Title: Reproductive modes of polyploid *Ludwigia grandiflora subsp. hexapetala* in western Europe: the effects of a late-acting self-incompatibility system and its absence on genetic diversity within populations

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 in populations 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.

Ludwigia grandiflora subsp. *hexapetala* (*Lgh*) is a partially clonal, polyploid, hermaphroditic, and heteromorphic plant that recently colonized multiple countries worldwide. In western Europe, individuals in this species are either self-incompatible caused by a late-acting self-incompatibility (LSI) system developing long-styled flowers, or self-compatible (SC) developing short-styled flowers. We do not yet know if the LSI system is effective in driving genetic diversity within populations, as previously showed for gametophytic and sporophytic SI systems. In addition, determining the reproductive mode of newly established *Lgh* populations in Europe will contribute to understanding their ecological and evolutionary roles at the margins of species distribution. However, we still lack of a full approach, including the essential step of unfolding population genetic indices to estimates rates of clonality, selfing and outcrossing in autopolyploids.

In this study, we proposed such an approach to measure genetic diversity to assess reproductive modes on 53 LSI and SC populations newly colonizing France and northern Spain, by assessing SNPs on a distinct autotetraploid part of the *Lgh* genome. SNPs are easily reproducible, adapted for intraspecific genetic studies and allow confident allele dosage to genotype autopolyploid individuals. In turn, such genotyping data made it possible to use recently developed methods for computing and interpreting genetic diversity to assess reproductive mode in each population.

We found that populations reproduced mainly clonally but with a high diversity of genotypes along with rates of sexual events up to 40%. We also found evidence for local admixture

between LSI and SC populations in a background of genetic structure between pairs of LSI and SC populations that was twice the level found among pairs of LSI populations or pairs of SC populations, arguing that SC and LSI populations may follow distinct expansion dynamics. These observations of sexual events in nearly all populations and of high diversity in clonal lineages imply to integrate their potential adaptation of populations to management plans.

We also took advantage of the spatial segregation of LSI and SC populations invading France to characterize the population genetic consequences of an LSI system. LSI and SC populations showed similar rates of clonality but significantly different rates of selfing, as expected considering their breeding system, and despite the small rates of failure in the LSI system. Within the 53 studied populations, the 13 SC-only populations were distinguished by fewer effective alleles, lower observed heterozygosity, and higher inbreeding coefficients, linkage disequilibrium and estimates of selfing than found in LSI populations. These results suggested that genetic structural differences found between LSI and SC populations mainly came from the increased selfing rates in SC populations rather than due to the genome-wide consequences of outcrossing in LSI populations or due to different rates of clonality. The overall maintenance of higher genetic diversity, with the possibility of resorting to clonality, selfing and outcrossing, may explain why LSI populations seem to be more prevalent in all newly expanding populations worldwide.

Beyond *Lgh*, our methodological approach may inspire future studies to assess the reproductive mode of other autopolyploid eukaryote populations, and our results emphasize the necessity to consider the variations of reproductive modes when managing invasive plant species.

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

70 # Introduction

71 Plants can reproduce using different mechanisms (i.e., breeding system), including clonality
72 and different types of sexuality (i.e., mating system, including allogamy and selfing;
73 Holsinger 2000). Reproductive mode corresponds to the actual balance between sexual and
74 clonal reproduction, measured by the rate of clonality (De Meeus et al. 2007; Stoeckel et al.
75 2021a), and, within the proportion of offspring produced sexuality, the balance between
76 autogamy and geitonogamy (hereafter *selfing*) and allogamy (*outcrossing*), measured by the
77 rate of selfing (Bürkli et al. 2017). Reproductive mode is one of the main drivers of genetic
78 diversity in populations and species (Duminil et al. 2007, Ellegren & Galtier 2016, Glémin et al.
79 2019). It determines the way hereditary material is transmitted over generations, and thus
80 constrains the range of genetic diversity that can evolve within populations and species (De
81 Meeus et al. 2007, Fehrer 2010, Stoeckel & Masson 2014, Rouger et al. 2016). Its genetic
82 consequences are deep enough to be discernible even in the presence of other affecting
83 factors (Charlesworth 2003). Considering its influence on genetic diversity and structure in
84 populations, reproductive mode is one of the major biological traits to decipher before
85 interpreting other biological and ecological forces using molecular data (Duminil et al. 2007,
86 Reichel et al. 2016, Bürkli et al. 2017, Orive & Krueger-Hadfield 2021). Given an adequate
87 theoretical framework, population geneticists can use genotypic data to infer reproductive
88 modes in populations (Halkett et al. 2005, Arnaud-Haond et al. 2007, Hardy 2016, Bürkli et al.
89 2017, Stoeckel et al. 2021a).

90 Approximately 80% of plant species are partially clonal (Barrett 2010, Vallejo-Marin et al.
91 2010). Clonality consists of a parent producing a new descendant that is a genetic copy of itself
92 with the exception of rare somatic mutations and mitotic recombinations (De Meeus et al.

93 2007). Clonality generates repeated genotypes in populations when parents can yield multiple
94 descendants (Arnaud-Haond et al. 2007). This results in an average excess of heterozygotes
95 compared to Hardy-Weinberg expectations (*i.e.,* negative F_{IS}), large variance of F_{IS} among
96 polymorphic markers along the genome and a small excess of linkage disequilibrium,
97 increasing with decreasing population size (Navascues et al. 2010, Stoeckel & Masson 2014,
98 Stoeckel et al. 2021a). These effects vary with the rates of clonality and, in turn, can be used
99 to infer the rates of clonality themselves within populations (De Meeus et al. 2006, Becheler
100 et al. 2020, Arnaud-Haond et al. 2020).
101 Approximately 70% of angiosperm species are hermaphroditic (Dellaporta & Calderon-
102 Urrea 1993) and thus have the possibility to sexually reproduce with themselves (selfing).
103 Selfing decreases effective population sizes and thus speeds up the effect of genetic drift,
104 leading gene diversity to decrease in populations and even to allele fixing (Wright et al. 2008,
105 Roze 2016, Glémin et al. 2019). Selfing also prohibits genetic mixing between different
106 ancestral lineages, resulting in decreased heterozygosity within individuals (David et al. 2007,
107 Hardy 2016) and increased linkage disequilibrium between genes along genomes (Golding &
108 Strobeck 1980, Norborg 2000, Lucek & Willi 2021).
109 Around half of hermaphrodite plants restrict self-fertilization using a variety of molecular
110 mechanisms grouped under the term "self-incompatibility (SI) systems" which favour
111 outcrossing within populations (De Nettancourt 2001, Castric & Vekemans 2004, Gibbs 2014,
112 Steinecke et al. 2022). Outcrossing is overall expected to increase genetic diversity, to limit
113 allele fixation and to reduce linkage disequilibrium when compared to selfing. However, such
114 expectations depend on the genetic linkage between the considered loci and the genomic
115 region carrying the genes involved in self-recognition (Glémin et al. 2001, Stoeckel et al. 2006,
116 Navascues et al. 2010). Within the different identified SI systems (Charlesworth et al. 2005,

117 Franklin-Tong 2008), the late-acting SI system (hereafter LSI) is still poorly studied despite
118 being identified in multiple taxonomic groups in angiosperms (De Nettancourt 1997,
119 Gibbs 2014). In LSI systems, self-pollen tubes grow and are only blocked shortly before
120 penetrating the ovule, or at the micropyle, or, in some species, the ovules are penetrated but
121 their syngamy fails. Such species may present reduced female fertility due to self-pollen
122 disabling ovules, favouring the clonal regeneration of populations (Vaughton et al. 2010). LSI
123 systems are also characterized by low but recurrent failures of the self-recognition system due
124 to its late mechanism, leading to the production of a low amount of selfed seeds in populations
125 (Seavey & Bawa 1986; Chen et al. 2012, Gibbs 2014). Gametophytic and sporophytic self-
126 recognition occurring very early in the pistil drastically limit self-pollination and thus selfed
127 seeds (Lawrence 1985, Gibbs 2014). Specific chemical or physical treatments are required to
128 overcome these mechanisms (Cabin et al. 1996, Ahmad et al. 2022). However, we do not yet
129 know if these rare selfed seeds contribute to the dynamics of genetic diversity of LSI
130 populations and species, especially in peripatric conditions or in any situation when
131 compatible partners may lack. Due to these selfed seeds, the LSI system may be relatively
132 ineffective at driving genetic diversity within populations, contrary to gametophytic and
133 sporophytic SI systems (Brennan et al. 2002, Stoeckel et al. 2006, Koelling et al. 2011, Busch &
134 Urban 2011). In such situations, the emergence and maintenance of LSI systems could become
135 a new evolutionary puzzle among the reproductive systems.
136 Autopolyploidy is pervasive in plants whether it derives from whole genome duplication or
137 from multiple hybridization events (Baduel et al. 2018). Clonality and selfing tend to occur
138 more frequently in polyploids than in their diploid counterparts (Husband et al. 2008, Herben
139 et al. 2017, Van Drunen & Husband 2018). Autopolyploid species bring two more main
140 challenges when tackling population genetics (Dufresne et al. 2014). The first issue is achieving

141 confident allele dosage to obtain a robust genotype for each individual at each locus. For
142 instance, discerning between AAAC, ACCC and AACC genotypes in an autotetraploid requires
143 a reproducible and robust quantitative approach to genotyping. The second issue is that
144 computing values of genetic diversity, identities and heterozygosity within autopolyploid
145 populations requires adapted methods (Hardy 2016, Meirmans et al. 2018). For instance, in
146 autopolyploids, heterozygosity, i.e., the probability that two random alleles are not identical-
147 in-state, can take more different values than in the case of diploids where they only are 0 for
148 homozygotes and 1 for heterozygotes (Haldane 1930, Hardy 2016). Autopolyploids require
149 dedicated computations of population genetic indices to not prevent artefactual population
150 genetic values. These can now be commonly computed using dedicated software such as
151 Spagedi (Hardy & Vekemans 2002) and GenAPoPop (Stoeckel et al. 2024).
152 Water primrose*, Ludwigia grandiflora* subsp. *hexapetala* (Hook. & Arn.) Nesom and Kartesz
153 (2000), hereafter *Lgh*, is one of the most invasive aquatic plants in the world (Thouvenot et al.
154 2013). *Lgh* is an alloautodecaploid species (2n=10x=80 chromosomes), resulting from
155 hybridization of different ancestral diploid species, some of which are represented more than
156 once in the total genome of *Lgh*, which belongs to the genus *Ludwigia* L. section Jussiaea (Hoch
157 et al. 2015, Barloy et al, 2024). Interestingly, *Lgh* includes an autotetraploid set of
158 chromosomes, shared with *L. peploides* subsp. *montevidensis* (2n=2x=16), hereafter *Lpm*, that
159 is clearly distinct from the other ancestral part of the *Lgh* genome (Barloy et al. 2024). *Lpm* is
160 a self-compatible-only diploid species with only one common floral morphology (Estes &
161 Thorp 1974, Grewel et al. 2016). In turn, *Lgh* presents heteromorphic flowers corresponding
162 to two floral morphs: L-morph individuals develop long-styled flowers and S-morph individuals
163 develop short-styled flowers (Portillo-Lemus et al. 2021). L-morph flowers express an active
164 LSI in western European populations (Portillo-Lemus et al. 2022). During the core flowering

165 season (summer), in experimental greenhouse conditions, L-morph individuals show a stable
166 seemingly insignificant rate of autogamy (around 0.2‰ of the available ovules) that increases
167 at the end of the flowering season, during autumn, to 1‰, which is common in LSI systems
168 (Gibbs 2014). Due to the massive blossoming of this species, growing in very dense
169 populations, these selfed seeds would add up yearly in field populations to thousands of seeds
170 per metre square (Portillo-Lemus et al. 2021). This pattern of low rates of autogamy, that
171 increase at the end of the flowering season, may provide the advantage of reproductive
172 assurance (Goodwillie & Weber 2018). In contrast, S-morph individuals are self-compatible
173 (SC) in western European populations but pollen tubes of the L-morph grow significantly faster
174 and are thus advantaged to fertilize ovules when in competition with self-pollen (Portillo-
175 Lemus et al. 2022). In addition, peripatric *Lgh* populations, including European populations,
176 were previously reported as exclusively clonal with few clonal lines (Dandelot 2004, Okada et
177 al. 2009). Recently established populations in France and northern Spain mostly present only
178 one of the two compatibility modes locally, sometimes with a population of the other type a
179 few to tens of kilometers away, which may result in effective allogamy (Portillo-Lemus et al.
180 2021, 2022). All these different breeding systems make possible very different reproductive
181 modes in populations, comprising all possible quantitative combinations of mainly clonal,
182 autogamous and allogamous modes. Genotyping an autotetraploid inherited genome
183 confidently, analyzing the genotyping to obtain population genetic measures to estimate the
184 respective and quantitative contributions of these different reproductive modes, still
185 constitute challenging tasks (Dufresne et al. 2014). Use of genetic diversity within a population
186 to estimate rates of clonality versus sexuality and of selfing versus allogamy is now commonly
187 achieved for diploids and haplodiplontics (David et al. 2007, Arnaud-Haond et al. 2020,
188 Krueger-Hadfield et al. 2021, Stoeckel et al. 2021a,b). Such a framework in the context of

189 autopolyploid populations, with an applied case study that would demonstrate its
190 effectiveness, is lacking.
191
192 In this study, we proposed a complete framework, from genotyping the autotetraploid part of
193 *Lgh* genome with allele dosage to the estimation of reproductive modes per population using
194 genetic data, compiling recent molecular and software developments (Besnard et al. 2023,
195 Hardy 2016, Stoeckel et al. 2024) to fill the gaps identified by Dufresne et al. (2014) for
196 confidently genotyping autopolyploid individuals. First, we developed a codominant and
197 reproducible set of single-nucleotide polymorphism (SNP) markers for this species that are
198 more adapted for intraspecific genetic studies, being able to distinguish between siblings and
199 different clonal lineages, and that allowed reproducible and confident allele dosage
200 assignment. Indeed, allele dosage at each marker within each individual is mandatory to
201 compute confident population genetic measures within populations that have no reason to
202 follow genotype frequencies expected under Hardy-Weinberg proportions, being partially
203 clonal, outcrossed and selfed. To be sure of the ploidy level of the studied markers and that
204 the genotyped markers did not interact with genomes of the other ancestral species of *Lgh*
205 genome, we genotyped SNPs only within the autotetraploid part of the *Lgh* genome shared
206 with the diploid species *Lpm*. The *Lpm* genome is sufficiently genetically distinct from the
207 other ancestral genomes of *Lgh* to avoid misgenotyping (Barloy et al. 2024).
208 We then used this approach to identify and study the variability of the reproductive modes of
209 *Lgh* populations within the invaded region of Western Europe considering their LSI and SC
210 status. To assess the reproductive modes of each sampled population and estimate their rates
211 of clonality and selfing, we measured gene diversity, heterozygosity, linkage disequilibrium
212 and genotypic diversities within 53 populations sampled in western Europe. Being able to

213 assess reproductive modes in each population, we will discuss the potential advantage of
214 uniparental reproduction including clonality and selfing, of recombination (selfing and
215 allogamy) and of heterozygosity in such peripatric invasive populations.
216 Second, due to the fact that recent local populations in France and northern Spain present
217 only one of the two compatibility modes, we took advantage of this contrasted situation
218 within the same species and in comparable ecological situations to compare genetic diversity
219 among LSI and SC populations, to assess the influence of LSI on genetic diversity as similarly
220 tackled in sporophytic self-incompatible and SC Brassicaceae species previously (Koelling et al.
221 2011). Indeed, genetic diversity underlies the evolvability of species and their survival and
222 adaptation to environmental change and ecosystem resilience (Booy et al. 2008, Hoban et al.
223 2021). The small proportions of selfed seeds allowed by the LSI mechanism may mitigate the
224 difference in genetic diversity between LSI and SC populations compared to sporophytic and
225 gametophytic self-incompatible systems. Considering the complex case of *Lgh* in western
226 European populations, we hypothesized that LSI populations should present typical genetic
227 footprints of dominant outcrossing, and perhaps higher rates of clonality due to the local lack
228 of compatible partners and self-pollen interferences, while SC populations should present
229 higher rates of sexuality prevailed by selfing. To understand the genetic impacts of
230 reproductive modes, we also quantified the covariations of genetic indices and their
231 importance to define the genetic diversity within the 53 genotyped populations, and
232 compared these observations to the theoretical expectations obtained from a Wright-Fisher-
233 like model extended to autotetraploids (Stoeckel et al. 2024). Finally, we interpreted the
234 genetic differences found within and between LSI and SC populations to understand how
235 reproductive mode(s) contributed to the genetic diversity in these peripatric *Lgh* populations.

236 # Materials and methods

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

238 We collected 795 stems of *Lgh* from 53 locations (52 locations in France and one in Catalonia,
239 Spain), corresponding to an area that spans 880 kilometres east-to-west and 1,100 kilometres
240 north-to-south (Figure SI1). At each location, we collected 15 stems (hereafter, 'individuals')
241 along a linear transect of 40 meters. Along each transect, we randomly collected 3 stems at
242 coordinates X_1 = 0m; X_2 = 10m; X_3 = 20m; X_4 = 30m and X_5 = 40m within a 1 meter-square
243 quadrat. The young leaves of each sampled individual were stored after lyophilization until
244 DNA extraction. We visually identified floral morphologies of flowers found along each
245 transect within the sampled *Lgh* populations.

246

247 **Definition of the SNP marker set**

248 As no molecular markers suitable for clonal discrimination were yet available for *Lgh* and *Lpm*,
249 we generated an original set of SNP markers via RAD-Seq (Baird et al., 2008) from five and
250 three *Lgh* and *Lpm* western European populations, respectively, using a pool of 15 individuals
251 sampled for each population. RAD DNA library generation, sequencing and the analysis
252 pipeline to identify *Lgh* SNP markers were carried out as described in Delord et al. (2018). In
253 brief, DNA of *Lgh* and *Lpm* were digested by *Sbfl* restriction enzyme and used to prepare DNA
254 libraries that were then sequenced using an Illumina HiSeq3000 (150bp paired-end reads). A
255 total of 14,233 and 34,287 RAD-Seq-determined SNPs were filtered to yield 340 and 326 SNPs
256 for *Lpm* and *Lgh*, respectively. Design of primers compatible with Hi-Plex multiplexing was
257 carried out by Melbourne Bioinformatics. Finally, sixty and fifty SNP markers were selected for
258 *Lpm* and *Lgh*, respectively. In our study, we only considered polymorphic SNPs that were

259 shared between *Lgh* and *Lpm* to be sure they belong to the tetraploid part of *Lgh* genome
260 derived from *Lpm* (Barloy et al. 2024). We finally kept a set of 36 polymorphic and stable SNP
261 markers to genotype *Lgh* samples and analyze genetic diversity in each sampled population
262 (primer sequences are openly listed in Stoeckel et al. 2023).
263

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

265 To genotype the 795 individually-tagged samples with these 36 SNPs, we extracted DNA from
266 25 to 30 mg of dried young leaf tissues. The DNA extractions were performed using the
267 NucleoSpin Plant II from MACHEREY-NAGEL kit, following the manufacturer's
268 recommendations. We used L1 buffer as lysis buffer. All individuals were genotyped from a
269 solution of 20ng/µl of *Lgh* DNA, using a modified version of the Hi-Plex protocol (Hammet et
270 al. 2019; Besnard et al. 2023). Hi-Plex is an amplicon sequencing technique (*sensu* Meek &
271 Larson 2019) in which all SNPs are co-amplified in a multiplex reaction before Illumina or Ion-
272 Torrent sequencing. Here, we used Illumina. Intermediate steps include dual indexing of
273 individual samples used for demultiplexing. Reads were then assigned to loci by aligning them
274 to reference sequences with BWA-MEM 0.7.15-r1140 (Li & Durbin 2010) and alleles were
275 counted with Samtools 1.9 (Danecek et al. 2021).

276

277 **Allele dosage**

278 The posterior probabilities of each single SNP genotype within each individual (hereafter
279 *single SNP genotype*) were computed using the likelihood of all possible genotypes considering
280 a multinomial distribution of the number of times each nucleobase was genotyped at one SNP
281 locus within one individual. In order to obtain the most confident dataset possible, we only

282 assigned a genotype when its posterior probability of allele dosage exceeded 70%. When one
283 SNP within one individual presented a posterior probability of allele dosage equal or lower
284 than 70%, we labelled and analysed it afterward as a missing genotype.

285

286 **Genotypic and genetic descriptors**

287 We expected that populations reproducing clonally may yield repeated multi-locus genotypes
288 (MLGs, i.e., the same genotype at all the 36 SNPs found in multiple individuals). By possibly
289 producing these repeated genotypes and by varying the relative distribution of the number of
290 samples of each of these distinct genotypes, rates of clonality impact genotypic richness and
291 evenness in populations (Halkett et al. 2005; Arnaud-Haond et al. 2007).
292 We measured genotypic richness using the R index (Dorken & Eckert 2001, Arnaud-Haond et
293 al. 2005), which is defined as:

294 $$R = \frac{G-1}{N_g-1}$$

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

299 $$N \geq X = aX - \beta$$

300 where N≥X is the number of sampled ramets belonging to genets containing X or more ramets
301 in the sample of the population studied, and the parameters *a* and *β* are fitted by regression
302 analysis. Low β-values indicate the dominance of one or few clonal lineages. Pareto β is less
303 biased by the sampling effort than clonal richness R (Stoeckel et al. 2021a). In our sample of
304 15 individuals, we expected a population with only one repeated MLG to present a Pareto β

305 value of 0.03 and a value of 3 in samples with no repeated genotypes. $\beta < 2$ indicates a
306 population reproducing with high rates of clonality (greater than 0.8 to 0.9). For a sample size
307 of 15 individuals, $2 < \beta < 3$ indicates intermediate to low rates of clonality (0.6 to 0.8). A β at
308 3, its maximum value, is indicative of a mainly sexual population, with rates of clonality ranging
309 from zero to 0.6, depending on the other genetic indices.
310 Rates of clonality and selfing confer different effects on the range of within-population genetic
311 polymorphism as well as on probabilities of genetic identities within individuals. We thus also
312 estimated expected and observed heterozygosity H_E and H_O, allelic richness A_E, inbreeding
313 coefficient F_{IS}, and linkage disequilibrium within each sampled population.
314 The Wright (1931, 1949) inbreeding coefficient F_{IS} accounts for intraindividual genetic
315 variation at one locus as a departure from Hardy-Weinberg assumptions of the genotyped
316 populations. We computed one F_{IS} value per locus per population. We then reported the mean
317 values of F_{IS} (MF_{IS}) and the inter-locus variance of F_{IS} ($VarF_{IS}$) for each population. Both MF_{IS}
318 and $VarF_{IS}$ are very informative about the underlying reproductive systems (Stoeckel et al.
319 2006, David et al. 2007). Clonality makes the MF_{IS} values tend toward -1 along with high
320 interlocus variance (Stoeckel & Masson 2014, Stoeckel et al. 2021a). A moderate amount of
321 sexual reproduction results in MF_{IS} values around 0 (Balloux et al. 2003). $VarF_{IS}$ varies with
322 rates of clonality, from very limited variance expected in sexual populations to high variance
323 in very clonal populations (Stoeckel & Masson 2014, Stoeckel et al. 2021a). Positive MF_{IS}
324 values are expected in populations reproducing using consanguinity and selfing (Castric et al.
325 2002, David et al. 2007). All sexual reproductive modes, allogamous and autogamous, result
326 in low $VarF_{IS}$ within a population as massive recombination tends to homogenize intra-
327 individual genetic identities along the genomes (Stoeckel et al. 2021a).

328 We measured linkage disequilibrium over all markers using the unbiased multilocus linkage
329 disequilibrium index, $\bar{r}_d$ (Agapow & Burt 2001). This mean correlation coefficient (*r*) of genetic
330 distances (*d*) between unordered alleles at *n* loci ranges from 0 to 1. This metric has the
331 advantage of limiting the dependency of the correlation coefficient on the number of alleles
332 and loci, and it is well suited to measure linkage disequilibrium in partially clonal populations
333 (De Meeus & Balloux 2004, De Meeus et al. 2006). In general, LD is only slightly affected by
334 clonality, except when clonality is high (c > 0.9) and/or when genetic drift dominates over
335 mutation rate (*e.g.,* when population sizes are small N<50 compared to u=0.001; Navascues
336 et al. 2010, Stoeckel et al. 2021a). In contrast, inbreeding and selfing are efficient processes
337 for quickly generating strong LD, after only few generations. Finally, we measured genetic
338 differentiation between populations of LSI, between populations of SC and between LSI and
339 SC populations using ρ_{ST}, an index adapted to study autopolyploid populations (Ronfort et al.
340 1998). All these indices were computed using GENAPOPOP (Stoeckel et al. 2024), software
341 dedicated to analysing genetic diversity and differentiation in autopolyploid populations
342 genotyped with confident allele dosage and reproducing through all possible rates of clonality
343 and selfing.
344 We also used SPAGEDI (V1.5, Hardy & Vekemans 2002) to estimate rates of selfing within
345 autopolyploid populations genotyped with confident allele dosage. This approach infers rate
346 of selfing (Sg) from identity disequilibrium coefficients (g_{2z} estimator), assuming that
347 populations reproduce by self-fertilization and random mating and are at inbreeding
348 equilibrium (David et al. 2007 for diploids, Hardy 2016 for autopolyploids). Identity
349 disequilibrium coefficients are measured as the correlation in heterozygosity of distinct loci
350 within the genome, and present the advantage of being more robust to null alleles and
351 genotyping errors than raw F_{IS} (David et al., 2007). The effect of partial clonality on g_{2z}

352 estimator is not yet defined. We thus considered for this study that clonality would only
353 marginally impact identity disequilibrium coefficients, the possibility to be at the inbreeding
354 equilibrium and the corresponding estimates of selfing. A synthesis of all these hypotheses
355 based on previous knowledge is provided in Table SI1.
356
357 **Ranking populations by clonality and selfing**
358 We proposed and computed a synthesis index ($\Sigma\mathrm{clon}$) calculated from Pareto β and $\mathrm{VarF_{IS}}$ to
359 rank the studied populations from the less to the more clonal. These two population genetic
360 indices are known to vary with rates of clonality being unbiased in samples (Stoeckel et al.
361 2021a). To avoid issues of scaling as the range of values of these descriptors are different by
362 several orders of magnitude, $\Sigma\mathrm{clon}$ is computed in each population i as the sum of Pareto $\beta(i)$
363 and $\mathrm{VarFIS}(i)$ values that were previously normalized over the whole dataset, respectively
364 $z_\beta(i)$ and $z_{\mathrm{VarFIS}}(i)$.

365 $$z_\beta(i) = 1 - \frac{\beta(i) - min(\mathrm{B})}{max(\mathrm{B}) - min(\mathrm{B})}$$

366 $$z_{\mathrm{VarFIS}}(i) = \frac{\mathrm{VarFIS}(i) - min(\mathrm{VF})}{max(\mathrm{VF}) - min(\mathrm{VF})}$$

367 $$\Sigma\mathrm{clon}(i) = {}^{1}/_{2}\left[z_\beta(i) + z_{\mathrm{VarFIS}}(i)\right]$$

368 where $\beta(i)$ and $\mathrm{VarFIS}(i)$ are the measured Pareto β and variance of $\mathrm{F_{IS}}$ over loci in the
369 population i, and where $\mathrm{B} = \{\beta(1), \beta(2), \dots, \beta(55)\}$ and $VF =$
370 $\{\mathrm{VarFIS}(1), \mathrm{VarFIS}(2), \dots, \mathrm{VarFIS}(55)\}$ are the respective sets of Pareto β and $\mathrm{VarF_{IS}}$ over loci
371 in the 53 populations used to obtain their minimum (min) and maximum (max) values.
372 Using the same approach as for $\Sigma\mathrm{clon}$, we computed a synthesis index ($\Sigma\mathrm{self}$) calculated from
373 $\bar{r}_d$, $\mathrm{MF_{IS}}$ and Sg to rank the studied populations from the less to the more selfed.

374 $$\Sigma\mathrm{self}(i) = {}^{1}/_{3}\left[z_{\bar{r}_d}(i) + z_{\mathrm{MFIS}}(i) + z_{\mathrm{Sg}}(i)\right]$$

375

376 **Statistical data analysis**

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

389 We analyzed the relationship between genetic diversity indices and their relationships with
390 $\Sigma\mathrm{clon}$ and $\Sigma\mathrm{self}$ measured in the 53 sampled populations using Kendal partial rank-order
391 correlation tests. We reported the corresponding matrices of correlation.

392 To detect differences in distribution of population genetic indices, $\Sigma\mathrm{clon}$ and $\Sigma\mathrm{self}$, among
393 floral morphs, we computed non-parametric Kruskal-Wallis tests that do not make any
394 assumptions about the type of distribution and about homogeneity of variances between the
395 tested distributions. When needed, we used post-hoc pairwise tests for multiple comparisons
396 of mean rank sums (Nemenyi's test).

All statistical tests were computed using PYTHON 3.11, SCIPY.STATS 1.9.3 (Virtanen et al. 2020) and SCIKIT-POSTHOCS (Terpilowski 2019), except the PCA that was performed using R V4.2.2 and the library FACTOMINER (Lê et al. 2008).

400 # Results

401 Among the 53 populations, we found 40 populations with only L-morph individuals and 13
402 populations with only S-morph individuals (Table SI2, Figure SI1). The 36 SNP markers we
403 developed for *Lgh* using RAD-Seq and used to genotype the 795 sampled individuals via Hi-
404 Plex allowed confident allele dosage and sufficient statistical power to identify repeated
405 genotypes due to clonal reproduction and to assess the respective proportions of clonality,
406 selfing and allogamy in each population.

407

408 **Allele dosage and missing genotypes**

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

415

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

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

423 (median=0.2).

424 These 36 SNPs in the autotetraploid part of *Lgh* would theoretically allow $4^{36}=4.7\times10^{21}$
425 different possible MLGs considering the four possible nucleobases and $2^{36}=6.9\times10^{10}$ different
426 possible MLGs assuming two possible nucleobases per locus. Considering allele frequencies in
427 the sampled populations, the probabilities of identities under panmixia ranged from $8.5\times10^{-}$
428 ^{12} to 5.2×10^{-5} (median=2.1×10^{-7}) and the unbiased probabilities of identity between sibs $P_{ID\text{-}SIB}$
429 ranged from 4.2×10^{-6} to 8.3×10^{-3} (median=5.6×10^{-4}; Table SI2). We then considered that the
430 SNP set we used showed sufficient statistical power to distinguish between true MLGs, and
431 that individuals with identical MLGs were true repeated genotypes (ramets) of a clonal lineage
432 (a genet).

433

434 **Genetic and genotypic diversity**

435 Across populations, we identified a total of 462 distinct MLGs (genets) within the 795 sampled
436 individuals genotyped. Among them, we found 404 genets (88%) with a single ramet and only
437 58 genets (12%) with more than two ramets (Table SI4). Forty-eight genets had two to seven
438 ramets distributed over one to six populations (median=2), seven genets with 10 to 33 ramets
439 distributed over two to 17 populations (median=9), and one large genet of 99 individuals
440 distributed over 24 populations (Figure 1).

441 Within populations, we found from three to 15 different genets (median=12) per population
442 among the 15 sampled individuals (Table SI4), implying the clear occurrence of repeated
443 genotypes but also a wide diversity of genets within most populations (Figure 1). Accordingly,
444 genotypic richness (R) ranged from 0.14 to 1 (median=0.79). The genotypic evenness, Pareto
445 β, ranged from 0.056 to 3 (median=1.478, Table SI2).

446 Observed heterozygosity H_O was also high in most populations, slightly above expected
447 heterozygosity, ranging from 0.13 to 0.38 (median=0.26). The mean standard inbreeding
448 coefficients (MF_{IS}) averaged over all genotyped loci within populations were negative in all 53
449 populations and ranged from -0.33 to -0.126 with a very negative median of -0.274 (Table SI2).
450 Variances of F_{IS} between loci within population were very high, ranging from 1.75 to 36.60
451 (median=27.57). All these measures argued for reproductive modes implying high rates of
452 clonality.
453 Estimates of selfing (S_g) were overall low with a handful of high values, ranging from 0 to 0.61
454 (median=0.06). Fifteen populations (28%) were estimated with no selfing. Nineteen
455 populations (36%) showed non-zero estimates under 0.1, 13 populations (25%) between 0.1
456 and 0.29, and six populations (11%) with estimates between 0.42 and 0.61.
457 Finally, linkage disequilibria within populations between genotyped SNPs were overall low,
458 with r_d values ranging from 0 to 0.51 with a median value of 0.12. Forty-two populations (79%)
459 were under 0.25, and only eight populations (21%) showed linkage disequilibrium between
460 0.25 and 0.51.
461
462 **Analyses of covariations between population genetic indices**
463 The first two components of the principal component analysis on the values of genetic
464 diversity found in the 53 genotyped populations accounted for 78.4% of the total variance
465 between populations (Figure SI2). Non-parametric Kendall partial rank-order correlations
466 between genetic indices (Table SI5) and correlations on the first two principal components
467 from the 17 population genetic indices measured showed three non-collinear groups of
468 associated genetic indices (Figure SI1.A) that are very similar to the theoretical groups of
469 population genetic indices expected to covary with different rates of clonality, selfing and

470 outcrossing in autopolyploids (Stoeckel et al. 2024). A first cluster regrouped G, D*, R, Pareto
471 β and $VarF_{IS}$, indices that are known to be sensitive to clonality (Stoeckel et al. 2021a) but also
472 VarHe, VarHo, and VarAe. This cluster largely explains the first dimension (50.2% of the total
473 variance) of the PCA and was collinear to Σ_{CLON} (also see Figure SI1.B). As expected under
474 partial clonality, $VarF_{IS}$ was negatively correlated to genotypic diversity indices (R and β; Figure
475 SI3). The second cluster regrouped $P_{ID\text{-}SIB}$, $P_{ID\text{-}u}$, H_E, A_E and H_O, indices that are linked to the
476 general genetic diversity of populations (Figure SI1). The third cluster regrouped S_g, SE.Sg, MF_{IS}
477 and r_D, indices that are usually used to identify, rank and estimate rates of selfing versus
478 outcrossing in sexual populations (Castric et al. 2002, Bürkly et al. 2017). The second
479 dimension of the PCA (23.4% of the total variance) was mostly correlated to Ho, Sg, He and r_D
480 which was collinear to Σ_{SELF} (Figure SI1.C). The clusters of genetic indices were corroborated
481 by the correlation between Σ_{CLON} and genetic indices (Figure SI2) and Σ_{SELF} and genetic indices
482 (Figure SI4).
483 The correlation across populations between Σ_{CLON} and Σ_{SELF} was negative and highly significant
484 (r_s = -0.66; $p < 0.001$; Figure 2). This correlation within LSI populations was negative and highly
485 significant (r_s = -0.65; $p < 0.001$) while it was nonsignificant in SC populations (r_s = -0.32; p
486 =0.289).
487

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

489 The number of ramets per genet was similar in LSI and SC populations (H=1.85, p=0.173,
490 Figure 1 & SI5), as were their genotypic richness (R, H=1.23, p=0.267) and evenness (Pareto's
491 β, H=1.60, p=0.206; Table SI2). The distributions of other population genetic indices related to

492 clonality were also not significantly different between LSI and SC populations ($VarF_{IS}$: H=2.94,
493 p=0.086; $P_{ID\text{-}SIB}$: H=1.98, p=0.160 and Σ_{CLON} H=2.28, p=0.131; Figure 3).
494 Mean observed heterozygosity (Ho, H=5.27, p=0.022) and their variances (VarHo, H=13.85,
495 p<0.001), mean effective number of alleles per locus (Ae, H=5.13, p=0.024) and its variance
496 over locus (VarAe, H=9.49, p=0.002), MF_{IS} (H=11.85, p<0.001), linkage disequilibrium (r_D,
497 H=5.55, p=0.018), estimate of selfing rates (Sg, H=12.39, p<0.001) and its standard error over
498 loci (SESg, H=9.58, p=0.002) significantly differed between LSI and SC populations. SC
499 populations, compared to LSI, showed lower mean observed heterozygosity (medians, LSI:
500 0.262, SC: 0.210) and less variance between loci (medians, LSI: 2.400, SC: 1.859), lower
501 effective number of alleles (medians, LSI: 1.341, SC: 1.283) and less variance between loci
502 (medians, LSI: 4.419, SC: 3.580), higher linkage disequilibrium (medians, LSI: 0.113, SC: 0.152),
503 higher estimates of selfing rate (medians, LSI: 0.045, SC: 0.175) even if with higher standard
504 error between loci (medians, LSI: 0.058, SC: 0.156) and higher Σ_{SELF} (medians, LSI: 0.222, SC:
505 0.495).
506 Genetic differentiation (ρ_{ST}) among pairs of SC populations (median value of ρ_{ST}=0.27) was
507 comparable to the genetic differentiation found among pairs of LSI populations (median value
508 of ρ_{ST}=0.26; H=319.37, post-hoc p=0.322; Figure 4). The median of ρ_{ST} between pairs of LSI and
509 SC populations reached 0.65. The distribution of inter-morph ρ_{ST} values differed both from the
510 distribution of pairs of intra-LSI ρ_{ST} values (post-hoc p<0.001) and from the distribution of pairs
511 of intra-SC ρ_{ST} values (post-hoc p<0.001). The minimum spanning tree of LSI and SC individuals
512 computed with *GenAPoPop* showed quite clustered distributions of LSI and SC individuals but
513 also with clear evidence of admixtures between LSI and SC lineages (Figure 5).
514

515 # Discussion

516 We used population genetics on autotetraploid SNPs with allele dosage to assess the rate of
517 clonality, of selfing and of outcrossing (i.e., reproductive mode) of 53 populations of *Ludwigia*
518 *hexapetala* subsp. *grandiflora* (*Lgh*) recently colonizing western European watersheds. To
519 achieve this goal, we developed reproducible codominant molecular markers (SNPs) that
520 enabled allele dosage-based genotyping of sampled individuals. We then used tailored
521 computational analyses to perform confident population genetic analyses in autopolyploids
522 (Dufresne et al. 2014). This framework can be applied to any autopolyploid species.
523 *Lgh* develops two floral morphs respectively associated with an active Late-acting self-
524 incompatible system (L-morph) and a self-compatible system (S-morph, Portillo-Lemus et al.
525 2021, 2022). Interestingly, the sampled populations in western Europe showed either L or S-
526 morph resulting in two groups of sampled populations: one group of 40 LSI (L-morph)
527 populations and one group of 13 SC (S-morph) populations. They live in close ecological
528 conditions (Portillo-Lemus et al. 2021) which thereby provides a rare opportunity to
529 characterize the genetic consequences of an LSI system compared to a group of SC populations
530 in the same species and ecological context.
531

532 **Nearly three quarters of the genetic variance is explained by different rates of clonality and**
533 **selfing among *Lgh* populations**

534 The correlations of population genetics indices on the 53 genotyped Lgh populations were
535 congruent with theoretical predictions on the variations of genetic indices with different rates
536 of clonality and selfing for autopolyploid populations (Stoeckel et al. 2024). Similarly, such
537 correlations were already predicted from Wright-Fisher-like models adapted for diploid

538 (Stoeckel et al. 2021a) and haplodiplontic life-cycles (Stoeckel et al. 2021b), and validated
539 respectively using multiple field populations of marine phanerogams for diploids (Arnaud-
540 Haond et al. 2020) and of brown, green and red algae and mosses for haplodiplontics (Krueger-
541 Hadfield et al. 2021). These predictions and observations for diploids, haplodiplontics and now
542 autopolyploids argue again for the primary importance of reproductive mode, especially of
543 high clonality, to drive variation in genetic diversity within populations (Duminil et al. 2007).
544 They support the methodological importance of accurate assessment of rates of clonality and
545 of selfing before starting to interpret genetic diversity with any other factor.
546

547 **Underestimated clonal diversity in the western European populations of *Lgh***

548 In our study, we chose to develop and use SNPs rather than AFLP markers. The latter can be
549 challenging in terms of precisely sizing fragments, leading to suboptimal reproducibility,
550 particularly across different platforms (Fry et al. 2009). These 36 SNPs are easily reproducible,
551 accurate enough to distinguish between offspring of sibling mating as evidenced by the
552 probabilities of identity we obtained, and less expensive in time and money. With these SNPs,
553 we found typical genetic signatures of high clonality, including clear occurrence of replicated
554 genotypes and mean negative F_{IS} values with high interlocus variances in all 53 western
555 European populations. Previous field and lab observations of peripatric *Lgh* populations
556 reported massive production of dispersal propagules, rapid expansions of patches and an
557 important capacity for spontaneous cutting-planting (Dandelot et al. 2005, Thouvenot et al.
558 2013, Grewell et al. 2016, Skaer-Thomason et al. 2018a, b). Okada et al. (2009) genotyped
559 around 800 individuals sampled in 27 *Lgh* populations in Californian wetlands using a set of
560 eight AFLP markers and reported an extremely reduced clonal diversity: 95% of the samples
561 had the same genotype and 18 populations over 27 (67%) supported a single AFLP-genotype.

562 All over the UK, only two haplotypes on 14 sampled stems were found using chloroplast
563 sequences (Armitage et al. 2013). These haplotypes were even shared with some *Lgh* samples
564 invading California. With no measure of the probability of identity to assess the marker set,
565 these results could be due to the lack of resolution of the markers used resulting in an
566 artificially elevated measure of clonality (Waits et al. 2001, Villate et al. 2010). In any case, all
567 these studies concluded that invasive populations of *Lgh* in Europe and in USA reproduce by
568 exclusive clonality, with a very narrow base of ancestral clones or being monoclonal (Dandelot
569 et al. 2005, Okada et al. 2009, Thouvenot et al. 2013, Grewell et al. 2016).
570 However, we found 462 distinct MLGs (genets) out of 795 individuals within these populations
571 and the large majority of these MLGs (404, thus 88%) were only sampled once in the 53
572 populations. Our results report a previously-underestimated genotype diversity in the invasive
573 populations in western Europe, yet congruent with Dandelot (2004) unpublished measures
574 obtained in three Mediterranean populations using inter-simple sequence repeats (ISSRs).
575 Dandelot (2004) reported three and seven genotypes over 11 sampling units within two *Lgh*
576 populations and two genotypes over nine sampling units in a third population, all sampled in
577 the south-east of France. These results could reveal contrasted situations between *Lgh*
578 populations introduced in the USA and continental Europe.
579 Among the 58 MLGs with replicates, seven MLGs were found in a single population and 51
580 MLGs were found in multiple sites. Clonal propagules of *Lgh* are known to disperse carried by
581 river current and by birds (zoochory: seedlings have the ability to stick to feathers; Grewel et
582 al. 2016). Even rare MLGs found in multiple sites are likely to blur the genetic structure among
583 *Lgh* populations if ignored as clonal dispersal. Preliminary detection of these MLGs distributed
584 among geographical locations will help to understand and interpret the invasive dynamics in
585 a subsequent study.

586

587 **Between 10 to 40% of sexual reproduction in *Lgh* populations in western Europe**

588 The large genotype diversity observed in western European populations may either result
589 from an initial or continuous introduction of a large clonal diversity that could have
590 propagated clonally in European watershed or be due to local sexual events mixing introduced
591 lineages into new local genotypes. Beyond the qualitative indication of clonality, we thus
592 aimed at estimating the quantitative rates of clonality versus sexuality using genetic indices in
593 this autotetraploid dataset.

594 Thirty-two populations showed Pareto β values of their distributions of clonal sizes under 2,
595 which are only found in theoretical populations with rates of clonality higher than 0.8 in
596 diploids, haplodiplontics and autopolyploids (Stoeckel et al. 2021a,b, 2024). All 53 populations
597 showed mean negative F_{IS} with high interlocus variance as expected in highly clonal
598 populations (Balloux et al. 2003, Stoeckel & Masson 2014, Reichel et al. 2016). These values
599 can be observed in theoretical populations with rates of clonality higher than 0.6 (Stoeckel et
600 al. 2021a, b, 2024). All 53 populations also showed small to medium values of linkage
601 disequilibrium between SNPs as expected in highly clonal populations with large population
602 sizes (Navascues et al. 2010, Stoeckel et al. 2021a). Globally, gene diversities measured in
603 these invasive *Lgh* populations were in the higher range of values commonly found for SNPs
604 (Fischer et al. 2017, Schmidt et al. 2021). Such levels of gene diversity are in line with strong
605 rates of clonality that buffer the loss of alleles due to genetic drift (Reichel et al. 2016, Becheler
606 et al. 2020, Stoeckel et al. 2021b). All these values of population genetic indices were
607 consistent with the interpretation that *Lgh* populations in western Europe must reproduce
608 with effective rates of clonality between 60% and 90% (De Meeus et al. 2006, Arnaud-Haond
609 et al. 2020, Stoeckel et al. 2021a, b), even if rates of clonality were variable within this range

610 among populations. Clonal reproduction by rhizomes at the local scale and, in a lesser
611 proportion, by the dispersal of clonal propagules at larger (*i.e.*, inter-locations) scales would
612 thus be the main source of population growth and invasive spread of *Lgh* out of its native
613 range.
614 Despite being predominant, clonality is, however, not the exclusive reproductive mode. The
615 measures of population genetic indices fit with 10% to 40% of sexuality (DeMeeus et al 2006,
616 Stoeckel et al. 2021a,b, Stoeckel et al. 2024). Estimates of the rates of sexuality could be
617 refined using a more resolutive method like ClonEstiMate using information from temporal
618 genotypings (Becheler et al. 2017). But our population genetic measures clearly rejected
619 exclusive clonality, even starting recently.
620 We also found local evidence of recombination between clonal lineages, and even rare crosses
621 between LSI and SC lineages as advocated by the genetic distances between individuals (Figure
622 5). The diversity of genotypes detected in western Europe would thus rather match rare but
623 significant sexual events than a large clonal diversity that would have maintain and propagate
624 by exclusive clonality since its introduction. Our results newly advocate that sexual
625 reproduction should not, therefore, be overlooked in these invasive populations, especially in
626 management plans.
627
628 **Genetic consequences of LSI compared to self-compatible populations**
629 The late-acting SI system remains one of the less studied breeding systems among the
630 mechanisms favouring outcrossing in plants (Gibbs 2014). Its efficiency to favour outcrossing
631 and its consequences on genetic diversity within populations, especially considering its low
632 but common failures, were not yet deciphered and not yet compared to SC populations in the

633 same ecological conditions, as already previously explored for gametophytic and sporophyte
634 SI systems (Busch 2005, Koelling et al. 2011).
635 The maintenance of SI systems is one of the most intriguing evolutionary puzzles (Porcher &
636 Lande 2005). Indeed, SI systems are fated to breakdown because SC individuals present the
637 advantage of reproductive assurance when compatible partners are limited, especially in
638 peripatric conditions (Eckert et al. 2006). This advantage is even absolute when no compatible
639 partners are available within pollination range, as we commonly found in *Lgh* populations in
640 western Europe (see Figure SI1). Conversely, outcrossing imposed by SI systems decreases the
641 probability of expressing deleterious mutations in descendants as compared to selfing with
642 the same genetic background (Rice 2004, Navascues et al. 2010).
643 We found that the 40 *Lgh* populations with L-morph individuals (LSI) showed relatively more
644 effective alleles, higher gene diversity, higher observed heterozygosity and less linkage
645 disequilibrium than found in the 13 populations with S-morph individuals (SC). However, these
646 genetic differences may either be due to the consequences of outcrossing in LSI populations,
647 or due to different rates of clonality, with potentially higher rates of clonality in LSI populations
648 due to the local lack of compatible partners and reduced female fertility due to self-pollen
649 disabling the available ovules, or due to the consequences of autogamy in SC populations. We
650 develop methods to differentiate these three potential mechanisms.
651
652 **Origin of the higher genetic diversity in LSI populations as compared to SC**
653 A first hypothesis to explain the genetic differences found between LSI and SC populations
654 may be due to the outcrossing consequences on LSI populations. Indeed, similar genetic
655 differences between SC and SI were recurrently reported in populations mating with
656 sporophytic and gametophytic SI systems (Igic et al. 2008, Mable 2008). For example,

657 *Leavenworthia alabamica* and *L. crassa* are two sister-species mating with a sporophytic SI
658 within a small genus of eight annual species in which self-fertilization has evolved at least four
659 times (Beck et al. 2006). In these species, comparisons between SI and SC populations also
660 showed higher mean genetic diversity in SI populations as compared to the SC, in an overall
661 background of high genetic diversity (Busch 2005, Koelling et al. 2011). High Ho, negative Fis
662 and high effective number of alleles per marker are also found in species mating with
663 gametophytic SI systems, such as in the *Prunus* genius (Stoeckel et al. 2006, Stoeckel et al.
664 2008, Abdallah et al. 2019). However, theoretical predictions argue that this kind of higher
665 genetic diversity favored and maintained in SI populations as compared to panmictic and SC
666 populations must only be found on polymorphic markers highly physically linked to the
667 genomic regions involved in the SI mechanisms that evolve under balancing selection, and
668 also, at a lesser magnitude, on physically unlinked markers but only in small population, with
669 less than 100 individuals including clones (Glémin et al. 2001, Navascues et al. 2010). The 36
670 SNPs we developed must randomly distribute over the 16 chromosomes inherited twice from
671 a common ancestor with *Lpm*, i.e., 32 chromosomes out of 80 total chromosomes of *Lgh*
672 (Barloy et al. 2024), with no reason that all aggregate around the genes involved in the LSI self-
673 recognition. In addition, individuals and populations of *Lpm* are reported to massively use self-
674 fertilization, thus with no SI system (Estes& Thorp 1974, Grewel et al. 2016). By parsimony,
675 there is less chance that genes involved in the LSI reaction would be stored on these
676 chromosomes than on some of the 48 other chromosomes; and therefore, very unlikely that
677 our SNPs would be physically linked to genes under balanced selection due to LSI. *Lgh*
678 populations in western Europe include far more than one thousand stems per local population
679 (Portillo-Lemus et al. 2021). It is thus unlikely that differences in genetic diversity observed

680 between LSI and SC populations would directly result from the consequences of the LSI system
681 hitchhiked to the whole genome (Navascues et al. 2010).
682 The second hypothesis is that high genetic diversity in LSI populations may arise and be
683 maintained due to higher rates of clonality (Stoeckel & Masson 2014, Reichel et al. 2016,
684 Stoeckel et al. 2021a) as compared to SC populations. In LSI, self-pollen tubes physiologically
685 blocked at the micropyle may prevent outcrossed-pollen tubes to reach the available ovules,
686 decreasing female fertility and increasing the importance of clonal regeneration in
687 populations (Vaughton et al. 2010). In addition, clonality is another uniparental reproductive
688 mode that allows the local maintenance of populations in the absence of sexually compatible
689 partners, which is more probable in peripatric conditions (Pannel et al. 2015). Dominant clonal
690 reproduction and propagation are commonly found in peripatric populations extending their
691 living range into new areas (Pyšek 1998, Wang et al. 2017). Furthermore, there is an
692 established intriguing co-occurrence of clonality in SI populations of flowering plants (Vallejo-
693 Marin et al. 2010). We know that dominant clonal reproduction theoretically favours the
694 accumulation of deleterious mutations and that these deleterious mutations contribute to the
695 advantage of outcrossing, which decreases the probability of inbreeding depression in
696 offsprings, as compared to SC reproducers (Navascues et al. 2010). The observed genetic
697 differences between LSI and SC populations may thus result from higher clonality in LSI
698 populations lacking of compatible partners. But we found here no difference in estimates of
699 clonality between LSI and SC populations, in Σ_{CLON} and in all indices sensible to clonality (R,
700 Pareto Beta and VarFis; Figure 3).
701 We rather found significant difference in estimates of selfing between LSI and SC populations,
702 in Sg (Hardy 2016), in Σ_{SELF} and in all indices sensible to selfing (Ae, linkage disequilibrium, Ho,
703 mean Fis) revealing typical signatures of higher selfing rates in SC populations. Using Hardy's

704 method (2016) extended from David et al. (2007) for autopolyploids, SC populations were
705 estimated to reproduce using a mean selfing rate of 0.18 while LSI populations with only a
706 mean selfing rate of 0.05, both estimated with similar rates of sexuality. Congruent with higher
707 estimates of selfing rates and Σ_{SELF} in SC populations, we found higher linkage disequilibrium,
708 lower observed heterozygosity and less effective number of alleles in SC populations (Figure
709 3). The observed differences in genetic diversity between LSI and SC populations must thus be
710 rather caused by the genetic consequences of higher selfing in SC populations, decreasing
711 their genetic diversity, than due to a direct outcrossing effect of the LSI hitchhiked over all the
712 *Lgh* genome within these large populations of *Lgh* (see Navascues et al. 2010). Interestingly,
713 the outcrossing breeding system allowed by LSI, with a small rate of selfing allowed especially
714 at the end of the flowering season, was measurable at the population level in genotyping (5%
715 of selfing in LSI versus 18% in SC) and thus may be efficient at an eco-evolutionary scale.
716 The recent arrival of *Lgh* in western Europe may also have contributed to shape these genetic
717 differences. Recent founding events in SC populations are indeed expected to decrease the
718 overall genetic diversity all along the genomes, including fewer effective alleles, lower gene
719 diversity, and potentially more linkage disequilibrium as compared to outcrossing populations
720 (Flanagan et al. 2021). Despite the ρ_{ST} values we observed between pairs of SI and SC
721 populations being approximately double those measured within pairs of SC populations and
722 within pairs of LSI populations, we also found clues of admixture in the genetic distances
723 between LSI and SC individuals in western Europe (Figure 5). This result would thus tend to
724 minimize the contribution of the foundation in the differences in genetic diversity we
725 observed between LSI and SC populations, and stress again the importance of selfing to shape
726 the genetic differences we found between SC and LSI populations.
727

728 **Advantages of partially clonal LSI populations over SC populations in *Lgh***

729 Interestingly, we found no correlation between estimates of rates of clonality and of selfing in
730 SC populations. Selfing rates rather decreased when outcrossing increased, while rates of
731 clonality remained similar in all populations.

732 In SC populations of *Lgh* (S-morph), pollen tubes of the L-morph (LSI type) grow significantly
733 faster in SC individuals (Portillo-Lemus et al. 2022). It seems a sufficient and efficient
734 mechanism to favour outcrossing when compatible pollens are available in western European
735 *Lgh* SC populations. The micro-physiological mechanism(s) impacting self-pollen tube growing
736 rather than blocking them in simultaneous (as opposed to dichogamy) monoecious or
737 hermaphrodite species seems to be currently overlooked in literature (e.g. Glover 2007,
738 Nasrallah 2017). Our result obtained on SC *Lgh* calls for a better appraisal of such mechanism
739 in species in which outcrossed offspring seems to be favoured without the occurrence of a
740 clear macro SI system.

741 Considering we found no differences in the estimated rates of clonality between LSI and SC
742 populations, our results thus argue for an overall stable background rate of clonal
743 propagation, common to the whole species in western Europe, and for an adjustment of the
744 mating system to the local availability of the compatible partners, shifting from selfing
745 allowing reproductive insurance, to outcrossing when possible. In such system, selfing
746 advantage, if any, would result from a transient reproductive insurance, decreasing the cost
747 of developing sexual organs when no outcrossed partners would be available at pollination
748 range, rather than as an efficient uniparental reproductive mode to expand into new areas
749 with no compatible partners, as compared to clonality. Heterozygosity rather than
750 recombination in *Lgh* would thus present some ecological and/or evolutive advantage. This
751 reproductive mode, dominant clonality with some allogamy, is commonly observed in

752 multiple plants (Vallejo-Marin et al. 2010). In such system, we should find higher vigour and
753 fertility of the outcrossed and most heterozygous individuals (heterosis) as compared to the
754 selfed and more homozygous individuals (Darwin 1876, Lippman & Zamir 2007, Birchleret al.
755 2010).
756 Reduced allele diversity, heterozygosity and increased linkage disequilibrium between
757 markers in SC populations due to the higher selfing rates as compared to LSI populations may
758 explain why only 24% of the western European populations were SC against 76% of LSI.
759 (Portillo et al. 2021). Based on an analysis of photo collected on the web, Portillo et al (2021)
760 reported that LSI populations of *Lgh* were more frequent in peripatric populations than in
761 natives in which SC populations seem in majority. Some successful invasive populations
762 develop with low genetic diversity (He et al. 2024), questioning on the biological and
763 environmental factors that may explain their successes in such peripatric conditions (called
764 the genetic paradox of invasions, Allendorf & Lundquist 2003). This seems to not be the case
765 of *Lgh* in western Europe: reproductive modes avoiding the effects of inbreeding depression
766 in *Lgh* populations may on the opposite facilitate their peripatric spreading, as commonly
767 observed in multiple other invasive plants (Roman & Darling 2007, Forsman 2014), mostly
768 when developing in harsh and stressful conditions (Fox & Reed, 2011) or when the costs of
769 inbreeding depression expressed by selfing are superior to the benefits of reproductive
770 insurance (Layman et al. 2017). Peripatric populations of *Lgh* in western Europe by mixing
771 clonality with preferential allogamy would thereby solve the problem of maintaining
772 populations locally without the presence of compatible or less related sexual partner while
773 reducing the probability to express inbreeding depression (uniparental clonal reproduction,
774 Baker's conjecture: Barrett et al. 2008, Pannel et al. 2015), and would maximize the probability
775 of developing offspring with hybrid vigour thanks to LSI and faster growth of crossed-pollen

776 tubes. In such peripatric conditions, selfing in SC populations would only have limited or
777 transient advantage to provide behedging reproductive insurance. To disentangle between
778 the genetic advantages of outcrossing versus reproductive insurance and better state if *Lgh*
779 populations belong to the genetic paradox of invasions genetic, our results call for measuring
780 the true proportion of LSI and SC in invasive versus native worldwide *Lgh* populations with
781 estimation of the rates of clonality and of selfing.
782
783 **Unusual selfing syndrome in *Lgh* populations**
784 Surprisingly, in *Lgh*, SC populations reproducing with higher selfing rates develop S-morph
785 flowers that are significantly larger than L-morph flowers that are allogamous controlled by
786 an LSI system (Portillo-Lemus et al. 2021). Yet, highly selfed populations of plant species tend
787 to share similar morphology and function resulting in a set of traits called the 'selfing
788 syndrome', including reduced flower size (Darwin 1876, Tsuchimatsu & Fujii 2022). In *Lgh*,
789 such high selfing rates may only have occurred recently and endemically in these isolated and
790 founding European invasive populations, explaining such anomaly trait. Indeed, pioneering
791 *Lgh* populations in western Europe only appeared during the last century (Dandelot 2004). But
792 selfing syndrome, including reduced flower size, seems to evolve rapidly, as observed for
793 example in five generations in *Mimulus guttatus* (Bodbyl-Roels & Kelly 2011), in four
794 generations in *Silene latifolia* (Delph et al. 2004), in three generations in *Phlox drummondii*
795 (Lendvai & Levin 2003) and after only two generations in *Eichhornia paniculata* (Worley &
796 Barrett 2000). However, in *Lgh*, SC populations with higher selfing rates also reproduce using
797 intermediate to high rates of clonality, which may delay or even impact the emergence of the
798 first steps of a selfing syndrome. Our results in *Lgh* thus call for further studies on the
799 emergence of selfing syndrome in partially clonal and selfed populations.

800

801 **Conclusion**

802 Using reproducible molecular markers enabling allele dosage to genotype sampled individuals
803 and tailored computational analyses that answer the remaining challenges to perform
804 population genetic analyses in autopolyploids (Dufresne et al. 2014), we found that peripatric
805 populations of *Lgh* in western Europe reproduced using dominant clonality with subtle but
806 significant sexuality, in nearly all populations, within a large clonal diversity.

807 *Lgh* is one of the most invasive aquatic plants in the world, and considerable efforts are made
808 to limit its deleterious effects on the newly colonized ecosystems and on their native species
809 (Thouvenot et al. 2013, Grewell et al. 2016, Portillo-Lemus et al. 2021). The rare sexual events,
810 allogamous when possible, occurring in invasive peripatric populations of *Lgh* may favour the
811 emergence of new genotypes, more adapted to local conditions. Managers should thus chiefly
812 concentrate their actions on the most sexual populations. We also found variations of the
813 rates of clonality and of selfing between *Lgh* populations. Considering the importance of
814 reproductive modes on the dynamics and evolution of populations (Duminil et al. 2007,
815 Ellegren& Galtier 2016, Glémin et al. 2019), our results advocate that management actions to
816 control *Lgh* should be built considering their effective reproductive modes, population by
817 population.

818 Finally, knowing the reproductive modes of populations, the distribution of clones and the
819 self-compatibility across western European populations will help to decipher and interpret
820 their current population structure in a next future, with the aim to identify the origin and the
821 routes of this recent invasion and to predict its possible short-term dynamics.

822

Data Accessibility Statement

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

Benefits Generated

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

Acknowledgements

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

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

Funding information

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

Conflict of interest

847 The authors declare that they have no financial conflicts of interest based on the content of
848 this article.
849
850 **Author contributions**
851 DB and SS laid the foundation of this work, conceived the study and were responsible for
852 funding applications. LPL, MH and DB collected samples and performed field observations. EP
853 and DB conceived and supervised RAD4SNP and Hi-Plex methodology, BJP and DJP contributed
854 to define the SNP set. LPL, MH and ALB performed RAD4SNP approach to develop the SNPs.
855 LPL and MH extracted DNA and genotyped the samples, SMC and RCV sequenced the data. GL
856 performed the bioinformatic analyses and SS computed the Bayesian allele dosage. DB, LPL
857 and SS performed early data explorations at the end of LPL PhD thesis that was supervised by
858 DB and SS. RB and SS performed computations and analyses, tracked and managed the
859 bibliography, produced tables, figures and interpretation. DB, RB and SS wrote the
860 manuscript. All authors read and approved the final manuscript.
861

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

887 Baduel, P., Bray, S., Vallejo-Marin, M., Kolář, F., & Yant, L. (2018). The "polyploid hop":
888 shifting challenges and opportunities over the evolutionary lifespan of genome
889 duplications. *Frontiers in Ecology and Evolution*, *6*.
890 https://www.frontiersin.org/articles/10.3389/fevo.2018.00117
891 Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E.
892 U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP Discovery and Genetic
893 Mapping Using Sequenced RAD Markers. *PLOS ONE*, 3(10), e3376.
894 https://doi.org/10.1371/journal.pone.0003376
895 Balloux, F., Lehmann, L., & de Meeûs, T. (2003). The population genetics of clonal and
896 partially clonal diploids. *Genetics*, *164*(4), 1635–1644.
897 https://doi.org/10.1093/genetics/164.4.1635
898 Barloy, D., Lemus, L. P.-, Krueger-Hadfield, S. A., Huteau, V., & Coriton, O. (2024). Genomic
899 relationships among diploid and polyploid species of the genus Ludwigia L. section
900 Jussiaea using a combination of cytogenetic, morphological, and crossing
901 investigations. *Peer Community in Evolutionary Biology*, 4, e8.
902 https://doi.org/10.24072/pci.evolbiol.100645.
903 Barrett, S. C. H., Colautti, R. I., & Eckert, C. G. (2008). Plant reproductive systems and
904 evolution during biological invasion. *Molecular Ecology*, 17(1), 373–383.
905 https://doi.org/10.1111/j.1365-294X.2007.03503.x
906 Barrett, S. C. H. (2010). Understanding plant reproductive diversity. *Philosophical*
907 *Transactions of the Royal Society B: Biological Sciences*, *365*(1537), 99–109.
908 https://doi.org/10.1098/rstb.2009.0199
909 Becheler, R., Guillemin, M.-L., Stoeckel, S., Mauger, S., Saunier, A., Brante, A., Destombe,
910 C., & Valero, M. (2020). After a catastrophe, a little bit of sex is better than nothing:
911 Genetic consequences of a major earthquake on asexual and sexual populations.

912 *Evolutionary Applications*, *13*(8), 2086–2100. https://doi.org/10.1111/eva.12967

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

917 Beck, J. B., Al-Shehbaz, I. A., & Schaal, B. A. (2006). *Leavenworthia* (Brassicaceae) revisited:
918 testing classic systematic and mating system hypotheses. *Systematic Botany*, *31*(1),
919 151–159. https://doi.org/10.1600/036364406775971732

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

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

926 Bodbyl Roels, S. A., & Kelly, J. K. (2011). Rapid evolution caused by pollinator loss in
927 *Mimulus guttatus*. *Evolution*, *65*(9), 2541–2552. https://doi.org/10.1111/j.1558-
928 5646.2011.01326.x

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

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

936 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., 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

962 Chen, X., Hao, S., Wang, L., Fang, W., Wang, Y., & Li, X. (2012). Late-acting self-
963 incompatibility in tea plant (*Camellia sinensis*). *Biologia*, *67*(2), 347–351.
964 https://doi.org/10.2478/s11756-012-0018-9
965 Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A.,
966 Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools
967 and BCFtools. GigaScience, 10, giab008. https://doi.org/10.1093/gigascience/giab008
968 Dandelot, S. (2004). *Les Ludwigia spp. Invasives du Sud de la France: Historique,*
969 *Biosystématique, Biologie et Ecologie* [PhD thesis, Aix-Marseille 3].
970 https://www.theses.fr/2004AIX30052
971 Dandelot, S., Verlaque, R., Dutartre, A., & Cazaubon, A. (2005). Ecological, Dynamic and
972 Taxonomic Problems Due to Ludwigia (Onagraceae) in France. Hydrobiologia, 551(1),
973 131–136. https://doi.org/10.1007/s10750-005-4455-0
974 Darwin, C. R. (1876). *The effects of cross and self-fertilization in the vegetable kingdom*. John
975 Murray (London).
976 David, P., Pujol, B., Viard, F., Castella, V., & Goudet, J. (2007). Reliable selfing rate estimates
977 from imperfect population genetic data. *Molecular Ecology*, *16*(12), 2474–2487.
978 https://doi.org/10.1111/j.1365-294X.2007.03330.x
979 De Meeûs, T., & Balloux, F. (2004). Clonal reproduction and linkage disequilibrium in
980 diploids: A simulation study. *Infection, Genetics and Evolution*, *4*(4), 345–351.
981 https://doi.org/10.1016/j.meegid.2004.05.002
982 De Meeûs, T., Lehmann, L., & Balloux, F. (2006). Molecular epidemiology of clonal diploids:
983 A quick overview and a short DIY (do it yourself) notice. *Infection, Genetics and*
984 *Evolution*, *6*(2), 163–170. https://doi.org/10.1016/j.meegid.2005.02.004
985 De Meeûs, T., Prugnolle, F., & Agnew, P. (2007). Asexual reproduction: Genetics and
986 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 cost-and-time effective procedure to develop SNP markers for multiple species: A support for community genetics. *Methods in Ecology and Evolution*, *9*(9), 1959–1974. https://doi.org/10.1111/2041-210X.13034

Delph, L. F., Gehring, J. L., Frey, F. M., Arntz, A. M., & Levri, M. (2004). Genetic constraints on floral evolution in a sexually dimorphic plant revealed by artificial selection. *Evolution*, *58*(9), 1936–1946. https://doi.org/10.1111/j.0014-3820.2004.tb00481.x

Dorken, M. E., & Eckert, C. G. (2001). Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodonverticillatus* (Lythraceae). *Journal of Ecology*, *89*(3), 339–350. https://doi.org/10.1046/j.1365-2745.2001.00558.x

Dufresne, F., Stift, M., Vergilino, R., & Mable, B. K. (2014). Recent progress and challenges in population genetics of polyploid organisms: An overview of current state-of-the-art molecular and statistical tools. *Molecular Ecology*, *23*(1), 40–69. https://doi.org/10.1111/mec.12581

Duminil, J., Fineschi, S., Hampe, A., Jordano, P., Salvini, D., Vendramin, G. G., & Petit, R. J. (2007). Can population genetic structure be predicted from life-history traits? *The*

American Naturalist, *169*(5), 662–672. https://doi.org/10.1086/513490

Eckert, C., Samis, K. & Dart, S. (2006). Reproductive Assurance and the Evolution of Uniparental Reproduction in Flowering Plants. In Harder & Barrett (Eds), *Ecology and Evolution of Flowers*, pp.183–200. Oxford, UK: Oxford University Press.

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), Article 7. https://doi.org/10.1038/nrg.2016.58

Estes, J. R., & Thorp, R. W. (1974). Pollination in Ludwigia peploides ssp. Glabrescens (Onagraceae). *Bulletin of the Torrey Botanical Club, 101*(5), 272–276. https://doi.org/10.2307/2484872

Fehrer, J. (2010). Unraveling the mysteries of reproduction. *Heredity*, *104*(5), Article 5. https://doi.org/10.1038/hdy.2010.12

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., 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.

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 pests. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *23*(5), 790–

803. https://doi.org/10.1002/aqc.2387

Tsuchimatsu, T., & Sota, F. (2022). The selfing syndrome and beyond: diverse evolutionary consequences of mating system transitions in plants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377, 20200510. https://doi.org/10.1098/rstb.2020.0510.

Vallejo-Marín, M., Dorken, M. E., & Barrett, S. C. H. (2010). The ecological and evolutionary consequences of clonality for plant mating. *Annual Review of Ecology, Evolution, and Systematics*, *41*(1), 193–213. https://doi.org/10.1146/annurev.ecolsys.110308.120258

Van Drunen, W. E., & Husband, B. C. (2019). Evolutionary associations between polyploidy, clonal reproduction, and perenniality in the angiosperms. *New Phytologist*, *224*(3), 1266–1277. https://doi.org/10.1111/nph.15999

Vaughton, G., Ramsey, M., & Johnson, S. D. (2010). Pollination and late-acting self-incompatibility in Cyrtanthus breviflorus (Amaryllidaceae): Implications for seed production. *Annals of Botany*, 106(4), 547–555. https://doi.org/10.1093/aob/mcq149

Villate, L., Esmenjaud, D., Van Helden, M., Stoeckel, S., & Plantard, O. (2010). Genetic signature of amphimixis allows for the detection and fine scale localization of sexual reproduction events in a mainly parthenogenetic nematode. *Molecular Ecology*, 19(5), 856–873. https://doi.org/10.1111/j.1365-294X.2009.04511.x

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., & van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), Article 3. https://doi.org/10.1038/s41592-019-0686-2

Waits, L. P., Luikart, G., & Taberlet, P. (2001). Estimating the probability of identity among

genotypes in natural populations: Cautions and guidelines. *Molecular Ecology*, 10(1), 249–256. 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

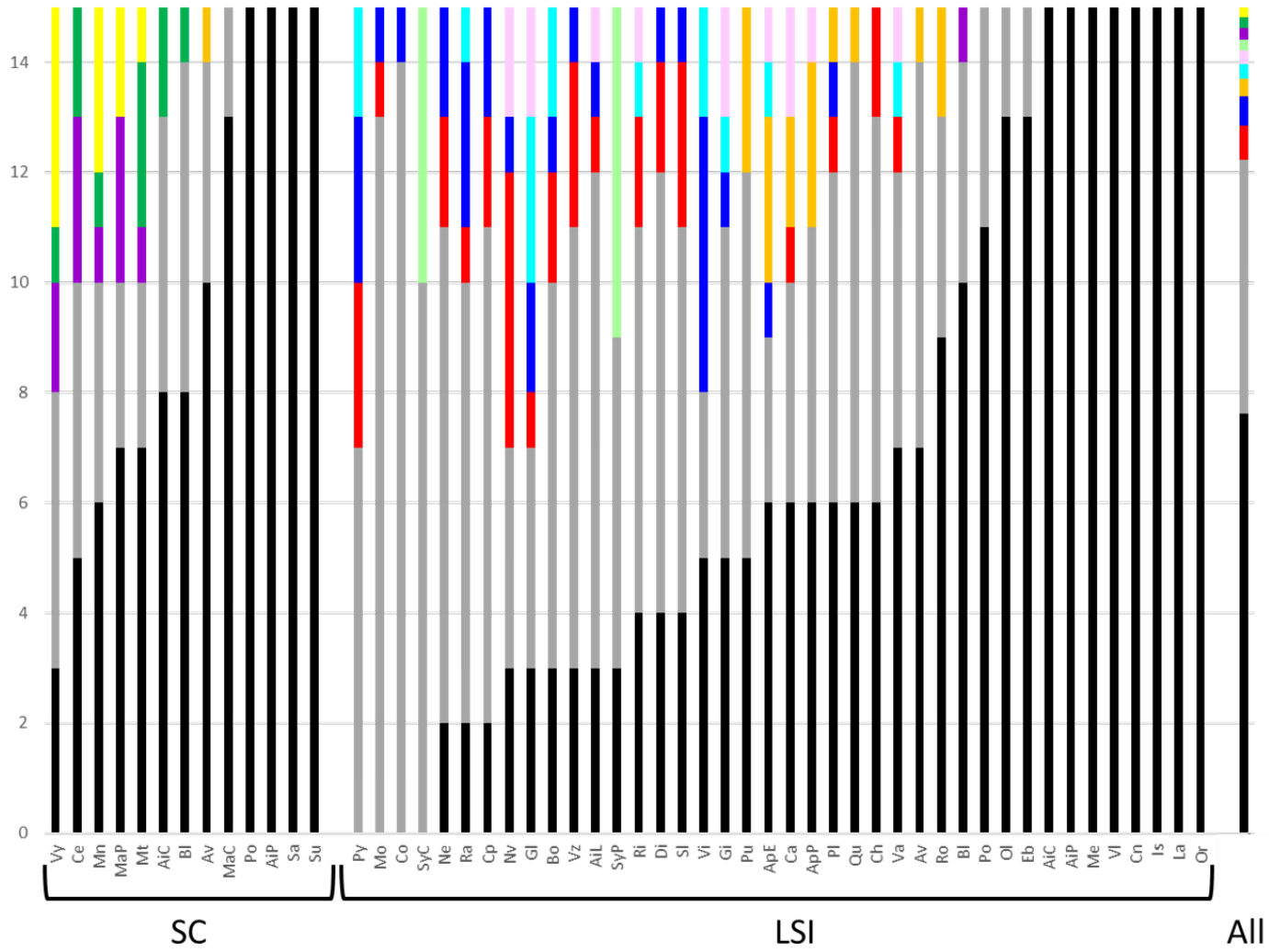

Figure 1: Distributions of Multi-Locus Genotypes (MLGs) among LSI and SC *Lgh* populations in western Europe. In black, the number of unique MLGs; In grey, the number of MLGs with less than 7 ramets; In color, MLGs with more than 10 ramets found in multiple populations. Proportions of each MLGs over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

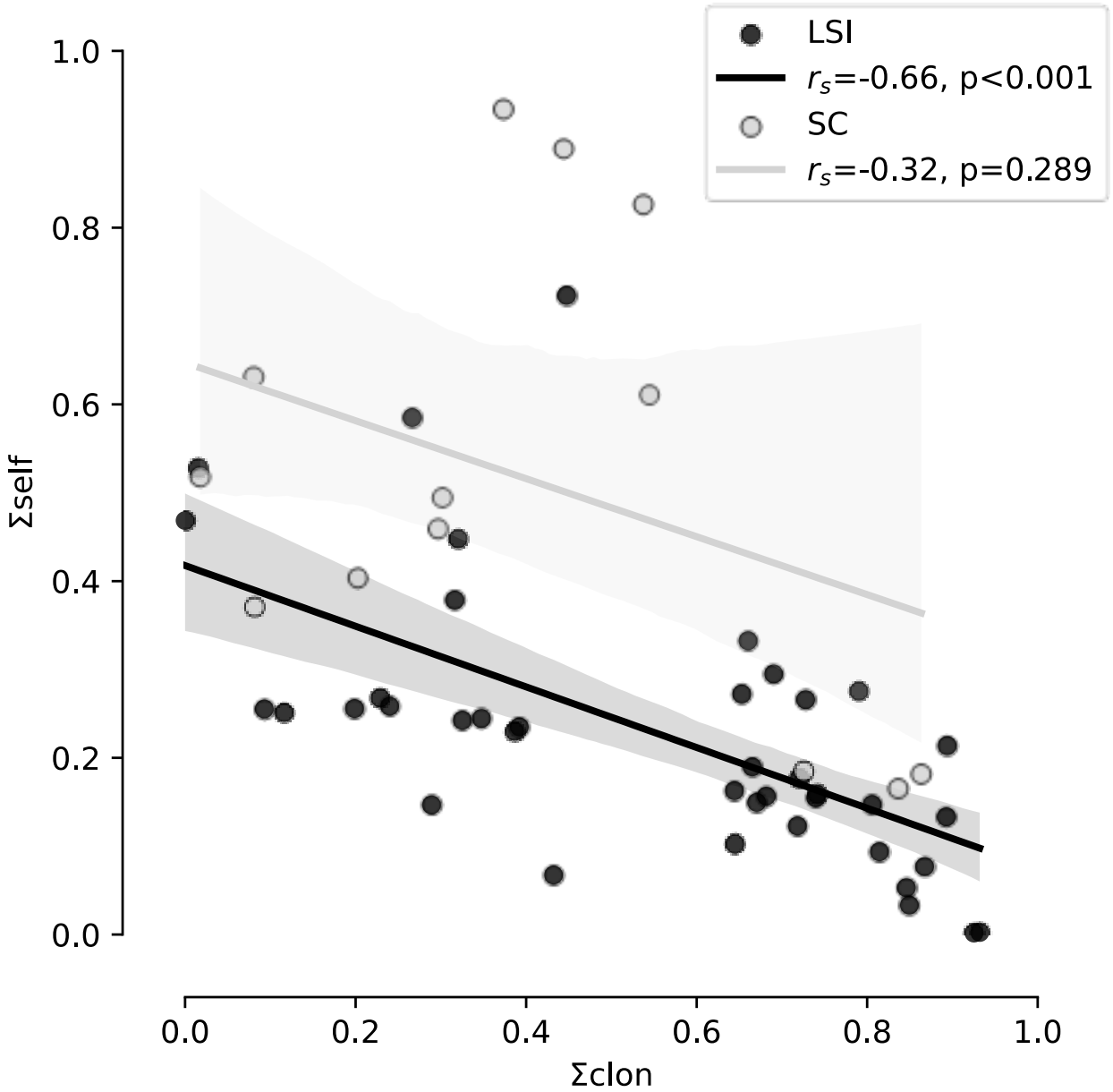

Figure 2: Correlations between Σ_{CLON} and Σ_{SELF} found in 13 SC populations (light grey points and regression line) and 40 LSI populations (black points and regression line). 95% confidence intervals are given for LSI and SC populations in dark grey and light grey, respectively. Spearman rank-order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

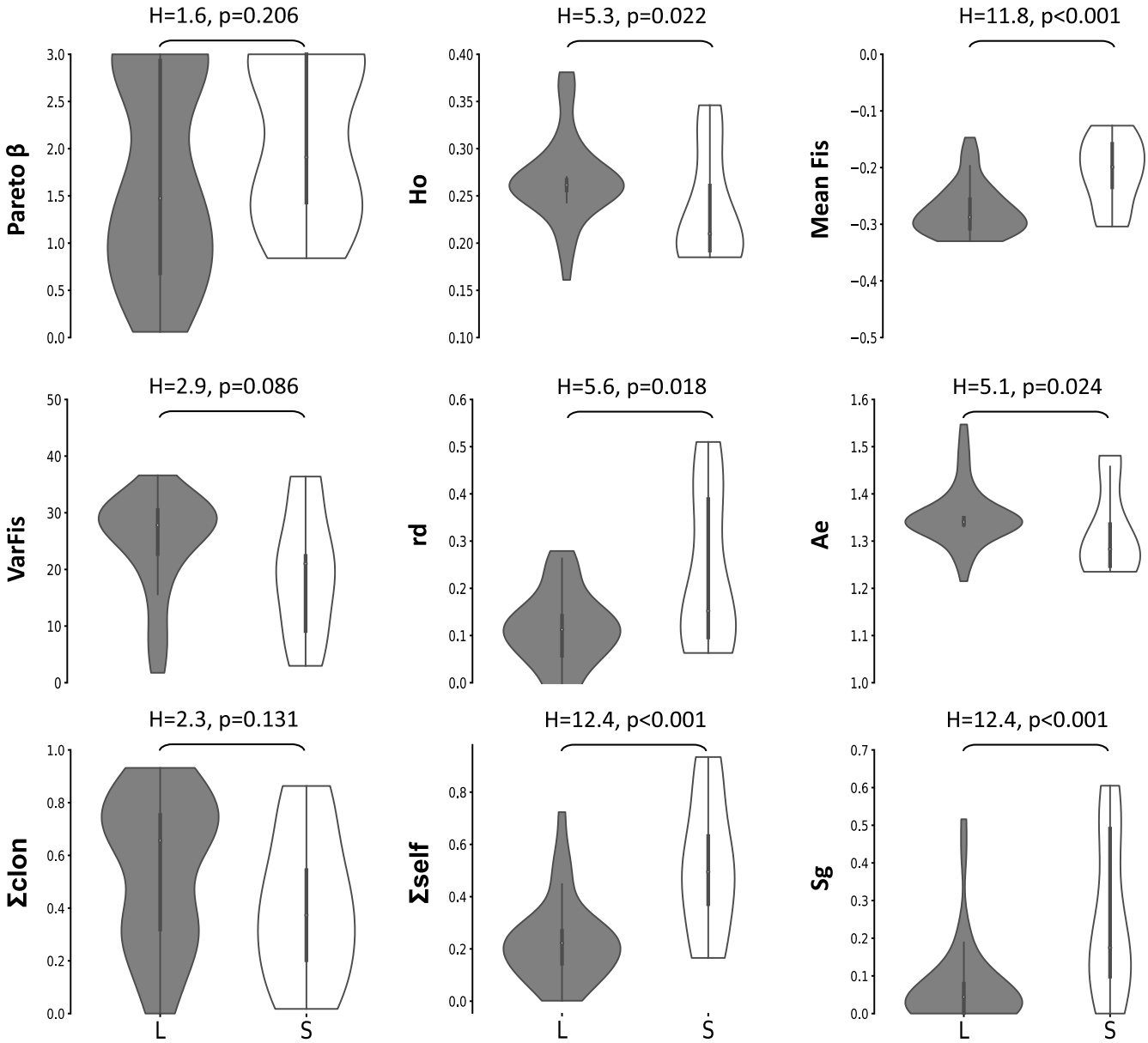

Figure 3: Violin plots comparing distributions of genetic indices in LSI (L, grey) versus SC (S, white) populations. Violin plots are cut for their minimum and maximum values. Kruskall-Wallis tests are reported as H-statistics as well as the probability that LSI and SC population would present the same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

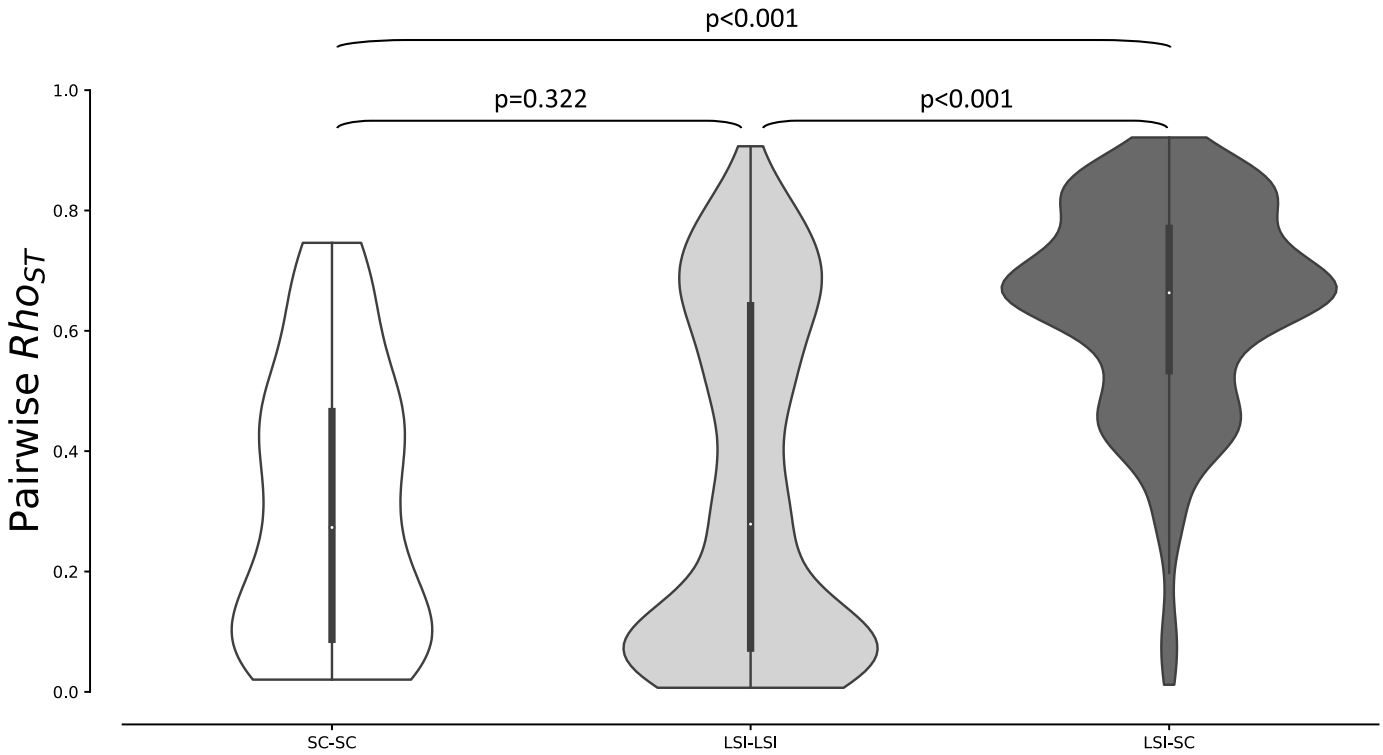

Figure 4: Distributions of pairwise rho_{st} within pairs of SC populations (white violin plot), pairs of LSI populations (light grey violin plot) and between pairs of LSI and SC populations (dark grey violin plot). Probabilities that pair of non-parametric distributions are equal are reported.

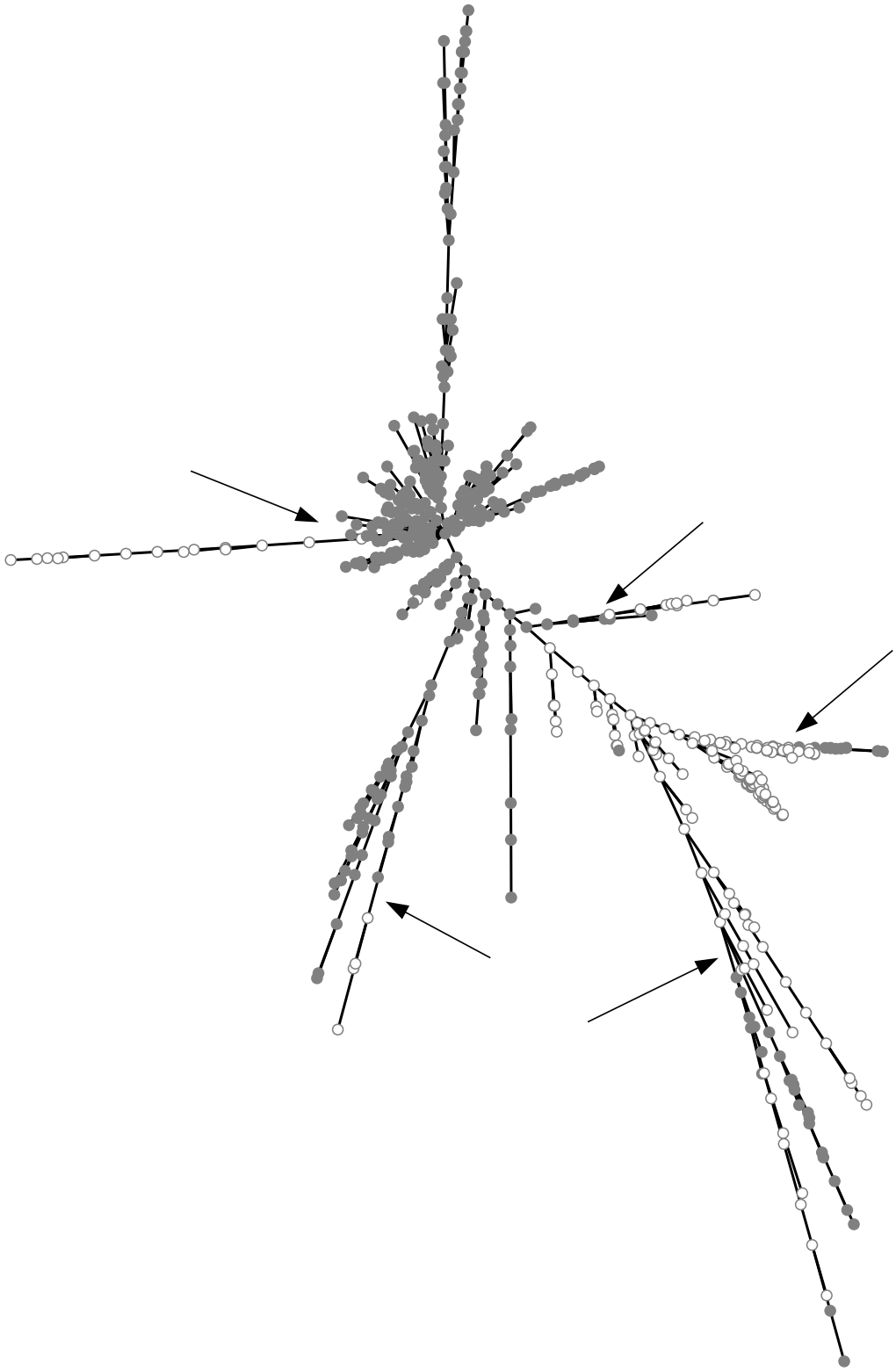

Figure 5: Minimum spanning tree of the genetic distances (number of different alleles) between LSI (grey points) and SC (white points) individuals. Arrows indicate some of the evidences of admixture between LSI and SC lineages.

1307 # Supplementary Information (SI)

1308

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 insurance 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?

1309

1310 **Supplementary information1:** Synthesis of the questionings of our study to estimate reproductive
1311 mode of 53 populations of invasive Lgh using genetic diversity on the autotetraploid part of its
1312 genome.

1313

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.

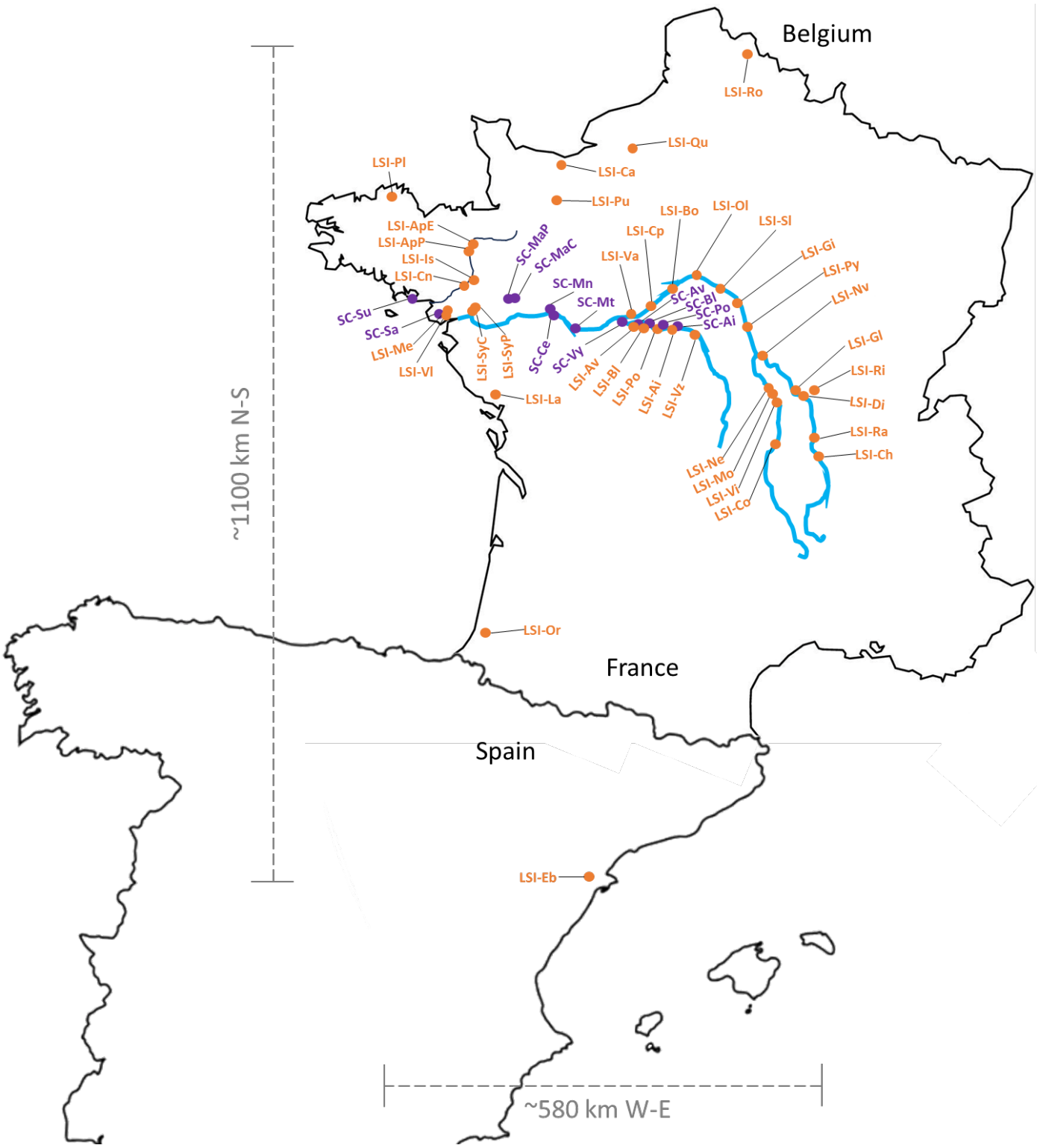

Figure SI1: Map of the locations of the 53 sampled populations across France and Spain. In orange, the late-acting self-incompatible (LSI) populations; In purple, the self-compatible (SC) populations. In light blue, the Loire river system. In grey, the North-to-South (N-S) and West-to-East (W-E) scale distances of the sampled populations.

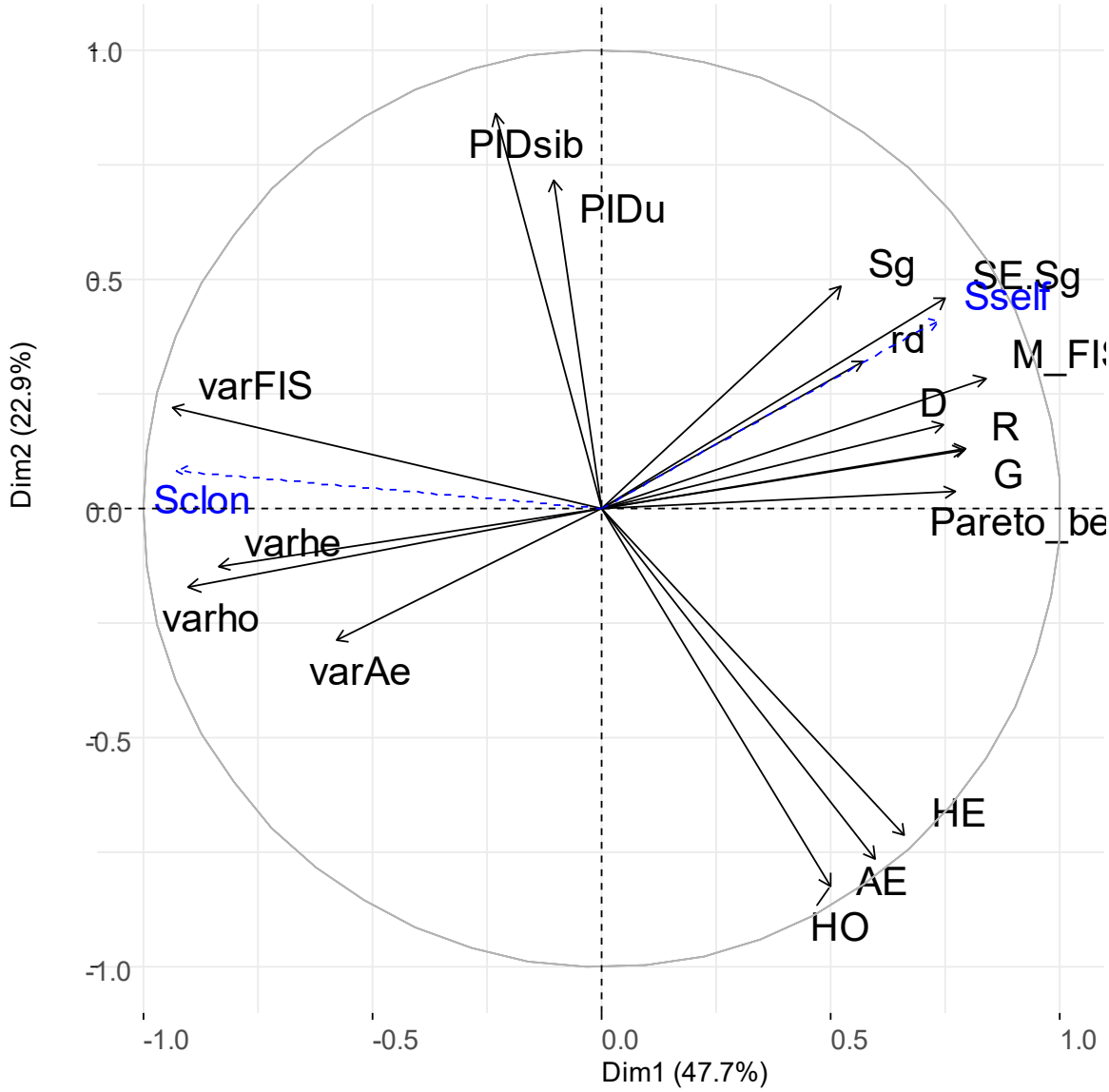

Figure SI2.A: Correlation circles of the principal component analysis on the 17 centered and scaled genetic indices measured in 53 *Lgh* populations in western Europe. The first horizontal dimension accounting for 50.24% of the variance regroups indices related to clonality (number of genotypes *G*, 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.

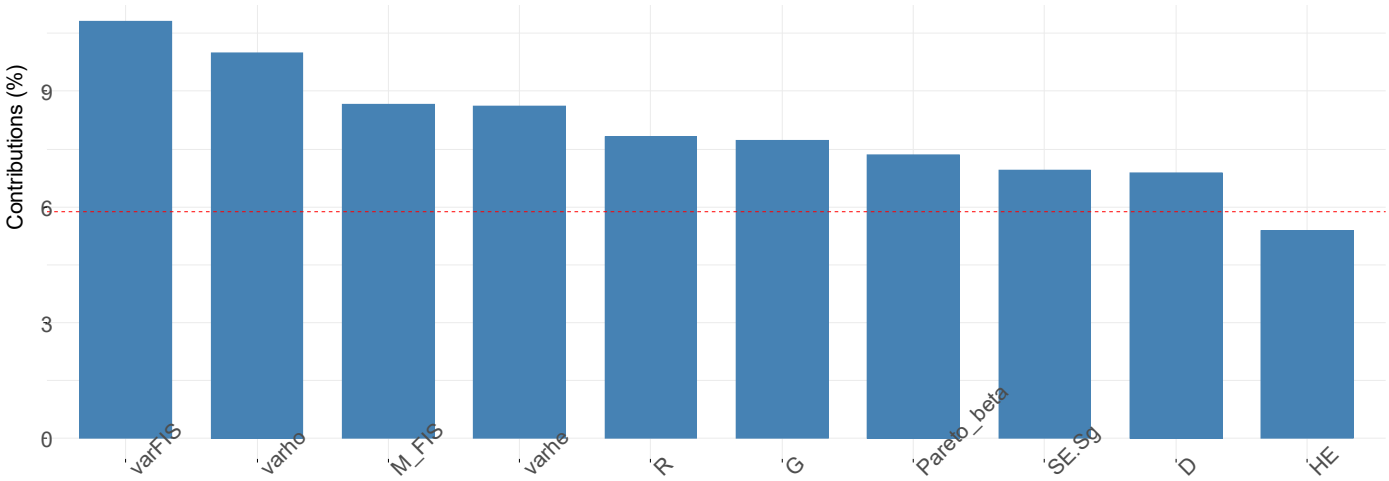

Figure SI2.B: Bar plot of the genetic index contributions to the first principal component. The expected average contribution *1/(number of genetic indices)* is plotted as the red dashed line. VarFis and indices of genotypic diversity (Pareto , R, G & D) are the main contributors to the first component.

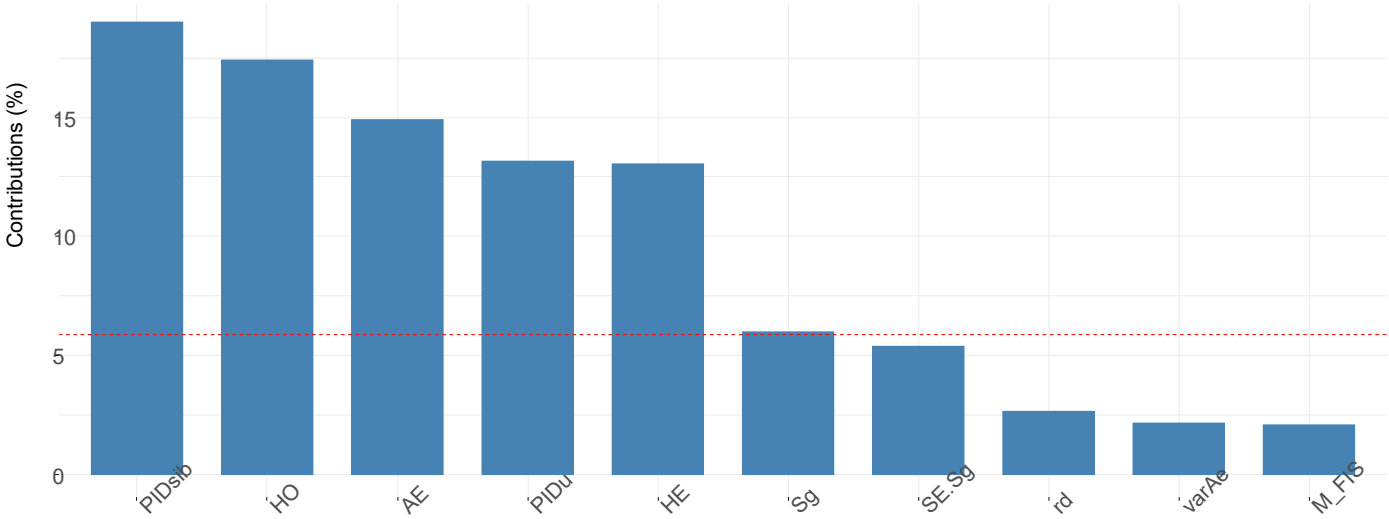

Figure SI2.C: Bar plot of the genetic index contributions to the second principal component. The expected average contribution *1/(number of genetic indices)* is plotted as the red dashed line. Estimates of selfing (Sg), mean observed heterozygosity (H_O), mean effective number of alleles (A_E) and probabilities of identities (pid_u and pid_{sib}) are the main contributors to the second component.

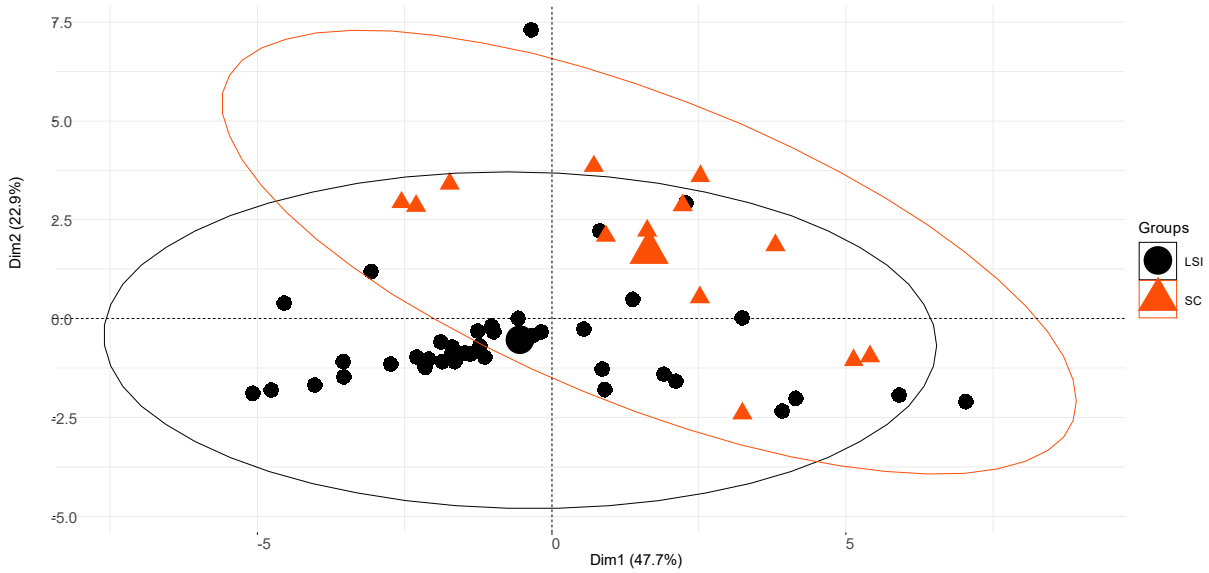

Figure SI2.D: Plot of the projection of each population onto the two first principal components. SC populations are plotted as red triangles and LSI populations as black diamonds. Concentration ellipses around LSI and SC populations are plotted in black and red, respectively. Barycenters of the two groups are represented as bigger black diamond and red triangle, respectively.

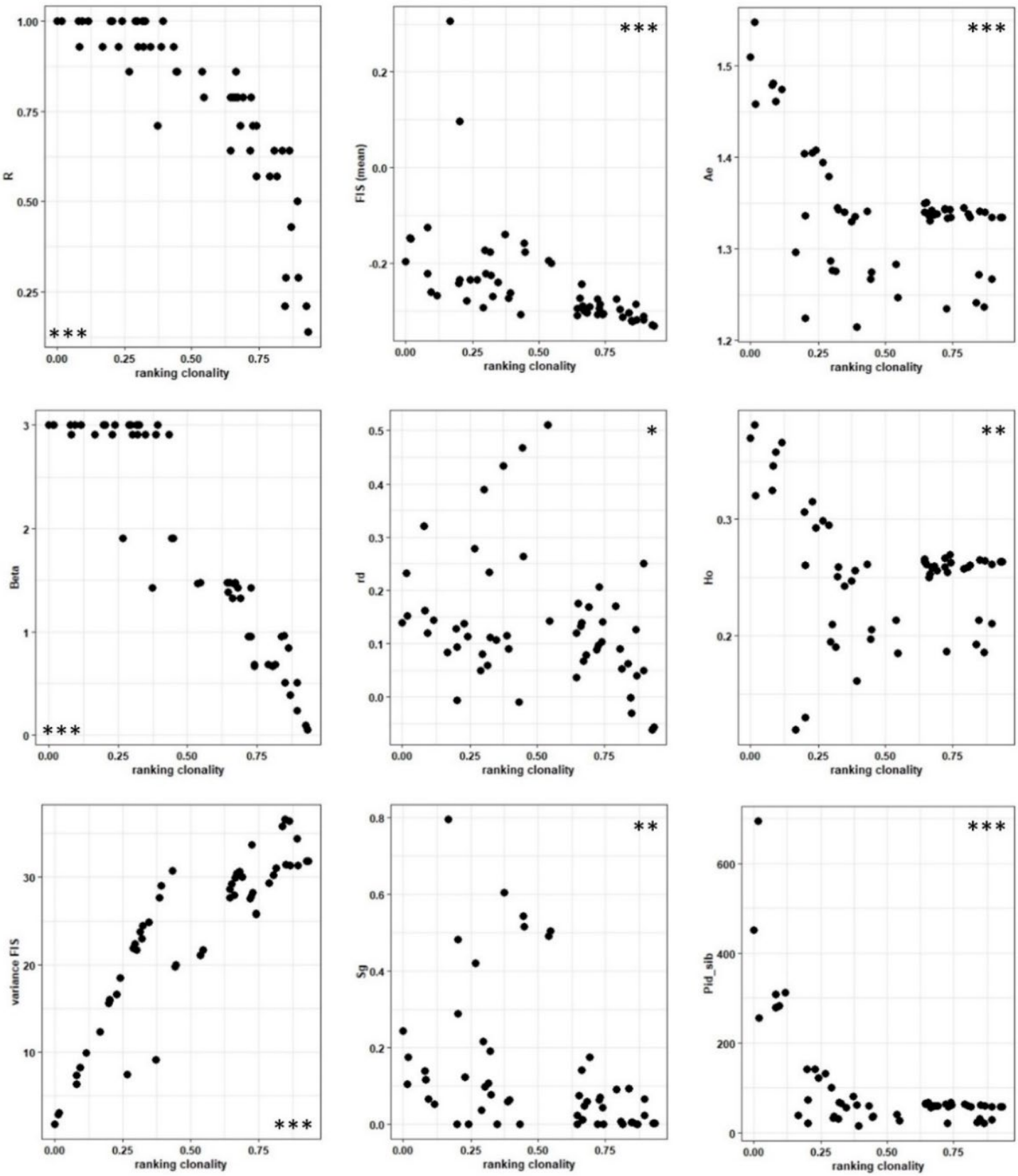

Figure SI3: Relationship and correlation between Σ_{CLON} and genetic indices (R, Pareto β, MFis, VarFis, Ho, Ae, Sg, pidsib and $\bar{r}_d$). Per graph, all the 53 populations were plotted as black dots. *: p<0.05, **: p<0.01, ***: p<0.001.

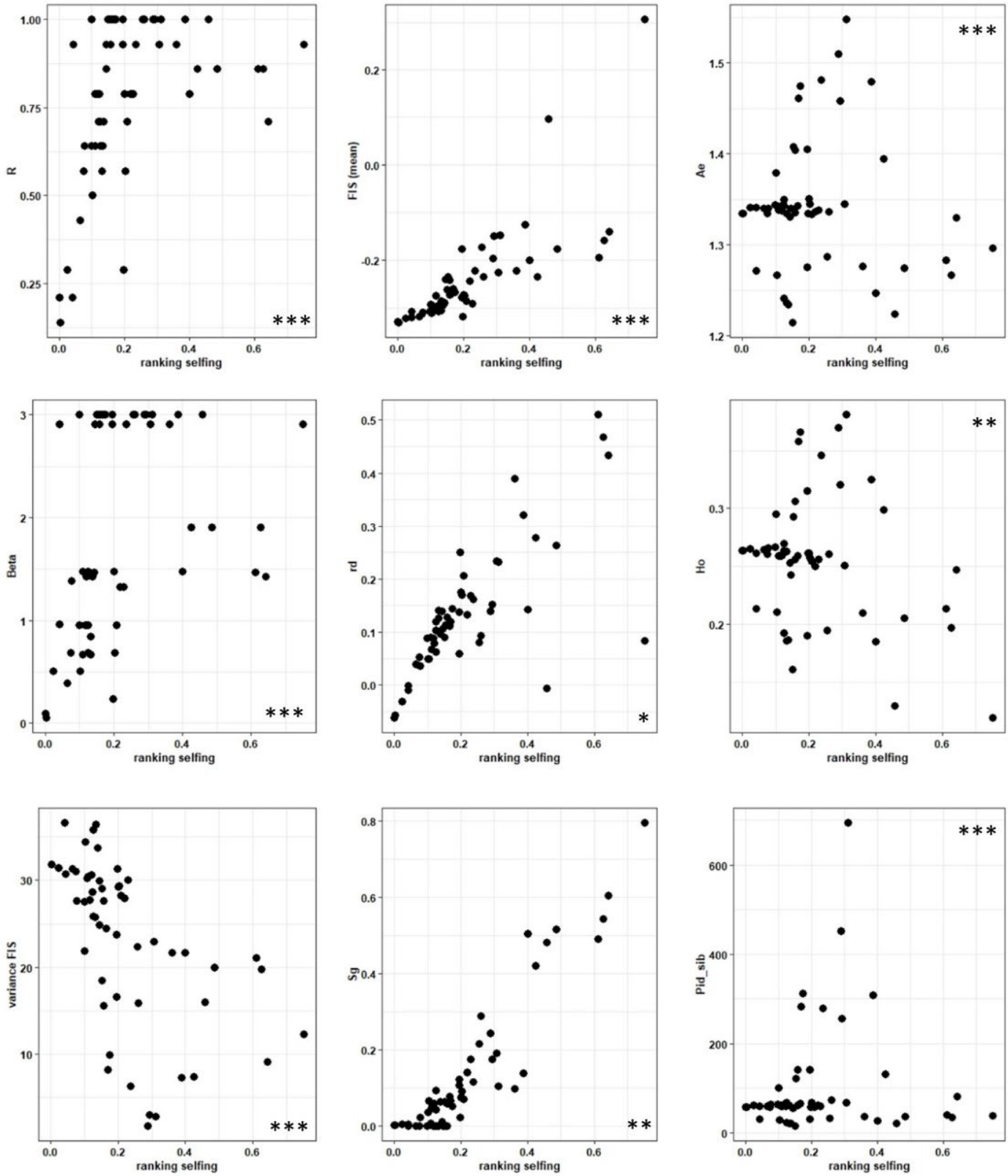

Figure SI4: Relationship and correlation between Σ_{SELF} and genetic indices (R, Pareto β, MFis, VarFis, Ho, Ae, Sg, pidsib and $\bar{r}_d$). Per graph, all the 53 populations were plotted as black dots. *: p<0.05, **: p<0.01, ***: p<0.001.

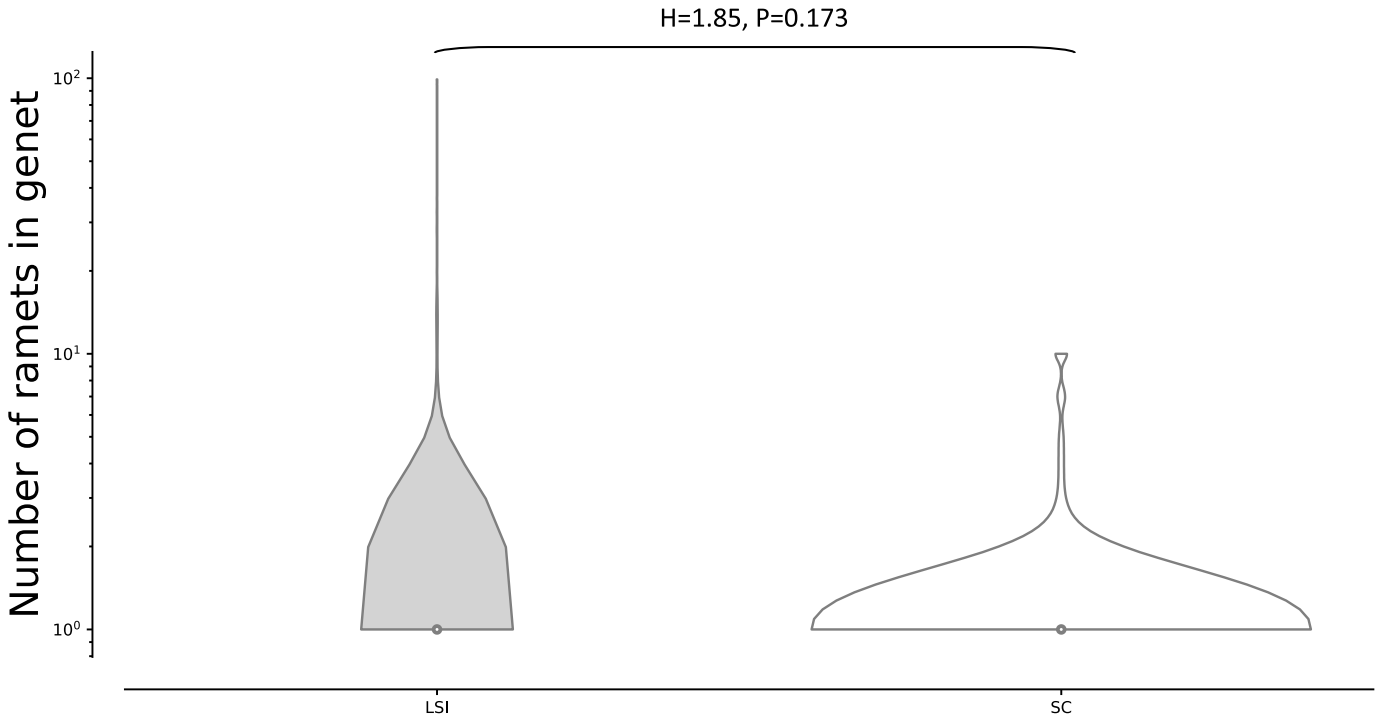

Figure SI4: Distributions of the logarithms of the number of ramets per genet in LSI (grey) and SC (white) populations. We report the Kruskal-Wallis probability that the two distributions are identical.