

ORIGINAL ARTICLE

Effective population size in a partially clonal plant is not predicted by the number of genetic individuals

Roberta Gargiulo¹  | Robin S. Waples^{2,3}  | Adri K. Grow⁴  | Richard P. Shefferson⁵  |
Juan Viruel¹  | Michael F. Fay^{1,6}  | Tiiu Kull⁷ 

¹Royal Botanic Gardens, Kew, Richmond, UK

²NOAA Fisheries, Northwest Fisheries Science Center, Seattle, Washington, USA

³University of Washington, Seattle, Washington, USA

⁴Department of Biological Sciences, Smith College, Northampton, Massachusetts, USA

⁵Organization for Programs in Environmental Sciences, University of Tokyo, Tokyo, Japan

⁶School of Biological Sciences, University of Western Australia, Crawley, Western Australia, Australia

⁷Estonian University of Life Sciences, Tartu, Estonia

Correspondence

Roberta Gargiulo, Royal Botanic Gardens, Kew, Richmond, TW9 3DS, UK.
Email: r.gargiulo@kew.org and robertaxgargiulo@gmail.com

Funding information

British Ecological Society, Grant/Award Number: SR18/1418

Abstract

Estimating effective population size (N_e) is important for theoretical and practical applications in evolutionary biology and conservation. Nevertheless, estimates of N_e in organisms with complex life-history traits remain scarce because of the challenges associated with estimation methods. Partially clonal plants capable of both vegetative (clonal) growth and sexual reproduction are a common group of organisms for which the discrepancy between the apparent number of individuals (ramets) and the number of genetic individuals (genets) can be striking, and it is unclear how this discrepancy relates to N_e . In this study, we analysed two populations of the orchid *Cypripedium calceolus* to understand how the rate of clonal versus sexual reproduction affected N_e . We genotyped >1000 ramets at microsatellite and SNP loci, and estimated contemporary N_e with the linkage disequilibrium method, starting from the theoretical expectation that variance in reproductive success among individuals caused by clonal reproduction and by constraints on sexual reproduction would lower N_e . We considered factors potentially affecting our estimates, including different marker types and sampling strategies, and the influence of pseudoreplication in genomic data sets on N_e confidence intervals. The magnitude of N_e/N_{ramets} and N_e/N_{genets} ratios we provide may be used as reference points for other species with similar life-history traits. Our findings demonstrate that N_e in partially clonal plants cannot be predicted based on the number of genets generated by sexual reproduction, because demographic changes over time can strongly influence N_e . This is especially relevant in species of conservation concern in which population declines may not be detected by only ascertaining the number of genets.

KEYWORDS

conservation genetics, *Cypripedium calceolus*, double-digest RAD sequencing, effective population size, microsatellites, partially clonal plants

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Monitoring effective population size (N_e) is a powerful strategy for predicting and preventing the loss of genetic diversity in natural populations (Frankham, 1995; Jamieson & Allendorf, 2012). N_e is a pivotal parameter in evolutionary biology and conservation biology because of its direct relationship with genetic drift and its effects on other evolutionary forces, including selection and gene flow (Allendorf et al., 2013; Charlesworth, 2009; Crow & Kimura, 1970; Frankham et al., 2010; Wright, 1931). At the same time, N_e is one of the most difficult parameters to estimate directly from demographic data (Caballero, 1994; Campbell & Husband, 2005; Nunney & Elam, 1994; Orive, 1993; Yonezawa et al., 2004), and thus researchers have commonly preferred to estimate it using genetic methods (Gilbert & Whitlock, 2015; Luikart et al., 2010; Nadachowska-Brzyska et al., 2021; Palstra & Ruzzante, 2008; Wang, 2005; Wang et al., 2016; Waples, 1989, 2005).

Among the factors that complicate the estimation of N_e are gene flow, nonrandom reproductive success among individuals, and overlapping generations (e.g., Hill, 1972; Jorde & Ryman, 1995; Neel et al., 2013; Nunney, 1993; Trask et al., 2017; Wang & Whitlock, 2003; Waples, 2002, 2016a, 2022; Waples & England, 2011). The research focused on evaluating biases in N_e estimates has been particularly active and has produced crucial insights for conservation applications (Luikart et al., 2010, 2021; Marandel et al., 2020; Palstra & Fraser, 2012; Ryman et al., 2019; Santiago et al., 2020; Wang et al., 2016; Waples, 2016b; Waples et al., 2014; Waples & England, 2011). Given the global biodiversity crisis, there is an urgency to obtain estimates of N_e (\hat{N}_e) for conservation practice and policy (Frankham, 2021; Hoban et al., 2021; Laikre et al., 2021).

Partial clonality is a widespread reproductive feature across all eukaryotes (Halkett et al., 2005; Krueger-Hadfield et al., 2021; Stoeckel, Arnaud-Haond, & Krueger-Hadfield, 2021; Vallejo-Marín et al., 2010). Partially clonal plants, in particular, represent one common, but relatively understudied, group of organisms for which there is often a striking discrepancy between the observed number of plants and the number of genetic individuals in a population (Alberto et al., 2005; Arnaud-Haond et al., 2020; Honnay & Jacquemyn, 2008; Lozada-Gobilard et al., 2021; Tepedino, 2012; Widén et al., 1994). The observed number of plants in partially clonal plants reflects the contribution of both sexual reproduction, which generates genetic individuals or genets, and clonal reproduction, which generates ramets that are replicates of the same genet (as in the most common type of clonal or vegetative growth, occurring in 80% of angiosperms; Klimeš et al., 1997; Vallejo-Marín et al., 2010). Just as N_e is imperfectly predicted by population census size (N_C) (Frankham, 1995; Palstra & Ruzzante, 2008; Waples et al., 2016), the number of ramets is notoriously a poor surrogate for N_e in a partially clonal plant population (Chung et al., 2004; Tepedino, 2012), as well as a poor indicator of levels of genetic diversity (Mandel, 2010; Raabová et al., 2015). It is worth noting that the number of ramets also encompasses juveniles, and thus may not equal N_C , because the latter is most commonly defined as the number of mature individuals (e.g., Frankham, 1995;

Luikart et al., 2010; Nunney, 1991, 1995). The number of genets, which can only be ascertained through genetic analysis, is generally expected to be closer to N_e than N_C , albeit not equivalent.

It has been suggested that the relationship between clonality and N_e is not straightforward because of confounding factors linked to other life-history traits (Campbell & Husband, 2005), such as lifespan and generation time (Nunney, 1993; Yonezawa, 1997), rate of selfing, and especially the variance of clonal and sexual reproductive contributions of individuals (Orive, 1993; Yonezawa et al., 2004). However, clonal reproduction alone should not cause any significant change in N_e , unless it occurs at extremely high rates and generates fixed heterozygosity (Balloux et al., 2003).

A review based on 63 iteroparous (i.e., capable of reproducing multiple times in a lifetime) species showed that only two traits, namely age at maturity and adult lifespan, explained half of the variance in N_e/N_C , demonstrating that the evolutionary implications of these two traits are consistent across taxa (Waples et al., 2013; Waples et al., 2016; see also Lee et al., 2011). Most of the investigations of N_e in partially clonal plants are based on demographic estimators of N_e , whereas exhaustive empirical comparisons of the number of ramets, number of genets, and \hat{N}_e obtained through genetic analyses are rare (Chung et al., 2004). In general, genetic estimates of N_e are expected to be lower than demographic estimates because they combine the influence of all demographic factors (Nunney & Elam, 1994; Palstra & Ruzzante, 2008) which are difficult to account for simultaneously.

The partially clonal orchid *Cypripedium calceolus* L. (lady's slipper orchid) is a good model system to investigate how N_e changes depending on the balance between clonal and sexual reproduction. Some populations of the species are characterized by a ratio of sexual reproduction to clonal reproduction equal to 1:200, mainly as a result of limitations in seed germination or the absence of pollinators (Devillers-Terschuren, 1999; Kull, 1998, 1999; but see García et al., 2010). Moreover, seedling survival is generally low, as in other terrestrial orchid populations (Shefferson et al., 2020), with a probability of seeds reaching maturity estimated as 10^{-7} in Polish populations (Nicolé et al., 2005). Nevertheless, genetic analyses employing traditional molecular markers (allozymes, AFLPs, and microsatellites) have shown moderately high levels of genetic diversity even in small populations ($N_C < 500$), and this has been mainly attributed to vegetative growth, genet longevity (30–100 years; Kull, 1999; 110–350 years according to Nicolé et al., 2005; with an age at reproductive maturity: 6–10 years old; Kull, 1999), and mating system by outcrossing (Brzosko et al., 2002; Fay et al., 2009; Gargiulo, Adamo et al., 2021; Kull & Paaver, 1997; Minasiewicz et al., 2018).

Tremblay et al. (2005) suggested that N_e/N_C in orchids is particularly low because of pollinator-related limitations, and this is consistent with the observation that an increase in the variance in reproductive success decreases N_e (Frankham, 1995; Nunney, 1991, 1993; Waples, 2016a). All else (i.e., generation length and age at maturity) being equal within a single species, populations in which sexual reproduction is less limited by pollinators (i.e., populations with a higher rate of sexual reproduction) should have a higher N_e/N_C ratio than populations with little sexual reproduction. Although effective population



FIGURE 1 A clump of *Cypripedium calceolus* in Kõrgessaare. Note the juvenile (different genetic individual, or genet) in the bottom-left corner.

size should not be significantly affected by clonal reproduction unless sexual reproduction is very rare (Balloux et al., 2003), the vegetative spread will imply that larger individuals may sexually reproduce more, thus increasing variance in reproductive success and lowering N_e .

In the present study, we asked how clonal versus sexual reproduction affected the effective population size of two populations of *C. calceolus* with different demographic histories. We start from the theoretical expectation that clonal reproduction lowers N_e by increasing variance in reproductive success among individuals, and the constraints on sexual reproduction also lower N_e by causing only a few plants to reproduce. In two populations of the same species, we expect that when sexual reproduction is less constrained, N_e/N_C would be larger. We used an exhaustive sampling strategy and analysed microsatellites and SNPs derived from double-digest restriction site-associated DNA sequencing (ddRADseq) to compare different sets of genetic estimates (which are influenced by different mutation rates and errors associated with different molecular marker types). We first assessed whether genetic data support the observation of different rates of clonal and sexual reproduction in the two populations. We then estimated contemporary N_e with the linkage disequilibrium method (Hill, 1981; Waples & Do, 2008) using both microsatellites and ddRADseq-SNPs. We improved the precision of our \hat{N}_e confidence interval by subsampling the number of loci, and we corrected the bias in \hat{N}_e point estimates due to physical linkage among loci following Waples et al. (2016).

2 | MATERIALS AND METHODS

2.1 | Population sampling

The two Estonian populations of *C. calceolus* selected for the present study have been monitored annually since 1978 and 1985, respectively (Hurskainen et al., 2017; Kull, 1995, 1998, 2003). The

continental population, hereafter “Ussisoo,” is characterized by a generally stable demography (Table S1) and little fruit set. The insular population, hereafter “Kõrgessaare,” occurs in a coastal forest on the Baltic Island of Hiiumaa at the border of a lagoon system (1–2 m above sea level) and includes abundant seedlings. Kõrgessaare is thought to have originated more recently than Ussisoo, probably around 100 years ago or less, after changes in the habitat type (Gargiulo et al., 2018; Kull & Paaver, 1997), with substantial population growth in the last few decades (Table S1).

Clonal growth in *C. calceolus* follows a phalanx strategy, with all ramets from the same clone (i.e., a clump) close to each other in a rounded shape (Figure 1). However, especially when understory vegetation is abundant, different clones may be difficult to distinguish; recruitment within a clump may also occur (Nicolé et al., 2005). All emerging ramets from every single clump were sampled for leaf tissue (nondestructive sampling) and stored in silica gel (Chase & Hills, 1991). In Ussisoo, we collected 451 ramets from 35 putative clumps (exhaustive sampling of all visible plants in the population), and in Kõrgessaare, we collected ~700 ramets from >40 putative clumps (exhaustive sampling of all visible plants at the random coordinates chosen). Sampling was carried out in June 2019, and the Ussisoo population was translocated in August 2019 due to the expansion of the adjacent road.

2.2 | Microsatellite genotyping and analyses of multilocus genotypes

Genomic DNA was extracted with a modified CTAB method (Doyle & Doyle, 1987) and purified with a QIAquick PCR purification kit (QIAGEN, Manchester, UK). All samples were genotyped for 11 nuclear microsatellite (or simple sequence repeat, SSR) loci (Gargiulo et al., 2018; Minasiewicz & Znaniecka, 2014) following the protocol in Gargiulo et al. (2019) for amplification, allele calling and evaluation

of scoring errors/null alleles. Samples belonging to the same populations/clumps were randomized at different steps of the analysis (DNA extraction, polymerase chain reaction, and capillary electrophoresis) to avoid batch effects (Bonin et al., 2004; Meirmans, 2015). The total SSR data set obtained (i.e., including all ramets) is hereafter indicated as “raw data set.” Multilocus genotypes (MLGs) were analysed in the *R* v4.0.5 (R Core Team, 2020) packages *poppr* v2.8.6 (Kamvar et al., 2014, 2015) and *adegenet* v2.1.3 (Jombart, 2008; Jombart & Ahmed, 2011) to obtain indices of genotypic diversity. In *poppr*, we identified identical MLGs and kept one representative for each MLG to generate the “MLG-based clone-corrected data set.”

Before proceeding with further analyses, we performed some checks aimed to avoid the overestimation of either clonal or sexual reproduction in the two populations. In *poppr*, we assessed (1) whether all replicates of the same MLG (i.e., putative clones) were truly part of the same genet and not randomly generated by sexual reproduction (*psex* probability; Parks & Werth, 1993; Arnaud-Haond et al., 2007; Figure S2) and (2) whether each distinct MLG actually belonged to a distinct genet and was not an artefact deriving from scoring errors (Arnaud-Haond et al., 2007; Halkett et al., 2005). To assess the second point, we estimated a genetic distance threshold for collapsing MLGs potentially deriving from scoring errors, using the function *cutoff_predictor* based on Bruvo's distances (see details in Figure S3). After establishing the genetic threshold, we recomputed indices of genotypic diversity in the data set obtained by collapsing potentially identical MLGs in multilocus lineages (MLLs; the related data set is hereafter indicated as “MLL-based clone-corrected data set”) and produced a minimum spanning network to visualize relationships among MLLs.

In the package *RClone* v1.0.2 (Baillieu et al., 2016), we computed Pareto (power law) distributions (β) for the two populations. The Pareto distribution is used to describe the distribution of ramets into MLLs and is influenced by both genotypic (clonal) richness and evenness. For example, when MLLs have comparable sizes (high evenness), the Pareto plot will result in a steeper slope (high β) (see Arnaud-Haond et al., 2007; Stoeckel, Porro, & Arnaud-Haond, 2021).

We used exact tests in *Genepop* v4.5.1 (Raymond & Rousset, 1995; Rousset, 2008) with default parameters to assess deviations from Hardy–Weinberg proportions (potentially indicative of deviations from random mating), and we compared the departures towards heterozygosity excess and deficit and calculated locus-level F_{IS} . The number of private alleles, observed heterozygosity (H_O) and unbiased expected heterozygosity (uH_E), were computed in *GeneAIEx* v6.5 (Peakall & Smouse, 2006, 2012), and private allelic richness was computed in *HP-Rare* (Kalinowski, 2004, 2005) on the MLG-based and the MLL-based clone-corrected data sets. To test the hypothesis that outbred (i.e., highly heterozygous) plants have higher fitness (Alberto et al., 2005; Hämmerli & Reusch, 2003; but see Shefferson et al., 2018), we evaluated the correlation between individual heterozygosity and the number of ramets representing each MLL, as we can assume that plants with more ramets have lived for a long time. Individual heterozygosity was computed as the proportion of typed loci for which an individual clone was heterozygous

(Hämmerli & Reusch, 2003). The allelic richness and F_{IS} values per population were computed in *FSTAT* v2.9.3 (Goudet, 2001). To assess the occurrence of nonrandom associations among loci, we computed the index of association (r_d) in *poppr*.

To check for the occurrence of recent migrants and internal population structure, which can bias \hat{N}_e , we analysed the “MLG-based clone-corrected data set” through Bayesian clustering in *Structure* v2.3.4 (Pritchard et al., 2000) using the Admixture model and no prior on sampling sites. We ran the analysis with 10^5 burn-in, 10^5 MCMC replicates, and 20 iterations, and tested K values (number of genetic clusters) ranging from 1 to 5. We evaluated the most likely K using the $\text{LnPr}(X|K)$ method (Pritchard & Wen, 2003) and the ΔK Evanno method (Evanno et al., 2005) in *Structure* Harvester (Earl & vonHoldt, 2012). The results were summarized in *CLUMPAK* (Kopelman et al., 2015).

2.3 | Analysis of double-digest RAD sequencing data

We used the double-digest RAD sequencing (*ddRADseq*) protocol and data set obtained as in Gargiulo, Kull, et al. (2021). The data set included 31 ramets from Ussisoo, each collected from a different clump, and 32 ramets from Kõrgessaare (Table S2) all representing different clumps except three pairs of putative “biological replicates.” Each pair of biological replicates includes ramets at short distances that may belong to the same genet (pairs: EK308-EK549, EK333-EK336, and EK471-EK206). We used one sample from Ussisoo as a technical replicate throughout *ddRADseq* library preparation and sequencing. *De novo* locus assembly was conducted in *Stacks* v2.4 (Catchen et al., 2013; Rochette et al., 2019) as detailed in Gargiulo, Kull, et al. (2021). We used the *populations* program of *Stacks* to filter *ddRADseq* data depending on the different assumptions and software programmes required in our downstream analyses, as detailed in Table S3, in addition to filtering mostly aimed at reducing the influence of repetitive and paralogous loci expected in the large genome of *C. calceolus* (Gargiulo, Kull, et al., 2021). We checked the occurrence of loci potentially under the effect of selection using *BayeScan* v2.1 (Foll et al., 2010; Foll & Gaggiotti, 2008), to exclude them from the subsequent analyses focused on neutral demographic processes. To avoid that the reduction of informative sites (due to our filtering strategies) determined the detection of false positives (Lotterhos & Whitlock, 2014), we performed the analysis on the “r80 data set” (i.e., the data set including loci shared by 80% of the samples in each population; see Table S3). Only the first SNP at each locus was included in the analysis (option in the *Stacks populations* program: *write-single-snp*). We set the prior odds of neutrality at 1000, the false discovery rate at 0.05, and the chain parameters at default values. Potential deviations from the Hardy–Weinberg proportions were evaluated in *Stacks populations* and in *vcftools* v0.1.16 (Danecek et al., 2011) using exact tests, after excluding F_{ST} -outlier loci (see also Table S3). P-values for the multiple

comparisons were corrected using the *p.adjust* function in R, using the false discovery rate method. After excluding the loci deviating from the Hardy-Weinberg proportions, we estimated the average nucleotide diversity (π), H_D , H_E , F_{IS} , fixation index (F_{ST}), and the number of private alleles in Stacks.

Fine-scale genetic structure associated with the SNPs data set was evaluated in fineSTRUCTURE (Lawson et al., 2012) and visualized in RADpainter (Malinsky et al., 2018; i.e., fineRADstructure). The model implemented in fineRADstructure assumes linkage disequilibrium (LD) among SNPs, although a possible limitation of our data set is the relatively large size of our loci (>250 bp), so we cannot exclude historical recombination. Samples with high percentages of missing data were removed from the data set, as they may bias the estimation of coancestry coefficients (sensu Malinsky et al., 2018). We used the Rscript *reorderLD.R* to reorder loci based on their LD, sampling RAD-tags variants 500 times, and then we ran fineSTRUCTURE with 100,000 burn-in and 100,000 MCMC iterations, keeping every 1000th sample. Nearest neighbour haplotype relationships based on the coancestry values among pairs of samples were visualized in R using the script provided by Malinsky et al. (2018).

2.4 | Estimation of effective population sizes N_e

We employed the software NeEstimator v2.1 (Do et al., 2014) to estimate contemporary N_e using the linkage disequilibrium method (LDN_e ; Hill, 1981; Waples & Do, 2008). Confidence intervals for \hat{N}_e were obtained by jackknifing over samples (Do et al., 2014; Jones et al., 2016; Waples et al., 2021). Both marker types (SSRs and SNPs) were analysed, on genet-level data only, because linkage disequilibrium at the ramet-level would be mainly affected by clonal reproduction and not by genetic drift. Below, we detail the analytical procedure for both marker types, including some caveats and the corrections aimed to improve both the precision and the accuracy on \hat{N}_e .

2.4.1 | SSRs

We analysed both the MLL-based and the MLG-based clone-corrected data sets by excluding singletons (i.e., alleles only occurring in one heterozygote), and using different thresholds to screen out rare alleles (*p*-crit for allele frequencies equal to 0.05, 0.02, 0.01, and 0) (Waples et al., 2016; Waples & Do, 2010). As we used an exhaustive sampling strategy (i.e., of all visible ramets/genets, except very young seedling stages and protocorm stages that are not visible overground), we assumed that the sampled cohorts approached the generation length of the species. Therefore, our \hat{N}_e estimated from microsatellites should be close to the true N_e , although a downward bias of at least 10% associated with mixed-age adult sampling cannot be ruled out (Waples et al., 2014). Note that we consider the genet age in the case of a partially clonal plant such as *C. calceolus*.

TABLE 1 Genotypic parameters associated with the SSR data set for the two populations of *Cyrtopodium calceolus* analysed in this study.

Population	N (raw data set)	#MLGs - #MLLs			eMLGs			G			E.5					
		#MLGs	#MLLs threshold 0.011	#MLLs threshold 0.028	MLGs	MLGs threshold 0.011	MLGs threshold 0.028	MLGs	MLGs threshold 0.011	MLGs threshold 0.028	MLGs	MLGs threshold 0.011	MLGs threshold 0.028	Lambda	rbarD	Genotypic richness
Ussisoo	491	66	66	59	66	66	59	29.3	29.3	28.3	0.76	0.76	0.78	0.97	0.11	0.13
Körgessaare	632	191	179	157	162	153	136	47.4	47.4	44.6	0.54	0.56	0.59	0.98	0.05	0.30
Total	1123	257	245	216	161	157	145	75.8	75.3	72.2	0.63	0.64	0.67	0.99	0.14	

Abbreviations: E.5, evenness; G, Stoddart and Taylor's Index of MLG diversity; Lambda, Lambda-Simpson's index; MLGs, multilocus genotypes; MLLs, multilocus lineages; eMLG, the number of expected MLG at the smallest sample size ≥ 10 based on rarefaction; rbarD, the standardized index of association.

Note: No identical multilocus genotypes were detected across populations.

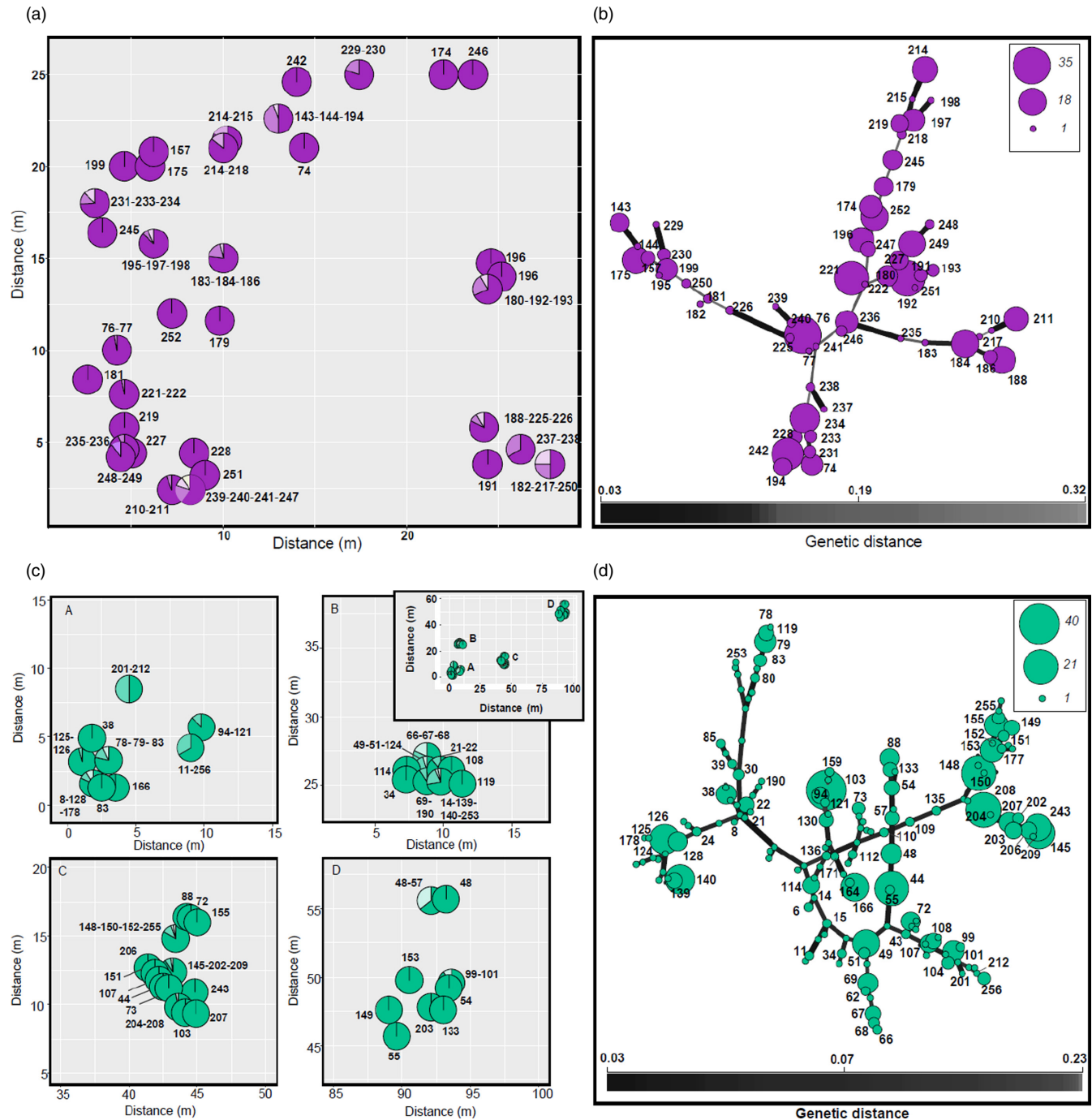


FIGURE 2 Maps representing the spatial distribution of the sampled clumps in Ussisoo (a) and Kõrgessaare (c); minimum spanning network (MSN) among multilocus lineages (MLLs) in Ussisoo (b) and Kõrgessaare (d). Each MLL is indicated by a code; pie sizes in the MSN are proportional to the number of ramets representing each MLL. Codes for the MLLs represented by isolated, single ramets are not shown for Kõrgessaare (and not reported in the spatial map). The inset in (c) refers to the entire sampling area in Kõrgessaare.

Estimates based on a single cohort have been shown to represent the effective number of breeders (N_b) rather than N_e (Waples, 2005; Waples et al., 2014). Unfortunately, we could not obtain an estimate of N_b by analysing single cohorts as, even if the clump size (in a number of ramets) of *C. calceolus* can be predictive of the age of the clump, some clumps only produce one or two ramets over the years (T. Kull, pers. obs.). The occurrence of migration and population structure may be another source of bias in the estimation of

N_e (Palstra & Ruzzante, 2008; Ryman et al., 2014, 2019). The LD method is robust to some population structure and to migration rates <0.1 , whereas for higher migration rates, \hat{N}_e approaches the N_e of the metapopulation (Waples & England, 2011; see also Gilbert & Whitlock, 2015). To evaluate the influence of potential migrants on \hat{N}_e , we removed the admixed individuals as detected in Structure for both $K = 2$ and $K = 3$ (displaying a proportion of admixture $>20\%$, see Section 3) and recomputed \hat{N}_e .

TABLE 2 Genetic diversity and differentiation in the two populations of *Cypripedium calceolus* analysed in this study. For SSRs, values for MLG-based and the MLL-based clone-corrected data sets are identical when not reported otherwise.

Population	SSRs				Private alleles (rarefied ^a)	Allelic richness ^b
	H_E (SE)		H_O (SE)			
	MLG-based clone-corrected data set	MLL-based clone-corrected data set	MLG-based clone-corrected data set	MLL-based clone-corrected data set		
Ussisoo	0.587 (0.071)	0.591 (0.070)	0.563 (0.079)	0.556 (0.078)	26 (1.86)	5.9
Körgessaare	0.411 (0.059)	0.414 (0.059)	0.393 (0.058)	0.390 (0.058)	4 (0.51)	3.4

Abbreviations: π , average nucleotide diversity; F_{IS} , inbreeding coefficient; F_{ST} , fixation index; H_E , expected heterozygosity; H_O , observed heterozygosity; SE, standard error.

Note: Refer to Table S3 for the filtering strategies applied to the SNPs data set ("reduced"; genetic indices for SNPs are computed excluding sites significantly out of the Hardy–Weinberg proportions, after a false discovery rate correction for multiple comparisons).

^aRarefied private alleles (private allelic richness in HP-RARE) are based on 128 gene copies.

^bAllelic richness is based on the minimum sample size: 64 samples for the MLG-based clone-corrected data set and 57 samples for the MLL-based clone-corrected data set.

2.4.2 | SNPs

The SNPs data set only included one SNP at each locus (option in the *Stacks populations* program: *write-random-snp*, see Table S3) to reduce the influence of physical linkage among sites. We used different p-crit values for allele frequencies (i.e., 0.05, 0.02, 0.01, and 0) and excluded singletons. The LDN_e method implemented in NeEstimator is particularly robust with RADseq data when the number of samples is ≥ 30 (Nunziata & Weisrock, 2018). As nonrandom missingness due to allele dropout may bias the N_e estimation (Marandel et al., 2020), we previously evaluated the correlation between the proportion of missing data and F_{IS} at each locus, using Spearman's rank correlation test.

To compare confidence intervals for \hat{N}_e generated by jackknifing over samples with parametric confidence intervals, which generally will be too narrow when thousands of loci are used (Waples, 2021; Waples et al., 2021), we subsampled the SNPs data set by generating 40 random whitelists of 800 loci in *Stacks populations* (with one SNP per locus, see Table S3) and analysed these subsets in NeEstimator. In addition, to further reduce potential biases due to the physical linkage among loci, we divided the point \hat{N}_e by $0.098 + 0.219 \times \ln(10)$, as detailed in Waples et al. (2016), where 10 represents the haploid chromosome number of *C. calceolus*.

To account for the confounding effect of migrants and/or population structure, we removed a potentially admixed individual detected in fineRADstructure from the data set and recomputed \hat{N}_e .

3 | RESULTS

3.1 | Microsatellite genotyping and analyses of multilocus genotypes

The raw data set was composed of 1123 samples, including 451 samples from Ussisoo and 632 from Körgessaare, with a negligible percentage of missing data (Figure S1). When replicated MLGs

were excluded, the MLG-based clone-corrected data set counted 66 MLGs in Ussisoo and 191 in Körgessaare (Table 1).

To avoid the overestimation of clones in the two populations, we evaluated whether all replicates of the same MLG (i.e., putative clones) were truly part of the same genet and not generated by chance due to sexual reproduction. The probability of encountering a genotype more than once by chance (single method) indicated that multiple MLGs are part of a single genet (p -value associated with $psex \ll 0.05$), except in five cases in Körgessaare (Figure S2a). Results obtained by employing the multiple method corroborated the discrepancy between Ussisoo and Körgessaare, with single MLGs arising by sexual reproduction more frequently represented in the second population (Figure S2b). To avoid the overestimation of sexual reproduction, on the other hand, we evaluated the thresholds for collapsing MLGs potentially representing scoring errors, based on genetic distances. The threshold was ~ 0.011 when genotypes with missing data were included in the data set and 0.028 when genotypes with missing data were excluded (Figure S3). The number of MLLs obtained with this approach was 59 in Ussisoo and 157 in Körgessaare. Most of the clumps included ramets with identical MLGs (i.e., truly clonal ramets; Figure 2). In a few cases, for example, when boundaries among clumps were less clear, different MLLs deriving from sexual recruitment were found at small distances (i.e., < 50 cm; Figure 2). Clump sizes were similar between the two populations, ranging from 1 to ~ 40 ramets. The Pareto β obtained in RClone (Figure S4) was low, reflecting the occurrence of clonal reproduction. Pareto plots showed a similar trend in both populations, reflecting lower genotypic richness/higher evenness in Ussisoo and higher genotypic richness/lower evenness in Körgessaare (Table 1).

After applying a Bonferroni correction for multiple testing, we found evidence for deviations from Hardy–Weinberg proportions and high variance in F_{IS} , differentially occurring depending on the data set considered (i.e. raw data set, MLL-based clone-corrected data set, and MLG-based clone-corrected data set, see Table S4). When removing clones, most of these deviations disappeared. In particular, Ussisoo still exhibited an excess of heterozygotes at locus Ccal_25,

F_{IS}		SNPs						
		F_{ST} (SE)	H_E (SE)	H_O (SE)	Private alleles	π (SE)	F_{IS} (SE)	F_{ST}
MLG-based clone-corrected data set	MLL-based clone-corrected data set							
0.04	0.06	0.12 (0.03)	0.254 (0.0004)	0.241 (0.0005)	40,113	0.259 (0.0004)	0.059 (0.005)	0.10
0.04	0.06		0.227 (0.0005)	0.214 (0.0006)	8695	0.232 (0.0005)	0.0497 (0.005)	

whereas Kõrgessaare showed an excess of heterozygotes at locus IM30B and a deficit of heterozygotes at locus IK9. One heterozygote deficit was locus-specific (Ccal_53) in both populations, possibly resulting from allele dropout. Twenty-six private alleles were found in Ussisoo (rarefied private allelic richness: 1.86), in contrast with the four private alleles found in Kõrgessaare (rarefied private allelic richness: 0.51; Table 2); a comparison with the data set including other Eurasian populations (Gargiulo et al., 2019) showed that all private alleles found in Ussisoo occur elsewhere in Eurasia (Table S5; except the alleles at locus Ccal_50, for which large-scale data are not available). In Kõrgessaare, allele 127 at Ccal_25 was found in one isolated ramet (single-ramet plant). Observed and expected heterozygosities at the genet-level were 0.56 and 0.59 in Ussisoo and 0.39 and 0.41 in Kõrgessaare (Table 2). There was no correlation between individual-genet heterozygosity and the number of ramets representing that genet ($R = 0.13$, $p = 0.34$ at Ussisoo and $R = 0.061$, $p = 0.45$ in Kõrgessaare; Figure S5). Allelic richness based on the minimum sample size (64 diploid individuals) was 5.9 in Ussisoo and 3.4 in Kõrgessaare; F_{IS} was 0.04 and 0.06 when considering, respectively, the MLG-based and the MLL-based clone-corrected data sets, in both populations (Table 2). The r_d revealed associations among loci in each population that did not disappear when removing clones and when considering MLLs (p -value = 0.001) (Figure S6), especially in Ussisoo. However, these associations mostly affected different loci across populations, suggesting that they are not related to physical linkage among loci.

The analysis of genetic differentiation carried out in Structure showed that the most likely number of genetic clusters was $K = 2$ (Figure S7), with a few admixed individuals occurring in Kõrgessaare (Figure S8). We also report the results for $K = 3$, in which some internal structure emerges in Kõrgessaare (Figure S8).

3.2 | Analysis of double-digest RAD sequencing data

The results of the ddRADseq data analysis using the different filtering strategies are summarized in Table S3. In BayeScan, we detected

two F_{ST} -outlier loci for which we did not recover any significant correspondence on the NCBI database (Figure S9), and we excluded them from the subsequent analyses (Table S3). For the estimation of genetic diversity parameters, we removed all possible clones (i.e., putative biological replicates) and samples with low coverage and a high percentage of missing data. This filtering strategy produced 34,484 polymorphic loci, with ~108,000 SNPs for Ussisoo and ~86,000 for Kõrgessaare (Table S3). After correcting the P-values for the test of Hardy-Weinberg proportions, 77 SNPs significantly deviated from the Hardy-Weinberg proportions in Ussisoo and 184 SNPs in Kõrgessaare in the main data set (Table S3). Population-level F_{IS} values were 0.06 in Ussisoo and 0.05 in Kõrgessaare (Table 2); at the locus-level, most of the deviations from Hardy-Weinberg proportions were associated with positive F_{IS} values, suggesting a deficit of heterozygotes (results not shown). Removing the out-of-equilibrium sites, however, did not produce changes in the values of genetic diversity and differentiation. Genetic diversity indices for Ussisoo were slightly higher than for Kõrgessaare: H_O was 0.24 for Ussisoo and 0.21 for Kõrgessaare, H_E was 0.25 for Ussisoo and 0.23 for Kõrgessaare, and π was 0.26 for Ussisoo and 0.23 for Kõrgessaare (Table 2). The most marked differences between the two populations were observed in terms of private alleles (~40K in Ussisoo and ~9 K in Kõrgessaare). The fixation index F_{ST} was ~0.1 (Table 2).

Fine-scale relationships as obtained in fineRADstructure revealed a clear differentiation between populations, with only one sample in Kõrgessaare (EK538) showing higher similarity with samples at Ussisoo (Figure 3). In Kõrgessaare, we observed a further population subdivision which reflected only partially the spatial distribution of samples. In Ussisoo, many sample pairs were highly related (with a coancestry coefficient sensu Malinsky et al., 2018 > 500). In general, most of the closest coancestry coefficients occurred between samples at short distances. The biological replicates EK549-EK308 were among the sample pairs with the highest coancestry coefficients. However, six pairs (five in Ussisoo and one in Kõrgessaare) had a coancestry coefficient higher than the biological replicates, possibly because of missing data affecting samples differently (Figure S10; see also Gargiulo, Kull, et al., 2021). Among the sample pairs with

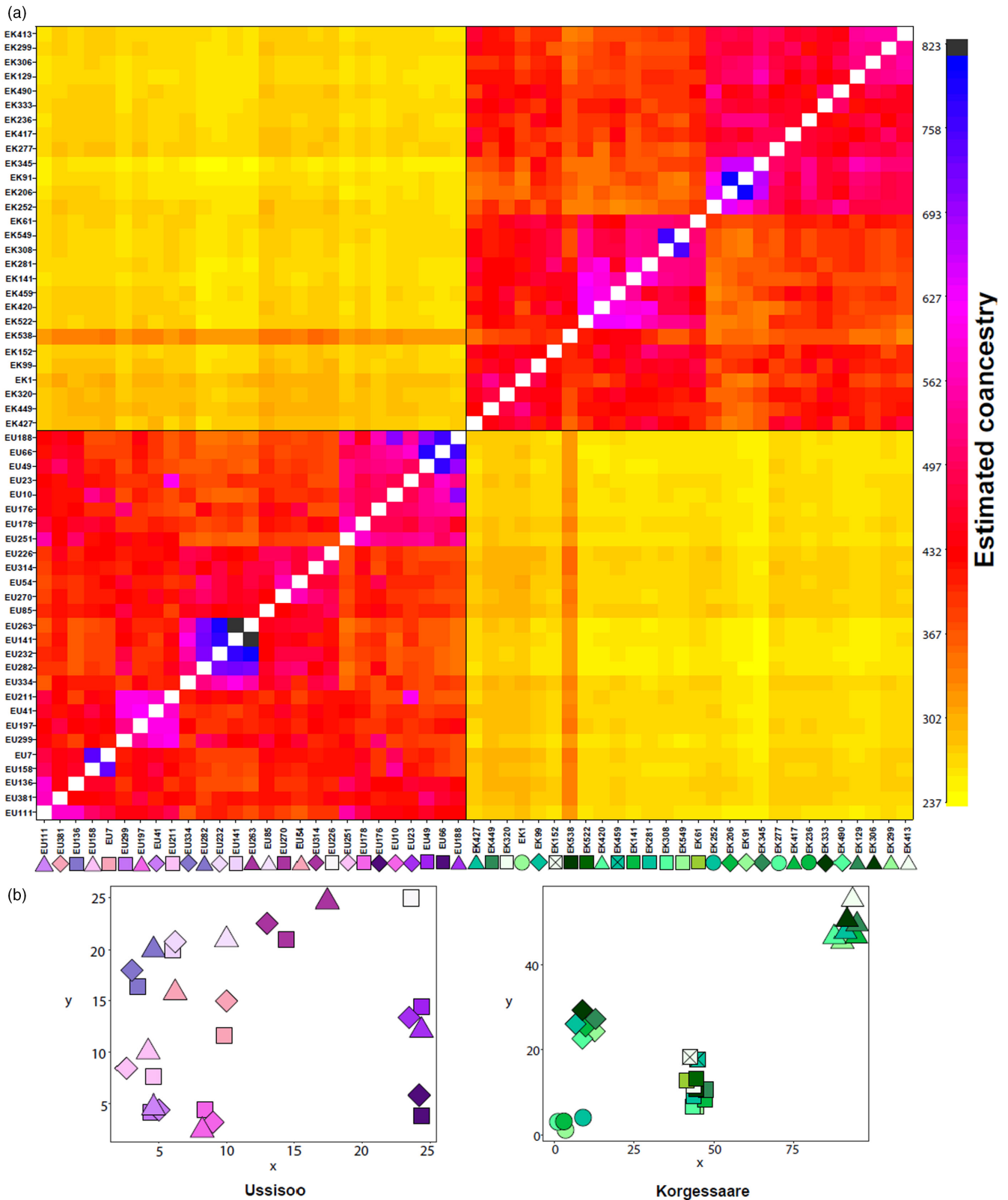


FIGURE 3 (a) Heatmap of the estimated coancestry between pairs of individuals obtained in fineRADstructure. Correspondence between colour gradient and coancestry value is indicated in the sidebar. Symbols located along the x-axis in (a) indicate the spatial position of the genes as represented in the population maps (b).

the highest coancestry coefficients, none exhibited an identical MLG for the SSR loci (as opposite to EK549-EK308, which exhibited an identical microsatellite genotype). In particular, each pair differed

at four to seven loci, and most of the time they shared one allele at such divergent loci (the raw data is available on Dryad: <https://doi.org/10.5061/dryad.6wwpzgn37>).

3.3 | Estimation of effective population sizes N_e

Estimates of N_e obtained in NeEstimator differed somewhat according to the marker type used, especially in Kõrgessaare (Table 3). The analysis of the microsatellite data set (after excluding singletons) in Ussisoo produced an \hat{N}_e equal to 16.9 (CI: 11–25.8) for the MLG-based clone-corrected data set and 20.6 (CI: 13.5–32.3) for the MLL-based clone-corrected data set. In Kõrgessaare, \hat{N}_e was 24 (CI: 13.9–39.5) and 24.1 (CI: 13–42.9) when considering respectively the MLG-based clone-corrected data set and the MLL-based clone-corrected data set. When 16 admixed individuals were removed, \hat{N}_e was 32.8 (CI: 18–59.5). In addition, accounting for the putative internal population structure (as observed for the second most-likely K value, $K = 3$) determined different estimates: 6.1 (CI: 2.0–20.2) for the purple cluster (53 individuals) and 15.5 (CI: 8–28.7) for the blue cluster (74 individuals) (Figure S8; Table 3). Therefore, ratios of N_e/N_C considering either the number of ramets or genets as an estimate for N_C were $N_e/N_{\text{ramets}} = 0.03$ and $N_e/N_{\text{genets}} = 0.26$ in Ussisoo, and $N_e/N_{\text{ramets}} = 0.04$ and $N_e/N_{\text{genets}} = 0.13$ in Kõrgessaare.

For the SNPs data set, Spearman's rank correlation ρ showed a very weak positive correlation between the proportion of missing data and F_{IS} at each locus, for both Ussisoo and Kõrgessaare ($\rho = 0.108$ for Ussisoo, $\rho = 0.087$ for Kõrgessaare, $p < 0.001$), suggesting that although allele dropout may be the cause for positive F_{IS} values, nonrandom missing data do not strongly affect genetic indices and thus \hat{N}_e . The analysis of the total SNPs data set (after excluding singletons) produced an \hat{N}_e equal to 25.5 (CI: 15.8–49.1) in Ussisoo and 5.7 (CI: 2.9–9.8) in Kõrgessaare. The latter estimate was still very small after removing the admixed sample EK538: 5.2 (CI: 2.8–9.0) (Table 3). We also evaluated how precision changed, depending on the LD among different loci, and reported the results obtained by analysing 40 different subsets of 800 loci in NeEstimator (Figure 4). In Ussisoo, the analysis of the 40 subsets (after excluding singletons) produced a median \hat{N}_e equal to 25.7 (empirical CI: 21.1–29.3), whereas in Kõrgessaare, the analysis of the 40 subsets produced a median \hat{N}_e equal to 5.9 (empirical CI: 4.3–6.7) (Table 3). For all estimates based on SNPs, a 40% downward bias due to physical linkage among loci can be expected in a species with 10 chromosomes (haploid number), according to the formula in Waples et al. (2016), although we expect that using only the first SNP in the data set was already mitigating some of the downward bias (Table 3).

4 | DISCUSSION

4.1 | Effective population size (N_e) in partially clonal plants

Our study represents the first exhaustive comparison among the number of ramets, number of genets, and genetic contemporary estimate of N_e , based on two molecular marker types (microsatellites

and ddRADseq-SNPs). Our findings show that N_e in a partially clonal plant cannot reliably be predicted based on the number of genets because demographic events (e.g., bottlenecks or declines) strongly influence N_e .

In particular, we observed that despite the smaller number of ramets and genets in Ussisoo, and its lower rate of sexual reproduction, this population has higher genetic diversity and higher \hat{N}_e than Kõrgessaare, and this is consistent with the stable demography of Ussisoo over the years. In Kõrgessaare, contemporary \hat{N}_e is extremely small when computed on the SNPs data set; considering the long generation time of this species: $T = \alpha + [s/(1-s)] = 30$ years [according to the formula in Lande et al. (2003), where T is the generation time in years, α is the age at first flowering (6 years), and s is the adult survival rate, which was estimated as ~ 0.96 (Shefferson et al., 2012)] and the fact that only adult individuals were included in the SNPs data set, \hat{N}_e still maintains the signature of a strong population bottleneck (which occurred during the founder event, probably less than 100 years ago). Nevertheless, this population has recently undergone an expansion, and this is also suggested by the larger \hat{N}_e obtained from the microsatellites data set including juveniles.

The differences between Ussisoo and Kõrgessaare show that generalization at the species level is not possible without a full understanding of the demographic changes different populations undergo, and that reproductive patterns alone cannot explain a small N_e . In different species, differences among N_e/N_C ratios can be dictated by generation length and age at maturity (the longer these are, the more N_e increases) and lifetime variance in reproductive success (higher variance decreases N_e) (Frankham, 1995; Nunney, 1991, 1993; Waples, 2016a). Within the same species, with generation length and age at maturity being similar, differences in N_e among populations will especially depend on variance in reproductive success and demographic events.

In terms of variance in reproductive success, larger clumps (genetic individuals with more ramets, which are not always older than other genetic individuals) may sexually reproduce more than smaller clumps, and may also inbreed via geitonogamy (i.e., mating among ramets of the same clone). These factors may reduce N_e by favouring larger individuals with more flowers and by causing inbreeding, respectively. The behaviour of pollinators may partially prevent geitonogamy, as pollinators tend to abandon clumps when they discover sexual deception or food deception in orchids that do not offer such rewards (Jersáková et al., 2006; Tremblay et al., 2005; Whitehead et al., 2015), but this may not be the case in other species. Moreover, the advantage of larger or older clumps may be reduced if some of the ramets die over the years, and this would counterbalance the reduction of N_e via lifetime variance in reproductive success. Another phenomenon that could potentially counterbalance the reduction of N_e via lifetime variance in reproductive success is vegetative dormancy (Davison et al., 2013; Lesica & Steele, 1994; Shefferson et al., 2001, 2020), which may cause some plants to skip breeding in some years, spreading lifetime reproductive success among individuals (Waples & Antao, 2014).

TABLE 3 \hat{N}_e for the two populations of *Cypripedium calceolus* analysed in this study.

Population	SSRs									
	MLG-based clone-corrected data set		MLL-based clone-corrected data set		MLG-based no admixed individuals (K = 2)		MLG-based data set for the purple cluster in K = 3 (no admixed individuals)		MLG-based data set for the blue cluster in K = 3 (no admixed individuals)	
	N	\hat{N}_e (CI)	N	\hat{N}_e (CI)	N	\hat{N}_e (CI)	N	\hat{N}_e (CI)	N	\hat{N}_e (CI)
Ussisoo	66	16.9 (11.0–25.8)	59	20.6 (13.5–32.3)	66	16.9 (11.0–25.8)	66	16.9 15.4 (11.0–25.8)	66	16.9 15.4 (11.0–25.8)
Kõrgessaare	191	24.0 (13.9–39.5)	157	24.1 (13.0–42.9)	175	32.8 (18.0–59.5)	53	6.1 (2.0–20.2)	74	15.5 (8.0–28.7)

Note: N, sample size; \hat{N}_e , effective population size estimate, calculated after excluding singletons (p-crit depending on sample size). In parentheses, confidence intervals (CI) are obtained by jackknifing over loci, except when empirical CI are specified. Refer to Figure S8 for the Structure results (purple and blue clusters). Refer to Table S3 for the filtering strategies applied to the SNPs data set (only one SNP per locus was included in all NeEstimator analyses). Removing sites significantly deviating from the Hardy-Weinberg proportions did not produce any change in the N_e estimates. LD-related bias correction refers to the formula $0.098 + 0.219 \times \ln(10) = 0.60$, used to correct the potential bias caused by physical linkage among a high number of loci (see Waples et al., 2016): $\hat{N}_e/0.60$.

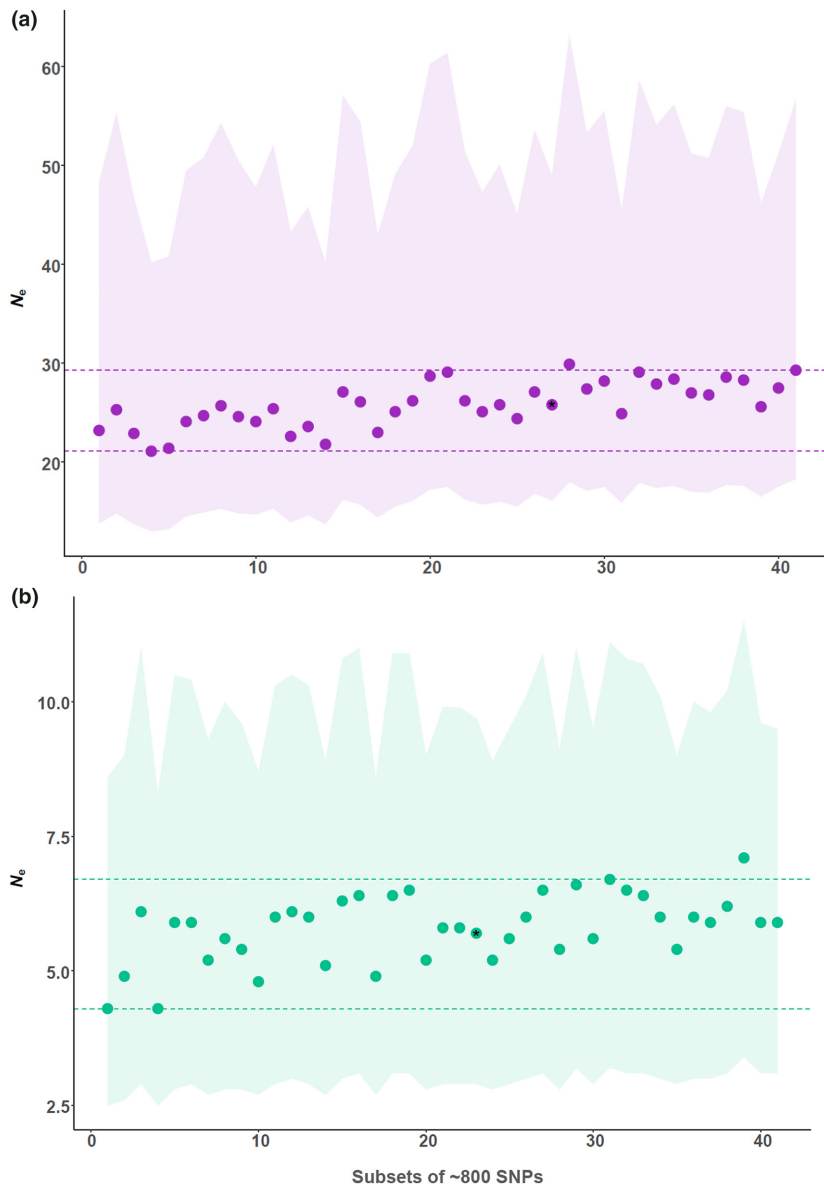


FIGURE 4 Estimates of the effective population size in the populations of *Cypripedium calceolus* at (a) Ussisoo and (b) Kõrgessaare, obtained from 40 different subsets of ~800 SNPs in NeEstimator (results reported in Table 3 and filtering strategies explained in Table S3). Dots indicate point estimates for N_e , ordered by effective degrees of freedom, ranging between 560 and 941 in Ussisoo and 221 and 391 in Kõrgessaare. Shading indicates jackknife confidence intervals (Jones et al., 2016), whereas dashed lines indicate empirical 95% confidence intervals. Dots with asterisks indicate estimates obtained from the full data set (27,136 SNPs), included for comparison. Note the different scales of y-axes for the two populations.

SNPs								
Data set "forNe" - 27,136 SNPs			Median \hat{N}_e on 40 subsets - ~800 SNPs			Without the admixed sample (EK538)		
N	\hat{N}_e (CI)	LD-related bias corrected	N	median \hat{N}_e (empirical CI)	LD-related bias corrected	N	\hat{N}_e (CI)	LD-related bias corrected
31	25.5 (15.8–49.1)	42.5	31	25.7 (21.1–29.3)	42.8	31	25.4 (15.8–49.1)	42.3
31	5.7 (2.9–9.8)	9.5	31	5.9 (4.3–6.7)	9.8	30	5.2 (2.8–9.0)	8.7

4.2 | Factors affecting estimates of effective population size (\hat{N}_e)

Biological and technical factors may strongly influence the estimation of \hat{N}_e (Frankham, 2019; Luikart et al., 2010; Palstra & Ruzzante, 2008). For example, gene flow and population structure can have strong confounding effects on \hat{N}_e (Luikart et al., 2010; Waples & England, 2011). Although the long lifespan of *C. calceolus* implies more opportunities for migration to have an impact on the patterns of LD, ecological observations in orchids suggest that local recruitment, especially at short distances from parental plants, predominates over recruitment of nonlocal seeds (Chung et al., 2009; Duffy et al., 2020; Hedrén et al., 2021; Jacquemyn et al., 2009; Zhang et al., 2019), and that long-distance dispersal, although important in the colonization stage, contributes less to gene flow once a stable population has been established (Hedrén et al., 2018). Such observations are corroborated by our genetic results, as we found clusters of similar genets in Ussisoo (Figure 3) and most juveniles were related to local adults in Kõrgessaare (Figure 2b). Therefore, we believe migration does not substantially influence our \hat{N}_e (see also Table 3), although some internal population structure may influence \hat{N}_e in Kõrgessaare (Figure S8).

The influence of pseudoreplication on confidence intervals is also important when estimating N_e for species of conservation concern, because it may lead to adopting measures that are too severe if confidence intervals are much narrower than the true ones (Waples, 2021; Waples et al., 2021). We investigated the influence of pseudoreplication (i.e., the lack of independence among thousands of loci, as they occur within a much smaller number of chromosomes) on \hat{N}_e confidence intervals by evaluating the differences in precision between a Snip data set including ~30K SNPs and subsets of 800 SNPs (Waples et al., 2021). The empirical CI across the subsets of 800 loci was narrower than the jackknife CI, and this is because empirical CIs do not account for the uncertainty associated with sampling individuals, whereas jackknife CIs do (Waples et al., 2021).

Regardless, results were consistent and point \hat{N}_e were almost identical between the two sets, showing that \hat{N}_e and their confidence intervals did not depend on the subsample of loci included in the analysis. A similar analysis performed by Moran et al. (2019) produced much narrower empirical CIs than jackknife CIs, which extended to positive infinite values.

Statistical biases may also arise as a result of mixed-age adult sampling, which may cause a 10%–50% underestimation from microsatellites and from SNPs (Waples et al., 2014). Age-structure may also be the cause for the Wahlund effect and the discrepancy between microsatellite and SNP estimates observed, especially in Kõrgessaare, where young individuals were abundant and were included in the SSRs data set but not in the SNPs data set. Such observations are particularly important as they suggest that adults in Kõrgessaare show a strong signal of the founder effect, but this signal largely disappears when offspring of recent reproduction are sampled. This also shows, once again, the importance of sampling strategies for both study design and interpretations.

4.3 | N_e/N_c ratios and implications for conservation

Although a small N_e/N_c ratio may be perceived as a typical feature of orchids, even in the absence of clonal reproduction (Trapnell et al., 2022; Tremblay & Ackerman, 2001; Tremblay et al., 2005), and the equilibrium among life-history traits can buffer genetic drift, populations are still susceptible to environmental and genetic stochasticity (Palstra & Ruzzante, 2008). In the terrestrial orchid *Cremastra appendiculata*, Chung et al. (2004) concluded that genetic diversity is maintained despite a small local N_e , possibly because of metapopulation dynamics or because genetic diversity reflected past levels of diversity, whereas their \hat{N}_e reflected a contemporary estimate.

Similarly, we observed that indices of genetic diversity (e.g., heterozygosity) contrast with the \hat{N}_e found (see also Siol et al., 2007).

In particular, \hat{N}_e is significantly smaller than the thresholds signaling critical genetic erosion ($N_e < 50$; Jamieson & Allendorf, 2012; Frankham et al., 2014). In Kõrgessaare, we have explained the extremely small \hat{N}_e as the legacy of a founder effect and, if the recent population expansion continues, N_e will probably increase in a few generations and the genetic signal of the founder effect will disappear. In Ussisoo, it is possible that moderate levels of genetic diversity despite small \hat{N}_e persist because of the long lifespan of genets, and thus reflect past levels of gene flow. Such time lags between the occurrence of demographic changes and the attainment of new equilibrium values of genetic parameters (Epps & Keyghobadi, 2015) are also known in the literature as “genetic extinction debt” (Honnay et al., 2006; Vranckx et al., 2012) and mostly depend on life-history traits. In the partially clonal species *Arnica montana*, Van Rossum and Raspé (2018) observed that populations maintain high genetic diversity (at microsatellite loci) despite a small population size, because of genet longevity, inbreeding depression in early development stages, or no recruitment. Although they did not estimate N_e , they recognized that in such conditions, the number of ramets misrepresents the conservation status of populations. We argue that estimating contemporary N_e is crucial for evaluating the conservation status of a population when there is a time lag between the occurrence of a perturbation and the appearance of a measurable response in the genetic structure of the population.

Very few reliable N_e/N_C estimates are available for plants based on genetic data over multiple generations (Frankham, 2019), and none of these plants have life history traits relatable to *C. calceolus* (e.g., Siol et al., 2007). We found the ratio $N_e/N_{\text{ramets}} < 0.05$ in the two populations of *C. calceolus*, whereas the ratio N_e/N_{genets} was < 0.3 (Table 3, when considering the estimates based on microsatellites, as SNPs were analysed only in a subset of samples). We assume that estimating N_e/N_C ratios using the number of genets (ideally flowering genets to avoid including juveniles and immature individuals) as a surrogate for the number of mature individuals in a population is more appropriate than the number of ramets for comparisons among species. Obtaining multi-generational N_e estimates in a species with a long generation time, such as *C. calceolus*, remains challenging, but our results may be used for this purpose if complemented with data available in the future. Including populations from further sites that may have different demographic histories and balance between clonal and sexual reproduction may also offer a more robust picture of N_e in this species.

In summary, we used a single-sample, contemporary estimate of N_e to understand how effective population size changes depending on the balance between clonal and sexual reproduction. We provided the first exhaustive comparison between the number of ramets, number of genets, and contemporary \hat{N}_e . The N_e/N_{ramets} and N_e/N_{genets} ratios we provide may be used as reference points by researchers and practitioners interested in the magnitude of these ratios in their partially clonal species of interest, and in general for meta-analyses across species (Frankham, 2019, 2021).

As N_e is notoriously difficult to calculate, we considered the factors potentially affecting our estimates, and we showed how estimates can differ when using different molecular marker types and different sampling strategies. In addition, the influence of pseudoreplication

on confidence intervals concerns modern data sets regardless of the study species, and we showed that N_e point estimates obtained from ~30K loci and from subsets of ~800 loci were comparable despite slightly narrower confidence intervals for the larger data set.

Most importantly, we found that effective population size in partially clonal plants cannot be predicted based on the number of genetic individuals (or genets) because demographic events (i.e., changes in the number of individuals over time) strongly influence N_e and this influence can last for decades, depending on the generation time of the species and other life-history traits. Our findings are especially relevant in partially clonal species of conservation concern, in which population declines may not be detected by only counting individuals or by only ascertaining the number of genets using genetic methods. Estimating contemporary effective population size in partially clonal species is therefore crucial for evaluating their conservation status, and results should be interpreted in light of population-specific demographic changes over time.

ACKNOWLEDGEMENTS

This study was funded by the British Ecological Society Research Grant SR18/1418 awarded to Roberta Gargiulo. The Estonian Ministry of Education and Research and the Herbarium of the Estonian University of Life Sciences (TAA) provided Tiiu Kull with funding for fieldwork. We thank László Csiba for his invaluable support in the laboratory. We thank Lydia Soifer and Maurizio Conza for their assistance during fieldwork, Robyn Cowan, Ana Juan, and Maycon Jordan Costa for their assistance in the laboratory. Matthew Clarke and the Linux Team are also acknowledged for their support in maintaining the HPC cluster at RBG Kew. We thank Gordon Luikart and two anonymous reviewers for providing valuable comments on the manuscript.

CONFLICT OF INTEREST

The authors declare no competing interests.

DATA AVAILABILITY STATEMENT

Double-digest RAD sequencing raw data are available under the BioProject PRJNA682386 on National Center for Biotechnology Information (NCBI). Microsatellite multilocus genotypes and filtered SNPs data are available on Dryad at: <https://doi.org/10.5061/dryad.6wwpzgn37>.

BENEFIT-SHARING STATEMENT

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

ORCID

Roberta Gargiulo  <https://orcid.org/0000-0001-8663-6568>

Robin S. Waples  <https://orcid.org/0000-0003-3362-7590>

Adri K. Grow  <https://orcid.org/0000-0003-1262-8412>

Richard P. Shefferson  <https://orcid.org/0000-0002-5234-3131>

Juan Viruel  <https://orcid.org/0000-0001-5658-8411>

Michael F. Fay  <http://orcid.org/0000-0003-3491-9093>

Tiiu Kull  <https://orcid.org/0000-0003-2841-7361>

REFERENCES

- Alberto, F., Gouveia, L., Arnaud-Haond, S., Pérez-Lloréns, J. L., Duarte, C. M., & Serrao, E. A. (2005). Within-population spatial genetic structure, neighbourhood size and clonal subrange in the seagrass *Cymodocea nodosa*. *Molecular Ecology*, 14(9), 2669–2681.
- Allendorf, F. W., Luikart, G. H., & Aitken, S. N. (2013). *Conservation and the genetics of populations*. John Wiley and Sons.
- Arnaud-Haond, S., Duarte, C. M., Alberto, F., & Serrao, E. A. (2007). Standardizing methods to address clonality in population studies. *Molecular Ecology*, 16(24), 5115–5139.
- 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.
- Bailleul, D., Stoeckel, S., & Arnaud-Haond, S. (2016). RClone: A package to identify multilocus clonal lineages and handle clonal data sets in R. *Methods in Ecology and Evolution*, 7(8), 966–970.
- Balloux, F., Lehmann, L., & de Meeûs, T. (2003). The population genetics of clonal and partially clonal diploids. *Genetics*, 164(4), 1635–1644.
- Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C., & Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, 13(11), 3261–3273.
- Brzosko, E., Wróblewska, A., & Ratkiewicz, M. (2002). Spatial genetic structure and clonal diversity of Island populations of lady's slipper (*Cypripedium calceolus*) from the Biebrza National Park (Northeast Poland). *Molecular Ecology*, 11, 2499–2509.
- Caballero, A. (1994). Developments in the prediction of effective population size. *Heredity*, 73(6), 657–679.
- Campbell, L. G., & Husband, B. C. (2005). Impact of clonal growth on effective population size in *Hymenoxys herbacea* (Asteraceae). *Heredity*, 94(5), 526–532.
- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140.
- Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, 10(3), 195–205.
- Chase, M. W., & Hills, H. G. (1991). Silica gel: An ideal material for field preservation of leaf samples for DNA studies. *Taxon*, 40, 215–220.
- Chung, J. M., Park, K. W., Park, C. S., Lee, S. H., Chung, M. G., & Chung, M. Y. (2009). Contrasting levels of genetic diversity between the historically rare orchid *Cypripedium japonicum* and the historically common orchid *Cypripedium macranthos* in South Korea. *Botanical Journal of the Linnean Society*, 160(2), 119–129.
- Chung, M. Y., Nason, J. D., & Chung, M. G. (2004). Implications of clonal structure for effective population size and genetic drift in a rare terrestrial orchid, *Cremastra appendiculata*. *Conservation Biology*, 18(6), 1515–1524.
- Crow, J. F., & Kimura, M. (1970). *An introduction to population genetics theory*. Burgess Publishing Company.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158.
- Davison, R., Nicolé, F., Jacquemyn, H., & Tuljapurkar, S. (2013). Contributions of covariance: Decomposing the components of stochastic population growth in *Cypripedium calceolus*. *The American Naturalist*, 181(3), 410–420.
- Devillers-Terschuren, J. (1999). *Action plan for Cypripedium calceolus in Europe* (No. 100-104). Council of Europe.
- Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator V2: re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, 14(1), 209–214.
- Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation method for small quantities of fresh tissues. *Phytochemical Bulletin, Botanical Society of America*, 19, 11–15.
- Duffy, K. J., Cafasso, D., Ren, M. X., & Cozzolino, S. (2020). High haplotypic diversity with fine-scale structure in a recently established population of an endangered orchid. *Plant Species Biology*, 35(3), 224–232.
- Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Epps, C. W., & Keyghobadi, N. (2015). Landscape genetics in a changing world: Disentangling historical and contemporary influences and inferring change. *Molecular Ecology*, 24(24), 6021–6040.
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Fay, M. F., Bone, R., Cook, P., Kahandawala, I., Greensmith, J., Harris, S., Pedersen, H. A., Ingrouille, M. J., & Lexer, C. (2009). Genetic diversity in *Cypripedium calceolus* (Orchidaceae) with a focus on North-Western Europe, as revealed by plastid DNA length polymorphisms. *Annals of Botany*, 104(3), 517–525.
- Foll, M., Fischer, M. C., Heckel, G., & Excoffier, L. (2010). Estimating population structure from AFLP amplification intensity. *Molecular Ecology*, 19(21), 4638–4647.
- Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180(2), 977–993.
- Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: A review. *Genetics Research*, 66(2), 95–107.
- Frankham, R. (2019). *A practical guide for genetic management of fragmented animal and plant populations*. Oxford University Press. <http://www.oup.co.uk/companion/FrankhamPG>
- Frankham, R. (2021). Suggested improvements to proposed genetic indicator for CBD. *Conservation Genetics*, 22, 531–532. <https://doi.org/10.1007/s10592-021-01357-y>
- Frankham, R., Ballou, J. D., & Briscoe, D. A. (2010). *Introduction to conservation genetics*. Cambridge University Press.
- Frankham, R., Bradshaw, C. J., & Brook, B. W. (2014). Genetics in conservation management: Revised recommendations for the 50/500 rules, red list criteria and population viability analyses. *Biological Conservation*, 170, 56–63.
- García, M. B., Goni, D., & Guzman, D. (2010). Living at the edge: Local versus positional factors in the long-term population dynamics of an endangered orchid. *Conservation Biology*, 24(5), 1219–1229.
- Gargiulo, R., Adamo, M., Cribb, P. J., Bartolucci, F., Sarasan, V., Alessandrelli, C., & Fay, M. F. (2021). Combining current knowledge of *Cypripedium calceolus* with a new analysis of genetic variation in Italian populations to provide guidelines for conservation actions. *Conservation Science and Practice*, 3(10), e513. <https://doi.org/10.1111/csp.2.513>
- Gargiulo, R., Ilves, A., Kaart, T., Fay, M. F., & Kull, T. (2018). High genetic diversity in a threatened clonal species, *Cypripedium calceolus* (Orchidaceae), enables long-term stability of the species in different biogeographical regions in Estonia. *Botanical Journal of the Linnean Society*, 186(4), 560–571.
- Gargiulo, R., Kull, T., & Fay, M. F. (2021). Effective double-digest RAD sequencing and genotyping despite large genome size. *Molecular Ecology Resources*, 21, 1037–1055. <https://doi.org/10.1111/1755-0998.13314>
- Gargiulo, R., Pironon, S., Zheleznyaya, E., Sanchez, M. D., Balázs, Z. R., Podar, D., & Fay, M. F. (2019). Phylogeography and post-glacial dynamics in the clonal-sexual orchid *Cypripedium calceolus* L. *Journal of Biogeography*, 46(3), 526–538.

- Gilbert, K. J., & Whitlock, M. C. (2015). Evaluating methods for estimating local effective population size with and without migration. *Evolution*, 69(8), 2154–2166.
- Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices, version 2.9.3. <https://www2.unil.ch/popgen/softwares/fstat.htm>
- 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.
- Hämmerli, A., & Reusch, T. B. H. (2003). Inbreeding depression influences genet size distribution in a marine angiosperm. *Molecular Ecology*, 12(3), 619–629.
- Hedrén, M., Birkedal, S., De Boer, H., Ghorbani, A., Gravendeel, B., Hansson, S., Svensson, Å., & Zarre, S. (2021). Asymmetric contributions of seed and pollen to gene dispersal in the marsh orchid *Dactylorhiza umbrosa* in Asia minor. *Molecular Ecology*, 30, 1791–1805.
- Hedrén, M., Olofsson, S. N., & Paun, O. (2018). Orchid colonization: Multiple parallel dispersal events and mosaic genetic structure in *Dactylorhiza majalis* ssp. *lapponica* on the Baltic Island of Gotland. *Annals of Botany*, 122(6), 1019–1032.
- Hill, W. G. (1972). Effective size of populations with overlapping generations. *Theoretical Population Biology*, 3(3), 278–289.
- Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetics Research*, 38(3), 209–216.
- Hoban, S., Paz-Vinas, I., Aitken, S., Bertola, L. D., Breed, M. F., Bruford, M. W., & Laikre, L. (2021). Effective population size remains a suitable, pragmatic indicator of genetic diversity for all species, including forest trees. *Biological Conservation*, 253, 108906.
- Honnay, O., Coart, E., Butaye, J., Adriaens, D., Van Glabeke, S., & Roldán-Ruiz, I. (2006). Low impact of present and historical landscape configuration on the genetics of fragmented *Anthyllis vulneraria* populations. *Biological Conservation*, 127(4), 411–419.
- Honnay, O., & Jacquemyn, H. (2008). A meta-analysis of the relation between mating system, growth form and genotypic diversity in clonal plant species. *Evolutionary Ecology*, 22(3), 299–312.
- Hurskainen, S., Jäkäläniemi, A., Kaitala, V., Kull, T., Motlep, M., Ramula, S., & Tuomi, J. (2017). Temporal cycles and spatial asynchrony in the reproduction and growth of a rare nectarless orchid, *Cypripedium calceolus*. *Botanical Journal of the Linnean Society*, 183(2), 316–326.
- Jacquemyn, H., Wiegand, T., Vandepitte, K., Brys, R., Roldán-Ruiz, I., & Honnay, O. (2009). Multigenerational analysis of spatial structure in the terrestrial, food-deceptive orchid *Orchis mascula*. *Journal of Ecology*, 97(2), 206–216.
- Jamieson, I. G., & Allendorf, F. W. (2012). How does the 50/500 rule apply to MVPs? *Trends in Ecology & Evolution*, 27(10), 578–584.
- Jersáková, J., Johnson, S. D., & Kindlmann, P. (2006). Mechanisms and evolution of deceptive pollination in orchids. *Biological Reviews*, 81(2), 219–235.
- Jombart, T. (2008). ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405.
- Jombart, T., & Ahmed, I. (2011). ADEGENET 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27(21), 3070–3071.
- Jones, A. T., Ovenden, J. R., & Wang, Y. G. (2016). Improved confidence intervals for the linkage disequilibrium method for estimating effective population size. *Heredity*, 117(4), 217–223.
- Jorde, P. E., & Ryman, N. (1995). Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics*, 139(2), 1077–1090.
- Kalinowski, S. T. (2004). Counting alleles with rarefaction: Private alleles and hierarchical sampling designs. *Conservation Genetics*, 5(4), 539–543.
- Kalinowski, S. T. (2005). Hp-rare 1.0: A computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, 5(1), 187–189.
- Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, 6, 208.
- Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, 2, e281.
- Klimeš, L., Klimešová, J., Hendriks, R., & van Groenendael, J. (1997). Clonal plant architecture: A comparative analysis of form and function. In H. Kroon & J. Groenendael (Eds.), *The ecology and evolution of clonal plants* (pp. 1–29). Backhuys Publishers.
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15(5), 1179–1191. <https://doi.org/10.1111/1755-0998.12387>
- 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.
- Kull, T. (1995). Genet and ramet dynamics of *Cypripedium calceolus* in different habitats. *Abstracta Botanica*, 19, 95–104.
- Kull, T. (1998). Fruit-set and recruitment in populations of *Cypripedium calceolus* L. in Estonia. *Botanical Journal of the Linnean Society*, 126(1–2), 27–38.
- Kull, T. (1999). *Cypripedium calceolus* L. *Journal of Ecology*, 87(5), 913–924.
- Kull, T. (2003). *Cypripedium calceolus* L. - Interpreting population trends through short-term and long-term monitoring. In T. Rytteri, Ü. Kukk, T. Kull, A. Jäkäläniemi, & M. Reitalu (Eds.), *Monitoring of threatened vascular plants in Estonia and Finland - Methods and experiences* (pp. 71–75). Finnish Environment Institute. https://helda.helsinki.fi/bitstream/handle/10138/40538/FE_659.pdf?sequence=1#page=73
- Kull, T., & Paaver, T. (1997). Patterns of aspartate aminotransferase variation in relation to population size, founder effect, and phylogeographic. *Proceedings of the Estonian Academy of Sciences, Biology and Ecology*, 46(1–2), 4–11.
- Laikre, L., Hohenlohe, P. A., Allendorf, F. W., Bertola, L. D., Breed, M. F., Bruford, M. W., Funk, W. C., Gajardo, G., González-Rodríguez, A., Grueber, C. E., Hedrick, P. W., Heuertz, M., Hunter, M. E., Johannesson, K., Liggins, L., MacDonald, A. J., Mergeay, J., Moharre, F., O'Brien, D., ... Hoban, S. (2021). Authors' reply to letter to the editor: Continued improvement to genetic diversity indicator for CBD. *Conservation Genetics*, 22, 533–536. <https://doi.org/10.1007/s10592-021-01359-w>
- Lande, R., Engen, S., & Sæther, B.-E. (2003). *Stochastic population dynamics in ecology and conservation*. Oxford University Press.
- Lawson, D. J., Hellenthal, G., Myers, S., & Falush, D. (2012). Inference of population structure using dense haplotype data. *PLoS Genetics*, 8(1), e1002453.
- Lee, A. M., Engen, S., & Sæther, B. E. (2011). The influence of persistent individual differences and age at maturity on effective population size. *Proceedings of the Royal Society B: Biological Sciences*, 278(1722), 3303–3312.
- Lesica, P., & Steele, B. M. (1994). Prolonged dormancy in vascular plants and implications for monitoring studies. *Natural Areas Journal*, 14, 209–212.
- Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, 23(9), 2178–2192.
- Lozada-Gobilard, S., Schwarzer, C., Dyer, R., Tiedemann, R., & Joshi, J. (2021). Genetic diversity and connectivity in plant species differing in clonality and dispersal mechanisms in wetland Island habitats. *Journal of Heredity*, 112(1), 108–121.
- Luikart, G., Antao, T., Hand, B. K., Muhlfeld, C. C., Boyer, M. C., Cosart, T., Trethewey, B., al-Chockhachy, R., & Waples, R. S. (2021). Detecting population declines via monitoring the effective number of breeders (Nb). *Molecular Ecology Resources*, 21, 379–393.

- Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K., & Allendorf, F. W. (2010). Estimation of census and effective population sizes: The increasing usefulness of DNA-based approaches. *Conservation Genetics*, 11, 355–373.
- Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: Population inference from RADseq data. *Molecular Biology and Evolution*, 35(5), 1284–1290.
- Mandel, J. R. (2010). Clonal diversity, spatial dynamics, and small genetic population size in the rare sunflower. *Helianthus Verticillatus*. *Conservation Genetics*, 11(5), 2055–2059.
- Marandel, F., Charrier, G., Lamy, J. B., Le Cam, S., Lorange, P., & Trenkel, V. M. (2020). Estimating effective population size using RADseq: Effects of SNP selection and sample size. *Ecology and Evolution*, 10(4), 1929–1937.
- Meirmans, P. G. (2015). Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, 24(13), 3223–3231.
- Minasiewicz, J., & Znaniecka, J. M. (2014). Characterization of 15 novel microsatellite loci for *Cypripedium calceolus* (Orchidaceae) using MiSeq sequencing. *Conservation Genetics Resources*, 6(3), 527–529.
- Minasiewicz, J., Znaniecka, J. M., Górniak, M., & Kawiński, A. (2018). Spatial genetic structure of an endangered orchid *Cypripedium calceolus* (Orchidaceae) at a regional scale: Limited gene flow in a fragmented landscape. *Conservation Genetics*, 19(6), 1449–1460.
- Moran, B. M., Hench, K., Waples, R. S., Höppner, M. P., Baldwin, C. C., McMillan, W. O., & Puebla, O. (2019). The evolution of microendemism in a reef fish (*Hypoplectrus maya*). *Molecular Ecology*, 28(11), 2872–2885.
- Nadachowska-Brzyska, K., Konczal, M., & Babik, W. (2021). Navigating the temporal continuum of effective population size. *Methods in Ecology and Evolution*, 13, 22–41. <https://doi.org/10.1111/2041-210X.13740>
- Neel, M. C., McKelvey, K., Ryman, N., Lloyd, M. W., Short Bull, R., Allendorf, F. W., Schwartz, M. K., & Waples, R. S. (2013). Estimation of effective population size in continuously distributed populations: There goes the neighborhood. *Heredity*, 111(3), 189–199.
- Nicolé, F., Brzosko, E., & Till-Bottraud, I. (2005). Population viability analysis of *Cypripedium calceolus* in a protected area: Longevity, stability and persistence. *Journal of Ecology*, 93(4), 716–726.
- Nunney, L. (1991). The influence of age structure and fecundity on effective population size. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 246(1315), 71–76.
- Nunney, L. (1993). The influence of mating system and overlapping generations on effective population size. *Evolution*, 47(5), 1329–1341.
- Nunney, L. (1995). Measuring the ratio of effective population size to adult numbers using genetic and ecological data. *Evolution*, 49, 389–392.
- Nunney, L., & Elam, D. R. (1994). Estimating the effective population size of conserved populations. *Conservation Biology*, 8(1), 175–184.
- Nunziata, S. O., & Weisrock, D. W. (2018). Estimation of contemporary effective population size and population declines using RAD sequence data. *Heredity*, 120(3), 196–207.
- Orive, M. E. (1993). Effective population size in organisms with complex life-histories. *Theoretical Population Biology*, 44, 316–340.
- Palstra, F. P., & Fraser, D. J. (2012). Effective/census population size ratio estimation: A compendium and appraisal. *Ecology and Evolution*, 2(9), 2357–2365.
- Palstra, F. P., & Ruzzante, D. E. (2008). Genetic estimates of contemporary effective population size: What can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, 17(15), 3428–3447.
- Parks, J. C., & Werth, C. R. (1993). A study of spatial features of clones in a population of bracken fern, *Pteridium aquilinum* (Dennstaedtiaceae). *American Journal of Botany*, 80(5), 537–544.
- Peakall, R., & Smouse, P. E. (2006). GenAIX 6: Genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288–295.
- Peakall, R., & Smouse, P. E. (2012). GenAIX 6.5: Genetic analysis in excel. Population genetic software for teaching and research—an update. *Bioinformatics*, 28, 2537–2539.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Pritchard, J. K., & Wen, W. (2003). *Documentation for STRUCTURE software: Version 2*. University of Chicago Press. http://web.stanford.edu/group/pritchardlab/software/structure2_1.html
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Raabová, J., Van Rossum, F., Jacquemart, A. L., & Raspé, O. (2015). Population size affects genetic diversity and fine-scale spatial genetic structure in the clonal distylous herb *Menyanthes trifoliata*. *Perspectives in Plant Ecology, Evolution and Systematics*, 17(3), 193–200.
- Raymond, M., & Rousset, F. (1995). GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86, 248–249.
- Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28(21), 4737–4754.
- Rousset, F. (2008). genepop'007: A complete re-implementation of the Genepop software for windows and Linux. *Molecular Ecology Resources*, 8(1), 103–106.
- Ryman, N., Allendorf, F. W., Jorde, P. E., Laikre, L., & Hössjer, O. (2014). Samples from subdivided populations yield biased estimates of effective size that overestimate the rate of loss of genetic variation. *Molecular Ecology Resources*, 14(1), 87–99.
- Ryman, N., Laikre, L., & Hössjer, O. (2019). Do estimates of contemporary effective population size tell us what we want to know? *Molecular Ecology*, 28(8), 1904–1918.
- Santiago, E., Novo, I., Pardiñas, A. F., Saura, M., Wang, J., & Caballero, A. (2020). Recent demographic history inferred by high-resolution analysis of linkage disequilibrium. *Molecular Biology and Evolution*, 37(12), 3642–3653.
- Shefferson, R. P., Jacquemyn, H., Kull, T., & Hutchings, M. J. (2020). The demography of terrestrial orchids: Life history, population dynamics and conservation. *Botanical Journal of the Linnean Society*, 192(2), 315–332.
- Shefferson, R. P., Kull, T., Tali, K., & Kellett, K. M. (2012). Linking vegetative dormancy to fitness in two long-lived herbaceous perennials. *Ecosphere*, 3(2), 1–19.
- Shefferson, R. P., Kull, T., Hutchings, M. J., Selosse, M. A., Jacquemyn, H., Kellett, K. M., Menges, E. S., Primack, R. B., Tuomi, J., Alahuhta, K., Hurskainen, S., Alexander, H. M., Anderson, D. S., Brys, R., Brzosko, E., Dostálík, S., Gregg, K., Ipser, Z., Jäkäläniemi, A., ... Whigham, D. F. (2018). Drivers of vegetative dormancy across herbaceous perennial plant species. *Ecology Letters*, 21(5), 724–733.
- Shefferson, R. P., Sandercock, B. K., Proper, J., & Beissinger, S. R. (2001). Estimating dormancy and survival of a rare herbaceous perennial using mark-recapture models. *Ecology*, 82(1), 145–156.
- Siol, M., Bonnin, I., Olivieri, I., Prospero, J. M., & Ronfort, J. (2007). Effective population size associated with self-fertilization: Lessons from temporal changes in allele frequencies in the selfing annual *Medicago truncatula*. *Journal of Evolutionary Biology*, 20(6), 2349–2360.
- Stoekel, S., Arnaud-Haond, S., & Krueger-Hadfield, S. A. (2021). 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.
- Stoekel, S., Porro, B., & Arnaud-Haond, S. (2021). 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.
- Tepedino, V. J. (2012). Overestimating population sizes of rare clonal plants. *Conservation Biology*, 26(5), 945–947.
- Trapnell, D. W., Smallwood, P. A., Dixon, K. W., & Phillips, R. D. (2022). Are small populations larger than they seem? Genetic insights into patchily distributed populations of *Drakaea glyptodon* (Orchidaceae). *Botanical Journal of the Linnean Society*, 198(1), 99–116.
- Trask, A. E., Bignal, E. M., McCracken, D. I., Piertney, S. B., & Reid, J. M. (2017). Estimating demographic contributions to effective population size in an age-structured wild population experiencing environmental and demographic stochasticity. *Journal of Animal Ecology*, 86(5), 1082–1093.
- Tremblay, R. L., & Ackerman, J. D. (2001). Gene flow and effective population size in *Lepanthes* (Orchidaceae): A case for genetic drift. *Biological Journal of the Linnean Society*, 72(1), 47–62.
- Tremblay, R. L., Ackerman, J. D., Zimmerman, J. K., & Calvo, R. N. (2005). Variation in sexual reproduction in orchids and its evolutionary consequences: A spasmodic journey to diversification. *Biological Journal of the Linnean Society*, 84(1), 1–54.
- Vallejo-Marín, M., Dorken, M. E., & Barrett, S. C. (2010). The ecological and evolutionary consequences of clonality for plant mating. *Annual Review of Ecology, Evolution, and Systematics*, 41, 193–213.
- Van Rossum, F., & Raspé, O. (2018). Contribution of genetics for implementing population translocation of the threatened *Arnica Montana*. *Conservation Genetics*, 19, 1185–1198.
- Vranckx, G. U. Y., Jacquemyn, H., Muys, B., & Honnay, O. (2012). Meta-analysis of susceptibility of woody plants to loss of genetic diversity through habitat fragmentation. *Conservation Biology*, 26(2), 228–237.
- Wang, J. (2005). Estimation of effective population sizes from data on genetic markers. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 360(1459), 1395–1409.
- Wang, J., Santiago, E., & Caballero, A. (2016). Prediction and estimation of effective population size. *Heredity*, 117(4), 193–206.
- Wang, J., & Whitlock, M. C. (2003). Estimating effective population size and migration rates from genetic samples over space and time. *Genetics*, 163(1), 429–446.
- Waples, R. K., Larson, W. A., & Waples, R. S. (2016). Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. *Heredity*, 117(4), 233–240.
- Waples, R. S. (1989). A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics*, 121(2), 379–391.
- Waples, R. S. (2002). Definition and estimation of effective population size in the conservation of endangered species. In S. R. Beissinger & D. R. McCullough (Eds.), *Population viability analysis* (pp. 147–168). University of Chicago Press.
- Waples, R. S. (2005). Genetic estimates of contemporary effective population size: To what time periods do the estimates apply? *Molecular Ecology*, 14(11), 3335–3352.
- Waples, R. S. (2016a). Life-history traits and effective population size in species with overlapping generations revisited: The importance of adult mortality. *Heredity*, 117(4), 241–250.
- Waples, R. S. (2016b). Making sense of genetic estimates of effective population size. *Molecular Ecology*, 25, 4689–4691.
- Waples, R. S. (2021). Relative precision of the sibship and LD methods for estimating effective population size with genomics-scale datasets. *Journal of Heredity*, 112(6), 535–539.
- Waples, R. S. (2022). What is Ne, anyway? *Journal of Heredity*, 113, 371–379.
- Waples, R. S., & Antao, T. (2014). Intermittent breeding and constraints on litter size: Consequences for effective population size per generation (Ne) and per reproductive cycle (Nb). *Evolution*, 68(6), 1722–1734.
- Waples, R. S., Antao, T., & Luikart, G. (2014). Effects of overlapping generations on linkage disequilibrium estimates of effective population size. *Genetics*, 197(2), 769–780.
- Waples, R. S., & Do, C. H. I. (2008). LDNE: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, 8(4), 753–756.
- Waples, R. S., & Do, C. H. I. (2010). Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3(3), 244–262.
- Waples, R. S., & England, P. R. (2011). Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics*, 189(2), 633–644.
- Waples, R. S., Luikart, G., Faulkner, J. R., & Tallmon, D. A. (2013). Simple life-history traits explain key effective population size ratios across diverse taxa. *Proceedings of the Royal Society B: Biological Sciences*, 280(1768), 20131339.
- Waples, R. S., Waples, R., & Ward, E. J. (2021). Pseudoreplication in genomics-scale datasets. *Molecular Ecology Resources*, 22, 503–518. <https://doi.org/10.1111/1755-0998.13482>
- Whitehead, M. R., Linde, C. C., & Peakall, R. (2015). Pollination by sexual deception promotes outcrossing and mate diversity in self-compatible clonal orchids. *Journal of Evolutionary Biology*, 28(8), 1526–1541.
- Widén, B., Cronberg, N., & Widén, M. (1994). Genotypic diversity, molecular markers and spatial distribution of genets in clonal plants, a literature survey. *Folia Geobotanica*, 29(2), 245–263.
- Wright, S. (1931). Evolution in mendelian populations. *Genetics*, 16(2), 97–159.
- Yonezawa, K. (1997). Effective population size of plant species propagating with a mixed sexual and asexual reproduction system. *Genetics Research*, 70(3), 251–258.
- Yonezawa, K., Ishii, T., & Nagamine, T. (2004). The effective size of mixed sexually and asexually reproducing populations. *Genetics*, 166(3), 1529–1539.
- Zhang, Z., Gale, S. W., Li, J. H., Fischer, G. A., Ren, M. X., & Song, X. Q. (2019). Pollen-mediated gene flow ensures connectivity among spatially discrete sub-populations of *Phalaenopsis pulcherrima*, a tropical food-deceptive orchid. *BMC Plant Biology*, 19(1), 1–16.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Gargiulo, R., Waples, R. S., Grow, A. K., Shefferson, R. P., Viruel, J., Fay, M. F., & Kull, T. (2023). Effective population size in a partially clonal plant is not predicted by the number of genetic individuals. *Evolutionary Applications*, 16, 750–766. <https://doi.org/10.1111/eva.13535>