia* accessions [54]. Ancestral components were
371 estimated using ADMIXTURE [55] with the assumed population number ranging from one to ten.
372 Optimal population composition with the least cross-validation error was selected to decompose
373 ancestral admixture. To obtain the genome-wide nucleotide variation pattern, nucleotide diversity (π)
374 and genetic differentiation (Weir and Cockerham's F_{ST}) were calculated for each 100 kb non-overlapping
375 sliding window using VCFtools [56, 57]. Pair-wise non-synonymous-to-synonymous (d_N/d_S) ratios of the
376 ten species were inferred by yn00 program in the Phylogenetic Analysis by Maximum Likelihood (PAML)
377 package [58]. Inter-lineage d_N/d_S value for each gene was derived by averaging d_N/d_S values obtained
378 from all pairwise comparisons of samples belonging to the two lineages under investigation. Candidate
379 genes with the 5% highest and 5% lowest d_N/d_S values were considered to have undergone strong
380 positive and purifying selection, respectively.

Cytosine methylation pattern and epigenetic population structure

Whole genome bisulfite sequencing data were pre-processed using TrimGalore (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/, accessed August 21, 2018). Trimmed paired-end reads were aligned to the reference genome using Bismark [59] with a moderately stringent minimum-score function (L,0,-0.3). De-duplicated alignments of the 36 *Aquilegia* accessions were used to report cytosine methylation level using bismark_methylation_extractor, on loci with a read depth ≥ 3. Genomic annotations of the methylated cytosine site were identified based on the reference genome using an in-house Python script. PCA was conducted to infer the CG-methylomic relatedness of the ten *Aquilegia* species. Differential cytosine methylation was determined at both the gene and chromosome levels, respectively. At the gene level, we determined DMRs for each 100 bp non-overlapping sliding window using Cochran-Mantel-Haenszel (CMH) test to account for imbalanced read depth (Supplementary Notes). Genomic regions that possess a Benjamini-Hochberg adjusted p value < 0.05 and show inter-specific or inter-lineage methylation divergence higher than 25% are defined as significant DMRs. Genes with > 20% of the genic region being DMR(s) were defined as DMGs. Chromosome-level methylation patterns were measured by chromosomal methylation discrepancy index (MDI) [60]. Methylation patterns of the identified DMGs were visually verified on Integrative Genomics Viewer [61] prior to downstream analyses and biological interpretation. In addition, we identified CG islands from the *A. coerulea* "Goldsmith" reference genome using EMBOSS cpgplot with default settings [62]. Only the identified CG-enriched genomic regions with > 200 bp were defined as CG islands. We then investigated inter-specific and inter-lineage methylation divergence in and around the CG islands.

Associations between the genetic variation and cytosine methylation

To address how the genetic and epigenetic variations interacted to promote species diversification, we tested for associations between the identified DMGs and genes under positive selection by a Chi-square test. Linear regression model was adopted to measure the direct causal effect of CG-loss variation on CG methylation. To further assess whether genetic variations drive the establishment of DMG, driving mutations of DMRs between the *A. japonica* and *A. oxysepala* were identified using an Eigenstrat method (See Supplementary Notes for more details) [63].

Identification of conservative clade-specific variant

CCVs were defined as these that were conserved across all samples belonging to the same species/lineage but not present in any sample of the other species/lineages. Since the biological consequences of heterozygous variants are less affirmable, only the homozygous point mutations and INDELs were considered in the characterization of CCVs, including frameshift, stop-gain, stop-loss, start-loss and splicing-alteration variations.

Functional enrichment analysis

The aforementioned genetic and epigenetic analyses identified candidate genes that might be associated with the rapid diversification of the *Aquilegia* species from different perspectives. These candidate genes were employed to conduct functional enrichment analyses using R package topGO [64]. Enriched GO terms that possessed a p value <0.05 and has an occurrence ≥ 5 in the background were considered statistically significant. Since the statistical tests performed by topGO are not independent, multiple testing correction does not apply [64]. Structures of functional domains of targeted genes were determined based on the InterPro database (https://www.ebi.ac.uk/interpro, accessed January 25, 2019). Distribution pattern of the identified candidate genes and their related functional pathways was visualized using the R package jvenn [65].

Author Contributions

L.F.L. conceived this project. T.L. and M.R.L. developed statistical analysis pipeline. T.L., M.R.L., N.D., Z.H.W., L.Z.L., and X.G. carried out experiments and analyzed the data. T.L., M.R.L., N.D., Z.H.W., L.Z.L., X.G., and L.F.L. participated in discussion and interpreted the data. T.L. and L.F.L. wrote the manuscript. All authors read and approved the manuscript.

Acknowledgments

We thank Aköz Gökçe for her constructive comments on our manuscript. This work was financially supported by National Natural Science Foundation of China (31670382), Shanghai Pujiang Program (19PJ1401500), Start-up funding at Fudan University (JIH1322105) and the Department of Science and Technology of Jilin Province (20190201299JC).

Financial Disclosure

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests

The authors have declared that no competing interests exist.

References

1. Schluter D. Ecological causes of adaptive radiation. American Naturalist. 1996; 148: S40-S64.
2. Schluter D. The Ecology of Adaptive Radiation. Heredity. 2000; 86(6): 749-50.
3. Givnish TJ. Adaptive radiation and molecular systematics: aims and conceptual issues. In: Givnish TJ, Systma KJ, eds. . Molecular Evolution and Adaptive Radiation. 1997: 1-54.
4. Givnish TJ. Adaptive radiation versus 'radiation' and 'explosive diversification': why conceptual distinctions are fundamental to understanding evolution. New Phytologist. 2015; 207(2): 297-303.
5. Simpson GG, Olson EC. The major features of evolution. Quarterly Review of Biology. 1953; 8(1): 87-8.
6. Losos JB. Adaptive radiation, ecological opportunity, and evolutionary determinism. American Society of Naturalists E. O. Wilson award address. American Naturalist. 2010; 175(6): 623-39.
7. Carlquist, Motley TJ. Tarweeds and Silverswords, Evolution of the Madiinae (*Asteraceae*). Economic Botany. 2003; 58(1): 123-4.
8. Losos JB, Ricklefs RE. Adaptation and diversification on islands. Nature. 2009;457(7231):830-6.
9. Lamichhaney S, Jonas B, Markus S, Liman A, Khurram M, Manfred G, Alvaro MB, et al. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature. 2015; 518(7539): 371-5.
10. Irisarri I, Singh P, Koblmüller S, Torresdowdall J, Henning F, Franchini P, et al. Phylogenomics uncovers early hybridization and adaptive loci shaping the radiation of Lake Tanganyika cichlid fishes. Nature Communications. 2018; 9(1): 3159-71.
11. Wagner CE, Harmon LJ, Seehausen O. Ecological opportunity and sexual selection together predict adaptive radiation. Nature. 2012; 487(7407): 366-U124.
12. Simpson EH. Measurement of Diversity. Nature. 1949; 163(4148): 688-688.
13. Berner D, Salzburger W. The genomics of organismal diversification illuminated by adaptive radiations. Trends in Genetics. 2015; 31(9): 491-9.
14. Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, Fan SH, et al. The genomic substrate for adaptive radiation in African cichlid fish. Nature. 2014; 513(7518): 375-81.

474 15. Ford AG, Rüber L, Newton J, Dasmahapatra KK, Balarin JD, Bruun K, et al. Niche divergence
475 facilitated by fine-scale ecological partitioning in a recent cichlid fish adaptive radiation. Evolution. 2016;
476 70(12): 2718-35.
477 16. Fior S, Mingai L, Bengt O, Roberto V, Hodges SA, Lino O, et al. Spatiotemporal reconstruction of
478 the *Aquilegia* rapid radiation through next-generation sequencing of rapidly evolving cpDNA regions.
479 New Phytologist. 2013; 198(2): 579-92.
480 17. Filiault DL, Ballerini ES, Mandakova T, Akoz G, Derieg NJ, Schmutz J, et al. The *Aquilegia* genome
481 provides insight into adaptive radiation and reveals an extraordinarily polymorphic chromosome with a
482 unique history. elife. 2018; 7.
483 18. Munz. Aquilegia: the cultivated and wild columbines. Gentes Herbarum. 1946; 7: 1-150.
484 19. Bastida JM, Herrera CM. Extended phylogeny of *Aquilegia*: the biogeographical and ecological
485 patterns of two simultaneous but contrasting radiations. Plant Systematics and Evolution. 2010; 284(3-
486 4): 171-85.
487 20. Whittall JB, Hodges SA. Pollinator shifts drive increasingly long nectar spurs in columbine
488 flowers. Nature. 2007; 447(7145): 706-9.
489 21. Alcántara JM, Bastida JM, Rey PJ. Linking divergent selection on vegetative traits to
490 environmental variation and phenotypic diversification in the Iberian columbines (*Aquilegia*). Journal of
491 Evolutionary Biology. 2010; 23(6): 1218-33.
492 22. Kramer EM, Hodges SA. *Aquilegia* as a model system for the evolution and ecology of petals.
493 Philosophical Transactions Biological Sciences. 2010; 365(1539): 477-90.
494 23. Sharma B, Yant L, Hodges SA, Kramer EM. Understanding the development and evolution of
495 novel floral form in *Aquilegia*. Current Opinion in Plant Biology. 2014; 17(1): 22-7.
496 24. Garrido JL, Alcántara JM, Rey PJ, Medrano M, Guitián J, Castellanos MC, et al. Geographic
497 genetic structure of Iberian columbines (gen. *Aquilegia*). Plant Systematics and Evolution. 2017; 303(9):
498 1145-60.
499 25. Tang LL, Yu Q, Sun JF, Huang SQ. Floral traits and isolation of three sympatric *Aquilegia* species
500 in the Qinling Mountains, China. Plant Systematics and Evolution. 2007; 267(1/4): 121-8.
501 26. Li LF, Wang HY, Pang D, Liu Y, Liu B, Xiao HX. Phenotypic and genetic evidence for ecological
502 speciation of *Aquilegia japonica* and A. *oxysepala*. New Phytologist. 2014; 204(4): 1028-40.
503 27. Li LF, Pang D, Liao QL, Xiao HX. Genomic and EST microsatellite markers for *Aquilegia* flabellata
504 and cross-amplification in *A. oxysepala* (Ranunculaceae). American Journal of Botany. 2011; 98(8):213-5.

28. Li MR, Wang HY, Ding N, Lu T, Huang YC, Xiao HX, et al. Rapid divergence followed by adaptation to contrasting ecological niches of two closely related columbine species *Aquilegia japonica* and A. *oxysepala*. Genome Bioligy and Evolution. 2019; 11(3): 919-30.

29. Zhang H, Dawe RK. Mechanisms of plant spindle formation. Chromosome Research. 2011; 19(3): 335-44.

30. Aguirre-Portoles C, Bird AW, Hyman A, Canamero M, de Castro IP, Malunnbres M. *Tpx2* controls spindle integrity, genome stability, and tumor development. Cancer Research. 2012; 72(6): 1518-28.

31. Fukui K, Nakagawa N, Kitamura Y, Nishida Y, Masui R, Kuramitsu S. Crystal structure of *MutS2* endonuclease domain and the mechanism of homologous recombination suppression. Journal of Biological Chemistry. 2008; 283(48): 33417-27.

32. Machado-Schiaffino G, Henning F, Meyer A. Species-specific differences in adaptive phenotypic plasticity in an ecologically relevant trophic trait: hypertrophic lips in Midas cichlid fishes. Evolution. 2014; 68(7): 2086-91.

33. Bossdorf O, Richards CL, Massimo P. Epigenetics for ecologists. Ecology Letters. 2008; 11(2): 106-15.

34. Diez CM, Roessler K, Gaut BS. Epigenetics and plant genome evolution. Current Opinion in Plant Biology. 2014; 18(1): 1-8.

35. Verhoeven KJ, Vonholdt BM, Sork VL. Epigenetics in ecology and evolution: what we know and what we need to know. Molecular Ecology. 2016; 25(8): 1631-8.

36. Richards CL, Oliver B, Verhoeven KJF. Understanding natural epigenetic variation. New Phytologist. 2010; 187(3): 562-4.

37. Richards CL, Alonso C, Becker C, Bossdorf O, Bucher E, Colome-Tatche M, et al. Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. Ecology Letters. 2017; 20(12): 1576-90.

38. Becker C, Hagmann J, Jonas M, Daniel K, Oliver S, Karsten B, et al. Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. Nature. 2011; 480(7376): 245.

39. Schmitz RJ, Schultz MD, Lewsey MG, O'Malley RC, Urich MA, Ondrej L, et al. Transgenerational epigenetic instability is a source of novel methylation variants. Science. 2011; 334(6054): 369-73.

40. Cortijo S, Wardenaar R, Colome-Tatche M, Gilly A, Etcheverry M, Labadie K, et al. Mapping the epigenetic basis of complex traits. Science. 2014; 343(6175): 1145-8.

535 41. Smith AK, Kilaru V, Kocak M, Almli LM, Mercer KB, Ressler KJ, et al. Methylation quantitative
536 trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. BMC
537 Genomics. 2014; 15(1): 145.
538 42. Drong AW, George N, Hedman AK, Eshwar M, Elin G, Small KS, et al. The presence of
539 methylation quantitative trait loci indicates a direct genetic influence on the level of DNA methylation in
540 adipose tissue. PloS ONE. 2013; 8(2): e55923.
541 43. Fisher VA, Wang L, Deng X, Sarnowski C, Cupples LA, Liu CT. Do changes in DNA methylation
542 mediate or interact with SNP variation? A pharmacoepigenetic analysis. BMC Genetics. 2018; 19(Suppl
543 1): 70.
544 44. Li L, Wang C, Lu T, Lin S, Hu YQ. Indirect effect inference and application to GAW20 data. BMC
545 Genetics. 2018; 19(Suppl 1): 67.
546 45. Bewick AJ, Schmitz RJ. Gene body DNA methylation in plants. Current Opinion in Plant Biology.
547 2017; 36: 103-10.
548 46. Takuno S, Gaut BS. Gene body methylation is conserved between plant orthologs and is of
549 evolutionary consequence. Proceedings of the National Academy of Sciences of the United States of
550 America. 2013; 110(5): 1797-802.
551 47. Bewick AJ, Ji L, Niederhuth CE, Willing EM, Hofmeister BT, Shi X, et al. On the origin and
552 evolutionary consequences of gene body DNA methylation. Proceedings of the National Academy of
553 Sciences of the United States of America. 2016; 113(32): 9111-6.
554 48. Zilberman D. An evolutionary case for functional gene body methylation in plants and animals.
555 Genome Biology. 2017; 18(1): 87.
556 49. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform.
557 Bioinformatics, 2010; 25(14): 1754-60.
558 50. Mckenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome
559 Analysis Toolkit: A map reduce framework for analyzing next-generation DNA sequencing data. Genome
560 Research. 2010; 20(9): 1297-303.
561 51. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map
562 format and SAMtools. Bioinformatics. 2009;2 5(16): 2078-9.
563 52. Cingolani P, Platts A, Le LW, Coon M, Nguyen T, Luan W, et al. A program for annotating and
564 predicting the effects of single nucleotide polymorphisms, SnpEff. Fly. 2012; 6(2): 80-92.
565 53. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for
566 bigger datasets. Molecular Biology and Evolution. 2016; 33(7): 1870.

54. Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics. 2012; 28(24): 3326-8.

55. Alexander DH, Novembre JK. Fast model-based estimation of ancestry in unrelated individuals. Genome Research. 2009; 19(9): 1655-64.

56. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics. 2011; 27(15): 2156-8.

57. Weir BS, Cockerham CC. Estimating F-Statistics for the Analysis of Population Structure. Evolution. 1984; 38(6): 1358-70.

58. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution. 2007; 24(8): 1586-91.

59. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011; 27(11): 1571-2.

60. O''Sullivan DE, Johnson KC, Skinner L, Koestler DC, Christensen BC. Epigenetic and genetic burden measures are associated with tumor characteristics in invasive breast carcinoma. Epigenetics. 2016; 11(5): 344-53.

61. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinformatics. 2013; 14(2): 178-92.

62. Rice P, Longden I, Bleasby A. EMBOSS: The European molecular biology open software suite. Trends in Genetics. 2000; 16(6): 276-7.

63. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, David R. Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics. 2006; 38: 904-9.

64. Alexa A, Rahnenfuhrer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics. 2006; 22(13): 1600-7.

65. Bardou P, Mariette J, Escudie F, Djemiel C, Klopp C. jvenn: an interactive Venn diagram viewer. BMC Bioinformatics. 2014; 15: 293.

595 **Figure legend**

596 **Figure 1**. **Phylogenetic relationship and population structure of the ten worldwide *Aquilegia* species**.
597 (a) Phylogenetic tree of the 36 accessions constructed by neighbor joining algorithm based on whole-
598 genome SNPs. (b) PCA reveals genetic similarity within each of the three lineages and genetic disparity
599 between lineages based on 15,988 LD-pruned SNPs. Curves for each lineage denote 99% probability
600 ellipses estimated from distribution of the first two principal components. (c) Population admixture of
601 the 36 *Aquilegia* accessions. Most accessions exhibit almost pure ancestry composition matching the
602 presumed evolutionary history of the corresponding species.

603

604 **Figure 2**. **Functional enrichment of genes harboring highly impactful CCVs and DMGs.** CCV-containing
605 genes specific to either of the two lineages/species being compared were merged to construct a target
606 gene set. Ratio denotes proportion of CCV-containing genes or DMGs in the corresponding gene set of
607 interest. Absence of dot indicates no significant enrichment.

608

609 **Figure 3**. **Patterns of cytosine methylation for the ten worldwide *Aquilegia* species.** (a) Genome-wide
610 cytosine methylation level distributions of 36 accessions. (b) MDI illustrates chromosome-level CG
611 methylation similarity. *Aquilegia viridiflora* was used as the reference. (c) CG methylation profiling in
612 genic region across the four *Aquilegia* groups. Each row represents one genic region starting at 5kb
613 upstream of its TSS and terminating at 5kb downstream of its TES, sorted by mean methylation level of
614 all analyzed CG sites. Length of gene body was normalized. (d) CG methylation profiling in and around
615 CG island.

616

617 **Figure 4**. **Association between CG-loss variations and epigenetic variability**. (a) Top 3,000 most variable
618 CG sites containing CG-loss variations. (b) Top 3,000 most variable non-variant CG sites across 36
619 accessions show clade-specific methylation patterns. CG methylation in the hybrid tends to be
620 neutralized possibly due to heterozygosity. (c) Linear regression demonstrates that CG-loss variations
621 explain a large proportion of CG methylation variation. (d) Summary of composition of each category
622 with regard to whether each CG site contains a CG-loss variation. Epigenetic variability was determined
623 by standard deviation in methylation β value across all 36 accessions. CG sites with top 10,000, 10,001-
624 50,000 and 50,001-150,000 largest standard deviation were ordinally labelled as possessing "very high",
625 "high" and "moderate" variability respectively. The rest CG sites were labelled as possessing "low"

variability. (e) Association test shows most DMRs are independent of *cis*-acting SNPs. Results under different significance levels are compared for this exploratory analysis.

629 **Supplementary file**

630 **Figure S1**. Per-chromosome phylogenetic trees reconstructed using neighbor-joining method.
631 Polymorphisms detected on each chromosome were retrieved separately to infer the phylogeny.

632 **Figure S2**. Distribution of nucleotide diversity (π) at the whole-genome level and the per-chromosome
633 level. Nucleotide diversity was estimated for each lineage pooling corresponding species, as well as for *A.*
634 *japonica* and *A. oxysepala*.

635 **Figure S3**. Spearman correlation of the genome-wide nucleotide variation pattern for each 100-kb
636 sliding window between European and Asian (a), North American and Asian (b), European and North
637 American (d), *Aquilegia japonica* and *A. oxysepala* (d). Each dot represents a 100-kb sliding window.
638 Values on the x- and y-axis are the nucleotide diversity (π) for each sliding window.

639 **Figure S4**. Overlapping of low diversity genomic region (LDGR) in the three lineages. 148 LDGRs with 5%
640 lowest nucleotide diversity were defined as LDNRs in each lineage, totaling 241 unique regions.

641 **Figure S5**. Pair-wise d_N/d_S ratio for all genes between the species within and between the Asian,
642 European and North American lineages.

643 **Figure S6**. PCA illustrates three distinct clusters corresponding to the three lineages. Asian species
644 further demonstrated higher inter-specific divergence than the American and the European species. PCA
645 was performed based on 588,659 loci with sufficiently high sequencing quality.

646 **Figure S7**. Illustration of differential methylation in two photosynthesis genes. CG methylation pattern
647 of two genes, *Aqcoe7G230600* photosystem I *PsaA*/*PsaB* (a) and *Aqcoe7G231300 CemA* (b) in *A.*
648 *japonica* and *A. oxysepala* throughout the gene body region. Red bars indicate methylation level (0-100)
649 at CG sites. Genomic coordinates on the chromosome 7 are annotated.

650 **Figure S8**. Overlapping of the CCVs (a) and DMGs (b) identified in inter-lineage/species comparisons. A
651 considerable proportion of these CCVs (84.1%) and DMGs (51.3%) were shared by two or more inter-
652 lineage/species comparisons.

653 **Figure S9**. Venn analyses of the candidate genes carrying CCVs and DMGs. Each subpanel indicates the
654 comparison between the *A. japonica* and *A. oxysepala* (a), *A. japonica* and North American (b), *A.*
655 *japonica* and European (c), *A. oxysepala* and North American (d), *A. oxysepala* and European (e), North
656 American and European (f).

657

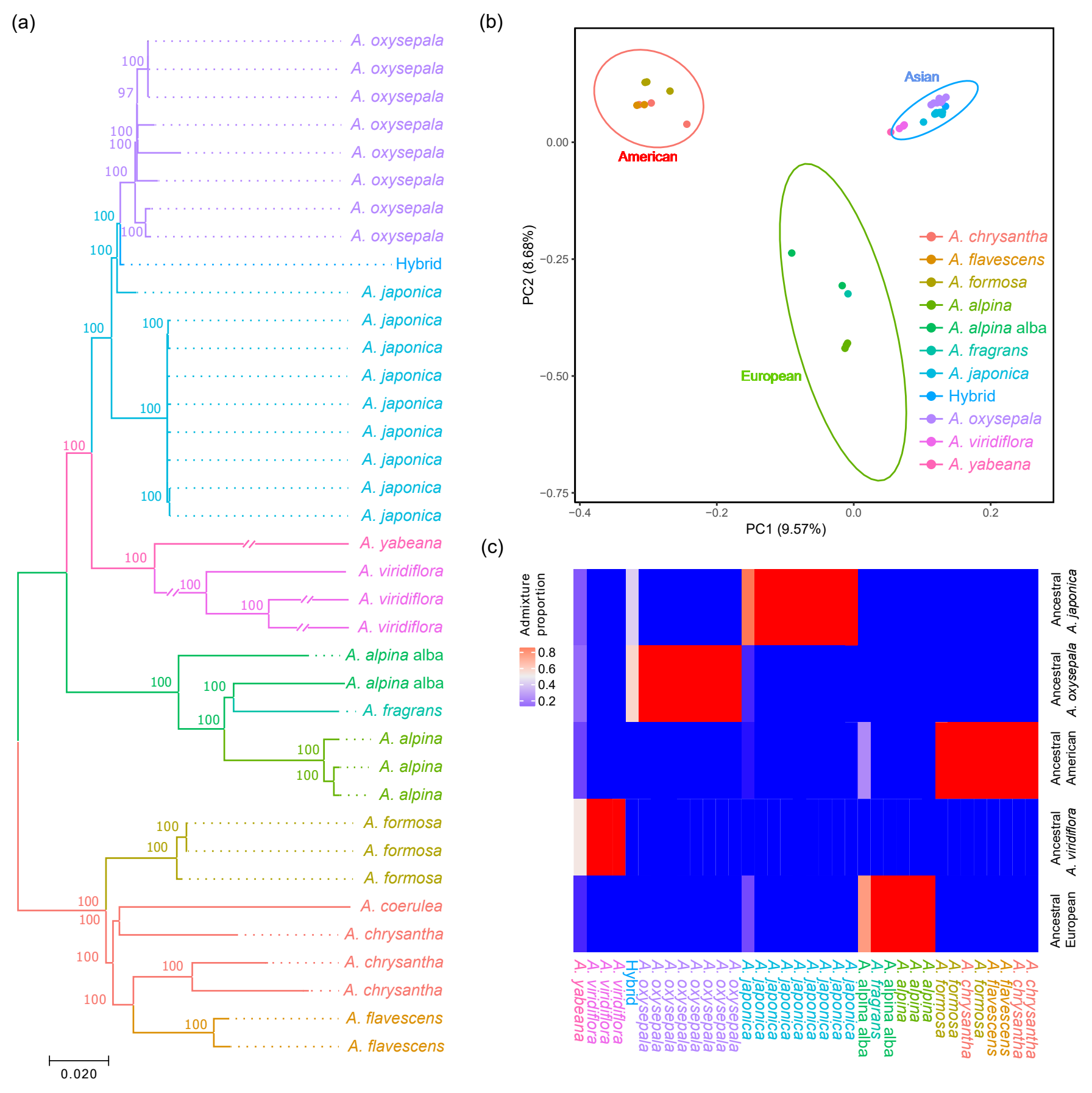
Table 1. Information of the high-impact conservative clade-specific variants (CCVs) in the cell reproduction related genes.

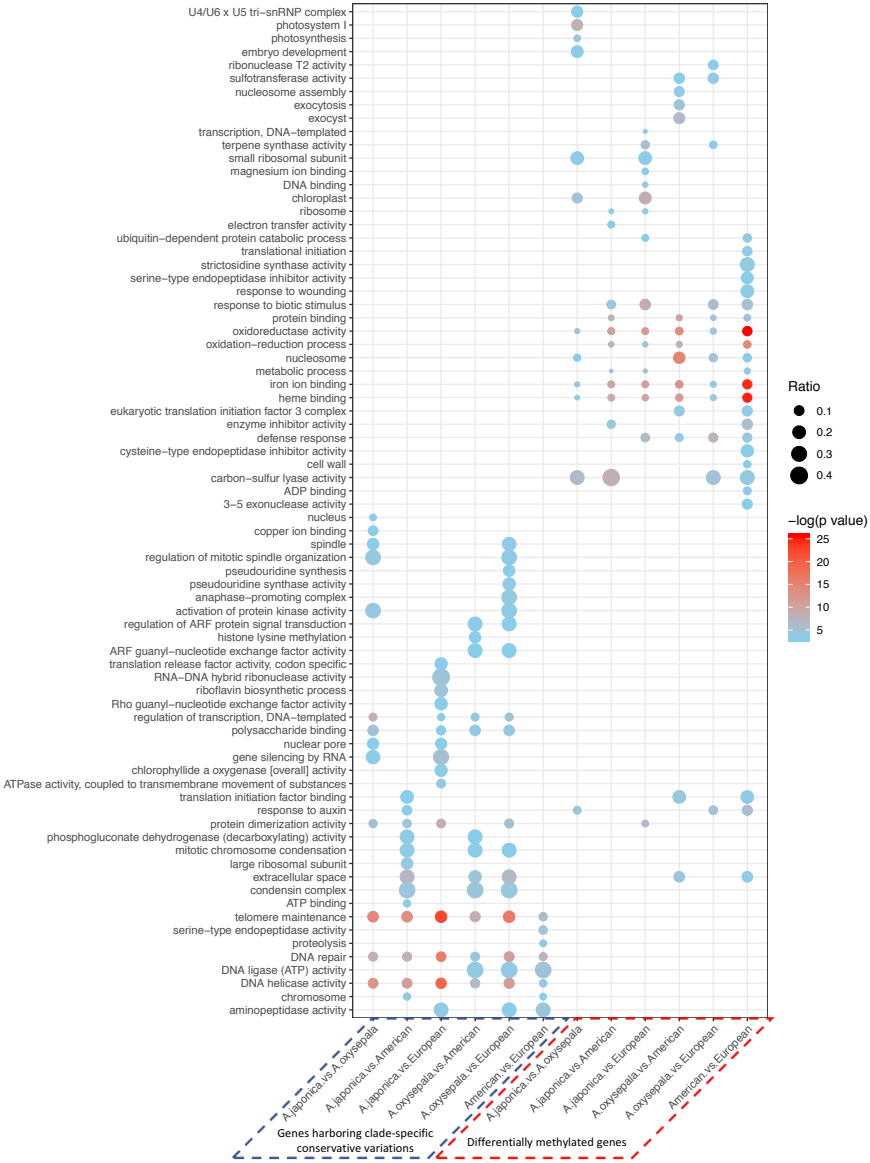
Gene	Variant-carrying	Reference	Chromosome	Position	Reference allele	Variant	Annotation	Gene function
Aqcoe1G273400	Asian	American	Chr1	18994915	GAA	GAAA	frameshift	DNA mismatch repair protein *MutS2*
Aqcoe2G151500	European	American	Chr2	15305837	A	G	splicing	PIF1-like helicase
	European	American		15307442	A	C	stop gain	
	European	American		15309865	AATATATAT	AATATATATAT	frameshift	
	European	Asian		15307442	A	C	stop gain	
	European	Asian		15309865	AATATATAT	AATATATATAT	frameshift	
	A. oxysepala	*A. japonica*		15305837	A	G	splicing	
	A. oxysepala	*A. japonica*		15309267	AT	A	frameshift	
Aqcoe2G177700	European	American	Chr2	21794397	TATGCACCAAAGGTATCACGATGC	TATGC	frameshift	PIF1-like helicase
	European	American		21794979	TT	TTGT	frameshift	
	European	Asian		21794397	TATGCACCAAAGGTATCACGATGC	TATGC	frameshift	
	A. oxysepala	*A. japonica*		21795089	CA	C	frameshift	
Aqcoe6G208600	European	American	Chr6	15364081	A	ATCTCTTCG	frameshift	PIF1-like helicase
	European	Asian		15364081	A	ATCTCTTCG	frameshift	
	A. japonica	*A. oxysepala*		15364330	TAA	TA	frameshift	
Aqcoe6G253800	European	American	Chr6	22789898	C	T	stop gain	DNA helicase
	European	American		22790012	G	A	splicing	
	European	Asian		22789898	C	T	stop gain	
	A. japonica	*A. oxysepala*		22790012	G	A	splicing	
Aqcoe2G276600	Asian	American	Chr2	33314422	AGGGGG	AGGGGGG	frameshift	DNA mismatch repair protein *Msh6*
Aqcoe6G160300	*A. japonica*	*A. oxysepala*	Chr6	9414625	G	A	stop gain	*TPX2*
Aqcoe7G062500	*A. oxysepala*	*A. japonica*	Chr7	3789055	G	A	stop gain	cell cycle regulated microtubule associated protein

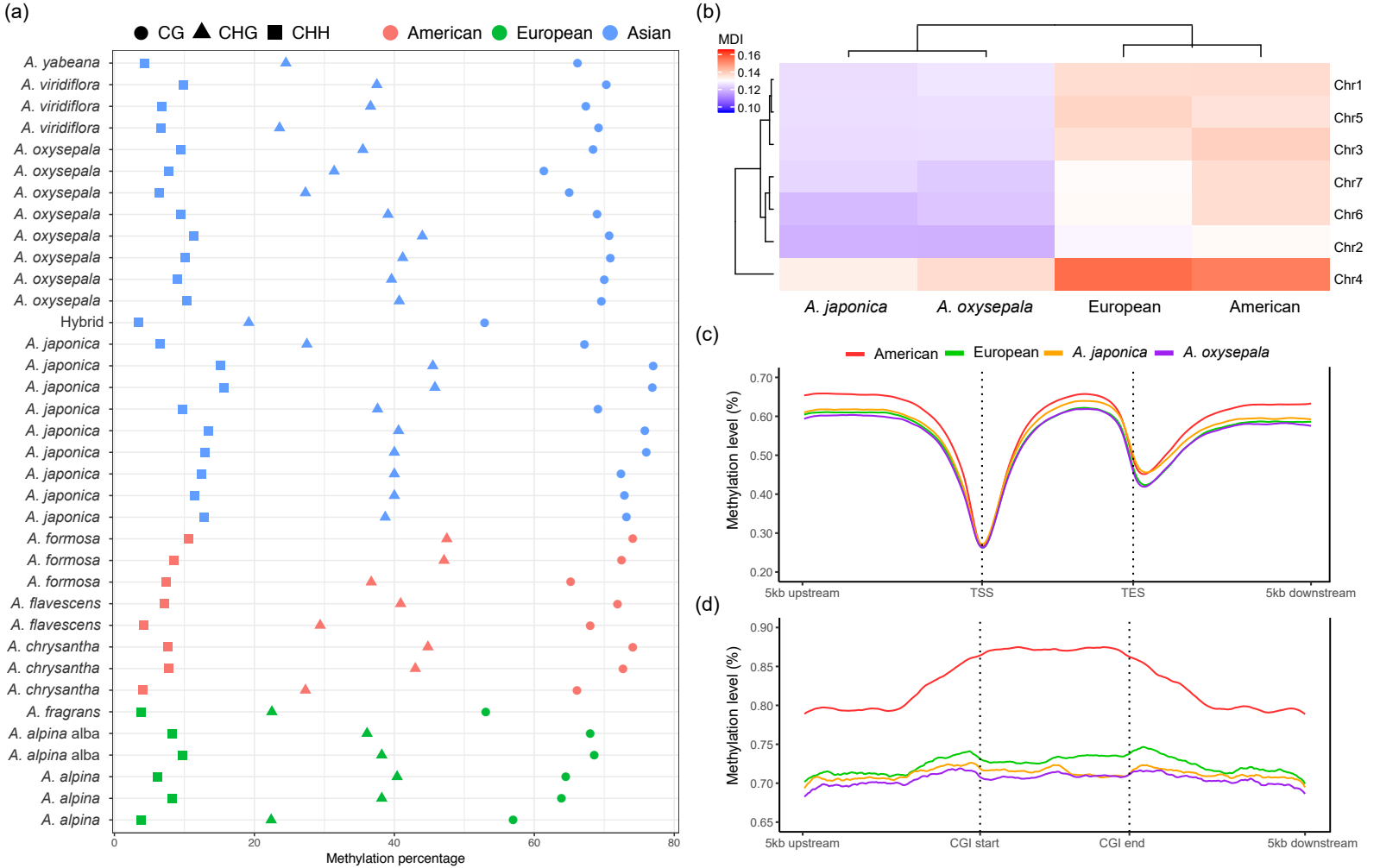
663 **Table 2.** Significant correlation between differential methylation and natural selection.
664

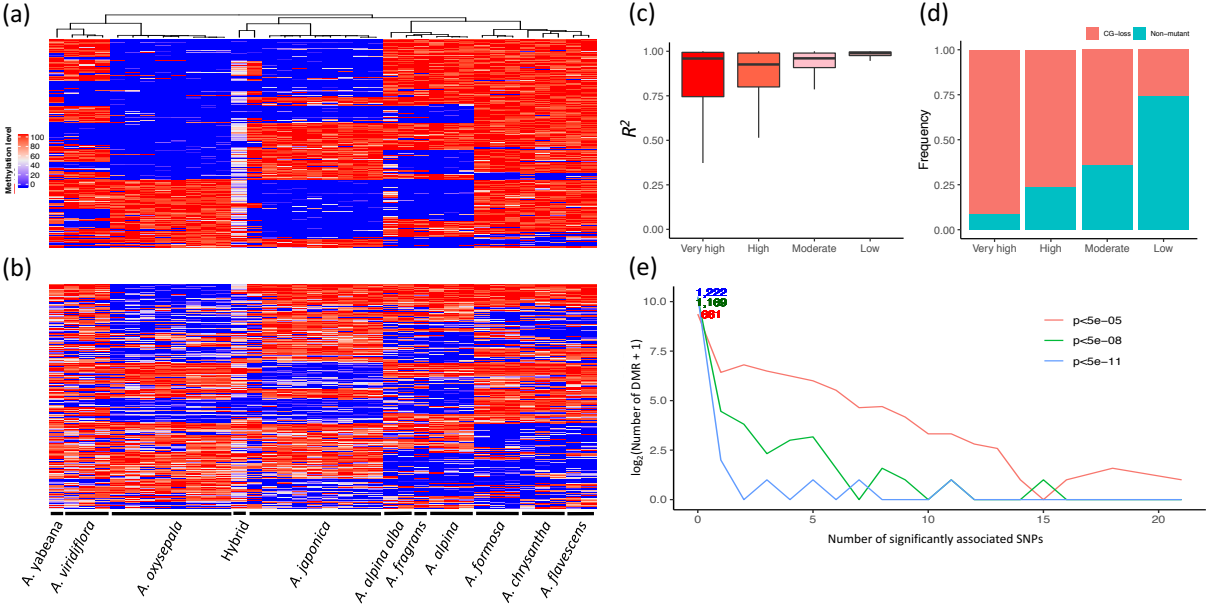
Type of selection	Differential methylation	Jap-Oxy*	Jap-Ame	Jap-Eur	Oxy-Ame	Oxy-Eur	Ame-Eur
Positive selection	DMG	7.2%	7.3%	11.9%	6.7%	8.4%	8.9%
	non-DMG	4.4%	5.0%	5.6%	4.7%	5.4%	5.4%
	p value	0.11	7.3e-02	3.9e-05	6.7e-02	1.8e-02	2.8e-04
Purifying selection	DMG	3.1%	1.8%	2.4%	2.3%	2.0%	1.9%
	non-DMG	4.3%	4.3%	4.7%	4.9%	5.1%	4.0%
	p value	0.53	3.2e-02	9.1e-02	1.3e-02	8.4e-03	1.0e-02

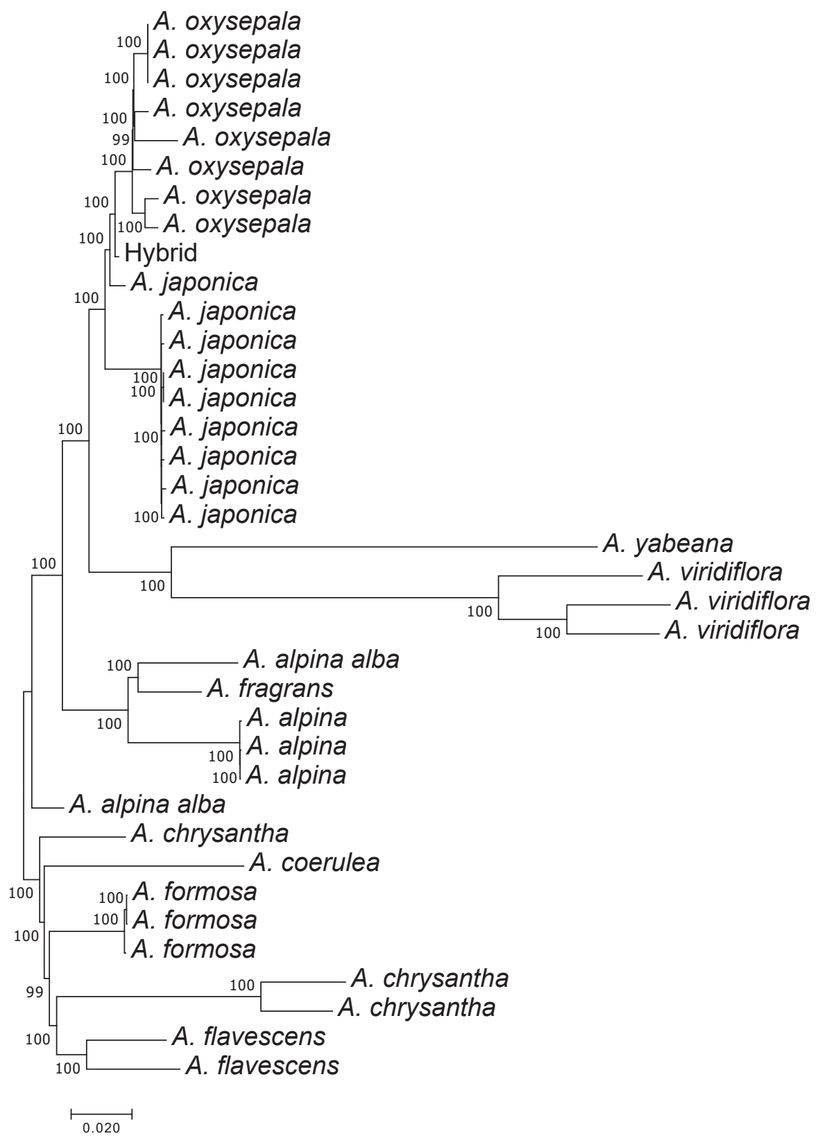
665 *: Each percentage represents the proportion of genes belonging to either DMGs or non-DMGs
666 compared between the two corresponding clades that are under corresponding or higher strength of
667 positive selection. For example, 7.2% indicates that 7.2% DMGs compared between *A. japonica* and *A.*
668 *oxysepala* are under strong selection; 4.4% indicates that 4.4% non-DMGs compared between these two
669 species are under strong selection. *p* values were obtained from Chi-square tests and were not adjusted
670 for multiple testing due to dependence arising from overlapping gene sets.
671 Jap: *A. japonica*; Oxy: *A. oxysepala*; Ame: American; Eur: European

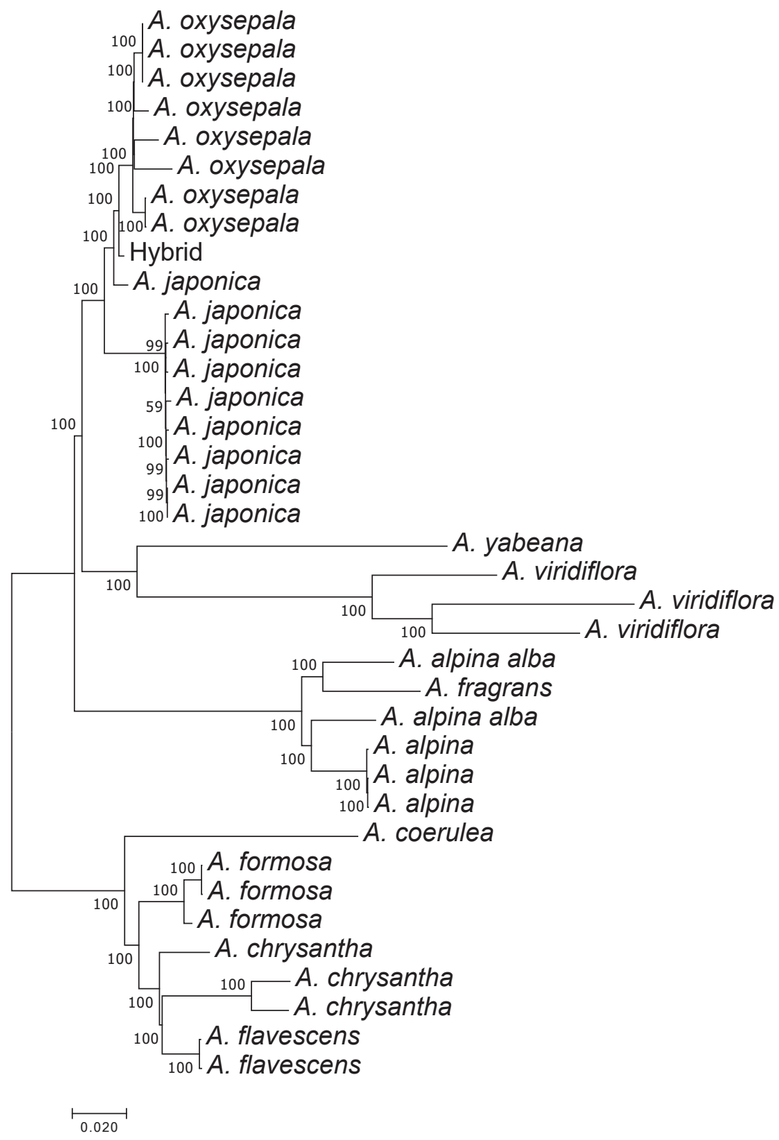

(a)
A. oxysepala
A. oxysepala
A. oxysepala
A. oxysepala
A. oxysepala
A. oxysepala
A. oxysepala
A. oxysepala
Hybrid
A. japonica
A. japonica
A. japonica
A. japonica
A. japonica
A. japonica
A. japonica
A. japonica
A. yabeana
A. viridiflora
A. viridiflora
A. viridiflora
A. alpina alba
A. alpina alba
A. fragrans
A. alpina
A. alpina
A. alpina
A. formosa
A. formosa
A. formosa
A. coerulea
A. chrysantha
A. chrysantha
A. chrysantha
A. flavescens
A. flavescens
0.020
(b)
Asian
American
European
PC2 (8.68%)
PC1 (9.57%)
0.00
−0.25
−0.50
−0.75
−0.4
−0.2
0.0
0.2
A. chrysantha
A. flavescens
A. formosa
A. alpina
A. alpina alba
A. fragrans
A. japonica
Hybrid
A. oxysepala
A. viridiflora
A. yabeana
(c)
Admixture proportion
0.8
0.6
0.4
0.2
Ancestral A. japonica
Ancestral A. oxysepala
Ancestral American
Ancestral A. viridiflora
Ancestral European

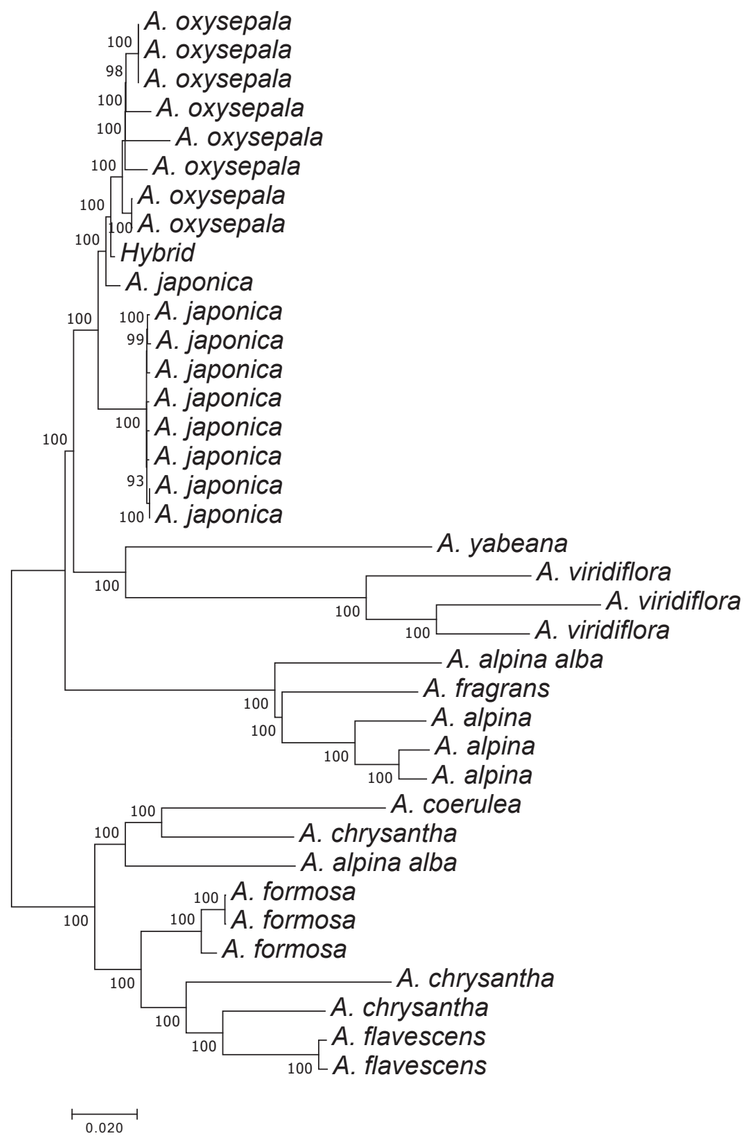

U4/U6 x U5 tri−snRNP complex
photosystem I
photosynthesis
embryo development
ribonuclease T2 activity
sulfotransferase activity
nucleosome assembly
exocytosis
exocyst
transcription, DNA−templated
terpene synthase activity
small ribosomal subunit
magnesium ion binding
DNA binding
chloroplast
ribosome
electron transfer activity
ubiquitin−dependent protein catabolic process
translational initiation
strictosidine synthase activity
serine−type endopeptidase inhibitor activity
response to wounding
response to biotic stimulus
protein binding
oxidoreductase activity
oxidation−reduction process
nucleosome
metabolic process
iron ion binding
heme binding
eukaryotic translation initiation factor 3 complex
enzyme inhibitor activity
defense response
cysteine−type endopeptidase inhibitor activity
cell wall
carbon−sulfur lyase activity
ADP binding
3−5 exonuclease activity
nucleus
copper ion binding
spindle
regulation of mitotic spindle organization
pseudouridine synthesis
pseudouridine synthase activity
anaphase−promoting complex
activation of protein kinase activity
regulation of ARF protein signal transduction
histone lysine methylation
ARF guanyl−nucleotide exchange factor activity
translation release factor activity, codon specific
RNA−DNA hybrid ribonuclease activity
riboflavin biosynthetic process
Rho guanyl−nucleotide exchange factor activity
regulation of transcription, DNA−templated
polysaccharide binding
nuclear pore
gene silencing by RNA
chlorophyllide a oxygenase [overall] activity
ATPase activity, coupled to transmembrane movement of substances
translation initiation factor binding
response to auxin
protein dimerization activity
phosphogluconate dehydrogenase (decarboxylating) activity
mitotic chromosome condensation
large ribosomal subunit
extracellular space
condensin complex
ATP binding
telomere maintenance
serine−type endopeptidase activity
proteolysis
DNA repair
DNA ligase (ATP) activity
DNA helicase activity
chromosome
aminopeptidase activity
A.japonica.vs.A.oxysepala
A.japonica.vs.American
A.japonica.vs.European
A.oxysepala.vs.American
A.oxysepala.vs.European
American.vs.European
A.japonica.vs.A.oxysepala
A.japonica.vs.American
A.japonica.vs.European
A.oxysepala.vs.American
A.oxysepala.vs.European
American.vs.European
Genes harboring clade-specific conservative variations
Differentially methylated genes
Ratio
0.1
0.2
0.3
0.4
−log(p value)
25
20
15
10
5

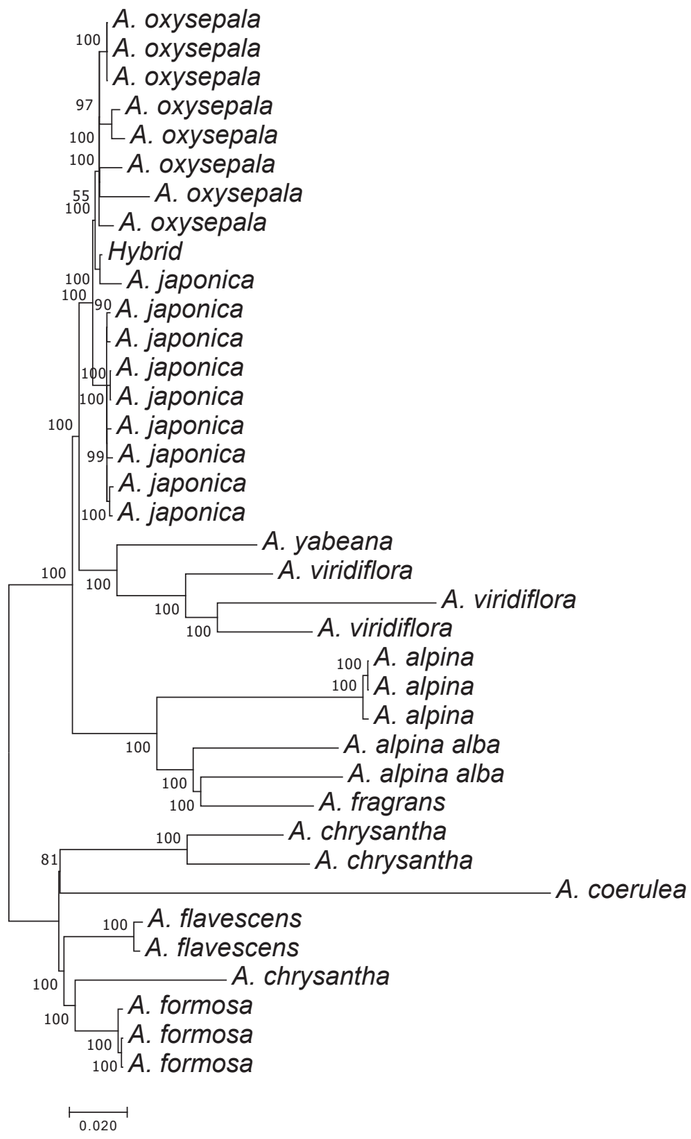

(a)
CG
CHG
CHH
American
European
Asian
A. yabeana
A. viridiflora
A. viridiflora
A. viridiflora
A. oxysepala
A. oxysepala
A. oxysepala
A. oxysepala
A. oxysepala
A. oxysepala
A. oxysepala
A. oxysepala
Hybrid
A. japonica
A. japonica
A. japonica
A. japonica
A. japonica
A. japonica
A. japonica
A. japonica
A. formosa
A. formosa
A. formosa
A. flavescens
A. flavescens
A. chrysantha
A. chrysantha
A. chrysantha
A. fragrans
A. alpina alba
A. alpina alba
A. alpina
A. alpina
A. alpina
0
20
40
60
80
Methylation percentage
(b)
MDI
0.16
0.14
0.12
0.10
Chr1
Chr5
Chr3
Chr7
Chr6
Chr2
Chr4
A. japonica
A. oxysepala
European
American
(c)
American
European
A. japonica
A. oxysepala
Methylation level (%)
0.70
0.60
0.50
0.40
0.30
0.20
5kb upstream
TSS
TES
5kb downstream
(d)
Methylation level (%)
0.90
0.85
0.80
0.75
0.70
0.65
5kb upstream
CGI start
CGI end
5kb downstream

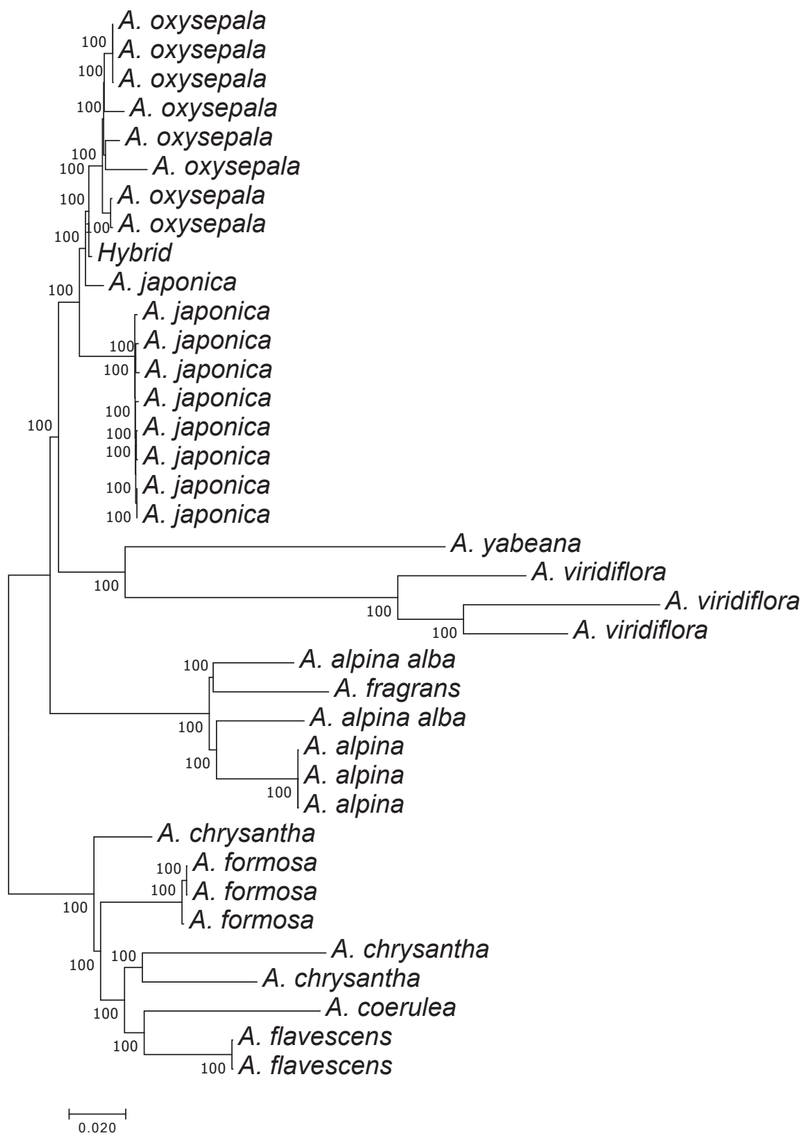

(a)
Methylation level
100
80
60
40
20
0
(b)
A. yabeana
A. viridiflora
A. oxysepala
Hybrid
A. japonica
A. alpina alba
A. fragrans
A. alpina
A. formosa
A. chrysantha
A. flavescens
(c)
R^2
1.00
0.75
0.50
0.25
0.00
Very high
High
Moderate
Low
(d)
CG-loss
Non-mutant
Frequency
1.00
0.75
0.50
0.25
0.00
Very high
High
Moderate
Low
(e)
1,222
1,169
581
p<5e−05
p<5e−08
p<5e−11
log2(Number of DMR + 1)
10.0
7.5
5.0
2.5
0.0
0
5
10
15
20
Number of significantly associated SNPs

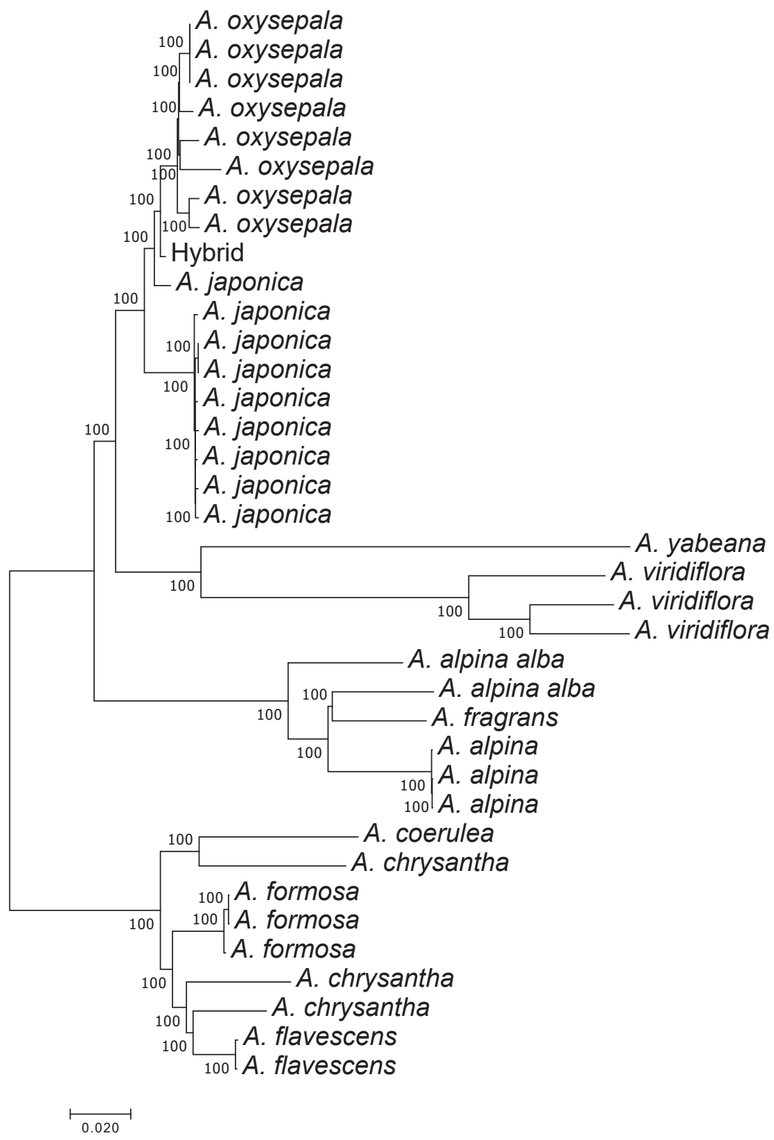

Chr1
A. oxysepala
Hybrid
A. japonica
A. yabeana
A. viridiflora
A. alpina alba
A. fragrans
A. alpina
A. coerulea
A. formosa
A. chrysantha
A. flavescens
0.020

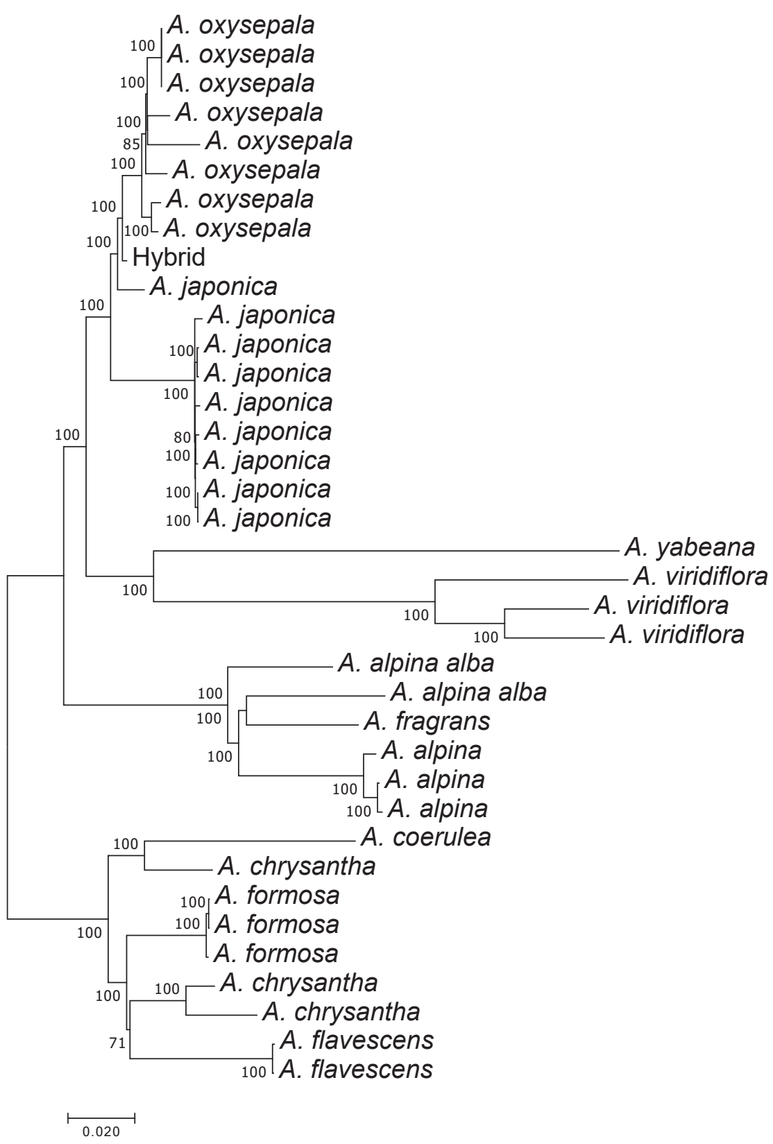

Chr2
A. oxysepala
Hybrid
A. japonica
A. yabeana
A. viridiflora
A. alpina alba
A. fragrans
A. alpina
A. coerulea
A. formosa
A. chrysantha
A. flavescens
0.020

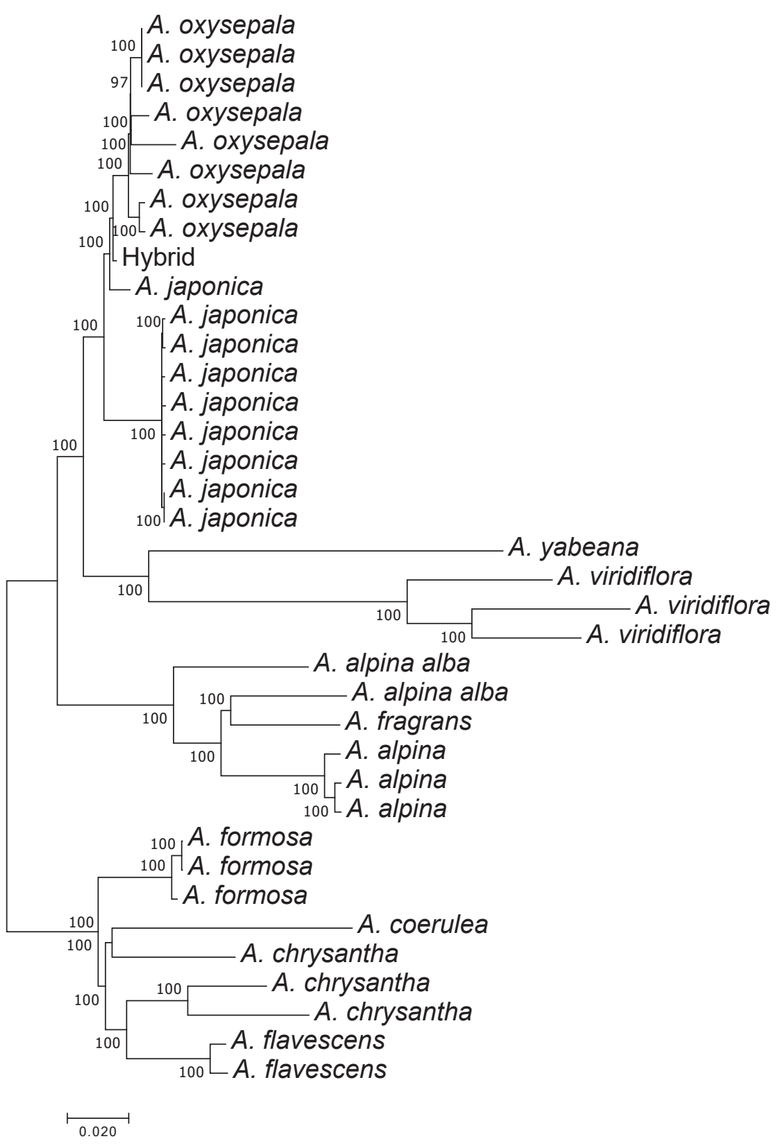

Chr3
A. oxysepala
Hybrid
A. japonica
A. yabeana
A. viridiflora
A. alpina alba
A. fragrans
A. alpina
A. coerulea
A. chrysantha
A. formosa
A. flavescens
0.020

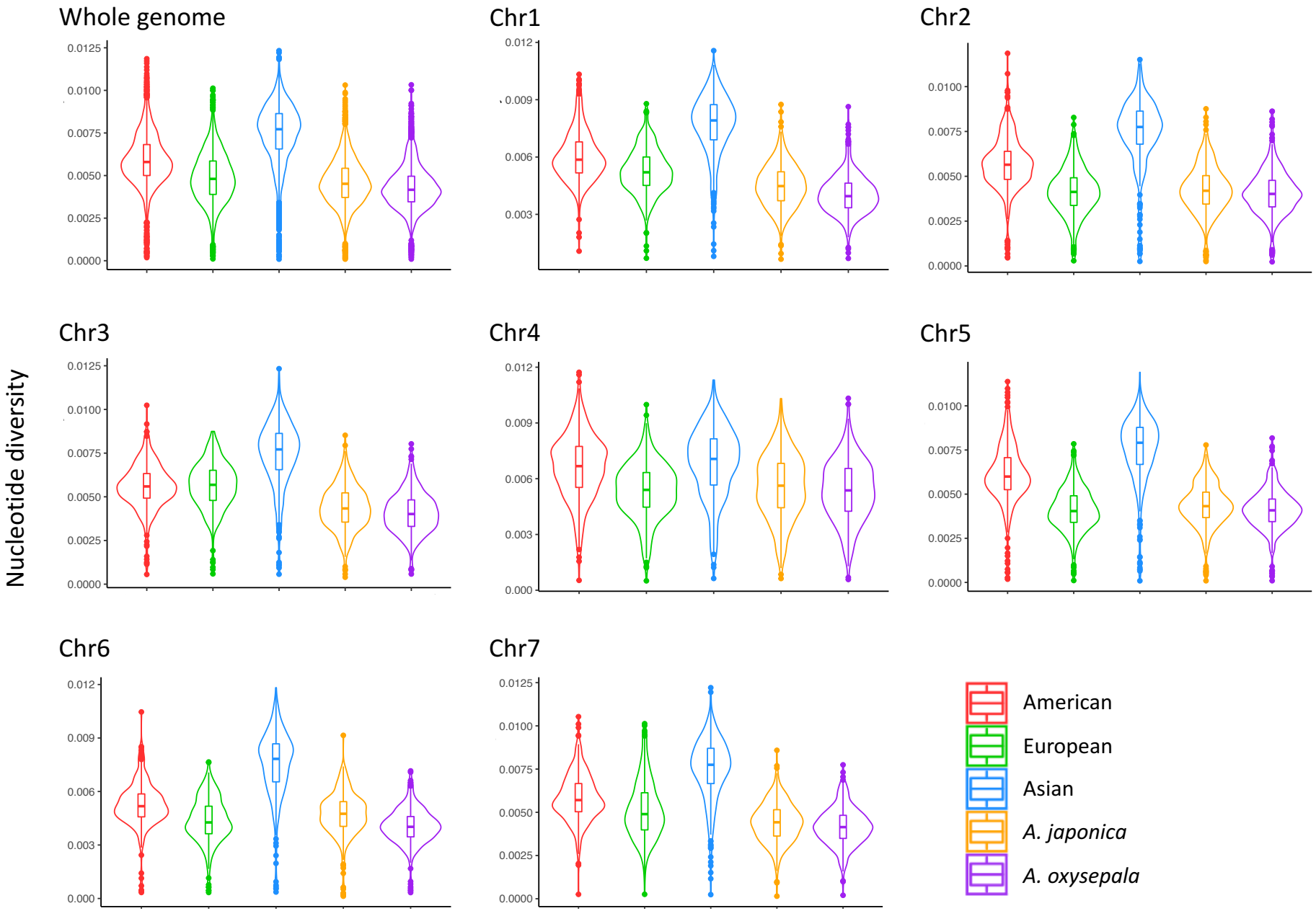

Chr4
A. oxysepala
Hybrid
A. japonica
A. yabeana
A. viridiflora
A. alpina
A. alpina alba
A. fragrans
A. chrysantha
A. coerulea
A. flavescens
A. formosa
0.020

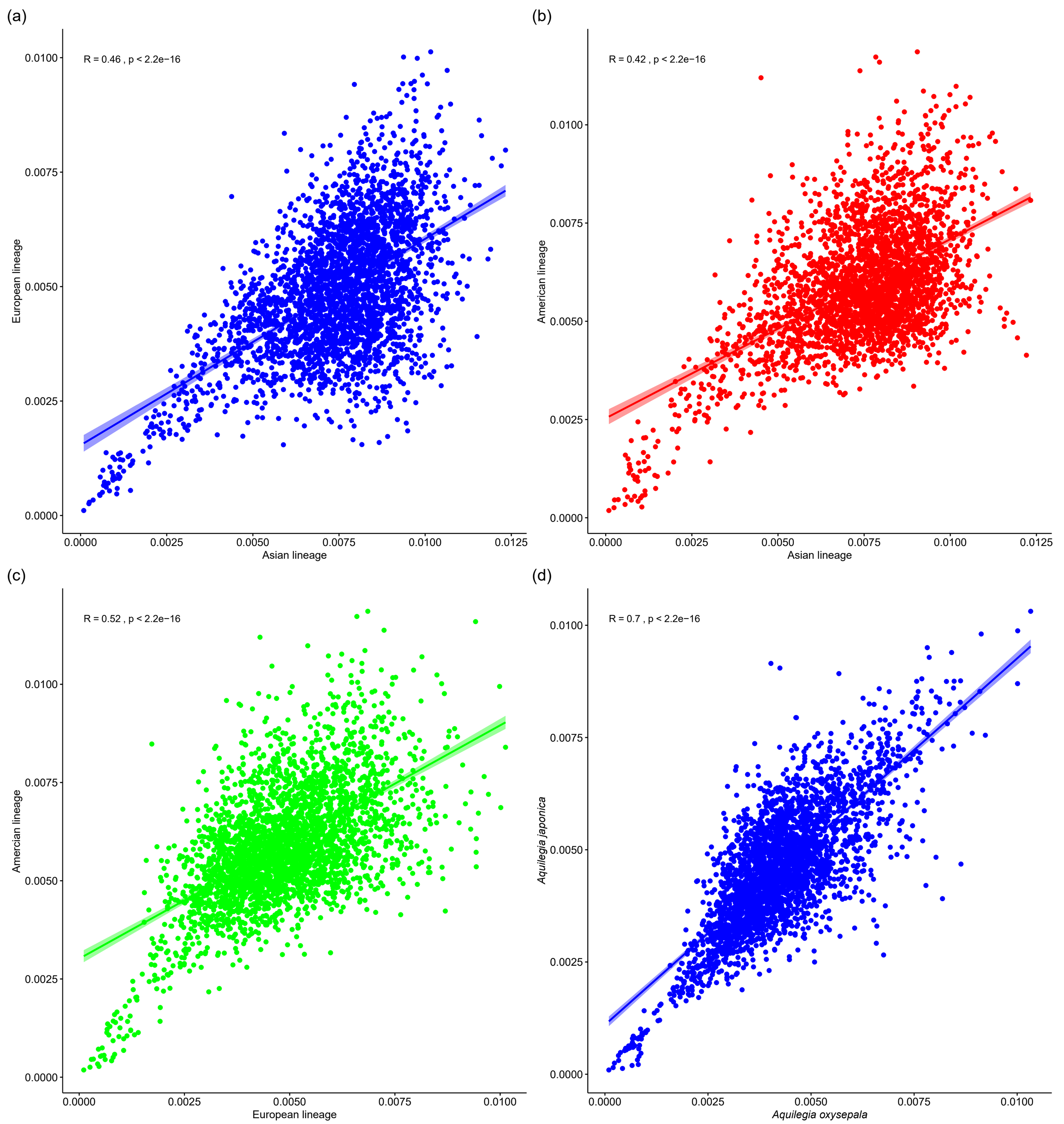

Chr5
A. oxysepala
Hybrid
A. japonica
A. yabeana
A. viridiflora
A. alpina alba
A. fragrans
A. alpina
A. chrysantha
A. formosa
A. coerulea
A. flavescens
0.020

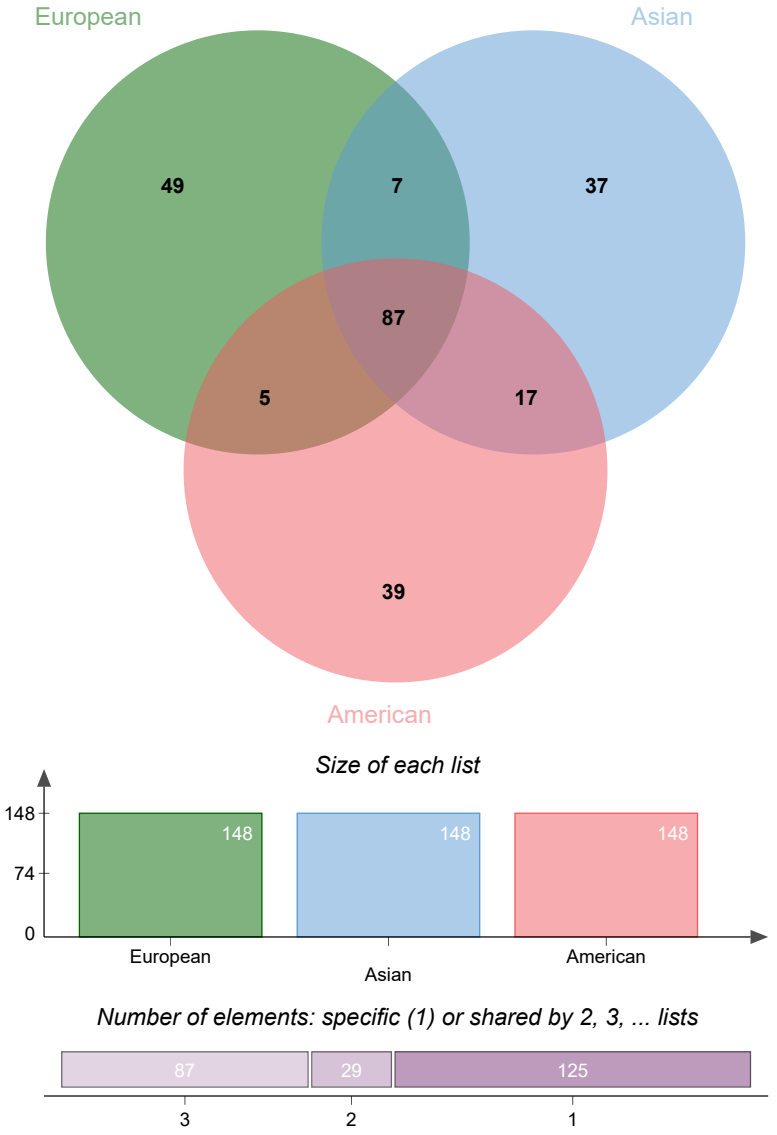

Chr6
A. oxysepala
Hybrid
A. japonica
A. yabeana
A. viridiflora
A. alpina alba
A. fragrans
A. alpina
A. coerulea
A. chrysantha
A. formosa
A. flavescens
0.020

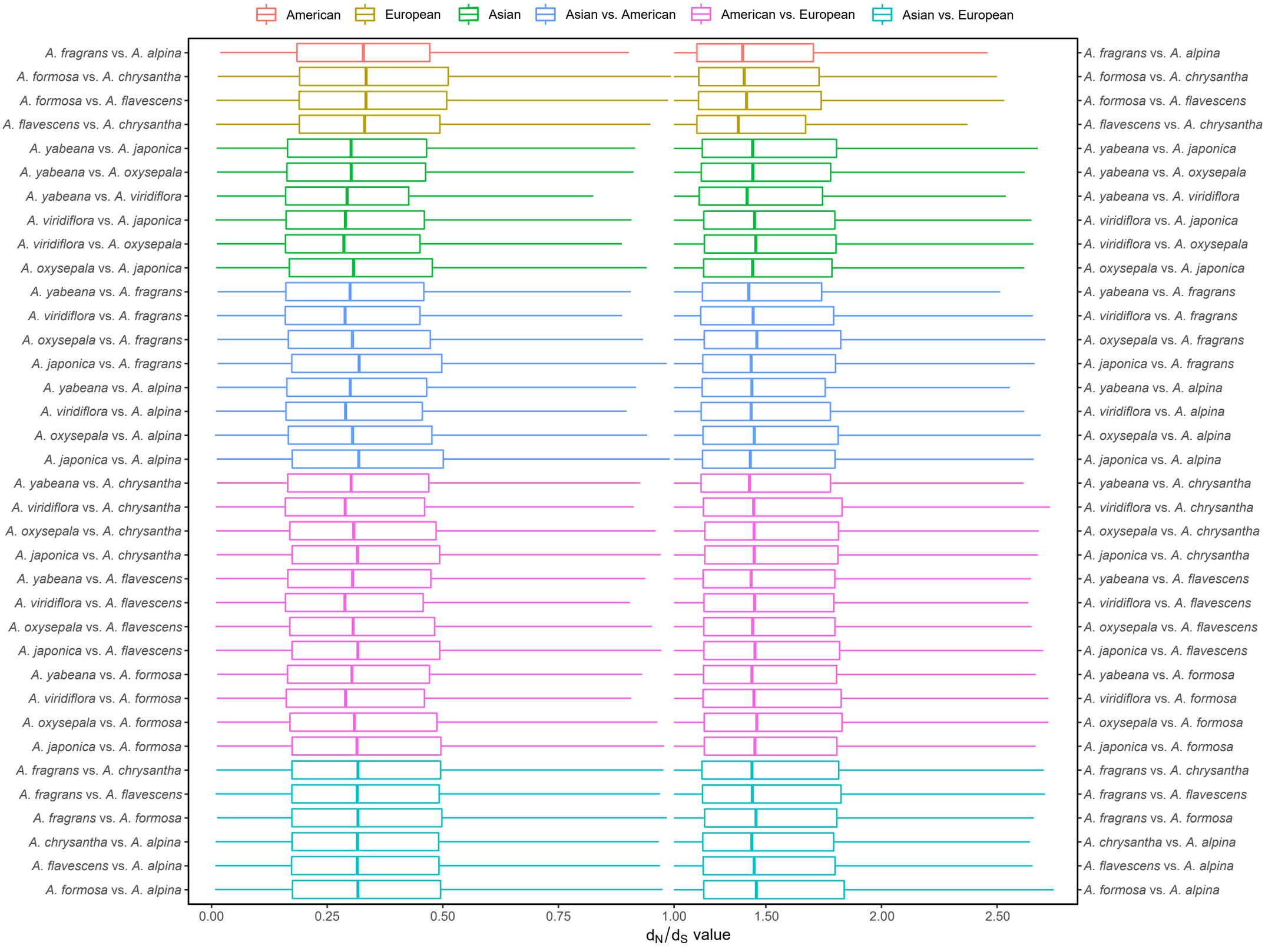

Chr7
A. oxysepala
Hybrid
A. japonica
A. yabeana
A. viridiflora
A. alpina alba
A. fragrans
A. alpina
A. coerulea
A. chrysantha
A. formosa
A. flavescens
0.020

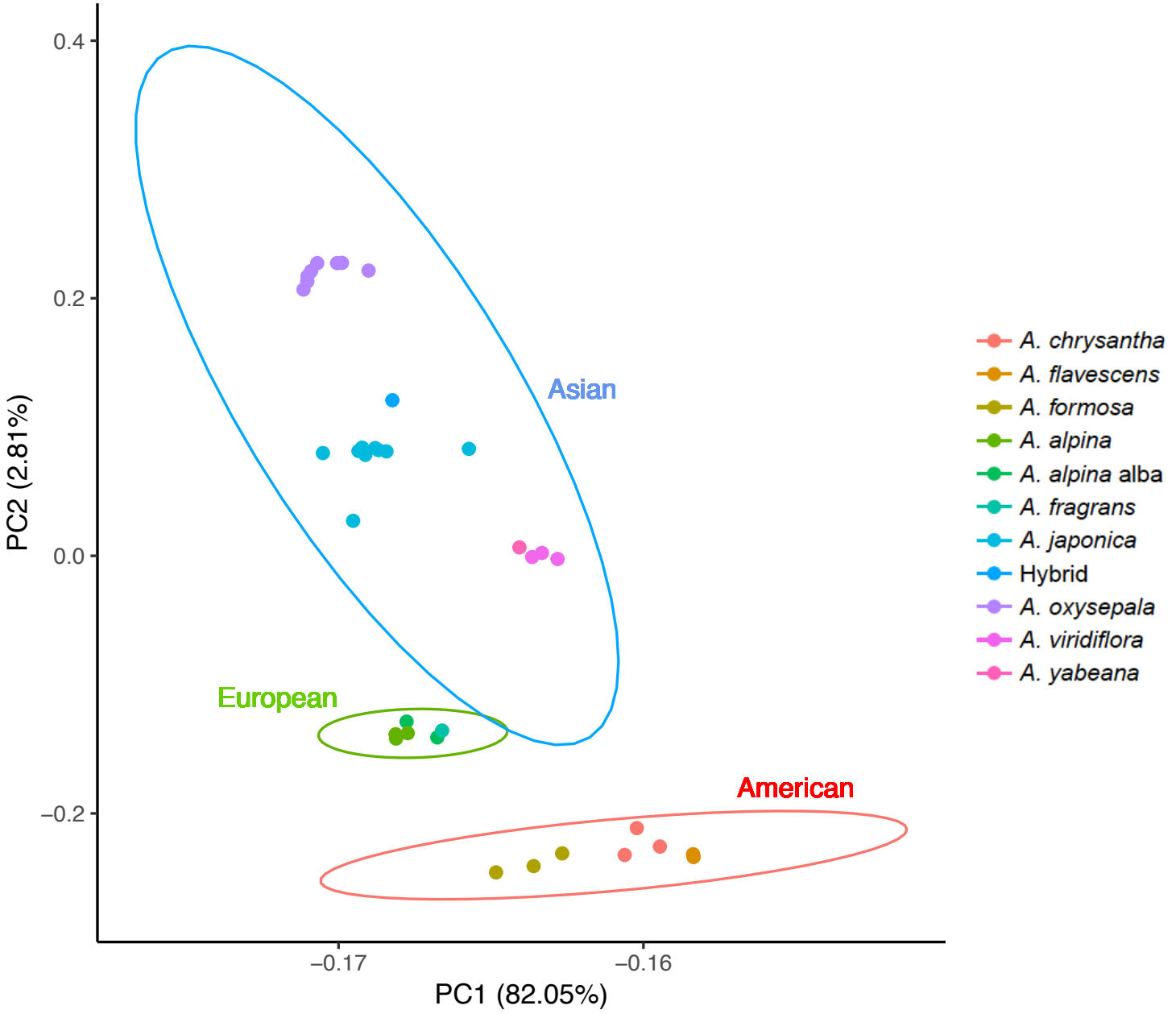

All
A. oxysepala
Hybrid
A. japonica
A. yabeana
A. viridiflora
A. alpina alba
A. fragrans
A. alpina
A. formosa
A. coerulea
A. chrysantha
A. flavescens
0.020


Whole genome
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Nucleotide diversity
American
European
Asian
A. japonica
A. oxysepala

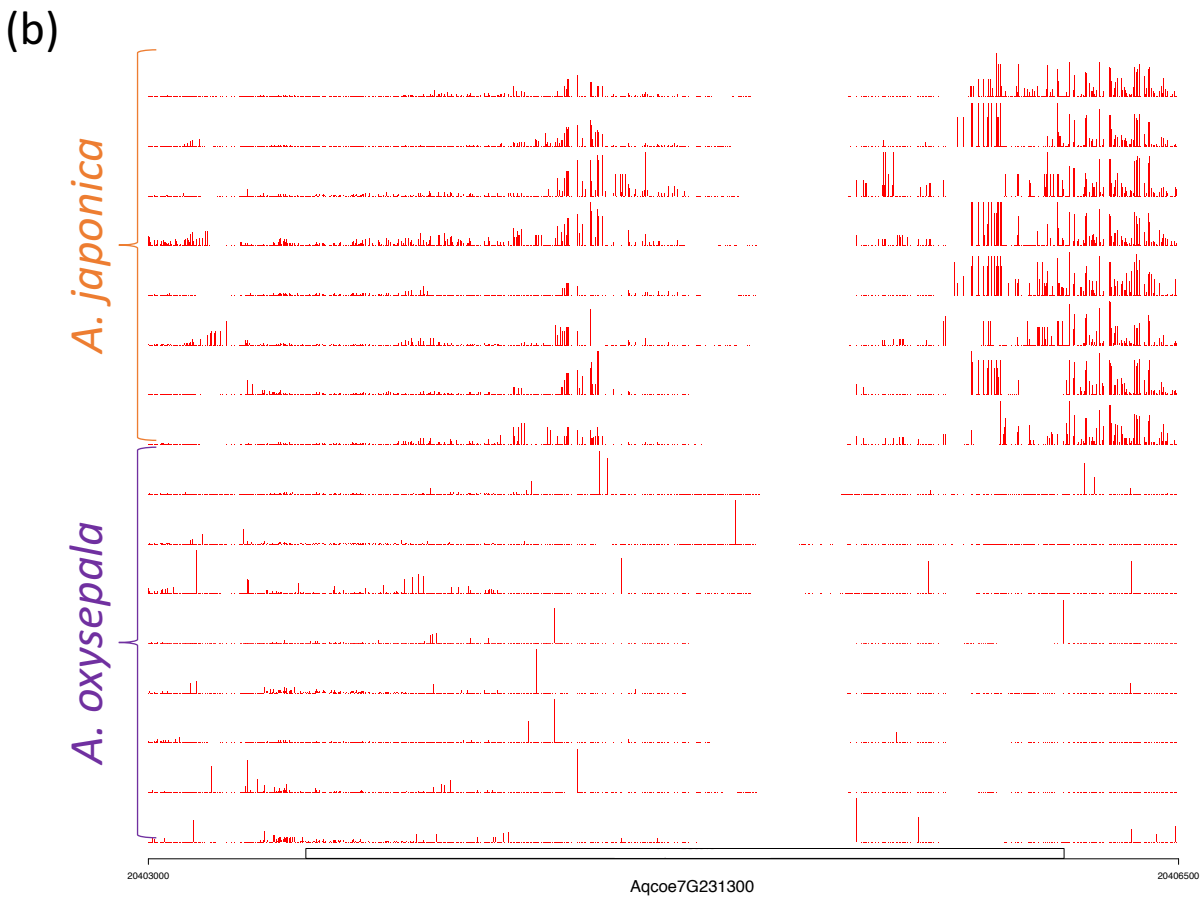

(a)
R = 0.46 , p < 2.2e−16
European lineage
Asian lineage
0.0000
0.0025
0.0050
0.0075
0.0100
0.0125
(b)
R = 0.42 , p < 2.2e−16
American lineage
Asian lineage
(c)
R = 0.52 , p < 2.2e−16
Amercian lineage
European lineage
(d)
R = 0.7 , p < 2.2e−16
Aquilegia japonica
Aquilegia oxysepala


European
Asian
American
49
7
37
87
5
17
39
Size of each list
148
74
0
148
148
148
European
Asian
American
Number of elements: specific (1) or shared by 2, 3, ... lists
87
29
125
3
2
1

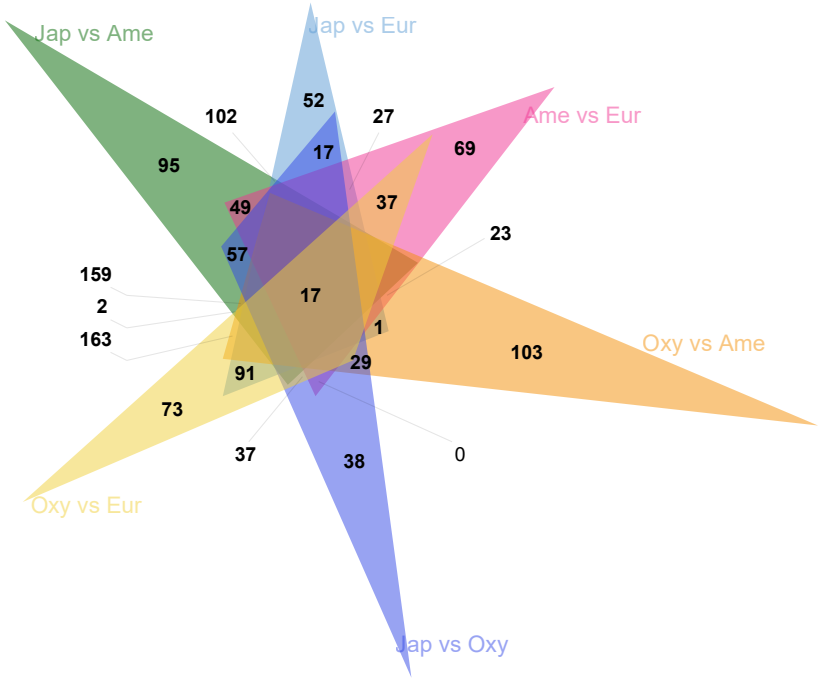

American
European
Asian
Asian vs. American
American vs. European
Asian vs. European
A. fragrans vs. A. alpina
A. formosa vs. A. chrysantha
A. formosa vs. A. flavescens
A. flavescens vs. A. chrysantha
A. yabeana vs. A. japonica
A. yabeana vs. A. oxysepala
A. yabeana vs. A. viridiflora
A. viridiflora vs. A. japonica
A. viridiflora vs. A. oxysepala
A. oxysepala vs. A. japonica
A. yabeana vs. A. fragrans
A. viridiflora vs. A. fragrans
A. oxysepala vs. A. fragrans
A. japonica vs. A. fragrans
A. yabeana vs. A. alpina
A. viridiflora vs. A. alpina
A. oxysepala vs. A. alpina
A. japonica vs. A. alpina
A. yabeana vs. A. chrysantha
A. viridiflora vs. A. chrysantha
A. oxysepala vs. A. chrysantha
A. japonica vs. A. chrysantha
A. yabeana vs. A. flavescens
A. viridiflora vs. A. flavescens
A. oxysepala vs. A. flavescens
A. japonica vs. A. flavescens
A. yabeana vs. A. formosa
A. viridiflora vs. A. formosa
A. oxysepala vs. A. formosa
A. japonica vs. A. formosa
A. fragrans vs. A. chrysantha
A. fragrans vs. A. flavescens
A. fragrans vs. A. formosa
A. chrysantha vs. A. alpina
A. flavescens vs. A. alpina
A. formosa vs. A. alpina
0.00
0.25
0.50
0.75
1.00
1.50
2.00
2.50
d_N/d_S value

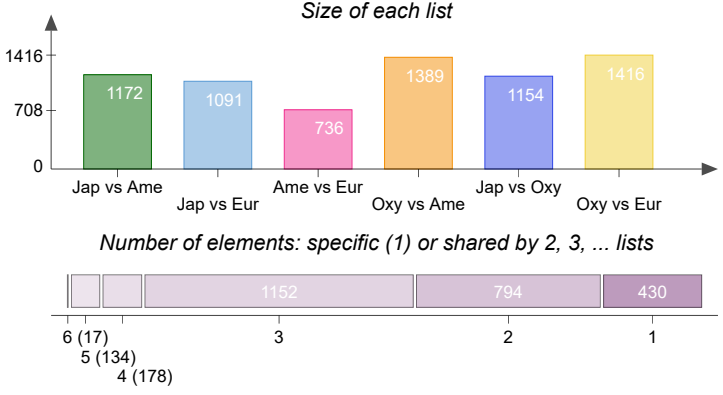

0.4
0.2
0.0
−0.2
PC2 (2.81%)
−0.17
−0.16
PC1 (82.05%)
Asian
European
American
A. chrysantha
A. flavescens
A. formosa
A. alpina
A. alpina alba
A. fragrans
A. japonica
Hybrid
A. oxysepala
A. viridiflora
A. yabeana

(a)
A. japonica
A. oxysepala
20392800
20394100
Aqcoe7G230600

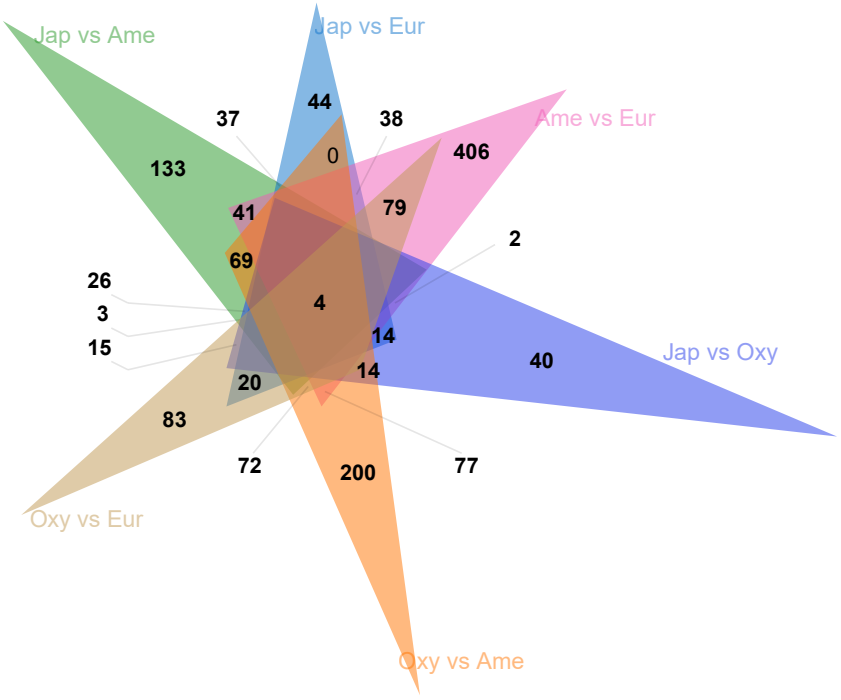

(b)
A. japonica
A. oxysepala
20403000
20406500
Aqcoe7G231300

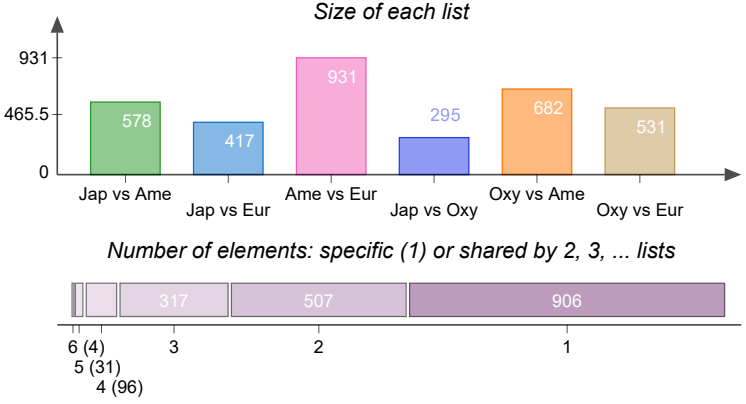
(a)


Jap vs Ame
Jap vs Eur
Ame vs Eur
Oxy vs Ame
Oxy vs Eur
Jap vs Oxy
102
52
27
95
17
69
49
37
23
57
159
17
2
1
163
103
91
29
73
37
38
0

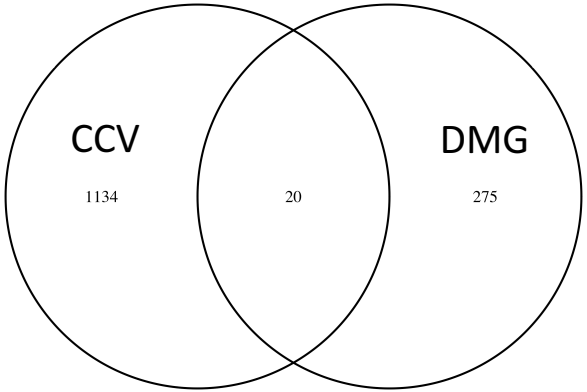

Size of each list
1416
708
0
1172
1091
736
1389
1154
1416
Jap vs Ame
Jap vs Eur
Ame vs Eur
Oxy vs Ame
Jap vs Oxy
Oxy vs Eur
Number of elements: specific (1) or shared by 2, 3, ... lists
1152
794
430
6 (17)
5 (134)
4 (178)
3
2
1

(b)

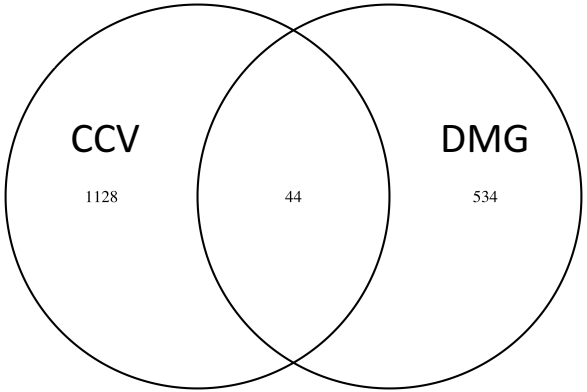

Jap vs Ame
Jap vs Eur
Ame vs Eur
Jap vs Oxy
Oxy vs Eur
Oxy vs Ame
37
44
38
133
0
406
41
79
2
69
26
4
3
14
15
40
14
20
83
72
200
77


Size of each list
931
465.5
0
578
417
931
295
682
531
Jap vs Ame
Jap vs Eur
Ame vs Eur
Jap vs Oxy
Oxy vs Ame
Oxy vs Eur
Number of elements: specific (1) or shared by 2, 3, ... lists
317
507
906
6 (4)
5 (31)
4 (96)
3
2
1

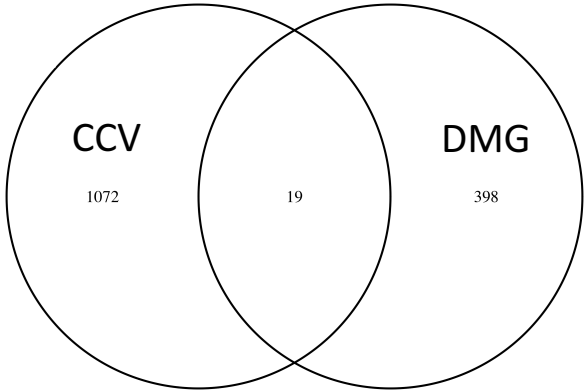
(a) *A. japonica* and *A. oxysepala*



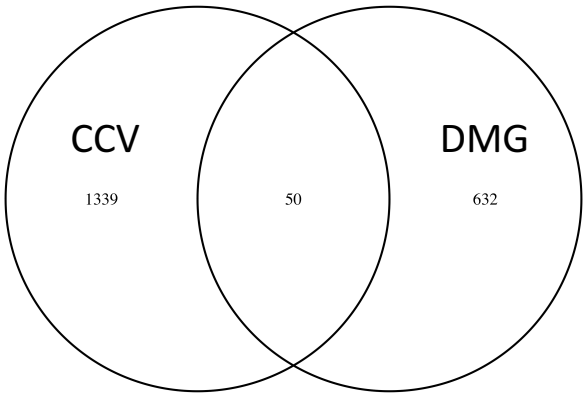
(b) *A. japonica* and North Amrican



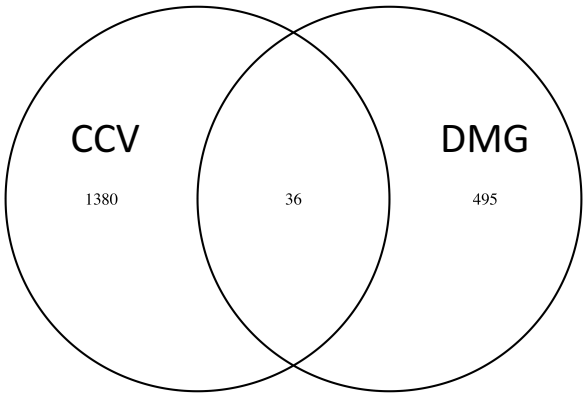
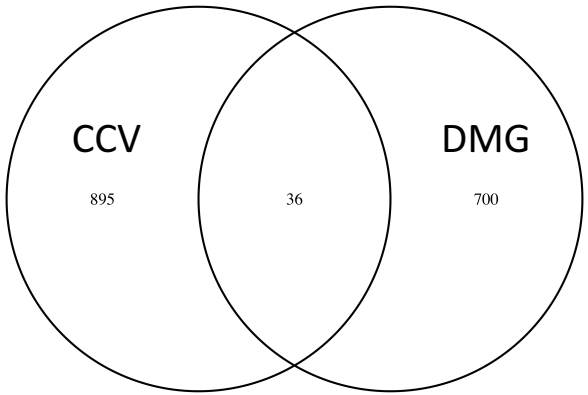
(c) *A. japonica* and European



(d) *A. oxysepala* and North American



(e) *A. oxysepala* and European



(f) North American and European