Title: Reproductive modes in populations of late-acting self-incompatible and self-compatible polyploid *Ludwigia grandiflora* subsp. *hexapetala* in western Europe

Short running title: Mixed reproductive modes in *Ludwigia*

Authors: Solenn Stoeckel[1,2], Ronan Becheler[1,2], Luis Portillo-Lemus[1], Marilyne Harang[1], Anne-Laure Besnard[1], Gilles Lassalle[1], Romain Causse-Védrines[3], Sophie Michon-Coudouel[3], Daniel J. Park[4], Bernard J. Pope[4], Eric J. Petit[1], Dominique Barloy[1]

Affiliations

1 DECOD (Ecosystem Dynamics and Sustainability), Institut Agro, IFREMER, INRAE, Rennes, France.

2 IGEPP, INRAE, Institut Agro, Université de Rennes, Le Rheu, France.

3 Plateforme EcogenO, UAR3343 OSUR, Université de Rennes, CNRS, Rennes, France.

4 Melbourne Bioinformatics, The University of Melbourne, Parkville, Australia.

Corresponding author: solenn.stoeckel@inrae.fr

Abstract

Reproductive mode, *i.e.,* the proportion of individuals produced by clonality, selfing and outcrossing in populations, determines how hereditary material is transmitted through generations. It shapes genetic diversity and its structure over time and space, which can be used to infer reproductive modes.

Ludwigia grandiflora subsp. *hexapetala* (*Lgh*) is a partially clonal, polyploid, hermaphroditic, and heteromorphic plant that recently colonized multiple countries worldwide. In western Europe, individuals are either self-incompatible caused by a late-acting self-incompatibility (LSI) system developing long-styled flowers, or self-compatible (SC), with short-styled flowers.

In this study, we genotyped 53 long- and short-styled populations newly colonizing France and northern Spain using SNPs to estimate rates of clonality, selfing and outcrossing. We found that populations reproduced mainly clonally but with a high diversity of genotypes along with rates of sexuality ranging from 10% up to 40%. We also found evidence for local admixture between long- and short-styled populations in a background of genetic structure between floral morphs that was twice the level found within morphs. Long- and short-styled populations showed similar rates of clonality but short-styled populations presented significantly higher rates of selfing, as expected considering their breeding system, and despite the small rates of failure of the LSI system. Within the 53 studied populations, the 13 short-styled populations had fewer effective alleles, lower observed heterozygosity, and higher inbreeding coefficients, linkage disequilibrium and estimates of selfing than what was found in long-styled populations. These results emphasize the necessity to consider the variation of reproductive modes when managing invasive plant species. The overall maintenance of higher genetic diversity with the possibility of maintaining populations clonally in the absence of compatible partners may explain why long-styled individuals seem to be more prevalent in all newly expanding populations worldwide. Beyond *Lgh*, our methodological approach may inspire future studies to assess the reproductive modes in other autopolyploid populations.

Keywords: Partial clonality, Mating system, Late-acting self-incompatibility system, Selfing, Outcrossing, Autopolyploidy, Water primrose

Introduction

Plants reproduce using different mechanisms (i.e., breeding system), including clonality and different types of sexuality (i.e., mating system, including allogamy and selfing; Holsinger 2000). Reproductive mode corresponds to the actual balance between sexual and clonal reproduction, measured by the rate of clonality (De Meeus et al. 2007; Stoeckel et al. 2021a), and, within the proportion of offspring produced by sexuality, the balance between autogamy and geitonogamy (hereafter *selfing*), and allogamy (*outcrossing*), measured by the rate of selfing (Bürkli et al. 2017). Reproductive mode is one of the main drivers of genetic diversity in populations and species (Duminil et al. 2007, Ellegren & Galtier 2016, Glémin et al. 2019). It determines the way hereditary material is transmitted over generations, and thus constrains the range of genetic diversity that can evolve within populations and species (De Meeus et al. 2007, Fehrer 2010, Stoeckel & Masson 2014, Rouger et al. 2016). Its genetic consequences are deep enough to be discernible even in the presence of other affecting factors (Charlesworth 2003). Considering its influence on population genetic diversity and structure, reproductive mode is one of the major biological traits to decipher before interpreting other biological and ecological forces using molecular data (Duminil et al. 2007, Reichel et al. 2016, Bürkli et al. 2017, Orive & Krueger-Hadfield 2021). Given an adequate theoretical framework, population geneticists can use genotypic data to infer reproductive modes in populations (Halkett et al. 2005, Arnaud-Haond et al. 2007, Hardy 2016, Bürkli et al. 2017, Stoeckel et al. 2021a).

Many plant species reproduce using different breeding systems they adapt to different ecological contexts resulting into different reproductive modes (Richards 1997, Charlesworth 2006).

69 Uniparental reproduction, including clonality and selfing, may help plants to spread into new areas
70 (Baker's conjecture: Barrett et al. 2008, Pannel et al. 2015). Clonality consists of a parent producing
71 a new descendant that is a genetic copy of itself with the exception of rare somatic mutations and
72 mitotic recombinations (De Meeus et al. 2007). It results into an average excess of heterozygotes
73 compared to Hardy-Weinberg expectations (*i.e.,* negative F_{IS}), large variance of F_{IS} along the genome
74 and a small excess of linkage disequilibrium, increasing with decreasing population size (Navascues
75 et al. 2010, Stoeckel & Masson 2014, Stoeckel et al. 2021a). When parents can yield multiple
76 descendants, it may generate repeated genotypes in populations (Arnaud-Haond et al. 2007). These
77 effects vary with the rates of clonality and, in turn, can be used to infer the rates of clonality
78 themselves within populations (De Meeus et al. 2006, Becheler et al. 2020, Arnaud-Haond et al.
79 2020). Hermaphroditic plants have the possibility to sexually reproduce with themselves (selfing).
80 Selfing decreases effective population sizes and gene diversity resulting into high probability of allele
81 fixing (Wright et al. 2008, Roze 2016, Glémin et al. 2019). Selfing also prohibits genetic mixing
82 between different ancestral lineages, which decreases heterozygosity within individuals (David et
83 al. 2007, Hardy 2016) and strongly increases linkage disequilibrium between genes along genomes
84 (Golding & Strobeck 1980, Norborg 2000, Lucek & Willi 2021). Around half of hermaphrodite plants
85 restrict self-fertilization using a variety of molecular mechanisms grouped under the term "self-
86 incompatibility (SI) systems" which favour outcrossing within populations (De Nettancourt 2001,
87 Castric & Vekemans 2004, Gibbs 2014, Steinecke et al. 2022). Outcrossing is overall expected to
88 increase genetic diversity, to limit allele fixation and to reduce linkage disequilibrium when
89 compared to selfing, but only if linked to the genes involved in self-recognition (Glémin et al. 2001,
90 Stoeckel et al. 2006, Navascues et al. 2010). Within the different identified SI systems (Charlesworth
91 et al. 2005, Franklin-Tong 2008), the late-acting SI system (hereafter LSI) is still poorly studied
92 despite being identified in multiple taxonomic groups in angiosperms (De Nettancourt 1997,
93 Gibbs 2014). In LSI systems, self-pollen tubes grow and are only blocked shortly before penetrating

the ovule. Such species may present reduced female fertility due to self-pollen disabling ovules, favouring the clonal regeneration of populations (Vaughton et al. 2010). LSI systems are also characterized by low but recurrent failures of the self-recognition system due to its late mechanism, leading to the production of a low amount of selfed seeds in populations (Seavey & Bawa 1986; Chen et al. 2012, Gibbs 2014). In contrast, gametophytic and sporophytic self-recognition occurring very early in the pistil drastically limit self-pollination and thus selfed seeds (Lawrence 1985, Gibbs 2014). However, we do not yet know if these rare selfed seeds contribute to the dynamics of genetic diversity of LSI populations and species, especially in peripatric conditions or in any situation when compatible partners may lack. Due to these selfed seeds, LSI system may be relatively ineffective at driving genetic diversity within populations, contrary to gametophytic and sporophytic SI systems (Brennan et al. 2002, Stoeckel et al. 2006, Koelling et al. 2011, Busch & Urban 2011). In such situations, the emergence and maintenance of LSI systems appear as new evolutionary puzzles among the reproductive systems.

Water primrose*, Ludwigia grandiflora* subsp. *hexapetala* (Hook. & Arn.) Nesom and Kartesz (2000), hereafter *Lgh*, is an insect pollinated, partially clonal, polyploid, hermaphroditic and heteromorphic plant. This species is one of the most invasive aquatic plants in the world (Thouvenot et al. 2013). *Lgh* is a decaploid species (2n=10x=80 chromosomes), resulting from hybridization of different ancestral diploid species, some of which are represented more than once in the total genome of *Lgh*, which belongs to the genus *Ludwigia* L. section Jussiaea (Hoch et al. 2015, Barloy et al, 2024). Interestingly, *Lgh* includes an autotetraploid set of chromosomes, shared with *L. peploides* subsp. *montevidensis* (2n=2x=16), hereafter *Lpm*, that is clearly distinct from the other ancestral part of the *Lgh* genome (Barloy et al. 2024). *Lpm* is a self-compatible-only diploid species with only one common floral morphology (Estes & Thorp 1974, Grewel et al. 2016). *Lgh* presents heteromorphic flowers corresponding to two floral morphs: L-morph individuals develop long-styled flowers and S-morph individuals develop short-styled flowers, that cross and result into 100% viable and fertile F1

119 and F2 descendants (Portillo-Lemus 2021, Portillo-Lemus et al. 2021) while inter-species crosses
120 only result in a low number of chlorotic and unfertile descendants (Barloy et al. 2024).
121 All tested L-morph flowers expressed an active LSI in western European populations (Portillo-Lemus
122 et al. 2022). During the core flowering season (summer), in experimental greenhouse conditions, L-
123 morph individuals show a stable seemingly insignificant rate of autogamy (around 0.2‰ of the
124 available ovules) that increases at the end of the flowering season, during autumn, to 1‰, which is
125 common in LSI systems (Gibbs 2014). Due to the massive blossoming of this species, growing in very
126 dense populations, these selfed seeds would add up yearly in field populations to thousands of
127 seeds per square meter (Portillo-Lemus et al. 2021). This pattern of low rates of autogamy, that
128 increases at the end of the flowering season, may provide the advantage of reproductive assurance
129 (Goodwillie & Weber 2018). In contrast, all tested S-morph individuals were self-compatible (SC) in
130 western European populations but in their pistils, pollen tubes of the L-morph growed significantly
131 faster and were thus advantaged to fertilize ovules when in competition with self-pollen (Portillo-
132 Lemus et al. 2022). In addition, peripatric *Lgh* populations, including European populations, were
133 previously reported as exclusively clonal with few clonal lines (Dandelot 2004, Okada et al. 2009).
134 Recently established populations in France and northern Spain mostly present only one of the two
135 compatibility modes locally, sometimes with a population of the other type a few to tens of
136 kilometers away, which may result in effective allogamy (Portillo-Lemus et al. 2021, 2022). All these
137 different breeding systems make possible very different reproductive modes in populations,
138 comprising all possible quantitative combinations of mainly clonal, autogamous and allogamous
139 modes.
140 Here, we assessed the reproductive modes of 53 invasive Lgh populations across western Europe.
141 Considering the complex case of *Lgh* in western European populations, we hypothesized that L-
142 morph populations supposed to express a LSI should present typical genetic footprints of dominant
143 outcrossing, and perhaps higher rates of clonality due to the local lack of compatible partners and

144 self-pollen interferences, while S-morph populations supposed to be SC should present higher rates
145 of sexuality prevailed by selfing. To understand the genetic impacts of reproductive modes, we also
146 quantified the covariations of genetic indices and their importance to define the genetic diversity
147 within the 53 genotyped populations, and compared these observations to the theoretical
148 expectations obtained from a Wright-Fisher-like model extended to autotetraploids (Stoeckel et al.
149 2024). Finally, we took advantage from the fact that recent local populations in France and northern
150 Spain still present only one of the two compatibility modes to compare and interpret the genetic
151 differences found within and between L-morph (LSI) and S-morph (SC) populations, with the aim to
152 assess the influence of LSI on genetic diversity, as similarly tackled in sporophytic self-incompatible
153 and SC Brassicaceae species previously (Koelling et al. 2011).
154

155 # Materials and methods

156 **Sampling strategy and floral morphology of populations**

157 We collected 795 stems of *Lgh* from 53 locations (52 locations in France and one in Catalonia, Spain),
158 corresponding to an area that spans 580 kilometres east-to-west and 1,100 kilometres north-to-
159 south (Figure SI1). At each location, we collected 15 stems (hereafter, 'individuals') along a linear
160 transect of 40 meters. Along each transect, we randomly collected three stems at coordinates X_1 =
161 0m; X_2 = 10m; X_3 = 20m; X_4 = 30m and X_5 = 40m within a one meter-square quadrat. The young
162 leaves of each sampled individual were stored after lyophilization until DNA extraction.
163 We visually identified floral morphologies of flowers found along each transect within the sampled
164 *Lgh* populations. In a previous study on seven populations among the 53 studied here (underlined
165 population names in Figure SI1), one hundred and five sampled individuals resulted to identify a
166 binary distribution of floral morphology with formally-identified self-incompatibility types: all L-
167 morph individuals were LSI typed and all S-morph individuals were SC typed (Portillo-Lemus et al

168 2021) while all these individuals succeeded to cross and give viable and fertile plants. Interestingly,
169 these two types of populations spatially distribute in monomorphic populations along different
170 rivers. We supposed for this study the LSI versus SC status of individuals and populations using their
171 floral morphologies: L-morph individuals were supposed to develop a LSI system and S-morph
172 individuals being SC. To support this conjecture, as done in Portillo-Lemus et al. (2021), we checked
173 the fruitset in each of the 53 sampled and genotyped populations: Low and even no fruitsets were
174 found in L-morph individuals and populations, while full fruitsets were found in S-morph individuals
175 and populations, in agreement with our conjecture.
176 In addition to the crossbreeding results in which all L- and S-morph individuals succeeded to cross,
177 giving full fruit set and 100% viable first- and second-generation descendants (Portillo-Lemus et al.
178 2021), we here counted the chromosome numbers on karyotypes of S-morph individuals sampled
179 in two fruitful populations and of L-morph individuals sampled in five fruitless populations to
180 validate that L- and S-morph individuals belong to *Lgh*. Between 50 to 150 kilometers separated two
181 consecutive samples (populations underlined Figure SI1). To prepare the karyotypes, we used the
182 method detailed in Barloy et al (2024) that already karyotyped a S-morph individual sampled near
183 the French Atlantic coast. The same method was used in Bou Manobens et al. (2019) to karyotype
184 a L-morph individual sampled in Catalunya.
185
186 **Definition of the autotetraploid SNP marker set**

187 As no molecular markers suitable for clonal discrimination were yet available for *Lgh* and *Lpm*, we
188 generated an original set of SNP markers via RAD-Seq (Baird et al., 2008) using two pools of 15
189 individuals each, respectively sampled across five *Lgh* and three *Lpm* western European populations.
190 RAD DNA library generation, sequencing and the analysis pipeline to identify *Lgh* SNP markers were
191 carried out as described in Delord et al. (2018). In brief, DNA of *Lgh* and *Lpm* were digested by *Sbfl*
192 restriction enzyme and used to prepare DNA libraries that were then sequenced using an Illumina

193 HiSeq3000 (150bp paired-end reads). A total of 14,233 and 34,287 RAD-Seq-determined SNPs were
194 filtered to yield 340 and 326 SNPs, one per aligned sequence, for *Lpm* and *Lgh*, respectively. Design
195 of primers compatible with Hi-Plex multiplexing was carried out by Melbourne Bioinformatics.
196 Finally, sixty and fifty SNP markers matched the quality criteria to design a Hi-Plex set of SNPs for
197 *Lpm* and *Lgh*, respectively (Hammet et al. 2019). In our study, we only considered polymorphic SNPs
198 that were shared between *Lgh* and *Lpm* to be sure they belong to the tetraploid part of *Lgh* genome
199 derived from *Lpm* (Barloy et al. 2024). We finally kept this set of 36 polymorphic and stable SNP
200 markers to genotype *Lgh* samples and analyze genetic diversity in each sampled population (primer
201 sequences are openly listed in Stoeckel et al. 2023).
202 To verify that this set of SNPs was really tetraploid, we computed the Akaike's information criterion
203 from the maximum likelihood of the best genotype considering the distribution of sequenced allele
204 countings among individuals and markers as a function of the ploidy level, using a similar approach
205 to that proposed by Burnham & Anderson (2002).

206 $$AIC(ploidy) = \sum_{l=1}^{l=36} \sum_{i=1}^{i=795} -2 \log \mathcal{L}\left(genotype_{i,l} \middle| ploidy, data_{i,l}\right) + 2K$$

207 where $data_{i,l}$ is the distribution of sequenced alleles among A, C, G and T within individual i at locus
208 l, and K is the number of possible genotypes given the ploidy and the four possible alleles (A, C, G
209 and T). The likelihood of the possible genotypes $\mathcal{L}\left(genotype_{i,l} \middle| ploidy, data_{i,l}\right)$ follows a
210 multinomial distribution of the sequenced allele countings distributed between the four different
211 possible nucleobases (A, C, G and T).
212

213 **DNA extraction, SNP marker production and genotyping**

214 To genotype the 795 individually-tagged samples with these 36 SNPs, we extracted DNA from 25 to
215 30 mg of dried young leaf tissues. The DNA extractions were performed using the NucleoSpin Plant II
216 from MACHEREY-NAGEL kit, following the manufacturer's recommendations. We used L1 buffer as

217 lysis buffer. All individuals were genotyped from a solution of 20ng/µl of *Lgh* DNA, using a modified
218 version of the Hi-Plex protocol (Hammet et al. 2019; Besnard et al. 2023). Hi-Plex is an amplicon
219 sequencing technique (*sensu* Meek & Larson 2019) in which all SNPs are co-amplified in a multiplex
220 reaction before Illumina or Ion-Torrent sequencing. Here, we used Illumina. Intermediate steps
221 include dual indexing of individual samples used for demultiplexing. Reads were then assigned to
222 loci by aligning them to reference sequences with BWA-MEM 0.7.15-r1140 (Li & Durbin 2010) and
223 alleles were counted with Samtools 1.9 (Danecek et al. 2021).

224

225 **Allele dosage**

226 The posterior probabilities of each single SNP genotype within each individual (hereafter *single SNP*
227 *genotype*) were computed using the likelihood of all possible genotypes considering a multinomial
228 distribution of the number of times each nucleobase was genotyped at one SNP marker within one
229 individual. In order to obtain the most confident dataset possible, we only assigned a genotype
230 when its posterior probability of allele dosage exceeded 70%. When one SNP within one individual
231 presented a posterior probability of allele dosage equal or lower than 70%, we assigned and
232 analysed it afterward as a missing genotype.

233

234 **Genotypic and genetic descriptors**

235 We expected that populations reproducing clonally may yield repeated multi-locus genotypes
236 (MLGs, i.e., the same genotype at all the 36 SNPs found in multiple individuals). By possibly
237 producing these repeated genotypes and by varying the relative distribution of the number of
238 samples of each of these distinct genotypes, rates of clonality impact genotypic richness and
239 evenness in populations (Halkett et al. 2005; Arnaud-Haond et al. 2007).

240 We measured genotypic richness using the R index (Dorken & Eckert 2001, Arnaud-Haond et al.
241 2005), which is defined as:

242 $$R = \frac{G-1}{N_g-1}$$

243 where G is the number of distinct genotypes (genets) and N_g is the number of genotyped individuals.
244 We also measured genotypic evenness as the parameter β of the Pareto distribution, which
245 describes the slope of the power-law inverse cumulative distribution of the size of MLGs (Arnaud-
246 Haond et al. 2007):

247 $$N \geq X = aX - \beta$$

248 where N≥X is the number of sampled ramets belonging to genets containing X or more ramets in
249 the sample of the population studied, and the parameters *a* and *β* are fitted by regression analysis.
250 Low β-values indicate the dominance of one or few clonal lineages. Pareto β is less biased by the
251 sampling effort than clonal richness R (Stoeckel et al. 2021a). In our sample of 15 individuals, we
252 expected a population with only one repeated MLG to present a Pareto β value of 0.03 and a value
253 of 3 in samples with no repeated genotypes. $\beta < 2$ indicates a population reproducing with high rates
254 of clonality (greater than 0.8 to 0.9). For a sample size of 15 individuals, $2 < \beta < 3$ indicates
255 intermediate to low rates of clonality (0.6 to 0.8). A β at 3, its maximum value, is indicative of a
256 mainly sexual population, with rates of clonality ranging from zero to 0.6, depending on the other
257 genetic indices.
258 Rates of clonality and selfing confer different effects on the range of within-population genetic
259 polymorphism as well as on probabilities of genetic identities within individuals. We thus also
260 estimated expected and observed heterozygosity H_E and H_O, allelic richness A_E, inbreeding
261 coefficient F_{IS}, and linkage disequilibrium within each sampled population.
262 The Wright (1931, 1949) inbreeding coefficient F_{IS} accounts for intraindividual genetic variation at
263 one locus as a departure from Hardy-Weinberg assumptions of the genotyped populations. We
264 computed one F_{IS} value per locus per population. We then reported the mean values of F_{IS} (MF_{IS})

265 and the inter-locus variance of F_{IS} ($VarF_{IS}$) for each population. Both MF_{IS} and $VarF_{IS}$ are very
266 informative about the underlying reproductive systems (Stoeckel et al. 2006, David et al. 2007).
267 Clonality makes the MF_{IS} values tend toward -1 along with high interlocus variance (Stoeckel &
268 Masson 2014, Stoeckel et al. 2021a). A moderate amount of sexual reproduction results in MF_{IS}
269 values around 0 (Balloux et al. 2003). $VarF_{IS}$ varies with rates of clonality, from very limited variance
270 expected in sexual populations to high variance in very clonal populations (Stoeckel & Masson 2014,
271 Stoeckel et al. 2021a). Positive MF_{IS} values are expected in populations reproducing using
272 consanguinity and selfing (Castric et al. 2002, David et al. 2007). All sexual reproductive modes,
273 allogamous and autogamous, result in low $VarF_{IS}$ within a population as massive recombination
274 tends to homogenize intra-individual genetic identities along the genomes (Stoeckel et al. 2021a).
275 We measured linkage disequilibrium over all markers using the unbiased multilocus linkage
276 disequilibrium index, $\bar{r}_d$ (Agapow & Burt 2001). This mean correlation coefficient (*r*) of genetic
277 distances (*d*) between unordered alleles at *n* loci ranges from 0 to 1. This metric has the advantage
278 of limiting the dependency of the correlation coefficient on the number of alleles and loci, and it is
279 well suited to measure linkage disequilibrium in partially clonal populations (De Meeus & Balloux
280 2004, De Meeus et al. 2006). In general, LD is only slightly affected by clonality, except when
281 clonality is high (c > 0.9) and/or when genetic drift dominates over mutation rate (*e.g.,* when
282 population sizes are small N<50 compared to u=0.001; Navascues et al. 2010, Stoeckel et al. 2021a).
283 In contrast, inbreeding and selfing are efficient processes for quickly generating strong LD, after only
284 few generations. Finally, we measured genetic differentiation between populations of LSI, between
285 populations of SC and between LSI and SC populations using ρ_{ST}, an index adapted to study
286 autopolyploid populations (Ronfort et al. 1998). All these indices were computed using GENAPOPOP
287 (Stoeckel et al. 2024), a software dedicated to analysing genetic diversity and differentiation in
288 autopolyploid populations genotyped with confident allele dosage and reproducing through all
289 possible rates of clonality and selfing.

We also used SPAGEDI (v1.5, Hardy & Vekemans 2002) to estimate rates of selfing within autopolyploid populations genotyped with confident allele dosage. This approach infers rate of selfing (Sg) from identity disequilibrium coefficients (g_{2z} estimator), assuming that populations reproduce by self-fertilization and random mating and are at inbreeding equilibrium (David et al. 2007 for diploids, Hardy 2016 for autopolyploids). Identity disequilibrium coefficients are measured as the correlation in heterozygosity of distinct loci within the genome, and present the advantage of being more robust to null alleles and genotyping errors than raw F_{IS} (David et al., 2007). The effect of partial clonality on g_{2z} estimator is not yet defined. We thus considered for this study that clonality would only marginally impact identity disequilibrium coefficients, the possibility to be at the inbreeding equilibrium and the corresponding estimates of selfing. A synthesis of all these hypotheses based on previous knowledge is provided in Table SI1.

Ranking populations by clonality and selfing

We proposed and computed a synthesis index ($\Sigma\mathrm{clon}$) calculated from Pareto β and $\mathrm{VarF_{IS}}$ to rank the studied populations from the less to the more clonal. These two population genetic indices are known to vary with rates of clonality being unbiased in samples (Stoeckel et al. 2021a). To avoid issues of scaling as the range of values of these descriptors are different by several orders of magnitude, $\Sigma\mathrm{clon}$ is computed in each population i as the sum of Pareto $\beta(i)$ and $\mathrm{VarFIS}(i)$ values that were previously normalized over the whole dataset, respectively $z_\beta(i)$ and $z_{\mathrm{VarFIS}}(i)$.

$$z_\beta(i) = 1 - \frac{\beta(i) - min(\mathrm{B})}{max(\mathrm{B}) - min(\mathrm{B})}$$

$$z_{\mathrm{VarFIS}}(i) = \frac{\mathrm{VarFIS}(i) - min(\mathrm{VF})}{max(\mathrm{VF}) - min(\mathrm{VF})}$$

$$\Sigma\mathrm{clon}(i) = {}^{1}/_{2}\left[z_\beta(i) + z_{\mathrm{VarFIS}}(i)\right]$$

where $\beta(i)$ and $\mathrm{VarFIS}(i)$ are the measured Pareto β and variance of F_{IS} over loci in the population i, and where $\mathrm{B} = \{\beta(1), \beta(2), \dots, \beta(55)\}$ and $VF = \{\mathrm{VarFIS}(1), \mathrm{VarFIS}(2), \dots, \mathrm{VarFIS}(55)\}$ are the

respective sets of Pareto β and $VarF_{IS}$ over loci in the 53 populations used to obtain their minimum (min) and maximum (max) values.

Using the same approach as for $\Sigma\mathrm{clon}$, we computed a synthesis index ($\Sigma\mathrm{self}$) calculated from $\bar{r}_d$, MF_{IS} and Sg to rank the studied populations from the less to the more selfed.

$$\Sigma\mathrm{self}(i) = {}^{1}/_{3}\left[z_{\bar{r}_d}(i) + z_{\mathrm{MFIS}}(i) + z_{\mathrm{Sg}}(i)\right]$$

Statistical data analysis

First, to better understand the correlation between population genetic indices and estimates of reproductive modes from a large field dataset and to compare with the theoretical correlations obtained from simulations (Stoeckel et al. 2024), we computed a Principal Component Analysis (PCA) to comprehend the covariations of genetic indices, the correlations and the redundancies between the 17 genetic diversity indices (described above: G, R, D, Pareto β, $\bar{r}_d$, Ae, varAe, He, varHe, Ho, varHo, MF_{IS}, $varF_{IS}$, PIDsib, PIDu, Sg, SE.Sg) measured among the 53 sampled populations, and their link with reproductive modes including self-compatibility. We reported the amount of variation retained by the first two principal components and the correlation circle on which we plotted predictions of $\Sigma\mathrm{clon}$ and $\Sigma\mathrm{self}$ as supplementary variables. We also reported the score plot to visualize how L-morph (LSI) and S-morph (SC) populations distribute along the principal components of population genetic diversity. To avoid scaling issues, all descriptors were normalized before analysis.

We then analyzed the relationship between genetic diversity indices and their relationships with $\Sigma\mathrm{clon}$ and $\Sigma\mathrm{self}$ measured in the 53 sampled populations using Kendal partial rank-order correlation tests. We reported the corresponding matrices of correlation.

To detect differences in distribution of population genetic indices, $\Sigma\mathrm{clon}$ and $\Sigma\mathrm{self}$, among floral morphs, we computed non-parametric Kruskal-Wallis tests that do not make any assumptions about the type of distribution and about homogeneity of variances between the tested distributions.

338 When needed, we used post-hoc pairwise tests for multiple comparisons of mean rank sums
339 (Nemenyi's test).
340 All statistical tests were computed using PYTHON 3.11, SCIPY.STATS 1.9.3 (Virtanen et al. 2020) and
341 SCIKIT-POSTHOCS (Terpilowski 2019), except the PCA that was performed using R V4.2.2 and the library
342 FACTOMINER (Lê et al. 2008).
343

344 # Results

345 Among the 53 populations, we found 40 populations with only L-morph individuals and 13
346 populations with only S-morph individuals (Table SI2, Figure SI1). Karyotypes of L-morph plants from
347 two populations and of S-morph plants from five populations all presented the same number of
348 chromosomes (2n=80, Figure SI2) confirming that L and S-morph individuals in France and L-morph
349 individuals in northern Spain (Bou Manobens et al. 2019) belong to *L. grandiflora* subsp. *hexapetala*
350 (2n=10X=80, Barloy et al. 2024). Akaike's information criterion on the distribution of allele counting
351 over all our data supported tetraploidy as the best ploidy level for this set of 36 SNPs (Figure SI3),
352 as expected from Barloy et al (2024).
353

354 **Allele dosage and missing genotypes**

355 Within the 795 individuals genotyped at 36 SNPs (resulting in a total of 28,620 single-SNP
356 genotypes), 99.97% (28,612) of SNPs were genotyped with posterior probability of allele dosage
357 superior to 70% (Table SI3). 785 individuals were genotyped with a full set of 36 SNPs with posterior
358 probabilities of allele dosage higher than 70%. Ten individuals distributed in nine populations
359 showed one of their SNP markers with posterior probabilities of allele dosage equal or lower than
360 70%, that we assigned therefore as missing genotypes.
361

362 **Statistical power of the developed SNP marker set**

363 Over the 36 polymorphic SNPs, we found an effective number of alleles (A_E) of 1.36 per SNP over all
364 populations (Table SI2). Among populations, mean A_E values over the 36 SNPs were homogenous,
365 ranging from 1.22 to 1.55 (median=1.34). We, however, found large standard deviation of A_E
366 between SNPs within populations, ranging from 1.6 to 2.3 (median=2.1). Some SNPs were
367 apparently fixed in some populations while polymorphic at the scale of the whole dataset. When
368 not fixed, gene diversity H_E in polymorphic SNPs ranged from 0.14 to 0.33 (median=0.2).

369 These 36 SNPs in the autotetraploid part of *Lgh* would theoretically allow $4^{36}=4.7\times10^{21}$ different
370 possible MLGs considering the four possible nucleobases and $2^{36}=6.9\times10^{10}$ different possible MLGs
371 assuming two possible nucleobases per locus. Considering allele frequencies in the sampled
372 populations, the probabilities of identities under panmixia ranged from 8.5×10^{-12} to 5.2×10^{-5}
373 (median=2.1×10^{-7}) and the unbiased probabilities of identity between sibs $P_{ID\text{-}SIB}$ ranged from $4.2\times10^{-}$
374 ^{6} to 8.3×10^{-3} (median=5.6×10^{-4}; Table SI2). We then considered that the SNP set we used to genotype
375 the 795 sampled individuals via Hi-Plex method showed sufficient statistical power to distinguish
376 between true MLGs, and that individuals with identical MLGs were true repeated genotypes
377 (ramets) of a clonal lineage (a genet).

378

379 **Genetic and genotypic diversity**

380 Across populations, we identified a total of 462 distinct MLGs (genets) within the 795 sampled
381 individuals genotyped. Among them, we found 404 genets (88%) with a single ramet and only 58
382 genets (12%) with more than two ramets (Table SI4). Forty-eight genets had two to seven ramets
383 distributed over one to six populations (median=2), seven genets with 10 to 33 ramets distributed
384 over two to 17 populations (median=9), and one large genet of 99 ramets distributed over 24
385 populations (Figure 1).

386 Within populations, we found from three to 15 different genets (median=12) per population among

387 the 15 sampled individuals (Table SI4), implying the clear occurrence of repeated genotypes but also
388 a wide diversity of genets within most populations (Figure 1). Accordingly, genotypic richness (R)
389 ranged from 0.14 to 1 (median=0.79). The genotypic evenness, Pareto β, ranged from 0.056 to 3
390 (median=1.478, Table SI2).

391 Observed heterozygosity H_O was also high in most populations, slightly above expected
392 heterozygosity, ranging from 0.13 to 0.38 (median=0.26). The mean standard inbreeding
393 coefficients (MF_{IS}) averaged over all genotyped loci within populations were negative in all 53
394 populations and ranged from -0.33 to -0.126 with a very negative median of -0.274 (Table SI2).
395 Variances of F_{IS} between loci within population were very high, ranging from 1.75 to 36.60
396 (median=27.57). All these measures argued for reproductive modes implying high rates of clonality.
397 Estimates of selfing (S_g) were overall low with a handful of high values, ranging from 0 to 0.61
398 (median=0.06). Fifteen populations (28%) were estimated with no selfing. Nineteen populations
399 (36%) showed non-zero estimates under 0.1, 13 populations (25%) between 0.1 and 0.29, and six
400 populations (11%) with estimates between 0.42 and 0.61.
401 Finally, linkage disequilibria within populations between genotyped SNPs were overall low, with r_d
402 values ranging from 0 to 0.51 with a median value of 0.12. Forty-two populations (79%) were under
403 0.25, and only eight populations (21%) showed linkage disequilibrium between 0.25 and 0.51.
404

405 **Analyses of covariations between population genetic indices**

406 The first two components of the principal component analysis on the values of genetic diversity
407 found in the 53 genotyped populations accounted for 78.4% of the total variance between
408 populations (Figure SI4). Non-parametric Kendall partial rank-order correlations between genetic
409 indices (Table SI5) and correlations on the first two principal components from the 17 population
410 genetic indices measured showed three non-collinear groups of associated genetic indices (Figure
411 SI4.A) that are very similar to the theoretical groups of population genetic indices expected to covary

412 with different rates of clonality, selfing and outcrossing in autopolyploids (Stoeckel et al. 2024). A
413 first cluster regrouped G, D*, R, Pareto β and $VarF_{IS}$, indices that are known to be sensitive to
414 clonality (Stoeckel et al. 2021a) but also VarHe, VarHo, and VarAe. This cluster largely explains the
415 first dimension (50.2% of the total variance) of the PCA and was collinear to Σ_{CLON} (also see Figure
416 SI4.B). As expected under partial clonality, $VarF_{IS}$ was negatively correlated to genotypic diversity
417 indices (R and β; Figure SI5). The second cluster regrouped P_{ID-SIB}, P_{ID-u}, H_E, A_E and H_O, indices that
418 are linked to the general genetic diversity of populations (Figure SI4.A). The third cluster regrouped
419 S_g, SE.Sg, MF_{IS} and r_D, indices that are usually used to identify, rank and estimate rates of selfing
420 versus outcrossing in sexual populations (Castric et al. 2002, Bürkly et al. 2017). The second
421 dimension of the PCA (23.4% of the total variance) was mostly correlated to Ho, Sg, He and r_D which
422 was collinear to Σ_{SELF} (Figure SI4.C). The clusters of genetic indices were corroborated by the
423 correlation between Σ_{CLON} and genetic indices (Figure SI5) and Σ_{SELF} and genetic indices (Figure SI6).
424 The correlation across populations between Σ_{CLON} and Σ_{SELF} was negative and highly significant (r_s = -
425 0.66; $p < 0.001$; Figure 2). This correlation within LSI populations was negative and highly significant
426 (r_s = -0.65; $p < 0.001$) while it was nonsignificant in SC populations (r_s = -0.32; p =0.289).
427

428 **Differences in genetic diversity and structure between L- and S-morph populations**

429 The number of ramets per genet was similar in L-morph (LSI) and S-morph (SC) populations (H=1.85,
430 p=0.173, Figure 1 & SI7), as were their genotypic richness (R, H=1.23, p=0.267) and evenness
431 (Pareto's β, H=1.60, p=0.206; Table SI2). The distributions of other population genetic indices
432 related to clonality were also not significantly different between in L-morph (LSI) and S-morph (SC)
433 populations ($VarF_{IS}$: H=2.94, p=0.086; P_{ID-SIB}: H=1.98, p=0.160 and Σ_{CLON} H=2.28, p=0.131; Figure 3).
434 Mean observed heterozygosity (Ho, H=5.27, p=0.022) and their variances (VarHo, H=13.85,
435 p<0.001), mean effective number of alleles per locus (Ae, H=5.13, p=0.024) and its variance over

436 locus (VarAe, H=9.49, p=0.002), MF_{IS} (H=11.85, p<0.001), linkage disequilibrium (r_D, H=5.55,
437 p=0.018), estimate of selfing rates (Sg, H=12.39, p<0.001) and its standard error over loci (SESg,
438 H=9.58, p=0.002) significantly differed between in L-morph (LSI) and S-morph (SC) populations. S-
439 morph (SC) populations, compared to L-morph (LSI), showed lower mean observed heterozygosity
440 (medians, L-morph: 0.262, S-morph: 0.210) and less variance between loci (medians, L-morph:
441 2.400, S-morph: 1.859), lower effective number of alleles (medians, L-morph: 1.341, S-morph:
442 1.283) and less variance between loci (medians, L-morph: 4.419, S-morph: 3.580), higher linkage
443 disequilibrium (medians, L-morph: 0.113, S-morph: 0.152), higher estimates of selfing rate
444 (medians, L-morph: 0.045, S-morph: 0.175) even if with higher standard error between loci
445 (medians, L-morph: 0.058, S-morph: 0.156) and higher Σ_{SELF} (medians, L-morph: 0.222, S-morph:
446 0.495).
447 Genetic differentiation (ρ_{ST}) among pairs of S-morph (SC) populations (median value of ρ_{ST}=0.27)
448 was comparable to the genetic differentiation found among pairs of L-morph (LSI) populations
449 (median value of ρ_{ST}=0.26; H=319.37, post-hoc p=0.322; Figure 4). The median of ρ_{ST} between pairs
450 of L-morph (LSI) and S-morph (SC) populations reached 0.65. The distribution of inter-morph ρ_{ST}
451 values differed both from the distribution of pairs of intra-L-morph (LSI) ρ_{ST} values (post-hoc
452 p<0.001) and from the distribution of pairs of intra-S-morph (SC) ρ_{ST} values (post-hoc p<0.001). The
453 minimum spanning tree of genetic distances between individuals computed with *GenAPoPop*
454 showed quite clustered distributions of L-morph (LSI) and S-morph (SC) individuals but also with
455 clear evidence of admixtures between their lineages (Figure 5).
456

457 **Discussion**

458 We used population genetics on autotetraploid SNPs with allele dosage to assess the rate of
459 clonality, of selfing and of outcrossing (i.e., reproductive mode) of 53 populations of *Ludwigia*
460 *hexapetala* subsp. *grandiflora* (*Lgh*) recently colonizing western European watersheds.
461 To achieve this goal, we developed reproducible codominant molecular markers (SNPs) that enabled
462 allele dosage-based genotyping of sampled individuals. We then used tailored computational
463 analyses to perform confident population genetic analyses in autopolyploids. These two
464 methodological steps allowed solving the remaining challenges to perform population genetic
465 analyses in autopolyploids (Dufresne et al. 2014), including *Lgh*. The resulting framework can be
466 applied to any autopolyploid species.
467 *Lgh* develops two floral morphs respectively associated with a Late-acting self-incompatible system
468 (L-morph) and a self-compatible system (S-morph, Portillo-Lemus et al. 2021, 2022). Interestingly,
469 the sampled populations in western Europe showed either L- or S-morph resulting in two groups of
470 sampled populations: one group of 40 L-morph (LSI) populations and one group of 13 S-morph (SC)
471 populations. They live in similar ecological conditions (Portillo-Lemus et al. 2021) which thereby
472 provides a rare opportunity to characterize the genetic consequences of an LSI system compared to
473 a group of SC populations in the same species and ecological context.
474 The five L-morph and the two S-morph water primrose populations we karyotyped, separated by 50
475 to 150 km from one to the other in France, all had 80 chromosomes corresponding to the species
476 *Ludwigia grandiflora* subsp. *hexapetala*. We didn't find yet individual with 48 chromosomes
477 corresponding to *Ludwigia grandiflora* subsp. *grandiflora*. These results agree with previous
478 observations of Dandelot et al (2005) and Barloy et al. (2024) in France, Bou Manobens et al. (2019)
479 in northern Spain and Armitage et al. (2013) in Great Britain.
480
481 **Most of the genetic variance explained by clonality and selfing among *Lgh* populations**

482 The correlations of population genetics indices on the 53 genotyped *Lgh* populations (Figure SI4)
483 were congruent with previous theoretical predictions on the variations of genetic indices with
484 different rates of clonality and selfing for autopolyploid populations (Stoeckel et al. 2024). Identical
485 covariations were already predicted by Wright-Fisher-like models adapted for diploid (Stoeckel et
486 al. 2021a) and haplodiplontic life-cycles (Stoeckel et al. 2021b), and validated respectively using
487 multiple field populations of marine phanerogams for diploids (Arnaud-Haond et al. 2020) and of
488 brown, green, red algae and mosses for haplodiplontics (Krueger-Hadfield et al. 2021). All these
489 predictions and observations argue for the primary importance of reproductive mode, especially of
490 high clonality, to drive variation in genetic diversity within populations (Duminil et al. 2007). They
491 support the methodological importance of accurate assessment of rates of clonality and of selfing
492 before starting to interpret genetic diversity.
493
494 **Dominant clonality within the western European *Lgh* populations**
495 Previous field and lab observations of peripatric *Lgh* populations reported massive production of
496 dispersing vegetative propagules, rapid expansions of patches, and an important capacity for
497 spontaneous cutting-planting (Dandelot et al. 2005, Thouvenot et al. 2013, Grewell et al. 2016,
498 Skaer-Thomason et al. 2018a,b). Okada et al. (2009) genotyped around 800 individuals sampled in
499 27 *Lgh* populations in Californian wetlands using a set of eight AFLP markers and reported an
500 extremely reduced clonal diversity: 95% of the samples had the same genotype and 18 populations
501 over 27 (67%) supported a single AFLP-genotype. All over the UK, only two haplotypes on 14
502 sampled stems were found using chloroplast sequences (Armitage et al. 2013). These haplotypes
503 were even shared with some *Lgh* samples invading California. With no measure of the probability
504 of identity to assess the marker set, these results could be due to the lack of resolution of the
505 markers used resulting in an artificially elevated measure of clonality (Waits et al. 2001, Villate et al.
506 2010). In any case, all these studies concluded that invasive populations of *Lgh* in Europe and in USA

507 reproduce by exclusive clonality, with a very narrow base of ancestral clones or being monoclonal
508 (Dandelot et al. 2005, Okada et al. 2009, Thouvenot et al. 2013, Grewell et al. 2016).
509 In our study, we chose to develop and use SNPs rather than AFLP markers. The latter can be
510 challenging in terms of precisely sizing fragments, leading to suboptimal reproducibility, particularly
511 across different platforms (Fry et al. 2009). The 36 SNPs we developed are easily reproducible,
512 accurate enough to distinguish between offspring of sibling mating as evidenced by the probabilities
513 of identity we obtained, and less expensive.
514 With these SNPs, we found typical genetic signatures of high clonality, including clear occurrence of
515 replicated genotypes and mean negative F_{IS} values with high interlocus variances, in all the 53
516 western European populations we genotyped. Globally, gene diversities measured in these invasive
517 *Lgh* populations were in the higher range of values commonly found for SNPs (Fischer et al. 2017,
518 Schmidt et al. 2021). Such levels of gene diversity are in line with strong rates of clonality that buffer
519 the loss of alleles due to genetic drift (Reichel et al. 2016, Stoeckel et al. 2021b). Among the 58 MLGs
520 with replicates, seven MLGs were found in a single population and 51 MLGs were found in multiple
521 sites. Clonal propagules of *Lgh* are known to disperse carried by river current and by birds (zoochory:
522 seedlings have the ability to stick to feathers; Grewel et al. 2016). Clonal reproduction by rhizomes
523 at the local scale and by the dispersal of clonal propagules would thus be the main source of
524 population growth and invasive spread of *Lgh* in western Europe.
525 Our results also report a previously-underestimated genotype diversity in the invasive populations
526 in western Europe. We found 462 distinct MLGs (genets) out of 795 individuals within these
527 populations and the large majority of these MLGs (404, thus 88%) were only sampled once in the 53
528 populations. Our results are still congruent with Dandelot (2004) unpublished measures obtained in
529 three Mediterranean populations using inter-simple sequence repeats (ISSRs). This data reports
530 three and seven genotypes over 11 sampling units within two *Lgh* populations and two genotypes
531 over nine sampling units in a third population, all sampled in the south-east of France. These results

532 could reveal contrasted situations between *Lgh* populations introduced in the USA and continental
533 Europe.
534
535 **Between 10 and 40% of sexual reproduction in *Lgh* populations in western Europe**
536 Beyond the qualitative indication of dominant clonality, we then aimed at estimating the rates of
537 sexuality in each of these 53 populations. All showed small to medium values of linkage
538 disequilibrium between SNPs. Such values are expected in highly but non-exclusive clonal
539 populations with large population sizes (Navascues et al. 2010, Stoeckel et al. 2021a). Thirty-two
540 populations showed Pareto β values of their distributions of clonal sizes under 2, which are only
541 found in theoretical populations with rates of clonality higher than 0.8 (Stoeckel et al. 2024). All 53
542 populations showed mean negative F_{IS} with high interlocus variance, with values observed in
543 theoretical populations with rates of clonality higher than 0.6 but under 0.9 (Balloux et al. 2003,
544 Stoeckel & Masson 2014, Reichel et al. 2016, Stoeckel et al. 2024). All these values of population
545 genetic indices were consistent with the interpretation that *Lgh* populations in western Europe must
546 reproduce with effective rates of clonality between 60% and 90%, thus with 10% to 40% sexuality
547 (De Meeus et al. 2006, Arnaud-Haond et al. 2020, Stoeckel et al. 2024). These estimates were also
548 supported by the local evidence of recombination between clonal lineages, and even between L-
549 morph (LSI) and S-morph (SC) lineages at the scale of a watershed (Figure 5). The diversity of
550 genotypes detected in western Europe is thus indicative of rare but significant local sexual events
551 rather than of a large clonal diversity that would have maintain and propagate by exclusive clonality
552 since its introduction. Our results newly advocate that sexual reproduction should not, therefore,
553 be overlooked in these invasive populations, especially in management plans.
554
555 **Genetic consequences of LSI compared to self-compatible populations**

556 The late-acting SI system remains one of the less studied breeding systems among the mechanisms
557 favouring outcrossing in plants (Gibbs 2014). Its efficiency to favour outcrossing and its
558 consequences on genetic diversity within populations, especially considering its low but common
559 failures, were not yet deciphered and not yet compared to SC populations in the same ecological
560 conditions, as already previously explored for gametophytic and sporophyte SI systems (Busch 2005,
561 Koelling et al. 2011).
562 The maintenance of SI systems is one of the most intriguing evolutionary puzzles (Porcher &
563 Lande 2005). Indeed, SI systems are fated to breakdown because SC individuals present the
564 advantage of reproductive assurance when compatible partners are limited, especially in peripatric
565 conditions (Eckert et al. 2006). This advantage is even absolute when no compatible partners are
566 available within pollination range, as we commonly found in *Lgh* populations in western Europe
567 (Figure SI1). Conversely, outcrossing imposed by SI systems decreases the probability of expressing
568 deleterious mutations in descendants as compared to selfing with the same genetic background
569 (Rice 2004, Navascues et al. 2010).
570 We found that the 40 populations with L-morph individuals (LSI) had a higher number of effective
571 alleles, higher gene diversity, higher observed heterozygosity and less linkage disequilibrium than
572 the 13 populations with S-morph individuals (SC; Figure 3).
573 These genetic differences are very unlikely to directly result from the consequences of the LSI
574 system hitchhiked to the whole genome in L-morph individuals and populations. Indeed, it would
575 imply either that all the 36 SNPs would be physically linked to the genes under negative frequency-
576 dependent selection coding for the LSI or that *Lgh* outcrossed for many generations in small
577 population sizes (Glémin et al. 2001, Navascues et al. 2010). *Lgh* in western Europe develops
578 including far more than thousands of stems per local population (Portillo-Lemus et al. 2021) and the
579 linkage disequilibrium values also argued for large effective population sizes. We found similar
580 estimated rates of sexuality and of clonality between L-morph and S-morph populations, using Σ_{CLON}

581 and all indices sensible to clonality (R, Pareto Beta and VarFis; Figure 3). We however found
582 significant difference in estimates of selfing between L-morph (LSI) and S-morph (SC) populations,
583 and in Σ_{SELF} and in all indices sensible to selfing (Ae, linkage disequilibrium, Ho, mean Fis) revealing
584 typical signatures of higher selfing rates in S-morph populations. Hardy (2016) method estimated a
585 median selfing rate (Sg) of 0.18 in S-morph populations versus 0.05 in L-morph populations.
586 Consequently, the genetic differences we found between S-morph and L-morph populations may
587 thus rather be due to the effects of selfing impacting S-morph (SC) populations than due to
588 outcrossing protecting the loss of genetic diversity or rates of clonality in L-morph (LSI) populations.
589

590 **Advantages of dominant clonality with preferential allogamy in invasive *Lgh* populations**

591 Uniparental clonal reproduction (including clonality and selfing) may help the demographical
592 maintenance of plants spreading out of their native range with limited or even without compatible
593 or less related sexual partner at pollination distance (Baker's conjecture: Barrett et al. 2008, Pannel
594 et al. 2015). If they rather reproduce using selfing, their descendants increase the probability to
595 express inbreeding depression and to lose heterozygosity. Some invasive populations develop with
596 low genetic diversity (He et al. 2024), questioning on the biological and environmental factors that
597 may explain their success (i.e., the genetic paradox of invasions, Allendorf & Lundquist 2003). But
598 many other plants spread out of their native ranges with substantial genetic diversity and using
599 reproductive modes that favour outcrossing (Roman & Darling 2007, Forsman 2014), like *Lgh*
600 populations in western Europe, mostly when developing in harsh and stressful conditions (Fox &
601 Reed, 2011) or when the costs of inbreeding depression expressed by selfing are superior to the
602 benefits of reproductive assurance (Layman et al. 2017).

603 Peripatric populations of *Lgh* in western Europe seem to solve all these problems and paradox by
604 mixing clonality with preferential allogamy: clonality allows the local maintenance and spreading of
605 population without losing heterozygosity and genetic diversity, and subtle but significant sexuality

606 with preferential allogamy, favoured by LSI and faster growth of crossed-pollen tubes, enables
607 recombination between lineages favouring the emergence of locally adapted genomes with
608 potential higher vigour and fertility (heterosis: Darwin 1876, Lippman & Zamir 2007, Birchleret al.
609 2010).
610 This reproductive mode, i.e. dominant clonality with preferential allogamy, is common in plants
611 (Vallejo-Marin et al. 2010). The micro-physiological mechanism(s) slowing down the growth of self-
612 pollen tubes, rather than blocking them, in simultaneous monoecious and hermaphrodite SC species
613 and resulting to favour allogamy when compatible pollen is available may also be common in plant,
614 although potentially overlooked (Glover 2007, Nasrallah 2017). In such peripatric conditions, selfing
615 would thus only present the limited or transient advantage to produce seeds that can maintain in
616 local seed banks and with different dispersal properties compared to clonal propagules. The limited
617 interest of selfing in this species may explain why, based on photos collected on the web, L-morph
618 (LSI) individuals seem more frequent in peripatric than in native populations, and why in western
619 Europe, around 76% of the populations are L-morph (LSI) against 24% of S-morph (SC; Portillo et al.
620 2021). Our results thus call for measuring the true proportions of LSI and SC in invasive versus native
621 worldwide *Lgh* populations with estimation of their rates of clonality and of selfing.
622
623 **Unusual selfing syndrome in *Lgh* populations**
624 Highly selfed populations of different plant species tend to share similar morphology and functions
625 resulting in a set of traits called *selfing syndrome*, including reduced flower size (Darwin 1876,
626 Tsuchimatsu & Fujii 2022). Selfing syndrome, including reduced flower size, seems to evolve rapidly,
627 as observed for example in five generations in *Mimulus guttatus* (Bodbyl-Roels & Kelly 2011), in four
628 generations in *Silene latifolia* (Delph et al. 2004), in three generations in *Phlox drummondii* (Lendvai
629 & Levin 2003) and after only two generations in *Eichhornia paniculata* (Worley & Barrett 2000).

630 On the contrary, S-morph (SC) individuals in *Lgh* develop larger flowers than allogamous L-morphs
631 (Portillo-Lemus et al. 2021). However, S-morph (SC) individuals and populations dominantly
632 reproduced using clonality and produced selfed offspring only when lacking crossed pollen. Both
633 clonality and faster growth of crossed-pollen tubes may delay or even compromise the emergence
634 of the first steps of a selfing syndrome. These two traits may call for further studies on the
635 emergence of selfing syndrome in partially clonal and selfed populations.
636
637 **Conclusion**
638 We found that peripatric populations of *Lgh* in western Europe reproduced using dominant clonality
639 with limited but stable significant sexuality with preferential outcrossing, in nearly all populations,
640 within a large clonal diversity. *Lgh* is one of the most invasive aquatic plants in the world, and
641 considerable efforts are made to limit its deleterious effects in the newly colonized ecosystems
642 (Thouvenot et al. 2013, Grewell et al. 2016, Portillo-Lemus et al. 2021). The rare sexual events,
643 allogamous when possible, occurring in invasive peripatric populations of *Lgh* may favour the
644 emergence of new genotypes, more adapted to local conditions. Managers should thus chiefly
645 concentrate their actions on the most sexual populations, and on the contact zones between S- and
646 L-morph populations in France. We also found variations of the rates of clonality and of selfing
647 among *Lgh* populations. Considering the importance of reproductive modes on the dynamics and
648 evolution of populations (Duminil et al. 2007, Ellegren& Galtier 2016, Glémin et al. 2019), our results
649 advocate that management actions should consider the local effective reproductive modes to
650 control *Lgh* population by population. Finally, knowing the reproductive modes of populations, the
651 distribution of clones and the self-compatibility across western European populations now allow
652 deciphering and interpreting their population structure, identifying their origin and routes, and
653 predicting their possible short-term dynamics.

654 **Data Accessibility Statement**

655 The data that support the findings of this study are openly available on Zenodo
656 (https://doi.org/10.5281/zenodo.10246264).

657

658 **Benefits Generated**

659 Benefits from this research accrue from the sharing of our data and results on public databases as
660 described above.

661

662 **Acknowledgements**

663 We thank Michel Bozec (UMR DECOD), Diane Corbin (FRAPNA Loire-Ecopôle du Forez), Guillaume
664 Le Roux (Reserve Naturel Val d'Alier Châtel-de-Neuvre), Fabrice Dejoux (Service Agriculture
665 Environnement Roannais Agglomération and Jordi Bou Manobens (University of Girona, Spain) for
666 helping to collect *Lgh* samples. We thank Olivier Coriton and Virginie Huteau for their help acquiring
667 karyotypes.

668 We warmly thank Ingrid Leveque, Marie-Therese Delaroche, Gervaise Fevrier and Patricia Nadan for
669 their administrative support and facilitation.

670

671 **Funding information**

672 This work was supported by FEDER Région Centre-Val de Loire, by Agence de l'eau Loire-Bretagne
673 (grant Nature 2045, programme 9025, AP 2015 9025) and by the French National Research Agency
674 (project Clonix2D ANR-18-CE32-0001). FEDER funded the doctoral grant of L. Portillo-Lemus and
675 salary of Marilyne Harang. Postdoctoral grant of Ronan Becheler was funded by Clonix2D ANR-18-
676 CE32-0001.

677

678 **Conflict of interest**

679 The authors declare that they have no financial conflicts of interest based on the content of this
680 article.
681
682 **Author contributions**
683 DB and SS laid the foundation of this work, conceived the study and were responsible for funding
684 applications. LPL, MH and DB collected samples and performed field observations. LPL and DB read
685 karyotypes and counted chromosomes. EP and DB conceived and supervised RAD4SNP and Hi-Plex
686 methodology, BJP and DJP contributed to define the SNP set. LPL, MH and ALB performed RAD4SNP
687 approach to develop the SNPs. LPL and MH extracted DNA and genotyped the samples, SMC and
688 RCV sequenced the data. GL performed the bioinformatic analyses and SS computed the Bayesian
689 allele dosages. DB, LPL and SS performed early data explorations at the end of LPL PhD thesis that
690 was supervised by DB and SS. RB and SS performed computations and analyses, tracked and
691 managed the bibliography, produced tables, figures and interpretation. DB, RB and SS wrote the
692 manuscript. All authors read and approved the final manuscript.
693

References

Abdallah, D., Baraket, G., Perez, V., Ben Mustapha, S., Salhi-Hannachi, A., & Hormaza, J. I. (2019). Analysis of self-incompatibility and genetic diversity in diploid and hexaploid plum genotypes. *Frontiers in Plant Science*, *10*. https://www.frontiersin.org/articles/10.3389/fpls.2019.00896

Agapow, P.-M., & Burt, A. (2001). Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, *1*(1–2), 101–102. https://doi.org/10.1046/j.1471-8278.2000.00014.x

Allendorf, F. W., & Lundquist, L. L. (2003). Introduction: Population Biology, Evolution, and Control of Invasive Species. *Conservation Biology*, 17(1), 24–30. https://doi.org/10.1046/j.1523-1739.2003.02365.x

Armitage, J. D., Könyves, K., Bailey, J. P., David, J. C., & Culham, A. (2013). A molecular, morphological and cytological investigation of the identity of non-native *Ludwigia* (Onagraceae) populations in Britain. *New Journal of Botany*, *3*(2), 88–95. https://doi.org/10.1179/2042349713Y.0000000023

Arnaud-Haond, S., Alberto, F., Teixeira, S., Procaccini, G., Serrão, E. A., & Duarte, C. M. (2005). Assessing genetic diversity in clonal organisms: low diversity or low resolution? combining power and cost efficiency in selecting markers. *Journal of Heredity*, *96*(4), 434–440. https://doi.org/10.1093/jhered/esi043

Arnaud-Haond, S., Duarte, C. M., Alberto, F., & Serrão, E. A. (2007). Standardizing methods to address clonality in population studies. *Molecular Ecology*, *16*(24), 5115–5139. https://doi.org/10.1111/j.1365-294X.2007.03535.x

Arnaud-Haond, S., Stoeckel, S., & Bailleul, D. (2020). New insights into the population genetics of partially clonal organisms: When seagrass data meet theoretical expectations. *Molecular Ecology*, *29*(17), 3248–3260. https://doi.org/10.1111/mec.15532

Baduel, P., Bray, S., Vallejo-Marin, M., Kolář, F., & Yant, L. (2018). The "polyploid hop": shifting challenges and opportunities over the evolutionary lifespan of genome duplications. *Frontiers*

720 *in Ecology and Evolution*, *6*. https://www.frontiersin.org/articles/10.3389/fevo.2018.00117

721 Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U.,
722 Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP Discovery and Genetic Mapping Using
723 Sequenced RAD Markers. *PLOS ONE*, 3(10), e3376.
724 https://doi.org/10.1371/journal.pone.0003376

725 Balloux, F., Lehmann, L., & de Meeûs, T. (2003). The population genetics of clonal and partially
726 clonal diploids. *Genetics*, *164*(4), 1635–1644. https://doi.org/10.1093/genetics/164.4.1635

727 Barloy, D., Lemus, L. P.-, Krueger-Hadfield, S. A., Huteau, V., & Coriton, O. (2024). Genomic
728 relationships among diploid and polyploid species of the genus Ludwigia L. section Jussiaea
729 using a combination of cytogenetic, morphological, and crossing investigations. *Peer*
730 *Community in Evolutionary Biology*, 4, e8. https://doi.org/10.24072/pci.evolbiol.100645.

731 Barrett, S. C. H., Colautti, R. I., & Eckert, C. G. (2008). Plant reproductive systems and evolution
732 during biological invasion. *Molecular Ecology*, 17(1), 373–383.
733 https://doi.org/10.1111/j.1365-294X.2007.03503.x

734 Barrett, S. C. H. (2010). Understanding plant reproductive diversity. *Philosophical Transactions of*
735 *the Royal Society B: Biological Sciences*, *365*(1537), 99–109.
736 https://doi.org/10.1098/rstb.2009.0199

737 Becheler, R., Guillemin, M.-L., Stoeckel, S., Mauger, S., Saunier, A., Brante, A., Destombe, C., &
738 Valero, M. (2020). After a catastrophe, a little bit of sex is better than nothing: Genetic
739 consequences of a major earthquake on asexual and sexual populations. *Evolutionary*
740 *Applications*, *13*(8), 2086–2100. https://doi.org/10.1111/eva.12967

741 Becheler, R., Masson, J.-P., Arnaud-Haond, S., Halkett, F., Mariette, S., Guillemin, M.-L., Valero,
742 M., Destombe, C., & Stoeckel, S. (2017). ClonEstiMate, a Bayesian method for quantifying
743 rates of clonality of populations genotyped at two-time steps. *Molecular Ecology Resources*,
744 17(6), e251–e267. https://doi.org/10.1111/1755-0998.12698

745 Beck, J. B., Al-Shehbaz, I. A., & Schaal, B. A. (2006). *Leavenworthia* (Brassicaceae) revisited:

testing classic systematic and mating system hypotheses. *Systematic Botany*, *31*(1), 151–159. https://doi.org/10.1600/036364406775971732

Besnard, A.-L., Park, D. J., Pope, B. J., Hammet, F., Michon-Coudouel, S., Biget, M., Krueger-Hadfield, S. A., Mauger, S., & Petit, E. J. (2023). Workflow for SNP genotyping using the Hi-Plex method. *Protocols.io*, https://doi.org/10.17504/protocols.io.8epv5jnnnl1b/v2

Birchler, J. A., Yao, H., Chudalayandi, S., Vaiman, D., & Veitia, R. A. (2010). Heterosis. *The Plant Cell*, *22*(7), 2105–2112. https://doi.org/10.1105/tpc.110.076133

Bodbyl Roels, S. A., & Kelly, J. K. (2011). Rapid evolution caused by pollinator loss in *Mimulus guttatus*. *Evolution*, *65*(9), 2541–2552. https://doi.org/10.1111/j.1558-5646.2011.01326.x

Booy, G., Hendriks, R.J.J., Smulders, M.J.M., Van Groenendael, J.M. and Vosman, B. (2000). Genetic Diversity and the Survival of Populations. *Plant Biology*, *2,* 379-395. https://doi.org/10.1055/s-2000-5958

Bou Manobens, J., Portillo-Lemus, L., & Curcó i Masip, A. (2019). New contributions to allochthonous Ludwigia species (Onagraceae) on Catalonia. Butlletí de La Institució Catalana d'Història Natural, 83, 45–48. https://doi.org/10.2436/20.1502.01.4

Brennan, A. C., Harris, S. A., Tabah, D. A., & Hiscock, S. J. (2002). The population genetics of sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae) I: S allele diversity in a natural population. *Heredity*, *89*(6), Article 6. https://doi.org/10.1038/sj.hdy.6800159

Bürkli, A., Sieber, N., Seppälä, K., & Jokela, J. (2017). Comparing direct and indirect selfing rate estimates: When are population-structure estimates reliable? *Heredity*, *118*(6), Article 6. https://doi.org/10.1038/hdy.2017.1

Busch, J. W. (2005). Inbreeding depression in self-incompatible and self-compatible populations of *Leavenworthia alabamica*. *Heredity*, *94*(2), Article 2. https://doi.org/10.1038/sj.hdy.6800584

Busch, J. W., & Urban, L. (2011). Insights gained from 50 years of studying the evolution of self-compatibility in *Leavenworthia* (Brassicaceae). *Evolutionary Biology*, *38*(1), 15–27. https://doi.org/10.1007/s11692-010-9104-5

Cabin, R. J., Evans, A. S., Jennings, D. L., Marshall, D. L., Mitchell, R. J., & Sher, A. A. (1996). Using bud pollinations to avoid self-incompatibility: Implications from studies of three mustards. *Canadian Journal of Botany*, 74(2), 285–289. https://doi.org/10.1139/b96-034

Castric, V., Bernatchez, L., Belkhir, K., & Bonhomme, F. (2002). Heterozygote deficiencies in small lacustrine populations of brook charr *Salvelinus fontinalis* Mitchill (Pisces, Salmonidae): A test of alternative hypotheses. *Heredity*, *89*(1), Article 1. https://doi.org/10.1038/sj.hdy.6800089

Castric, V., & Vekemans, X. (2004). Plant self-incompatibility in natural populations: a critical assessment of recent theoretical and empirical advances. *Molecular Ecology*, *13*(10), 2873–2889. https://doi.org/10.1111/j.1365-294X.2004.02267.x

Charlesworth, D. (2003). Effects of inbreeding on the genetic diversity of populations. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1434), 1071–1084. https://doi.org/10.1098/rstb.2003.1296

Charlesworth, D. (2006). Evolution of Plant Breeding Systems. Current Biology, *16*(17), 726–735. https://doi.org/10.1016/j.cub.2006.07.068

Charlesworth, D., Vekemans, X., Castric, V., & Glémin, S. (2005). Plant self-incompatibility systems: A molecular evolutionary perspective. *New Phytologist*, *168*(1), 61–69. https://doi.org/10.1111/j.1469-8137.2005.01443.x

Chen, X., Hao, S., Wang, L., Fang, W., Wang, Y., & Li, X. (2012). Late-acting self-incompatibility in tea plant (*Camellia sinensis*). *Biologia*, *67*(2), 347–351. https://doi.org/10.2478/s11756-012-0018-9

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience, 10, giab008. https://doi.org/10.1093/gigascience/giab008

Dandelot, S. (2004). *Les Ludwigia spp. Invasives du Sud de la France: Historique, Biosystématique, Biologie et Ecologie* [PhD thesis, Aix-Marseille 3]. https://www.theses.fr/2004AIX30052

Dandelot, S., Verlaque, R., Dutartre, A., & Cazaubon, A. (2005). Ecological, Dynamic and Taxonomic Problems Due to Ludwigia (Onagraceae) in France. Hydrobiologia, 551(1), 131–136. https://doi.org/10.1007/s10750-005-4455-0

Darwin, C. R. (1876). *The effects of cross and self-fertilization in the vegetable kingdom*. John Murray (London).

David, P., Pujol, B., Viard, F., Castella, V., & Goudet, J. (2007). Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology*, *16*(12), 2474–2487. https://doi.org/10.1111/j.1365-294X.2007.03330.x

De Meeûs, T., & Balloux, F. (2004). Clonal reproduction and linkage disequilibrium in diploids: A simulation study. *Infection, Genetics and Evolution*, *4*(4), 345–351. https://doi.org/10.1016/j.meegid.2004.05.002

De Meeûs, T., Lehmann, L., & Balloux, F. (2006). Molecular epidemiology of clonal diploids: A quick overview and a short DIY (do it yourself) notice. *Infection, Genetics and Evolution*, *6*(2), 163–170. https://doi.org/10.1016/j.meegid.2005.02.004

De Meeûs, T., Prugnolle, F., & Agnew, P. (2007). Asexual reproduction: Genetics and evolutionary aspects. *Cellular and Molecular Life Sciences*, *64*(11), 1355–1372. https://doi.org/10.1007/s00018-007-6515-2

De Nettancourt, D. (1997). Incompatibility in angiosperms. *Sexual Plant Reproduction*, *10*(4), 185–199. https://doi.org/10.1007/s004970050087

De Nettancourt, D. (2001). *Incompatibility and Incongruity in Wild and Cultivated Plants*. Springer. https://doi.org/10.1007/978-3-662-04502-2

Dellaporta, S. L., & Calderon-Urrea, A. (1993). Sex Determination in Flowering Plants. *The Plant Cell*, *5*(10), 1241–1251. https://doi.org/10.2307/3869777

Delord, C., Lassalle, G., Oger, A., Barloy, D., Coutellec, M.-A., Delcamp, A., Evanno, G., Genthon, C., Guichoux, E., Le Bail, P.-Y., Le Quilliec, P., Longin, G., Lorvelec, O., Massot, M., Reveillac, E., Rinaldo, R., Roussel, J.-M., Vigouroux, R., Launey, S., & Petit, E. J. (2018). A

824 cost-and-time effective procedure to develop SNP markers for multiple species: A support for
825 community genetics. *Methods in Ecology and Evolution*, *9*(9), 1959–1974.
826 https://doi.org/10.1111/2041-210X.13034
827 Delph, L. F., Gehring, J. L., Frey, F. M., Arntz, A. M., & Levri, M. (2004). Genetic constraints on
828 floral evolution in a sexually dimorphic plant revealed by artificial selection. *Evolution*, *58*(9),
829 1936–1946. https://doi.org/10.1111/j.0014-3820.2004.tb00481.x
830 Dorken, M. E., & Eckert, C. G. (2001). Severely reduced sexual reproduction in northern populations
831 of a clonal plant, *Decodonverticillatus* (Lythraceae). *Journal of Ecology*, *89*(3), 339–350.
832 https://doi.org/10.1046/j.1365-2745.2001.00558.x
833 Dufresne, F., Stift, M., Vergilino, R., & Mable, B. K. (2014). Recent progress and challenges in
834 population genetics of polyploid organisms: An overview of current state-of-the-art molecular
835 and statistical tools. *Molecular Ecology*, *23*(1), 40–69. https://doi.org/10.1111/mec.12581
836 Duminil, J., Fineschi, S., Hampe, A., Jordano, P., Salvini, D., Vendramin, G. G., & Petit, R. J. (2007).
837 Can population genetic structure be predicted from life-history traits? *The American*
838 *Naturalist*, *169*(5), 662–672. https://doi.org/10.1086/513490
839 Eckert, C., Samis, K. & Dart, S. (2006). Reproductive Assurance and the Evolution of Uniparental
840 Reproduction in Flowering Plants. In Harder & Barrett (Eds), *Ecology and Evolution of*
841 *Flowers*, pp.183–200. Oxford, UK: Oxford University Press.
842 Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7),
843 Article 7. https://doi.org/10.1038/nrg.2016.58
844 Estes, J. R., & Thorp, R. W. (1974). Pollination in Ludwigia peploides ssp. Glabrescens
845 (Onagraceae). *Bulletin of the Torrey Botanical Club, 101*(5), 272–276.
846 https://doi.org/10.2307/2484872
847 Fehrer, J. (2010). Unraveling the mysteries of reproduction. *Heredity*, *104*(5), Article 5.
848 https://doi.org/10.1038/hdy.2010.12
849 Fischer, M. C., Rellstab, C., Leuzinger, M., Roumet, M., Gugerli, F., Shimizu, K. K., Holderegger,

R., & Widmer, A. (2017). Estimating genomic diversity and population differentiation – an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *BMC Genomics*, *18*(1), 69. https://doi.org/10.1186/s12864-016-3459-7

Flanagan, B. A., Krueger-Hadfield, S. A., Murren, C. J., Nice, C. C., Strand, A. E., & Sotka, E. E. (2021). Founder effects shape linkage disequilibrium and genomic diversity of a partially clonal invader. *Molecular Ecology*, *30*(9), 1962–1978. https://doi.org/10.1111/mec.15854

Forsman, A. (2014). Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. *Proceedings of the National Academy of Sciences*, 111(1), 302–307. https://doi.org/10.1073/pnas.1317745111

Fox, C. W., & Reed, D. H. (2011). Inbreeding Depression Increases with Environmental Stress: An Experimental Study and Meta-Analysis. *Evolution*, 65(1), 246–258. https://doi.org/10.1111/j.1558-5646.2010.01108.x

Franklin-Tong, V. E. (2008). *Self-Incompatibility in Flowering Plants*. Springer. https://doi.org/10.1007/978-3-540-68486-2

Fry, N. K., Savelkoul, P. H., & Visca, P. (2009). Amplified Fragment Length Polymorphism Analysis. In: Caugant, D. (eds) Molecular Epidemiology of Microorganisms. *Methods in Molecular Biology,* vol 551, 89-104. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-60327-999-4_8

Gibbs, P. E. (2014). Late-acting self-incompatibility – the pariah breeding system in flowering plants. *New Phytologist*, *203*(3), 717–734. https://doi.org/10.1111/nph.12874

Glémin, S., Bataillon, T., Ronfort, J., Mignot, A., & Olivieri, I. (2001). inbreeding depression in small populations of self-incompatible plants. *Genetics*, *159*(3), 1217–1229. https://doi.org/10.1093/genetics/159.3.1217

Glémin, S., François, C. M., & Galtier, N. (2019). Genome Evolution in Outcrossing vs. Selfing vs. Asexual Species. In M. Anisimova (Ed.), *Evolutionary Genomics: Statistical and Computational Methods* (pp. 331–369). Springer. https://doi.org/10.1007/978-1-4939-9074-

0_11

Glover, B. J. (2007). Preventing Self-fertilization. In B. Glover (Ed.), Understanding Flowers and Flowering: An integrated approach (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198565970.003.0012

Golding, G. B., & Strobeck, C. (1980). Linkage disequilibrium in a finite population that is partially selfing. *Genetics*, *94*(3), 777–789. https://doi.org/10.1093/genetics/94.3.777

Goodwillie, C., & Weber, J. J. (2018). The best of both worlds? A review of delayed selfing in flowering plants. *American Journal of Botany*, *105*(4), 641–655. https://doi.org/10.1002/ajb2.1045

Grewell, B., Netherland, M., & Skaer Thomason, M. (2016). Establishing research and management priorities for invasive water primroses *(Ludwigia spp.)*. https://doi.org/10.13140/RG.2.1.5020.6482

Haldane, J. B. S. (1930). Theoretical genetics of autopolyploids. *Journal of Genetics*, 22(3), 359–372. https://doi.org/10.1007/BF02984197

Halkett, F., Simon, J.-C., & Balloux, F. (2005). Tackling the population genetics of clonal and partially clonal organisms. *Trends in Ecology & Evolution*, *20*(4), 194–201. https://doi.org/10.1016/j.tree.2005.01.001

Hammet, F., Mahmood, K., Green, T. R., Nguyen-Dumont, T., Southey, M. C., Buchanan, D. D., Lonie, A., Nathanson, K. L., Couch, F. J., Pope, B. J., & Park, D. J. (2019). Hi-Plex2: A simple and robust approach to targeted sequencing-based genetic screening. *BioTechniques*, 67(3), 118–122. https://doi.org/10.2144/btn-2019-0026

Hao, Y.-Q., Zhao, X.-F., She, D.-Y., Xu, B., Zhang, D.-Y., & Liao, W.-J. (2012). The Role of Late-acting self-incompatibility and early-acting inbreeding depression in governing female fertility in monkshood, *Aconitum kusnezoffii*. *PLOS ONE*, *7*(10), e47034. https://doi.org/10.1371/journal.pone.0047034

Harder, L. D., & Barrett, S. C. H. (2006). *Ecology and Evolution of Flowers*. OUP Oxford.

Hardy, O. J. (2016). Population genetics of autopolyploids under a mixed mating model and the estimation of selfing rate. *Molecular Ecology Resources*, *16*(1), 103–117. https://doi.org/10.1111/1755-0998.12431

Hardy, O. J., & Vekemans, X. (2002). SPAGEDI: A versatile computer program to analyze spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, *2*(4), 618–620. https://doi.org/10.1046/j.1471-8286.2002.00305.x

He, Z.-Z., Stotz, G. C., Liu, X., Liu, J.-J., Wang, Y.-G., Yang, J., Li, L.-F., Zhang, W.-J., Nan, P., & Song, Z.-P. (2024). A global synthesis of the patterns of genetic diversity in endangered and invasive plants. *Biological Conservation*, 291, 110473. https://doi.org/10.1016/j.biocon.2024.110473

Herben, T., Suda, J., & Klimešová, J. (2017). Polyploid species rely on vegetative reproduction more than diploids: A re-examination of the old hypothesis. *Annals of Botany*, *120*(2), 341–349. https://doi.org/10.1093/aob/mcx009

Hoban, S., Campbell, C. D., da Silva, J. M., Ekblom, R., Funk, W. C., Garner, B. A., Godoy, J. A., Kershaw, F., MacDonald, A. J., Mergeay, J., Minter, M., O'Brien, D., Vinas, I. P., Pearson, S. K., Pérez-Espona, S., Potter, K. M., Russo, I.-R. M., Segelbacher, G., Vernesi, C., & Hunter, M. E. (2021). Genetic diversity is considered important but interpreted narrowly in country reports to the Convention on Biological Diversity: Current actions and indicators are insufficient. *Biological Conservation*, *261*, 109233. https://doi.org/10.1016/j.biocon.2021.109233

Hoch, P. C., Wagner, W. L., & Raven, P. H. (2015). The correct name for a section of Ludwigia L. (Onagraceae). *PhytoKeys*, *50*, 31–34. https://doi.org/10.3897/phytokeys.50.4887

Holsinger, K. E. (2000). Reproductive systems and evolution in vascular plants. *Proceedings of the National Academy of Sciences*, *97*(13), 7037–7042. https://doi.org/10.1073/pnas.97.13.7037

Husband, B. C., Ozimec, B., Martin, S. L., & Pollock, L. (2008). Mating consequences of polyploid evolution in flowering plants: current trends and insights from synthetic polyploids.

International Journal of Plant Sciences, *169*(1), 195–206. https://doi.org/10.1086/523367

Igic, B., Lande, R., & Kohn, J. R. (2008). Loss of self-incompatibility and its evolutionary consequences. *International Journal of Plant Sciences*, *169*(1), 93–104. https://doi.org/10.1086/523362

Koelling, V. A., Hamrick, J. L., & Mauricio, R. (2011). Genetic diversity and structure in two species of *Leavenworthia* with self-incompatible and self-compatible populations. *Heredity*, *106*(2), 310-318. https://doi.org/10.1038/hdy.2010.59

Krueger-Hadfield, S. A., Guillemin, M.-L., Destombe, C., Valero, M., & Stoeckel, S. (2021). Exploring the Genetic Consequences of Clonality in Haplodiplontic Taxa. *Journal of Heredity*, 112(1), 92–107. https://doi.org/10.1093/jhered/esaa063

Lawrence, M. J., Marshall, D. F., Curtis, V. E., & Fearon, C. H. (1985). Gametophytic self-incompatibility re-examined: A reply. *Heredity*, 54(1), 131–138. https://doi.org/10.1038/hdy.1985.17

Layman, N. C., Fernando, M. T. R., Herlihy, C. R., & Busch, J. W. (2017). Costs of selfing prevent the spread of a self-compatibility mutation that causes reproductive assurance. *Evolution*, 71(4), 884–897. https://doi.org/10.1111/evo.13167

Lê, S., Josse, J. & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software, 25*(1), 1-18. https://doi.org/10.18637/jss.v025.i01

Lendvai, G., & Levin, D. A. (2003). Rapid response to artificial selection on flower size in *Phlox*. *Heredity*, *90*(4), Article 4. https://doi.org/10.1038/sj.hdy.6800249

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, *26*, 589-595. https://doi.org/10.1093/bioinformatics/btp698.

Lippman, Z. B., & Zamir, D. (2007). Heterosis: Revisiting the magic. *Trends in Genetics*, *23*(2), 60–66. https://doi.org/10.1016/j.tig.2006.12.006

Lucek, K., & Willi, Y. (2021). Drivers of linkage disequilibrium across a species' geographic range. *PLOS Genetics*, *17*(3), e1009477. https://doi.org/10.1371/journal.pgen.1009477

Mable, B. K. (2008). Genetic causes and consequences of the breakdown of self-incompatibility: Case studies in the Brassicaceae. *Genetics Research*, *90*(1), 47–60. https://doi.org/10.1017/S0016672307008907

Meek, M. H., & Larson, W. A. (2019). The future is now: Amplicon sequencing and sequence capture usher in the conservation genomics era. Molecular Ecology Resources, 19(4), 795–803. https://doi.org/10.1111/1755-0998.12998

Meirmans, P. G., Liu, S., & van Tienderen, P. H. (2018). The analysis of polyploid genetic data. *Journal of Heredity*, *109*(3), 283–296. https://doi.org/10.1093/jhered/esy006

Nasrallah, J. B. (2017). Plant mating systems: Self-incompatibility and evolutionary transitions to self-fertility in the mustard family. Current Opinion in Genetics & Development, 47, 54–60. https://doi.org/10.1016/j.gde.2017.08.005

Navascués, M., Stoeckel, S., & Mariette, S. (2010). Genetic diversity and fitness in small populations of partially asexual, self-incompatible plants. *Heredity*, *104*(5), Article 5. https://doi.org/10.1038/hdy.2009.159

Nesom, G.L. & Kartesz, J.T. (2000). Observations on the *Ludwigia uruguayensis* complex (Onagraceae) in the United States. Castanea 65(2): 123-125. http://www.jstor.org/stable/4034110

Nordborg, M. (2000). Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. *Genetics*, *154*(2), 923–929. https://doi.org/10.1093/genetics/154.2.923

Okada, M., Grewell, B. J., & Jasieniuk, M. (2009). Clonal spread of invasive *Ludwigia hexapetala* and *L. grandiflora* in freshwater wetlands of California. *Aquatic Botany*, *91*(3), 123–129. https://doi.org/10.1016/j.aquabot.2009.03.006

Orive, M. E., & Krueger-Hadfield, S. A. (2021). Sex and asex: a clonal lexicon. *Journal of Heredity*, *112*(1), 1–8. https://doi.org/10.1093/jhered/esaa058

Pannell, J. R., Auld, J. R., Brandvain, Y., Burd, M., Busch, J. W., Cheptou, P.-O., Conner, J. K.,

Goldberg, E. E., Grant, A.-G., Grossenbacher, D. L., Hovick, S. M., Igic, B., Kalisz, S., Petanidou, T., Randle, A. M., de Casas, R. R., Pauw, A., Vamosi, J. C., & Winn, A. A. (2015). The scope of Baker's law. *New Phytologist*, *208*(3), 656–667. https://doi.org/10.1111/nph.13539

Porcher, E., & Lande, R. (2005). The evolution of self-fertilization and inbreeding depression under pollen discounting and pollen limitation. *Journal of Evolutionary Biology*, *18*(3), 497–508. https://doi.org/10.1111/j.1420-9101.2005.00905.x

Portillo Lemus, L. O. (2021). Système de reproduction, polyploïdie et diversité génétique des populations invasives de Ludwigia grandiflora subsp hexapetala en France. *PhD Thesis*. http://www.theses.fr/2021NSARA088/document

Portillo Lemus, L. O., Bozec, M., Harang, M., Coudreuse, J., Haury, J., Stoeckel, S., & Barloy, D. (2021). Self-incompatibility limits sexual reproduction rather than environmental conditions in an invasive water primrose. *Plant-Environment Interactions*, *2*(2), 74–86. https://doi.org/10.1002/pei3.10042

Portillo Lemus, L. O., Harang, M., Bozec, M., Haury, J., Stoeckel, S., & Barloy, D. (2022). Late-acting self-incompatible system, preferential allogamy and delayed selfing in the heteromorphic invasive populations of *Ludwigia grandiflora subsp. Hexapetala*. *Peer Community Journal*, *2*. https://doi.org/10.24072/pcjournal.108

Pyšek, P. (1998). Is there a taxonomic pattern to plant invasions? *Oikos*, *82*(2), 282–294. https://doi.org/10.2307/3546968

Reichel, K., Masson, J.-P., Malrieu, F., Arnaud-Haond, S., & Stoeckel, S. (2016). Rare sex or out of reach equilibrium? The dynamics of FISin partially clonal organisms. *BMC Genetics*, *17*(1), 76. https://doi.org/10.1186/s12863-016-0388-z

Rice, S. H. (2004). *Evolutionary theory: Mathematical and conceptual foundations*. Sinauer Associates.

Richards, A. J. (1997). *Plant breeding systems*. Garland Science.

Roman, J., & Darling, J. A. (2007). Paradox lost: Genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution*, 22(9), 454–464. https://doi.org/10.1016/j.tree.2007.07.002

Ronfort, J., Jenczewski, E., Bataillon, T., & Rousset, F. (1998). Analysis of population structure in autotetraploid species. *Genetics*, *150*(2), 921–930. https://doi.org/10.1093/genetics/150.2.921

Rouger, R., Reichel, K., Malrieu, F., Masson, J. P., & Stoeckel, S. (2016). Effects of complex life cycles on genetic diversity: Cyclical parthenogenesis. *Heredity*, *117*(5), Article 5. https://doi.org/10.1038/hdy.2016.52

Roze, D. (2016). Background selection in partially selfing populations. *Genetics*, *203*(2), 937–957. https://doi.org/10.1534/genetics.116.187955

Schmidt, T. L., Jasper, M.-E., Weeks, A. R., & Hoffmann, A. A. (2021). Unbiased population heterozygosity estimates from genome-wide sequence data. *Methods in Ecology and Evolution*, *12*(10), 1888–1898. https://doi.org/10.1111/2041-210X.13659

Seavey, S. R., & Bawa, K. S. (1986). Late-acting self-incompatibility in angiosperms. *The Botanical Review*, *52*(2), 195–219. https://doi.org/10.1007/BF02861001

Skaer Thomason, M. J., Grewell, B. J., & Netherland, M. D. (2018a). Dynamics of *Ludwigia hexapetala* invasion at three spatial scales in a regulated river. *Wetlands*, *38*(6), 1285–1298. https://doi.org/10.1007/s13157-018-1053-2

Skaer Thomason, M. J., McCort, C. D., Netherland, M. D., & Grewell, B. J. (2018b). Temporal and nonlinear dispersal patterns of *Ludwigia hexapetala* in a regulated river. *Wetlands Ecology and Management*, *26*(5), 751–762. https://doi.org/10.1007/s11273-018-9605-z

Steinecke, C., Gorman, C. E., Stift, M., & Dorken, M. E. (2022). Outcrossing rates in an experimentally admixed population of self-compatible and self-incompatible Arabidopsis lyrata. *Heredity*, *128*(1), Article 1. https://doi.org/10.1038/s41437-021-00489-8

Stoeckel, S., Arnaud-Haond, S., & Krueger-Hadfield, S. A. (2021b). The combined effect of haplodiplonty and partial clonality on genotypic and genetic diversity in a finite mutating

population. *Journal of Heredity*, *112*(1), 78–91. https://doi.org/10.1093/jhered/esaa062

Stoeckel, S., Barloy, D., Portillo Lemus, L., & Becheler, R. (2023). Genotyping measures and population genetic indices for assesing reproductive modes of polyploid Ludwigia grandiflora subsp. hexapetala in western Europe [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10246264

Stoeckel, S., Becheler, R., Bocharova, E., & Barloy, D. (2023). GenAPoPop 1.0: A user-friendly software to analyse genetic diversity and structure from partially clonal and selfed autopolyploid organisms. Molecular Ecology Resources, 00, 1–11. https://doi.org/10.1111/1755-0998.13886

Stoeckel, S., Grange, J., Fernández-Manjarres, J. F., Bilger, I., Frascaria-Lacoste, N., & Mariette, S. (2006). Heterozygote excess in a self-incompatible and partially clonal forest tree species—Prunus avium L. *Molecular Ecology*, *15*(8), 2109–2118. https://doi.org/10.1111/j.1365-294X.2006.02926.x

Stoeckel, S., Castric, V., Mariette, S., & Vekemans, X. (2008). Unequal allelic frequencies at the self-incompatibility locus within local populations of Prunus avium L.: An effect of population structure? Journal of Evolutionary Biology, 21(3), 889–899. https://doi.org/10.1111/j.1420-9101.2008.01504.x

Stoeckel, S., & Masson, J.-P. (2014). The exact distributions of fis under partial asexuality in small finite populations with mutation. *PLOS ONE*, *9*(1), e85228. https://doi.org/10.1371/journal.pone.0085228

Stoeckel, S., Porro, B., & Arnaud-Haond, S. (2021a). The discernible and hidden effects of clonality on the genotypic and genetic states of populations: Improving our estimation of clonal rates. *Molecular Ecology Resources*, *21*(4), 1068–1084. https://doi.org/10.1111/1755-0998.13316

Terpilowski, (2019). scikit-posthocs: Pairwise multiple comparison tests in Python. Journal of Open Source Software, 4(36), 1169, https://doi.org/10.21105/joss.01169

Thouvenot, L., Haury, J., & Thiebaut, G. (2013). A success story: Water primroses, aquatic plant

1058 pests. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *23*(5), 790–803.
1059 https://doi.org/10.1002/aqc.2387
1060 Tsuchimatsu, T., & Sota, F. (2022). The selfing syndrome and beyond: diverse evolutionary
1061 consequences of mating system transitions in plants. *Philosophical Transactions of the Royal*
1062 *Society B: Biological Sciences*, 377, 20200510. https://doi.org/10.1098/rstb.2020.0510.
1063 Vallejo-Marín, M., Dorken, M. E., & Barrett, S. C. H. (2010). The ecological and evolutionary
1064 consequences of clonality for plant mating. *Annual Review of Ecology, Evolution, and*
1065 *Systematics*, *41*(1), 193–213. https://doi.org/10.1146/annurev.ecolsys.110308.120258
1066 Van Drunen, W. E., & Husband, B. C. (2019). Evolutionary associations between polyploidy, clonal
1067 reproduction, and perenniality in the angiosperms. *New Phytologist*, *224*(3), 1266–1277.
1068 https://doi.org/10.1111/nph.15999
1069 Vaughton, G., Ramsey, M., & Johnson, S. D. (2010). Pollination and late-acting self-incompatibility
1070 in Cyrtanthus breviflorus (Amaryllidaceae): Implications for seed production. *Annals of*
1071 *Botany*, 106(4), 547–555. https://doi.org/10.1093/aob/mcq149
1072 Villate, L., Esmenjaud, D., Van Helden, M., Stoeckel, S., & Plantard, O. (2010). Genetic signature
1073 of amphimixis allows for the detection and fine scale localization of sexual reproduction
1074 events in a mainly parthenogenetic nematode. *Molecular Ecology*, 19(5), 856–873.
1075 https://doi.org/10.1111/j.1365-294X.2009.04511.x
1076 Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E.,
1077 Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K.
1078 J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., & van Mulbregt, P. (2020).
1079 SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*,
1080 *17*(3), Article 3. https://doi.org/10.1038/s41592-019-0686-2
1081 Waits, L. P., Luikart, G., & Taberlet, P. (2001). Estimating the probability of identity among
1082 genotypes in natural populations: Cautions and guidelines. *Molecular Ecology*, 10(1), 249–
1083 256. https://doi.org/10.1046/j.1365-294X.2001.01185.x

Wang, Y.-J., Müller-Schärer, H., van Kleunen, M., Cai, A.-M., Zhang, P., Yan, R., Dong, B.-C., & Yu, F.-H. (2017). Invasive alien plants benefit more from clonal integration in heterogeneous environments than natives. *New Phytologist*, *216*(4), 1072–1078. https://doi.org/10.1111/nph.14820

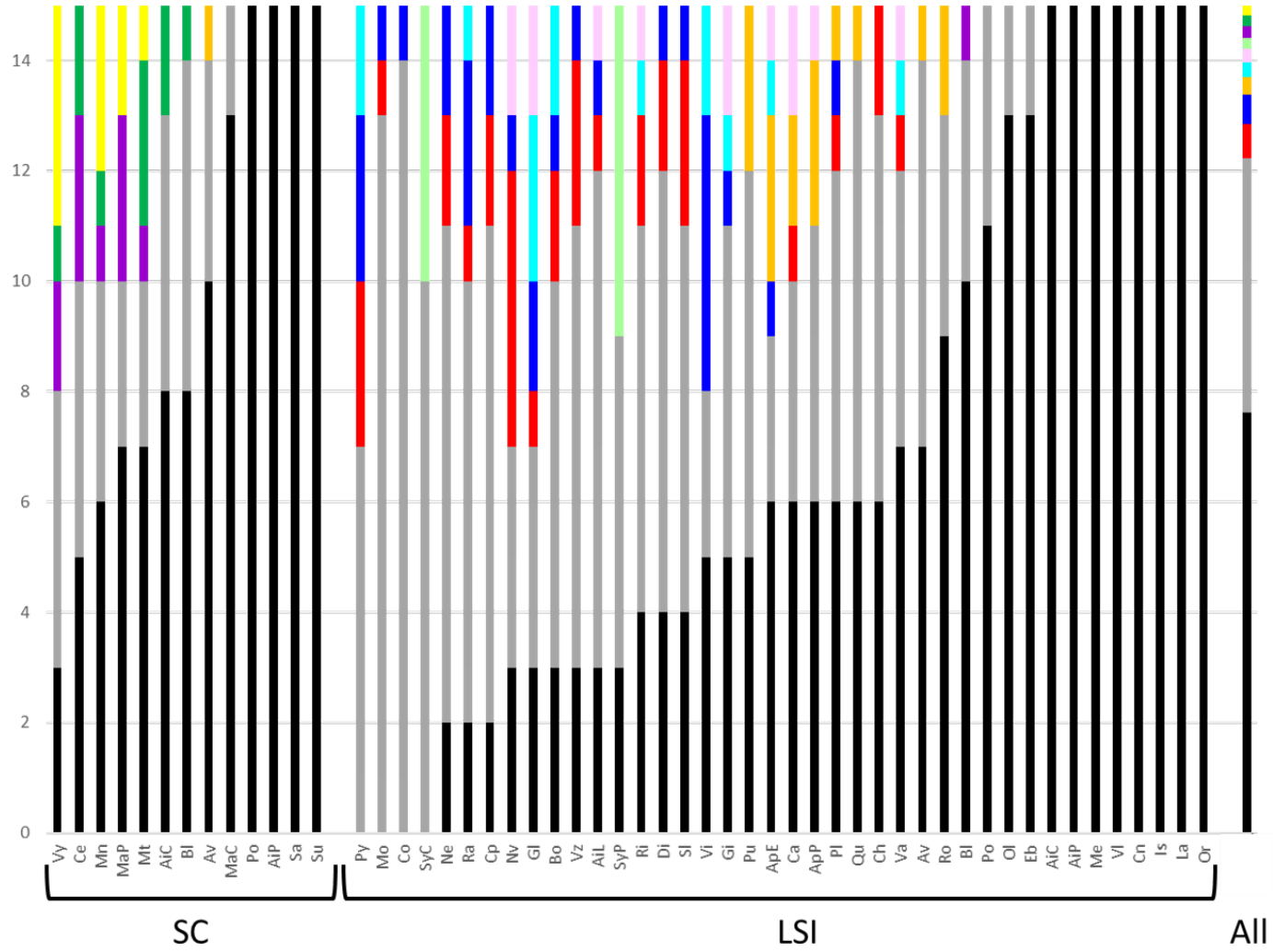
Worley, A. C., & Barrett, S. C. H. (2000). Evolution of floral display in *Eichhornia paniculata* (Pontederiaceae): direct and correlated responses to selection on flower size and number. *Evolution*, *54*(5), 1533–1545. https://doi.org/10.1111/j.0014-3820.2000.tb00699.x

Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, *16*(2), 97–159.

Wright, S. (1949). The Genetical Structure of Populations. *Annals of Eugenics*, *15*(1), 323–354. https://doi.org/10.1111/j.1469-1809.1949.tb02451.x

Wright, S. I., Ness, R. W., Foxe, J. P., & Barrett, S. C. H. (2008). Genomic consequences of outcrossing and selfing in plants. *International Journal of Plant Sciences*, *169*(1), 105–118. https://doi.org/10.1086/523366

1098



1099

1100 **Figure 1:** Distributions of Multi-Locus Genotypes (MLGs) among LSI and SC *Lgh* populations in western
1101 Europe. In black, the number of unique MLGs; In grey, the number of MLGs with less than 7 ramets; In
1102 color, MLGs with more than 10 ramets found in multiple populations. Proportions of each MLGs over all the
1103 populations are plotted in the last right bar (All). More than half the samples are unique genotypes.

1104

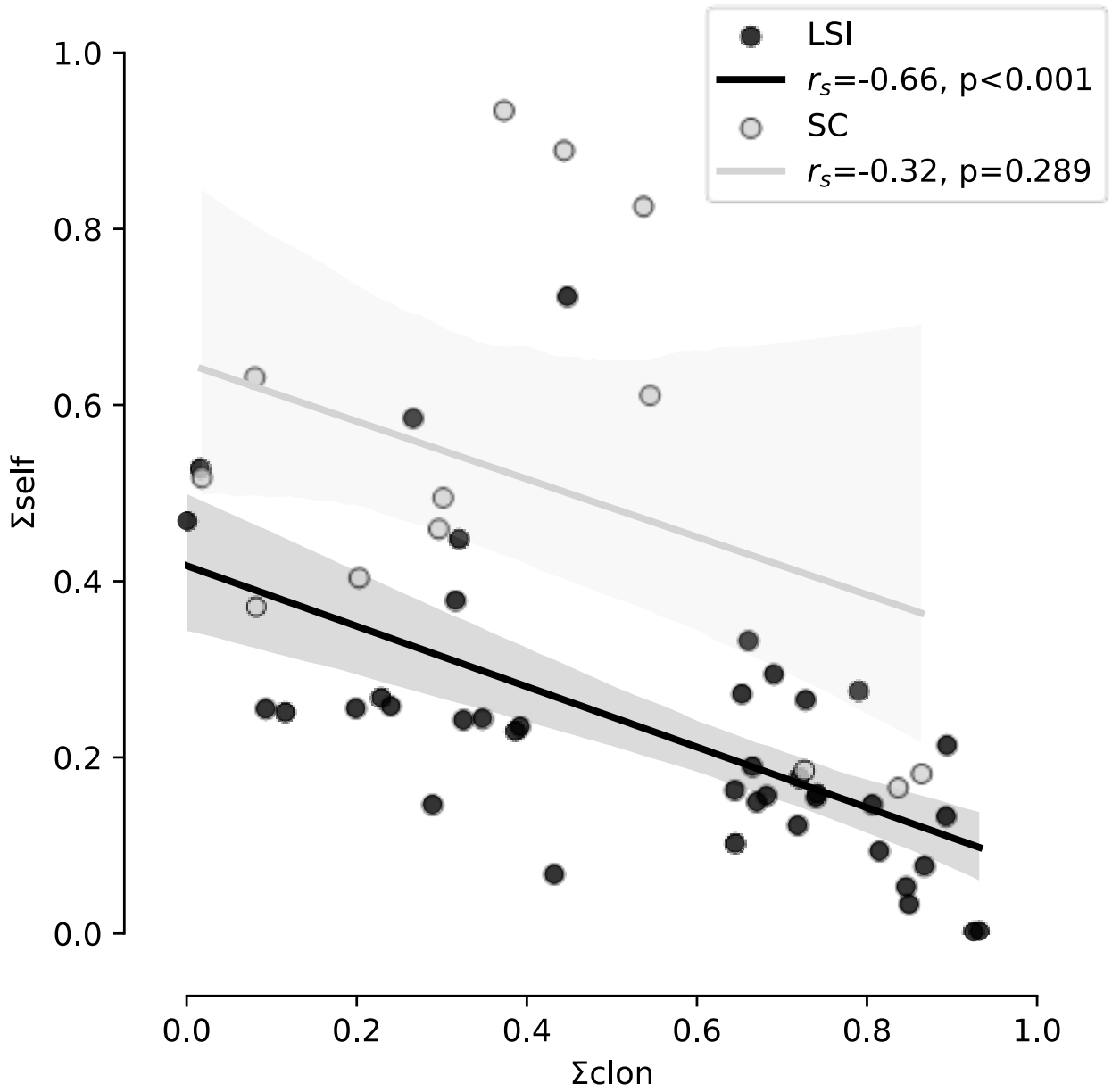


Figure 2: Correlations between Σ_{CLON} and Σ_{SELF} found in 13 SC populations (light grey points and regression line) and 40 LSI populations (black points and regression line). 95% confidence intervals are given for LSI and SC populations in dark grey and light grey, respectively. Spearman rank-order correlation coefficients (r_s) and probabilities that Σ_{CLON} and Σ_{SELF} would not be correlated (p) are reported for LSI and SC populations.

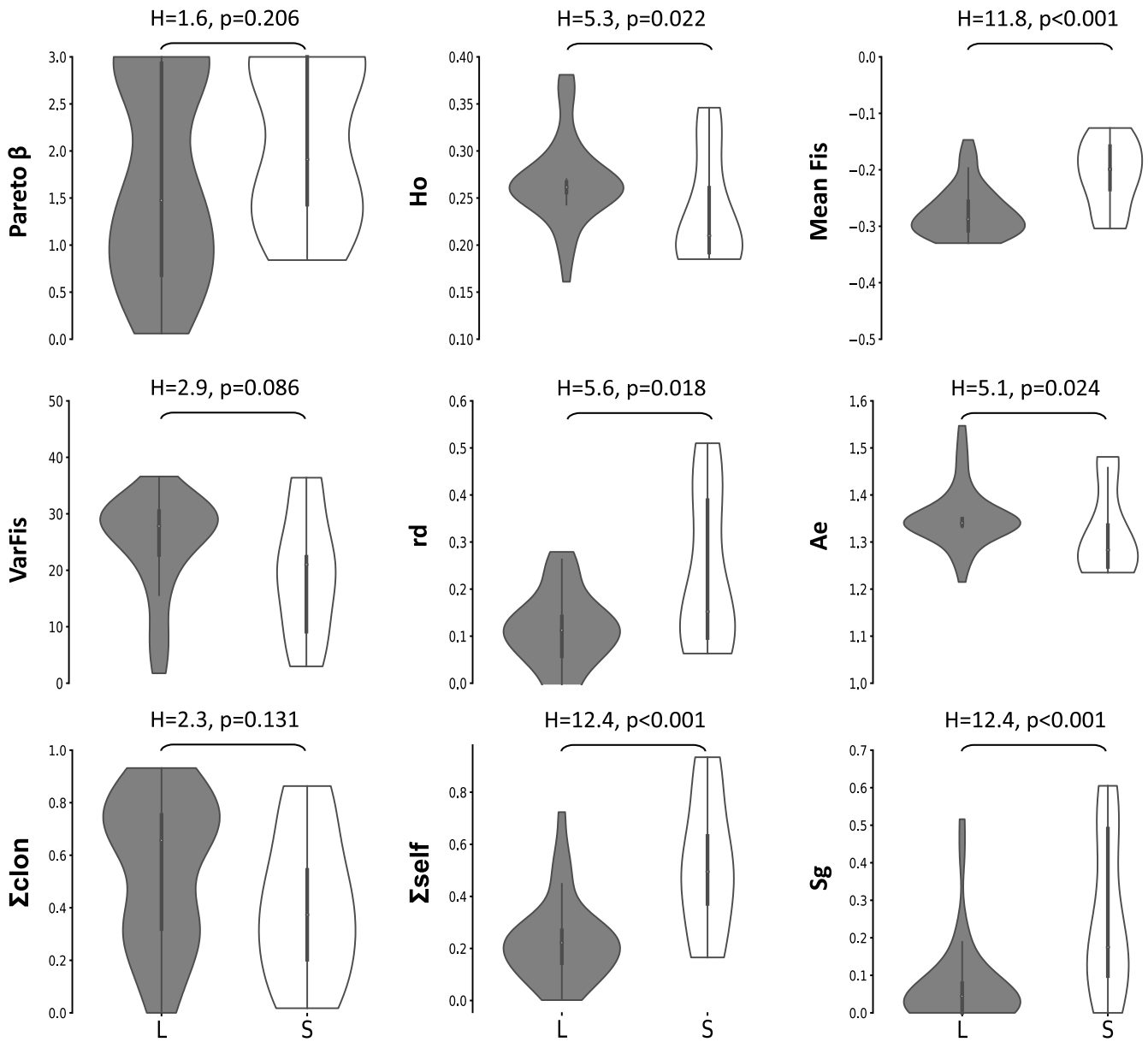


Figure 3: Violin plots comparing distributions of genetic indices in LSI (L, grey) versus SC (S, white) populations. Violin plots are cut for their minimum and maximum values. Kruskall-Wallis tests are reported as H-statistics as well as the probability that LSI and SC population would present the same distribution of genetic indices.

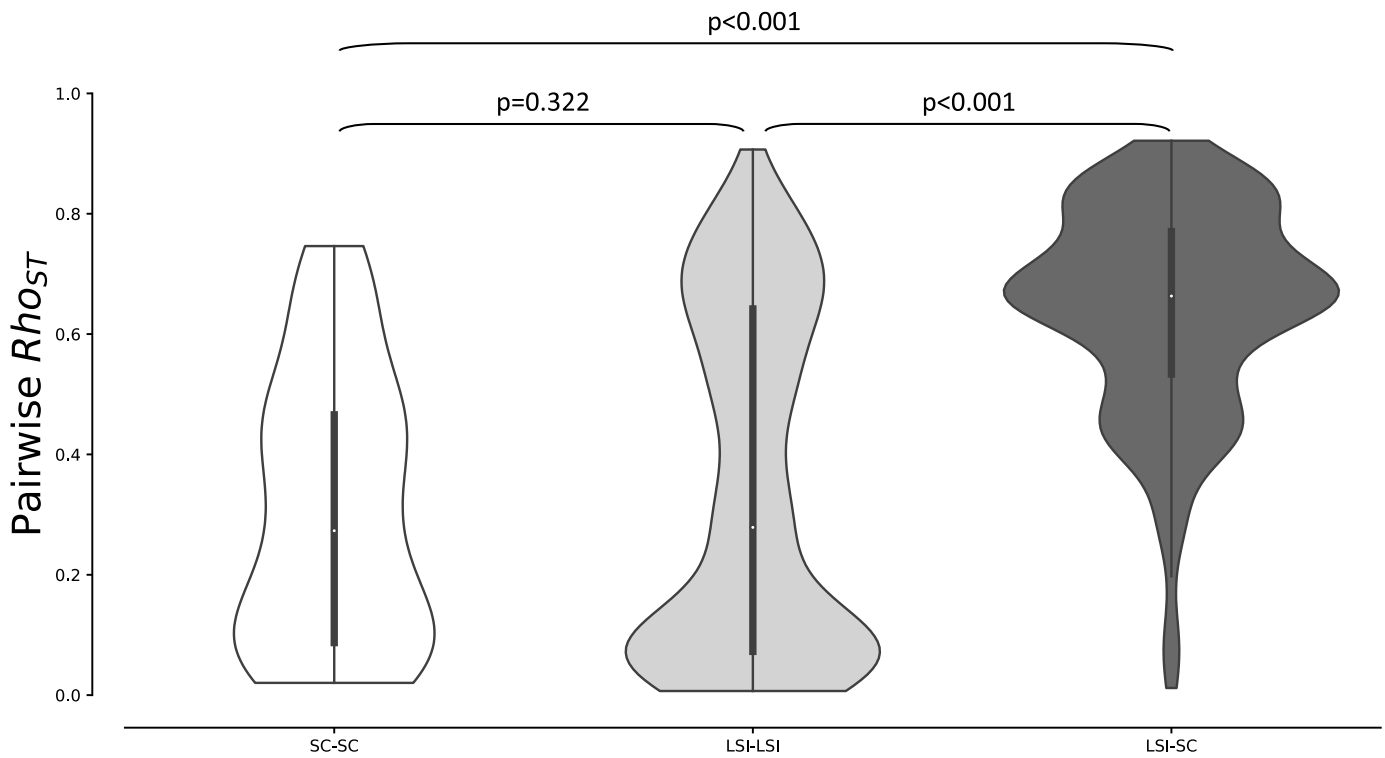


Figure 4: Distributions of pairwise rho_{st} within pairs of SC populations (white violin plot), pairs of LSI populations (light grey violin plot) and between pairs of LSI and SC populations (dark grey violin plot). Probabilities that pair of non-parametric distributions are equal are reported.

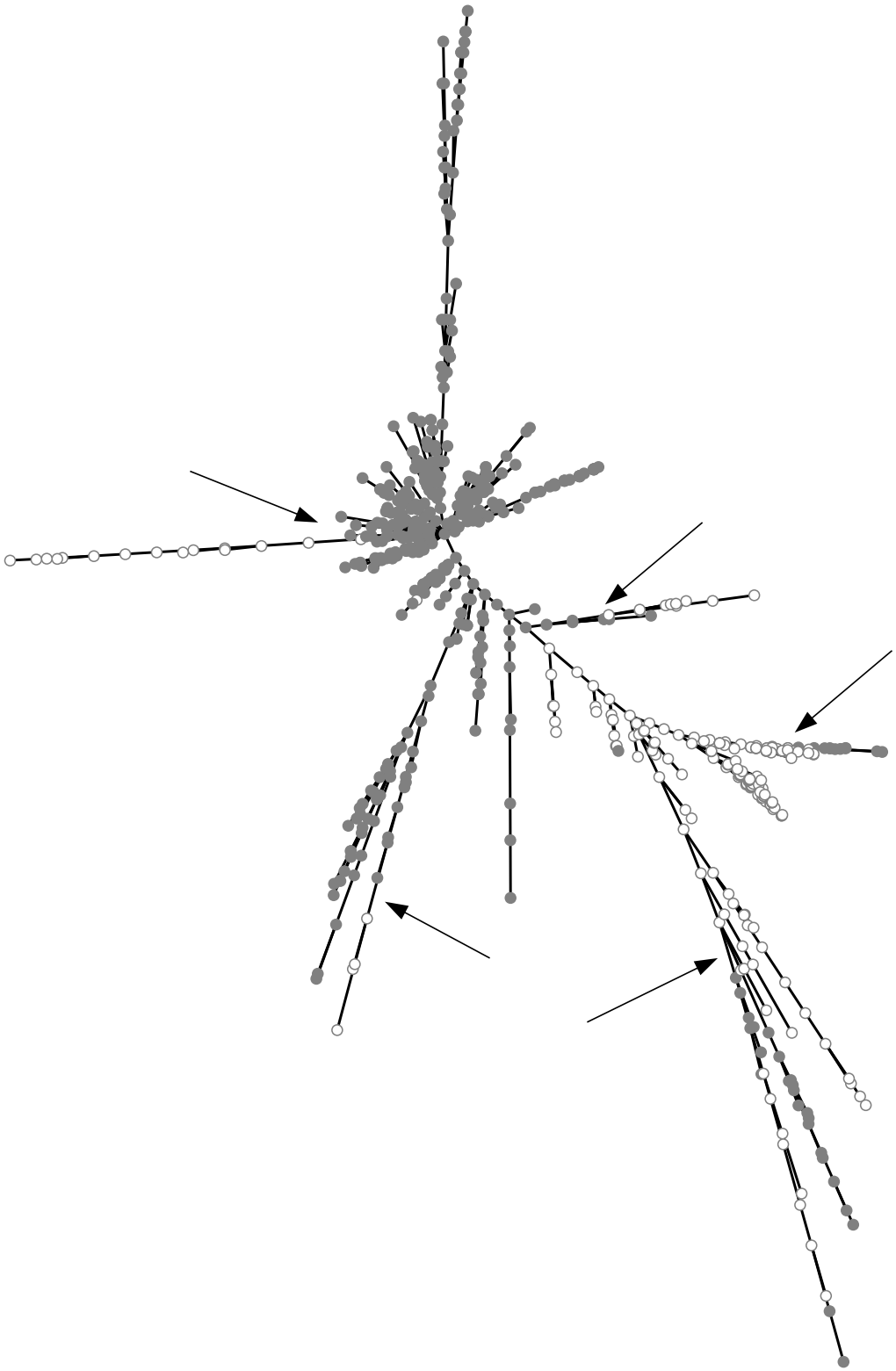


Figure 5: Minimum spanning tree of the genetic distances (number of different alleles) between LSI (grey points) and SC (white points) individuals. Arrows indicate some of the evidences of admixture between LSI and SC lineages.

Supplementary Information (SI)

Situation of *Lgh* in USA and Mediterranean populations in previous studies (Dandelot 2004, Okada et al. 2009, Armitage et al. 2013):
- nearly monoclonal populations arguing for the ecological advantage of clonality (fragmentation and rhizome propagation) in these invasive populations.

But
- using weakly polymorphic markers,
- before the demonstration of the heterostyly and of the Late-acting self-incompatibility system in *Lgh*

Situation of *Lgh* in western Europe, studied here:
- Spatial context: The other compatible type populations are away from hundreds of meters to several kilometers (Portillo-Lemus et al. 2021)
- Clonality through fragmentation and rhizomes, as in other populations (Portillo-Lemus 2021).
- Massive blossoming with a lot of pollinators harvesting flowers (Portillo-Lemus 2021).
- Faster pollen tube elongation of pollen from the other compatible type (Portillo-Lemus et al. 2022)

Our working hypotheses about the main, realized reproductive modes in population:

- **LSI populations:** A very limited production of seeds (Portillo-Lemus et al. 2021)

Hypothesis 1: **preferential clonality** due to lack of compatible partners to produce seeds → Baker's conjecture on reproductive modes (Pannell et al. 2015).

Hypothesis 2: The rare seeds really contribute to the populations → biological feature of seeds or recombination would have advantage over clones. If so, are they mainly:

Selfed seeds → reproductive assurance and advantage of recombination over clonality in a uniparental context?

Outcrossed seeds → advantage of outcrossed seeds due to heterosis and/or recombination despite the lack of locally-available compatible partners → selection of long pollination ranges

- **SC populations:** Production of a lot of seeds in natural populations

Hypothesis 1: despite a huge production of seeds, **preferential clonality** due to the physiological and/or ecological advantage of clones over seed germination process?

Hypothesis 2: the mass of seeds does contribute to the populations or some evolutionary advantage of recombination in offspring compared to clonal fragmentation and rhizome propagation to adapt local conditions? If so, are these seeds mainly:

Selfed seeds → within uniparental reproduction, biological feature of seeds or genetic recombination would provide advantage over clones?

Outcrossed seeds → Faster pollen tube elongation of pollen from the other type gives real advantage or advantage to allogamy to adapt new local conditions?

Supplementary information1: Synthesis of the questionings of our study to estimate reproductive mode of 53 populations of invasive Lgh using genetic diversity on the autotetraploid part of its genome.

Genotypic or genetic index	Range of values	Main driver	Behavior	references
β*	ϵ [0; +∞]	Rate of clonality	↘ when c ↗ $\beta \rightarrow 0$ when $c \rightarrow 1$	Halkett et al. 2005, Arnaud-Haond et al. 2007, 2020, Stoeckel et al. 2021
$varF_{IS}$	ϵ [0; +∞]	Rate of clonality	↗ when c ↗ $VarF_{IS} \rightarrow 0$ when $c \rightarrow 0$	Stoeckel & Masson 2014, Stoeckel et al. 2021
r_d	ϵ [0; 1]	Rate of selfing	slightly ↗ when c ↗ ↗ when s ↗ $r_d \rightarrow 1$ when $s \rightarrow 1$	Weir and Cockerham 1973, Navascues et al. 2010, Halkett et al. 2005
H_O	ϵ [0; 1]	Rate of selfing	↘ when s ↗ $H_O \rightarrow 0$ when $s \rightarrow 1$	David et al. 2007, Halkett et al. 2005
F_{IS}	ϵ [-1; 1]	Both rates	↘ when c ↗ $FIS \rightarrow -1$ when $c \rightarrow 1$ $FIS \rightarrow +1$ when $s \rightarrow 1$	Stoeckel & Masson 2014, Balloux et al. 2003, Stoeckel et al. 2021
H_E	ϵ [0; 1]	Rate of selfing	↘ when s ↗ $H_E \rightarrow 0$ when $s \rightarrow 1$	Glémin et al. 2001, Ritland & Ganders 1987
$P_{ID\text{-}SIB}$	ϵ [0; +∞]	Both rates	↘ when c ↗ ↘ when s ↗	Waits et al. 2001
S_g**	ϵ [0; 1]	Rate of selfing	↗ when s ↗	David et al. 2007, Hardy et al. 2016

* similar behavior of the clonal richness R
** inferred values of s, according to the method of Hardy (2016)

Table SI1: Hypotheses about the influence of clonality and selfing on the expected variations of genetic indices.

see the attached xls file

Table SI2: All genetic and genotypic indices computed on the 53 sampled populations. Each index was obtained from 15 individuals sampled in each population. (See spreadsheet supplementary material).

Population with missing genotype	nb_missing_genotypes	nb_individuals_concerned
Bléré_Ls	1	1
Chambéon_écopôle_du_Forez	2	2
Le_Port_Ls	1	1
Plougonver_Léguer	1	1
Quevillon	1	1
Rennes_apigné_prairie	1	1
Roost-Warendin	1	1
Saint_Aignan_Couflons_Ls	1	1
Vichy	1	1

Over single SNP genotypes	N	%age
single SNP genotype P(cad)>70%	28610	0.99965
single SNP genotype P(cad)≤70%	10	0.00035
total	28620	1

Over individuals	N	%age
nb_individuals_with_full_genotype	785	0.98742
nb_individuals_with_one_missing_genotype	10	0.01258
total	795	1

Table SI3: Summary of the allele dosage. First subtable lists populations with individual genotyped with at least one individual with missing genotype at one locus. Second subtable reports the frequency of single SNP genotypes with posterior probability of allele dosage strictly superior to 70%. Third subtable reports the frequency of individuals with their 36 SNP genotypes assigned with posterior probability superior to 70%.

Table SI4: Distribution of the different repeated Multi-Locus Genotypes (MLGs) found with 36 SNPs among populations considering their mating system (LSI or SC). Figure 1 presents a summarized plot of this table. (See spreadsheet supplementary material).

	G	H_O	H_E	A_E	R	β	F_{IS}	$VarF_{IS}$	r_d	PID_{SIB}	S_g	Σ_{clon}	Σ_{self}
G	-	0.438	**0.013**	**0.009**	***	***	***	***	0.057	**0.008**	**0.008**	***	***
H_O	0.076	-	***	***	0.397	0.462	0.122	**0.025**	0.653	***	**0.032**	0.065	0.142
H_E	**0.244**	**0.742**	-	***	**0.01**	**0.014**	0.212	***	**0.026**	***	0.844	***	0.372
A_E	**0.255**	**0.722**	**0.82**	-	**0.008**	**0.013**	0.353	***	0.094	***	0.347	***	0.845
R	**0.991**	0.083	**0.251**	**0.258**	-	***	***	***	0.052	**0.007**	**0.005**	***	***
β	**0.91**	0.071	**0.238**	**0.24**	**0.919**	-	***	***	0.126	**0.009**	**0.004**	***	***
F_{IS}	**0.602**	-0.144	0.116	0.086	**0.603**	**0.557**	-	***	***	0.163	***	***	***
$VarF_{IS}$	**-0.579**	**-0.209**	**-0.435**	**-0.346**	**-0.589**	**-0.551**	**-0.626**	-	***	***	***	***	***
r_d	0.187	0.042	**0.207**	0.156	0.19	0.148	**0.432**	**-0.365**	-	**0.021**	***	**0.006**	***
PID_{SIB}	**0.259**	**0.729**	**0.981**	**0.833**	**0.265**	**0.251**	0.13	**-0.443**	**0.214**	-	0.936	***	0.286
S_g	**0.264**	**-0.202**	-0.019	-0.089	**0.277**	**0.28**	**0.542**	**-0.379**	**0.353**	-0.008	-	***	***
Σ_{clon}	**-0.79**	**-0.171**	**-0.376**	**-0.347**	**-0.799**	**-0.809**	**-0.604**	**0.756**	**-0.254**	**-0.387**	**-0.34**	-	***
Σ_{self}	**0.388**	**-0.136**	0.083	0.018	**0.399**	**0.361**	**0.723**	**-0.524**	**0.631**	0.099	**0.69**	**-0.451**	-

Table SI5: Correlations between genetic indices using non-parametric Kendall partial rank-order correlation. Below the diagonal, the values of the coefficient of correlation τ are consigned. Above, the related p-values are provided. Values inferior to the threshold 0.05, and the related coefficient τ were bold. The parameters Σ_{CLON} and Σ_{SELF} correspond to the sum of normalized indices sensitive to clonality and selfing, respectively.

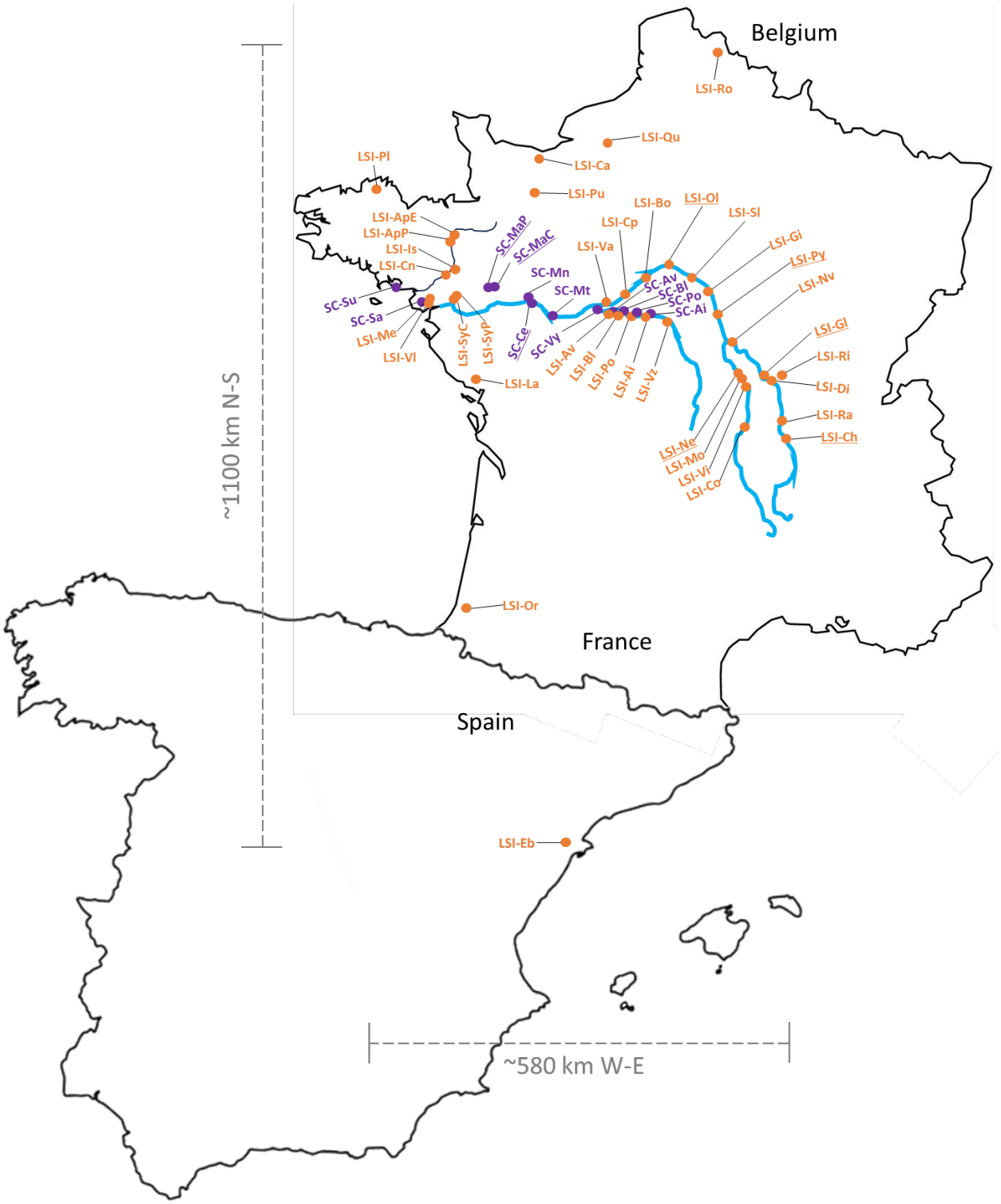
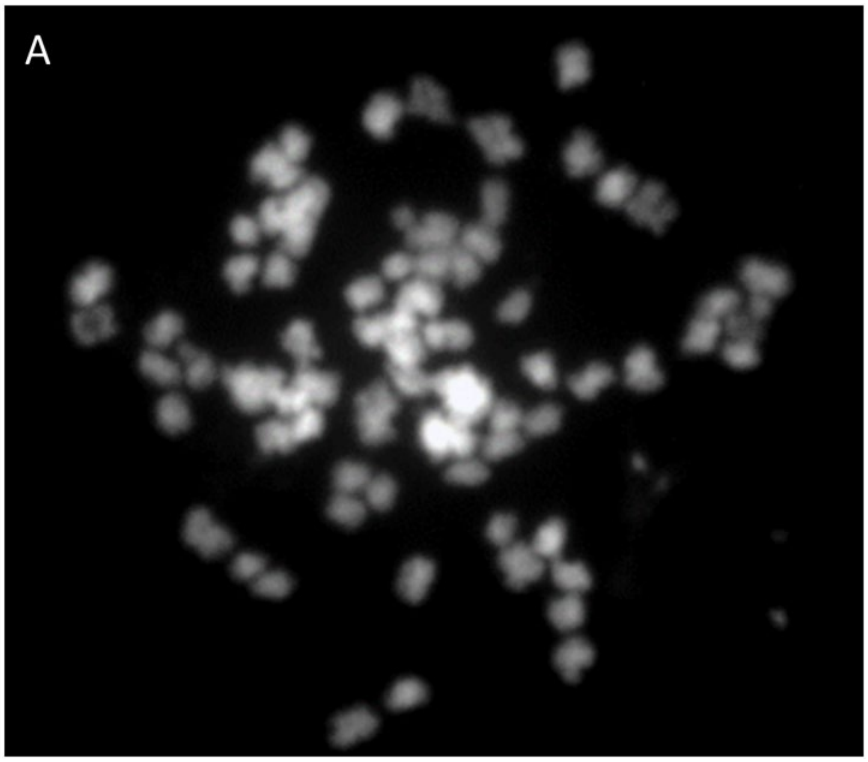


Figure SI1: Map of the locations of the 53 sampled populations across France and Spain. In orange, populations with only L-morph individuals, supposed to mate under the control of a late-acting self-incompatible (LSI); In purple, populations with only S-morph individuals, supposed to be self-compatible (SC). In light blue, the Loire river system. In grey, the North-to-South (N-S) and West-to-East (W-E) scale distances of the sampled populations. Underlined, the seven populations in which 15 individuals were tested for the self-incompatibility (Portillo-Lemus et al. 2021) and one individual was karyotyped to count the chromosome numbers.



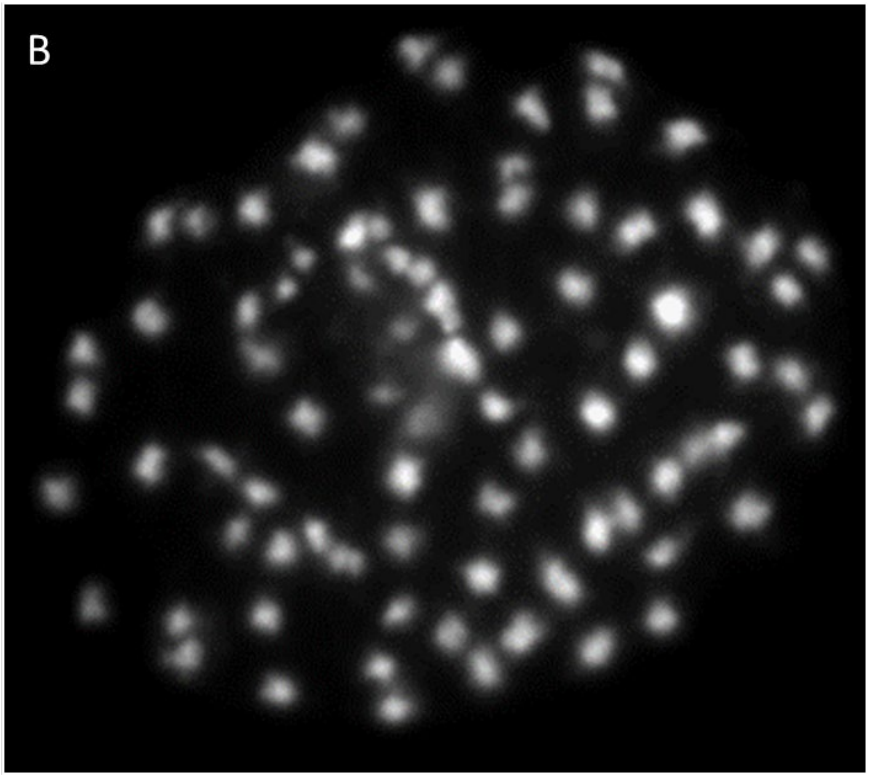


Figure SI2: Photos of somatic metaphase chromosomes indicating that L- and S-morph individuals that succeed to mate and to give full fruit set and 100% viable first- and second-generation descendants (Portillo-Lemus et al. 2021) also present the same number of chromosomes (2n=10X=80), belonging to the same taxon *Ludwigia grandiflora* subsp. *hexapetala*. We still didn't find *Ludwigia grandiflora* subsp. *grandiflora* karyotype in France (2n=6X=48, see Dandelot 2005, Barloy et al. 2024). **A**: somatic metaphase chromosomes from L-morph (2n=10X=80); **B**: somatic metaphase chromosomes from S-morph (2n=10X=80).

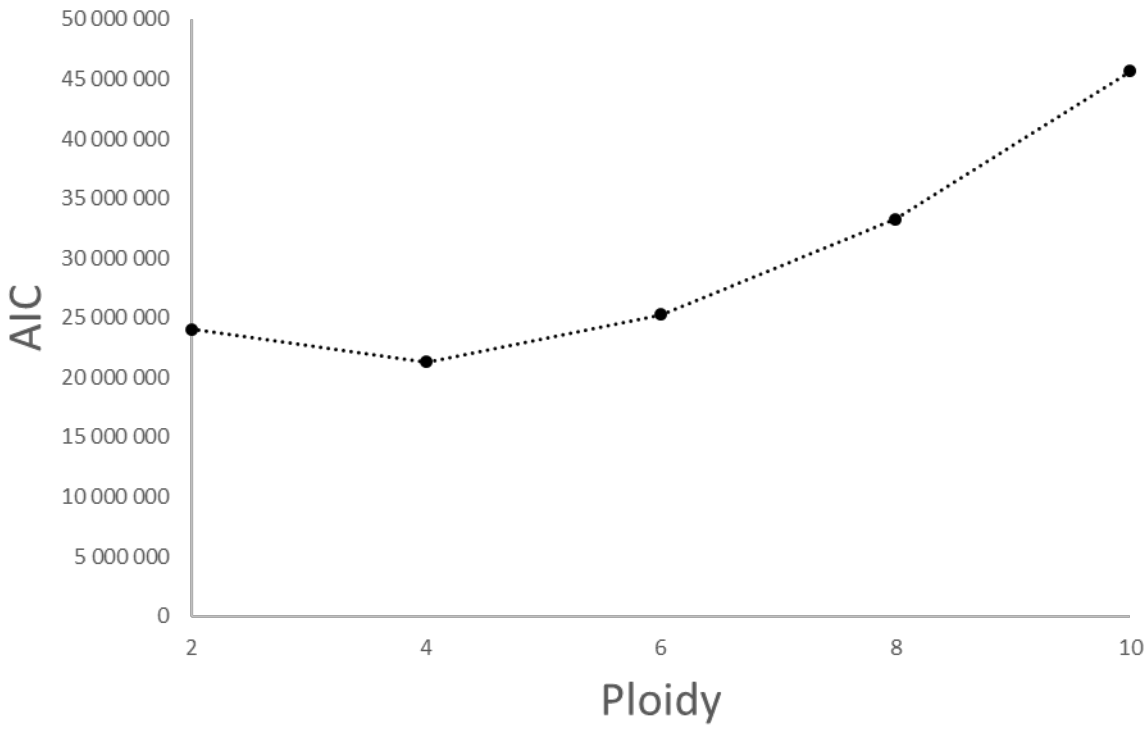


Figure SI3: Akaike's information criterion (AIC) as function of the ploidy level on the likelihoods of the sequenced allele countings over all individuals. Tetraploidy, a ploidy 4x, presents the lowest AIC, and thus the best supported model to explain distribution of sequenced allele countings over all individuals, which confirm our initial SNP development to be sure of their ploidy level.

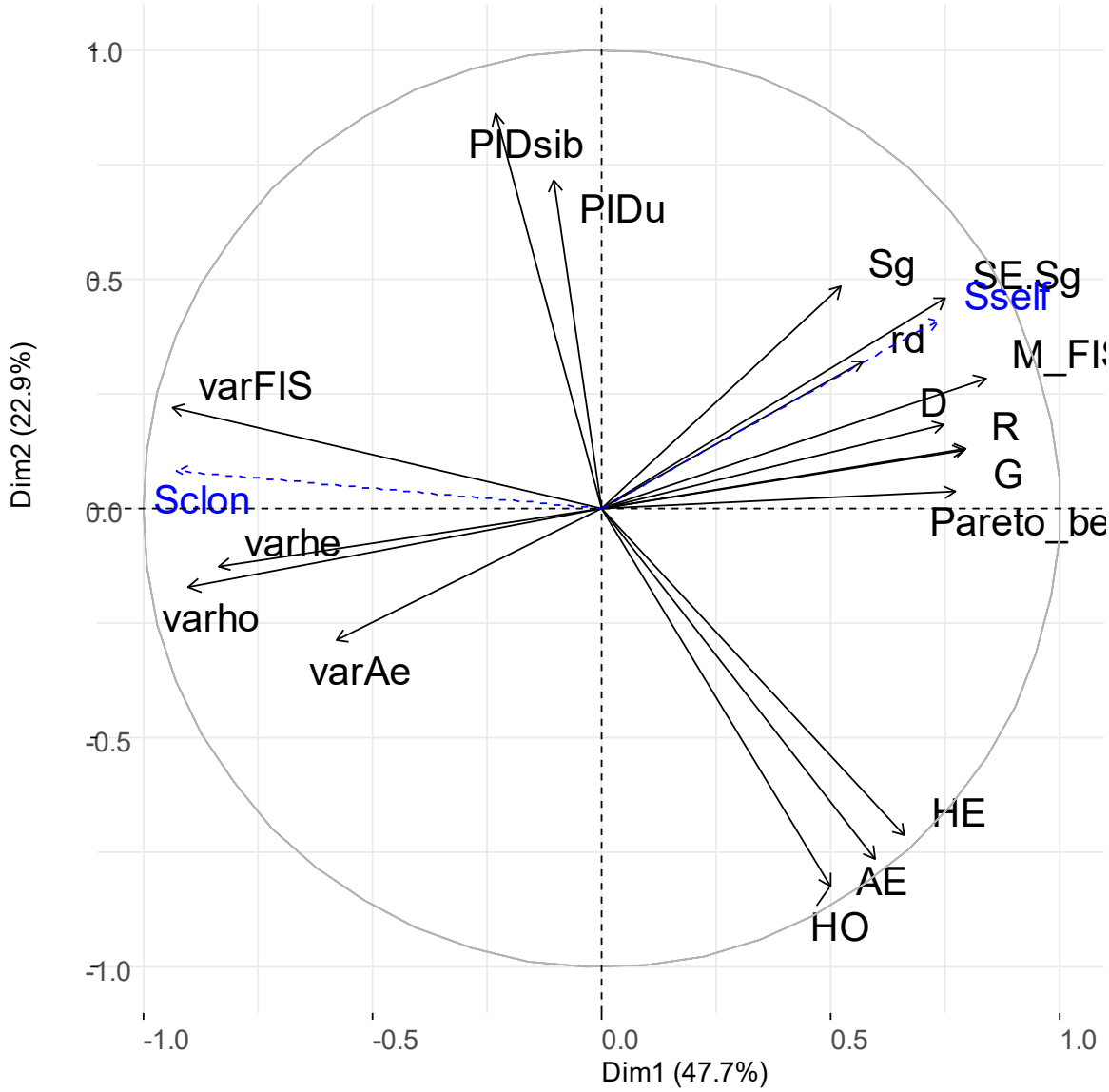


Figure SI4.A: Correlation circles of the principal component analysis on the 17 centered and scaled genetic indices measured in 53 *Lgh* populations in western Europe. The first horizontal dimension accounting for 50.24% of the variance regroups indices related to clonality (number of genotypes *G*, genotypic diversity *R*, clonal heterogeneity *D*, clonal evenness *Beta_Pareto*, variance of Fis *varFis*, variance of expected heterozygosity *varHe*, variance of observed heterozygosity *varHo* and variance of effective number of alleles varAe between loci). The second vertical dimension accounting for 23.39% of the variance with a main contribution of estimate of selfing Sg. Other indices, *i.e.*, mean observed heterozygosity *Ho*, mean expected heterozygosity *He*, effective number of alleles *Ae*, inbreeding coefficient *Mfis*, linkage disequilibrium *rd*, standard error of estimate of selfing *SE.Sg*, unbiased probability of identity under panmixia *PIDu* and between sibs *PIDsib*. In blue and dashed lines, the predictions of Σclon (*Sclon*) and Σself (*Sself*) provided as supplementary variables.

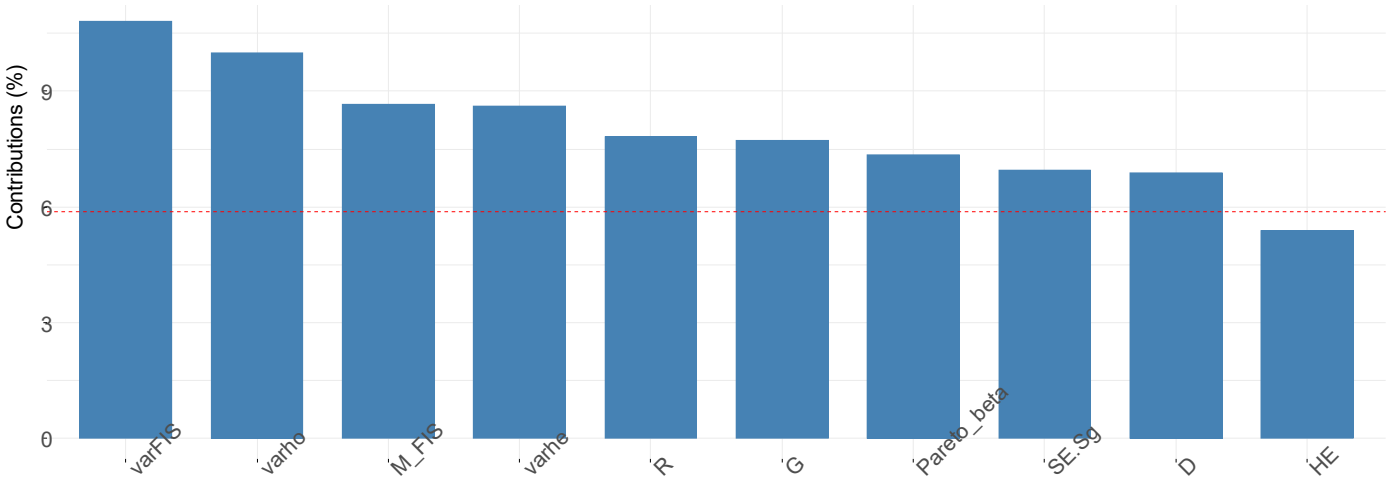


Figure SI4.B: Bar plot of the genetic index contributions to the first principal component. The expected average contribution *1/(number of genetic indices)* is plotted as the red dashed line. VarFis and indices of genotypic diversity (Pareto , R, G & D) are the main contributors to the first component.

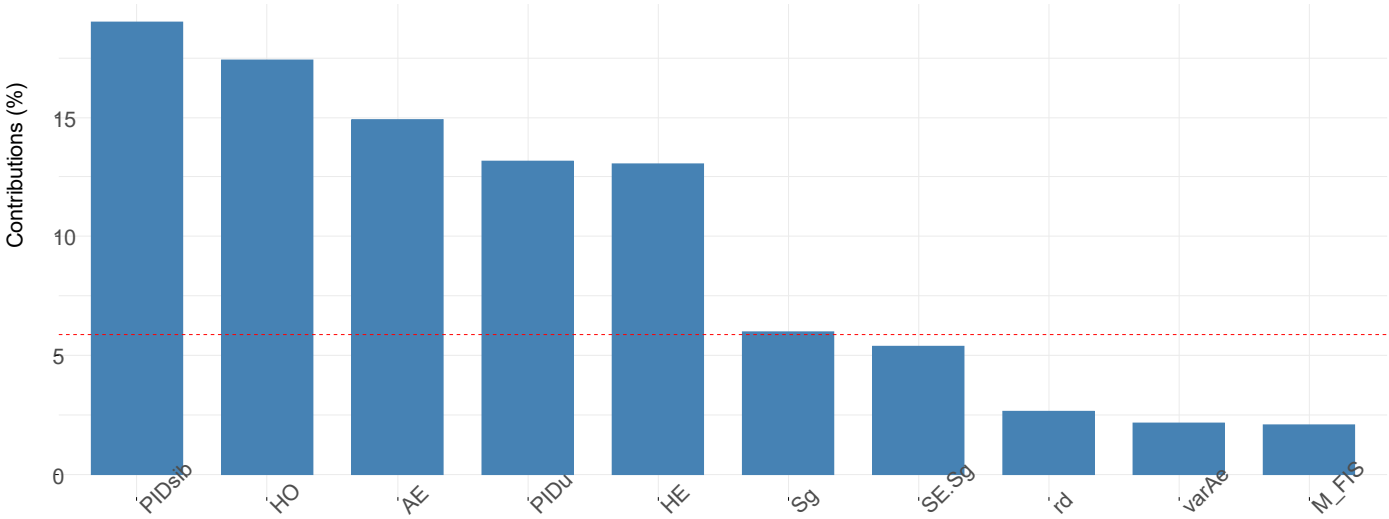


Figure SI4.C: Bar plot of the genetic index contributions to the second principal component. The expected average contribution *1/(number of genetic indices)* is plotted as the red dashed line. Estimates of selfing (Sg), mean observed heterozygosity (H_O), mean effective number of alleles (A_E) and probabilities of identities (pid_u and pid_{sib}) are the main contributors to the second component.

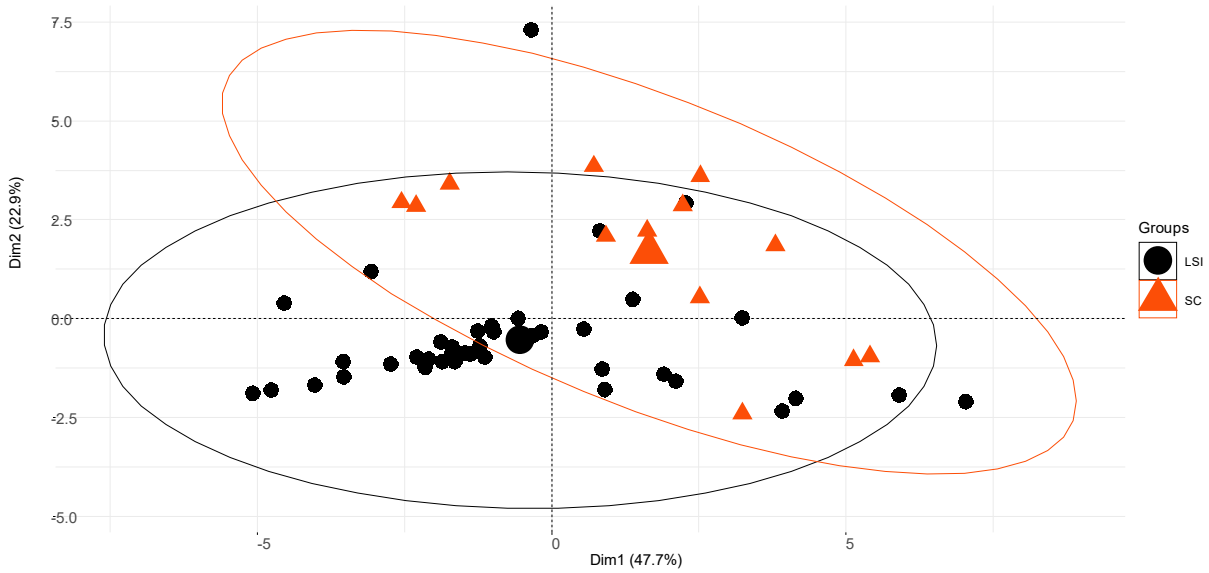


Figure SI4.D: Plot of the projection of each population onto the two first principal components. SC populations are plotted as red triangles and LSI populations as black diamonds. Concentration ellipses around LSI and SC populations are plotted in black and red, respectively. Barycenters of the two groups are represented as bigger black diamond and red triangle, respectively.

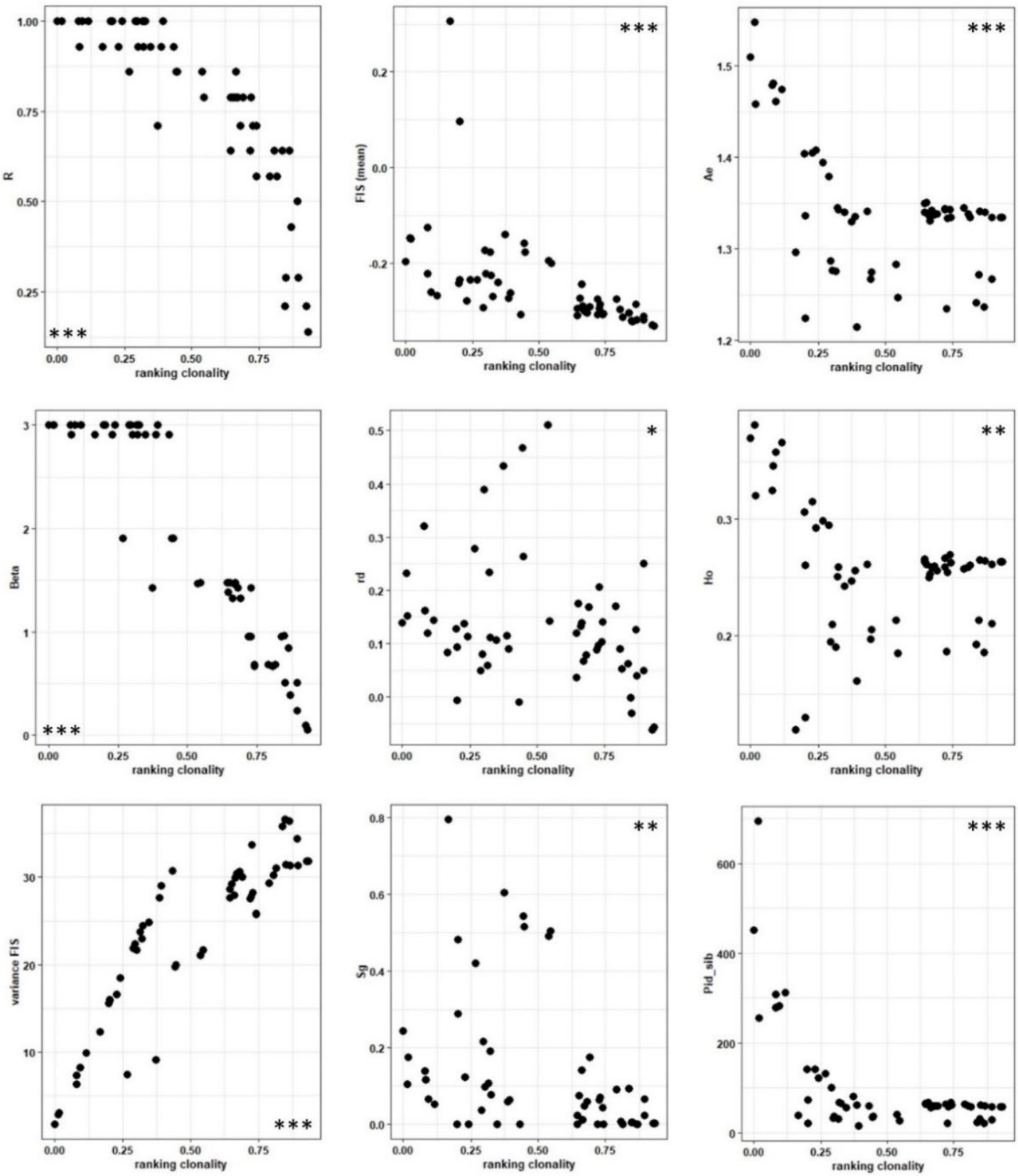


Figure SI5: Relationship and correlation between Σ_{CLON} and genetic indices (R, Pareto β, MFis, VarFis, Ho, Ae, Sg, pidsib and $\bar{r}_d$). Per graph, all the 53 populations were plotted as black dots. *: p<0.05, **: p<0.01, ***: p<0.001.

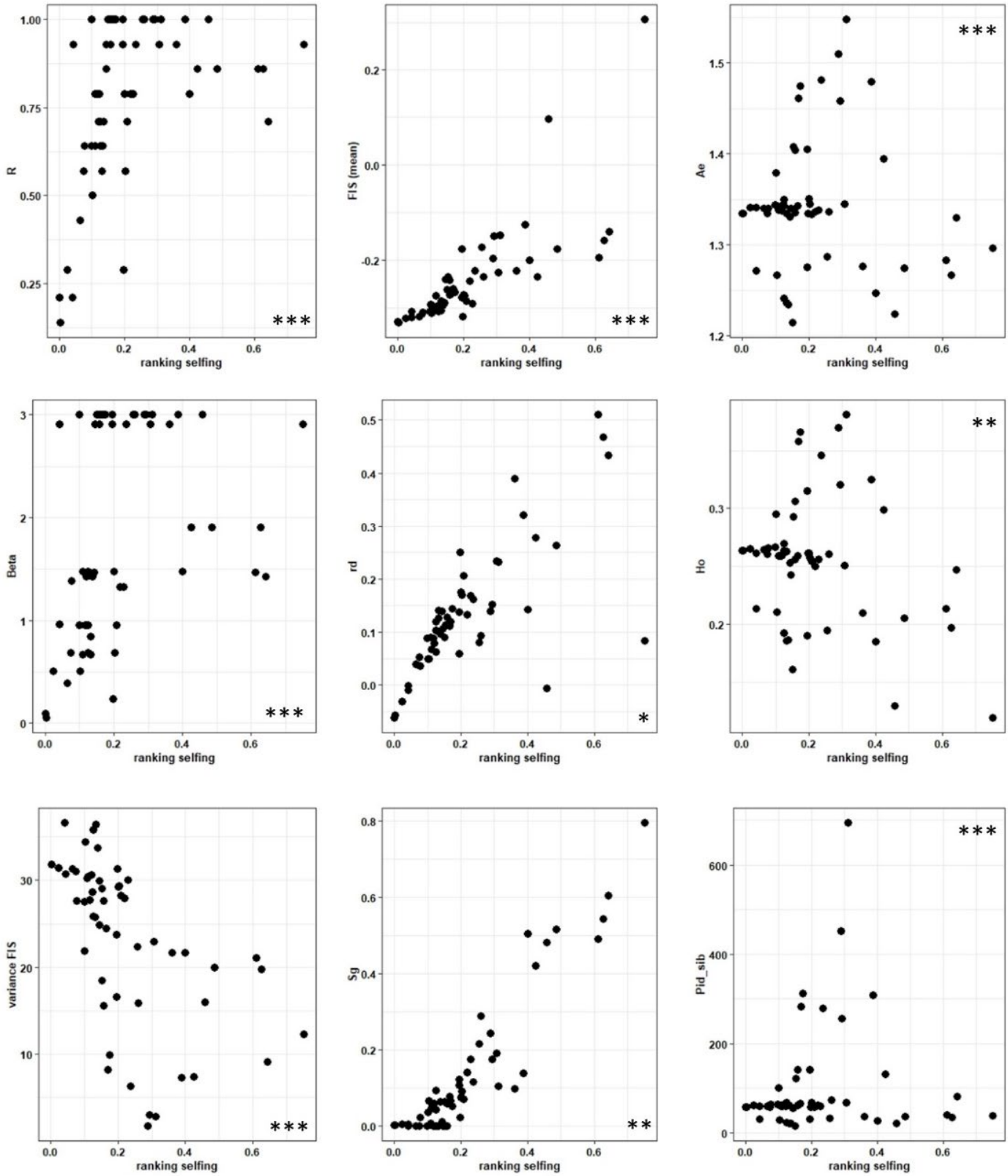


Figure SI6: Relationship and correlation between Σ_{SELF} and genetic indices (R, Pareto β, MFis, VarFis, Ho, Ae, Sg, pidsib and $\bar{r}_d$). Per graph, all the 53 populations were plotted as black dots. *: p<0.05, **: p<0.01, ***: p<0.001.

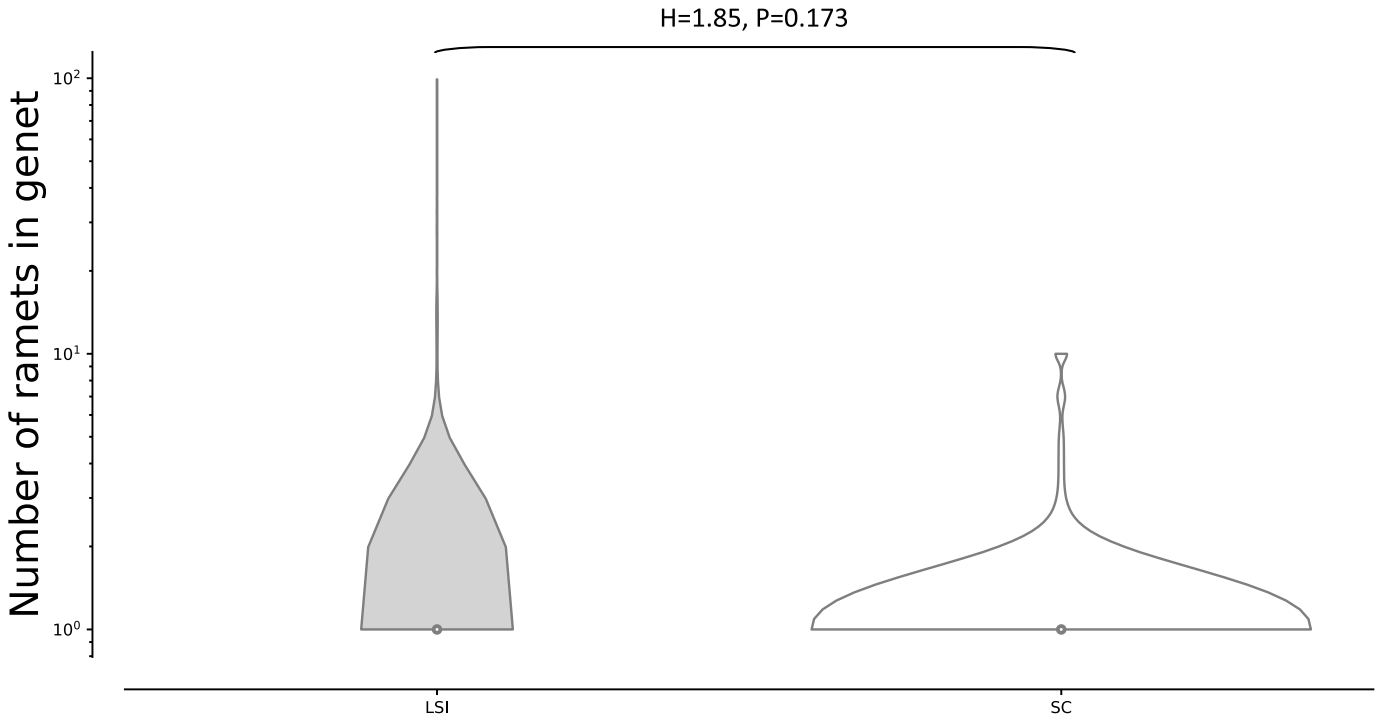


Figure SI7: Distributions of the logarithms of the number of ramets per genet in LSI (grey) and SC (white) populations. We report the Kruskal-Wallis probability that the two distributions are identical.

66 **Origin of the higher genetic diversity in LSI populations as compared to SC**

67 A first hypothesis to explain the genetic differences found between LSI and SC populations
68 may be due to the outcrossing consequences on LSI populations. Indeed, similar genetic
69 differences between SC and SI were recurrently reported in populations mating with
70 sporophytic and gametophytic SI systems (Igic et al. 2008, Mable 2008). For example,
71 *Leavenworthia alabamica* and *L. crassa* are two sister-species mating with a sporophytic SI
72 within a small genus of eight annual species in which self-fertilization has evolved at least four
73 times (Beck et al. 2006). In these species, comparisons between SI and SC populations also
74 showed higher mean genetic diversity in SI populations as compared to the SC, in an overall
75 background of high genetic diversity (Busch 2005, Koelling et al. 2011). High Ho, negative Fis
76 and high effective number of alleles per marker are also found in species mating with
77 gametophytic SI systems, such as in the *Prunus* genius (Stoeckel et al. 2006, Stoeckel et al.
78 2008, Abdallah et al. 2019). However, theoretical predictions argue that this kind of higher
79 genetic diversity favored and maintained in SI populations as compared to panmictic and SC
80 populations must only be found on polymorphic markers highly physically linked to the
81 genomic regions involved in the SI mechanisms that evolve under balancing selection, and
82 also, at a lesser magnitude, on physically unlinked markers but only in small population, with
83 less than 100 individuals including clones (Glémin et al. 2001, Navascues et al. 2010). The 36
84 SNPs we developed must randomly distribute over the 16 chromosomes inherited twice from
85 a common ancestor with *Lpm*, i.e., 32 chromosomes out of 80 total chromosomes of *Lgh*
86 (Barloy et al. 2024), with no reason that all aggregate around the genes involved in the LSI
87 self-recognition. In addition, individuals and populations of *Lpm* are reported to massively use
88 self-fertilization, thus with no SI system (Estes& Thorp 1974, Grewel et al. 2016). By
89 parsimony, there is less chance that genes involved in the LSI reaction would be stored on

these chromosomes than on some of the 48 other chromosomes; and therefore, very unlikely that our SNPs would be physically linked to genes under balanced selection due to LSI. *Lgh* populations in western Europe include far more than one thousand stems per local population (Portillo-Lemus et al. 2021). It is thus unlikely that differences in genetic diversity observed between LSI and SC populations would directly result from the consequences of the LSI system hitchhiked to the whole genome (Navascues et al. 2010).

The second hypothesis is that high genetic diversity in LSI populations may arise and be maintained due to higher rates of clonality (Stoeckel & Masson 2014, Reichel et al. 2016, Stoeckel et al. 2021a) as compared to SC populations. In LSI, self-pollen tubes physiologically blocked at the micropyle may prevent outcrossed-pollen tubes to reach the available ovules, decreasing female fertility and increasing the importance of clonal regeneration in populations (Vaughton et al. 2010). In addition, clonality is another uniparental reproductive mode that allows the local maintenance of populations in the absence of sexually compatible partners, which is more probable in peripatric conditions (Pannel et al. 2015). Dominant clonal reproduction and propagation are commonly found in peripatric populations extending their living range into new areas (Pyšek 1998, Wang et al. 2017). Furthermore, there is an established intriguing co-occurrence of clonality in SI populations of flowering plants (Vallejo-Marin et al. 2010). We know that dominant clonal reproduction theoretically favours the accumulation of deleterious mutations and that these deleterious mutations contribute to the advantage of outcrossing, which decreases the probability of inbreeding depression in offsprings, as compared to SC reproducers (Navascues et al. 2010). The observed genetic differences between LSI and SC populations may thus result from higher clonality in LSI populations lacking of compatible partners. But we found here no difference in estimates of

113 clonality between LSI and SC populations, in Σ_{CLON} and in all indices sensible to clonality (R,
114 Pareto Beta and VarFis; Figure 3).
115 We rather found significant difference in estimates of selfing between LSI and SC populations,
116 in Sg (Hardy 2016), in Σ_{SELF} and in all indices sensible to selfing (Ae, linkage disequilibrium, Ho,
117 mean Fis) revealing typical signatures of higher selfing rates in SC populations. Using Hardy's
118 method (2016) extended from David et al. (2007) for autopolyploids, SC populations were
119 estimated to reproduce using a mean selfing rate of 0.18 while LSI populations with only a
120 mean selfing rate of 0.05, both estimated with similar rates of sexuality. Congruent with higher
121 estimates of selfing rates and Σ_{SELF} in SC populations, we found higher linkage disequilibrium,
122 lower observed heterozygosity and less effective number of alleles in SC populations (Figure
123 3). The observed differences in genetic diversity between LSI and SC populations must thus be
124 rather caused by the genetic consequences of higher selfing in SC populations, decreasing
125 their genetic diversity, than due to a direct outcrossing effect of the LSI hitchhiked over all the
126 *Lgh* genome within these large populations of *Lgh* (see Navascues et al. 2010). Interestingly,
127 the outcrossing breeding system allowed by LSI, with a small rate of selfing allowed especially
128 at the end of the flowering season, was measurable at the population level in genotyping (5%
129 of selfing in LSI versus 18% in SC) and thus may be efficient at an eco-evolutionary scale.
130 The recent arrival of *Lgh* in western Europe may also have contributed to shape these genetic
131 differences. Recent founding events in SC populations are indeed expected to decrease the
132 overall genetic diversity all along the genomes, including fewer effective alleles, lower gene
133 diversity, and potentially more linkage disequilibrium as compared to outcrossing populations
134 (Flanagan et al. 2021). Despite the ρ_{ST} values we observed between pairs of SI and SC
135 populations being approximately double those measured within pairs of SC populations and
136 within pairs of LSI populations, we also found clues of admixture in the genetic distances

between LSI and SC individuals in western Europe (Figure 5). This result would thus tend to minimize the contribution of the foundation in the differences in genetic diversity we observed between LSI and SC populations, and stress again the importance of selfing to shape the genetic differences we found between SC and LSI populations.