

## RESEARCH ARTICLE

# Population genetic structure in a self-compatible hermaphroditic snail is driven by drift independently of its contemporary mating system

Cansu Çetin<sup>1,2,3</sup>  | Jukka Jokela<sup>1,2</sup> | Philine G. D. Feulner<sup>4,5</sup> | Tamara Schlegel<sup>1,2</sup> | Nadine Tardent<sup>1,2</sup> | Otto Seppälä<sup>1,2,6</sup>

<sup>1</sup>Department of Aquatic Ecology, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

<sup>2</sup>Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland

<sup>3</sup>Institute of Environmental Sciences, Faculty of Biology, Jagiellonian University, Kraków, Poland

<sup>4</sup>Department of Fish Ecology and Evolution, Centre of Ecology, Evolution and Biogeochemistry, EAWAG Swiss Federal Institute of Aquatic Science and Technology, Kastanienbaum, Switzerland

<sup>5</sup>Division of Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, Bern, Switzerland

<sup>6</sup>Research Department for Limnology, Universität Innsbruck, Mondsee, Austria

## Correspondence

Cansu Çetin, Institute of Environmental Sciences, Faculty of Biology, Jagiellonian University, Kraków, Poland.  
Email: [cansucetin@uj.edu.pl](mailto:cansucetin@uj.edu.pl)

## Funding information

Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 31003A 169531; Eidgenössische Technische Hochschule Zürich, Grant/Award Number: ETH-20 17-1

## Abstract

Genetic drift, gene flow, and natural selection commonly influence population genetic diversity. In populations of self-compatible hermaphrodites, the mating system (e.g., self-fertilization) further reduces individual heterozygosity. Furthermore, selfing, as a form of inbreeding, significantly impacts genetic drift by reducing effective population size ( $N_e$ ). This can potentially accelerate genetic drift, particularly in small populations where self-fertilization is likely during founder events. To investigate the roles of genetic drift and contemporary mating system in populations of the freshwater snail *Lymnaea stagnalis*, we examined their effective population sizes ( $N_e$ ) and Tajima's  $D$  values, which reflect genetic drift over extended time periods, as well as estimates of within-population selfing rates and pairwise relatedness reflecting contemporary mating system. We used 4054 SNP markers obtained using restriction site associated DNA (RAD) sequencing from individuals in five snail populations originating from geographically closely located ponds. We found strong population genetic structure and differences in genetic diversity among populations. Covariation between genetic diversity and  $N_e$  estimates and Tajima's  $D$  values suggested drift being an important determinant of genetic diversity and structure in these populations. However, this effect was independent of the contemporary mating system, as indicated by the similarity of selfing rates and relatedness estimates among populations. Thus, founder events (possibly including historical inbreeding) and/or drift due to small sizes of *L. stagnalis* populations are likely to explain their genetic structure and limit within-population genetic diversity.

## KEYWORDS

genetic structure, inbreeding, RAD sequencing, relatedness, small populations

## TAXONOMY CLASSIFICATION

Ecological genetics, Evolutionary ecology, Population genetics

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.

© 2024 The Author(s). *Ecology and Evolution* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Genetically diverse populations are assumed to have the highest evolutionary potential and, thus, the lowest probability of extinction in changing environments (Hoffmann et al., 2017). Directional natural selection reduces genetic diversity owing to the fixation of alleles with positive fitness effects (Charlesworth et al., 1993; Charlesworth & Charlesworth, 2010; Hohenlohe et al., 2010), while genetic drift reduces variation through the random loss of rare gene variants (Byrne & Nichols, 1999; Wright, 1931). The contemporary mating system and inbreeding history, including biparental inbreeding and self-fertilization, can further impact genetic diversity by reducing individual heterozygosity (Crnokrak & Roff, 1999; DeRose & Roff, 1999) and increasing levels of pairwise relatedness (Ingvarsson, 1998; Wang, 2007, 2014). Additionally, inbreeding creates identity disequilibrium (i.e., correlations of heterozygosity), increasing genetic differences among individuals, and may reduce individual fitness due to inbreeding depression (Coutellec & Lagadic, 2006; Henry et al., 2003; Noël et al., 2017).

The interplay of genetic drift and the mating system is especially important for populations experiencing extinction-recolonization dynamics (Ellegren & Galtier, 2016). This is because stochastic demographic processes referring to the random changes in the survival and reproductive rates of individuals can enhance the impact of genetic drift when the effective population size ( $N_e$ ) is small (Pannell & Charlesworth, 2000). Such effects could be even stronger in species that can self-fertilize because the number of founders could be as low as one, making populations extremely prone to increased homozygosity during (re)colonization events (Jarne, 1995). While past inbreeding events may have shaped genetic diversity (together with the effect of small population size), understanding how the contemporary mating system contributes to genetic variation in such populations remains elusive (but see Diaz-Martin et al., 2023; Ingvarsson, 1998). However, addressing this gap is essential for our ability to decipher the adaptive potential and extinction risks of organisms with small and/or fluctuating populations because the contemporary mating system directly influences, for example, individual heterozygosity and relatedness (Byrne & Nichols, 1999; Keller & Waller, 2002; Wright, 1931).

Modern genome-wide multilocus analyses allow for the estimation of different population genetic processes, such as genetic drift (reflected by  $N_e$  and linkage disequilibrium) and details of the mating system (e.g., relatedness estimates based on identity by descent (IBD) and selfing rate) in nature (Kardos et al., 2015; Puurtinen et al., 2004). We examined the roles of genetic drift and contemporary mating system in determining the genetic diversity and structure in natural populations of the self-compatible hermaphroditic freshwater snail, *Lymnaea stagnalis*. In this system, both processes (i.e., drift, mating system) could have large roles in determining genetic diversity. This is because, first, *L. stagnalis* commonly inhabits small ponds where extinctions and founder events can accentuate genetic drift. In fact, genetic diversity in spatially closely located *L. stagnalis* populations varies considerably, and genetic differentiation

among them is often high (Kopp et al., 2012; Puurtinen et al., 2004). However, earlier population genetic analyses of *L. stagnalis* used co-dominant markers covering fewer loci (e.g., microsatellites, allozyme polymorphism) (Coutellec-Vreto et al., 1994; Kopp et al., 2012; Puurtinen et al., 2004; Ritland & Jain, 1984), which provided less precise estimates for genetic diversity, linkage, and relatedness. Second, although preferring outcrossing (Cain, 1956), *L. stagnalis* can commit to the most extreme form of inbreeding, self-fertilization, which can help them while founding populations (Kopp et al., 2012). While selfing can be advantageous for colonizing new environments, it can also lead to reduced genetic diversity within populations. Additionally, population bottlenecks and founder events may have purged recessive deleterious mutations, which may be the reason why most *L. stagnalis* populations do not show strong inbreeding depression (Coutellec & Lagadic, 2006; Koene et al., 2008; Puurtinen et al., 2007). Therefore, our study system allows for testing the role of the contemporary mating system on genetic diversity without strong selection against selfing influencing the results.

Here, we examined both population- and individual-level genetic parameters in five *L. stagnalis* populations using 4054 SNP markers. SNP markers are powerful for estimating the linkage of alleles, selfing rates,  $N_e$ , Tajima's *D* (informing about deviations in rare allele frequencies from neutral expectations, which can indicate demographic events such as bottlenecks), and relatedness between individuals (Attard et al., 2018; Lemopoulos et al., 2019). Variation in the estimates reflecting genetic drift (e.g.,  $N_e$  and Tajima's *D*) and their covariation with genetic variation would suggest the effect of genetic drift (e.g., founder events including historical inbreeding due to selfing/biparental inbreeding) being an important determinant of genetic diversity. Furthermore, variation in mating system estimates (relatedness and selfing rates) and their covariation with genetic diversity would imply that the contemporary mating system plays a crucial role in shaping genetic diversity. Thus, we predicted that populations with more genetic drift (i.e., lower  $N_e$  and higher deficit of rare alleles indicated by Tajima's *D*) and/or with more contemporary inbreeding (high selfing rates and/or relatedness) would have the lowest genetic diversity.

## 2 | MATERIALS AND METHODS

### 2.1 | Sampling, DNA extraction, and sequencing

We collected a total of 92 *L. stagnalis* snails from five ponds near Zürich in northern Switzerland in 2017 (see Figure 1 and Table 1). Pairwise distances between the populations ranged from 2.5 to 34 km. Because larger ponds can contain larger populations and provide more migrants, we calculated the area of each pond to the nearest  $m^2$  using a topographic map ([map.geo.admin.ch](http://map.geo.admin.ch)). Furthermore, because landscape may influence the isolation of the populations, we also considered the elevation of the studied ponds.

We extracted genomic DNA from the head tissue of each collected snail using the CTAB method (modified from Doyle and

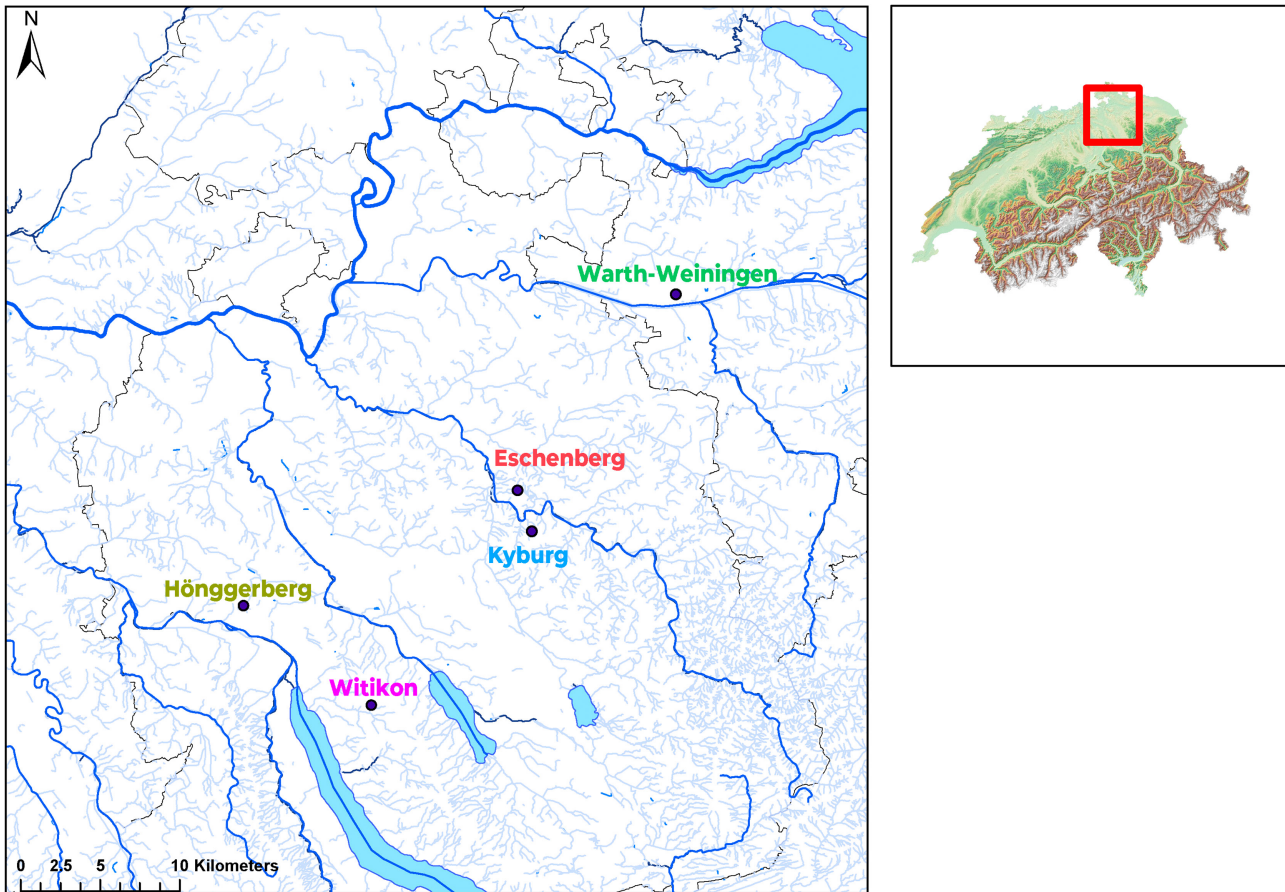


FIGURE 1 Locations of the investigated *Lymnaea stagnalis* populations in northern Switzerland. Data were provided by Geodata CH ([geo.admin.ch](https://geo.admin.ch)), and the map was produced using ArcGis Map (version 10.7.1, ESRI).

TABLE 1 Location information and genetic diversity values for the examined *L. stagnalis* populations based on the non-thinned dataset.

Population	Eschenberg	Hönggerberg	Warth-Weiningen	Kyburg	Witikon
Code	EB	HB	WW	KB	WT
Coordinates	47°28'29.9" N, 8°43'46.2" E	47°24'42.2" N, 8°29'57.9" E	47°35'01.4" N, 8°51'51.6" E	47°27'04.8" N, 8°44'26.2" E	47°21'15.1" N, 8°36'16.8" E
N	19	19	19	17	18
Elevation (m)	527	516.7	423	624	613
Area (m <sup>2</sup> )	429	218	307	NA	308
Polymorphic sites	2010	2685	2506	1454	1655
Private alleles	219	200	177	229	90
$H_e$	0.18	0.21	0.21	0.13	0.16
$H_o$	0.14	0.15	0.15	0.09	0.13
$\pi$	0.0033	0.0038	0.0038	0.0024	0.0029

Note: N refers to the number of individuals sampled from each population. Each pond's elevation (m) and area (m<sup>2</sup>) are indicated. Genetic diversity for each population, including the number of polymorphic sites, private alleles, expected heterozygosity ( $H_e$ ), observed heterozygosity ( $H_o$ ), and nucleotide diversity ( $\pi$ ) is displayed.

Doyle (1987)) followed by magnetic bead cleanup (see Text S1). We measured DNA concentrations and quality using Qubit 2.0 Fluorometer (dsDNA, HS, Invitrogen, Carlsbad, California, USA) and ND1000 spectrophotometer (Thermo Fisher, Waltham,

Massachusetts, USA), respectively. We stored DNA in RNase/DNase-free water (Sigma-Aldrich, MO) at  $-20^{\circ}\text{C}$  until library preparation. RAD library preparation and single digest RAD-sequencing (restriction enzyme SbfI) using a 100-bp single-end Illumina

HiSeq2000 platform were conducted by Floragenex (Oregon, USA) using the protocol of Baird et al. (2008). Briefly, sequencing adapters and multiplex sample indices (barcodes) were ligated to SbfI digested gDNA, and fragments were selected for a size of 200–400 bp before sequencing. To segregate sequences by individual barcodes (i.e., sample indices), we demultiplexed the FASTQ sequence file and simultaneously trimmed all sequences to 85 bp using *process\_radtags* module of STACKS v. 1.40 (Catchen et al., 2013).

## 2.2 | Genotyping and filtering

We mapped the 85 bp sequenced reads to the *L. stagnalis* reference genome (GCA\_900036025.1) using Bowtie2 v. 2/2.2.3 (Langmead & Salzberg, 2012), followed by genotype calling using FreeBayes v. 1.0.0 (Garrison & Marth, 2012). We used the dDocent pipeline (Puritz et al., 2014) with minor changes to filter variant calls from the raw vcf file (Table S1). We applied a minimum mean depth filter of 10 and a maximum mean depth filter of 550 using vcftools v.0.1.16 (Danecek et al., 2011). We removed indels, kept only biallelic sites, and restricted the data to variants called in 80% of individuals across the populations. We excluded 17 RAD loci (85 bp sequenced reads) containing more than five SNPs and removed sites that were present less than five times in the overall dataset to reduce erroneous calls. This resulted in a vcf file with 4054 SNPs across populations (average mean SNP depth: 252.8; average mean individual depth: 252.5). This “non-thinned dataset” kept as many loci as possible and was used to analyze the genetic diversity of the populations (heterozygosity, nucleotide diversity, number of private alleles). We created another dataset in which we randomly kept one SNP per RAD locus to avoid the inclusion of physically linked markers in analyses of population structure and relatedness that could be biased from this (see Table S1 for detailed filtering steps). This “thinned dataset” consisted of 2263 SNPs across populations and was used to analyze population structure ( $F_{ST}$ , PCA), genetic drift estimates (linkage disequilibrium and Tajima's  $D$ ), and mating system estimates (inbreeding, relatedness, and selfing rates).

## 2.3 | Population genetic diversity and differentiation

We examined genetic diversity in the studied snail populations using the nonthinned dataset. We calculated nucleotide diversity ( $\pi$ , the average number of pairwise nucleotide differences between sequences in a sample (Nei & Li, 1979)) for each population using vcftools and corrected for the total number of sequenced nucleotides. We then calculated the number of polymorphic sites, the number of private alleles, expected heterozygosity ( $H_e$ ), and observed heterozygosity ( $H_o$ ) using the “populations” module of the STACKS software v.2.41 (Catchen et al., 2013; Rochette et al., 2019). Then, we examined genetic differentiation among populations using the

thinned dataset. We conducted a principal components analysis (PCA) using snprrelate (Zheng et al., 2012). We also calculated fixation index  $F_{ST}$  using pairwise AMOVA directly from the output of the population script of the STACKS v.2.41. (Weir, 1996). A locus was included in the pairwise  $F_{ST}$  analyses only if it was present in at least 85% of individuals within a population and in at least 50% of individuals across populations. This filtering resulted in the exclusion of three of the 2263 loci. We then calculated  $F_{ST}$  values using the “populations” module of the STACKS software v.2.41.

## 2.4 | Estimates of genetic drift and historical inbreeding


To approximate the rate of genetic drift and potential historical inbreeding due to founder events, we first estimated effective population size ( $N_e$ ) using the nonthinned dataset with the  $LDN_e$  (linkage disequilibrium) method (Waples & Do, 2008) in NeEstimator v2.1 (Do et al., 2014). We analyzed each population separately, keeping only polymorphic loci for each population. We estimated confidence intervals for  $N_e$  values using a parametric approach implemented in the program. We used a “two allele” minimum for each locus within each population based on recommendations from Waples and Do (2010) for sample sizes smaller than 25. We also tested the effect of minor allele filtering (MAF) on  $LDN_e$  estimates using the settings in NeEstimator v2.1. We tested MAFs of 0.05 and no filter (0) per population. The  $N_e$  estimates did not change between the filters. The 95% CI was slightly larger for the 0.05 filter compared to no filter (0). We thus report the results of no additional MAF per population. Second, using the non-thinned dataset, we compared the observed number of segregating sites and nucleotide diversity ( $\pi$ ) using Tajima's  $D$  test (Tajima, 1989), which informs about deviations from neutral expectations. Positive values indicate a deficiency of rare alleles, while negative values suggest an excess of rare alleles relative to neutral expectations. We calculated Tajima's  $D$  in 10-kb sliding windows using VCFtools v. 0.1.16 (Danecek et al., 2011) and reported the genome-wide average values across windows for each population. We aimed for a window size that is small enough to minimize recombination in the region but large enough to contain enough data. Additionally, we performed Tajima's  $D$  test on the dataset with no MAF applied to assess if minor allele filtering significantly impacted Tajima's  $D$  estimates. Lastly, to test the null hypothesis of linkage equilibrium of SNP loci, we estimated the index of association,  $I_A$  (Brown et al., 1980), that measures the pairwise allele correlations in the population. We also calculated the standard index of association ( $r_d$ ) and corresponding  $p$ -values for  $I_A$  and  $r_d$  using 1000 permutations with the “poppr” package (Kamvar et al., 2014).

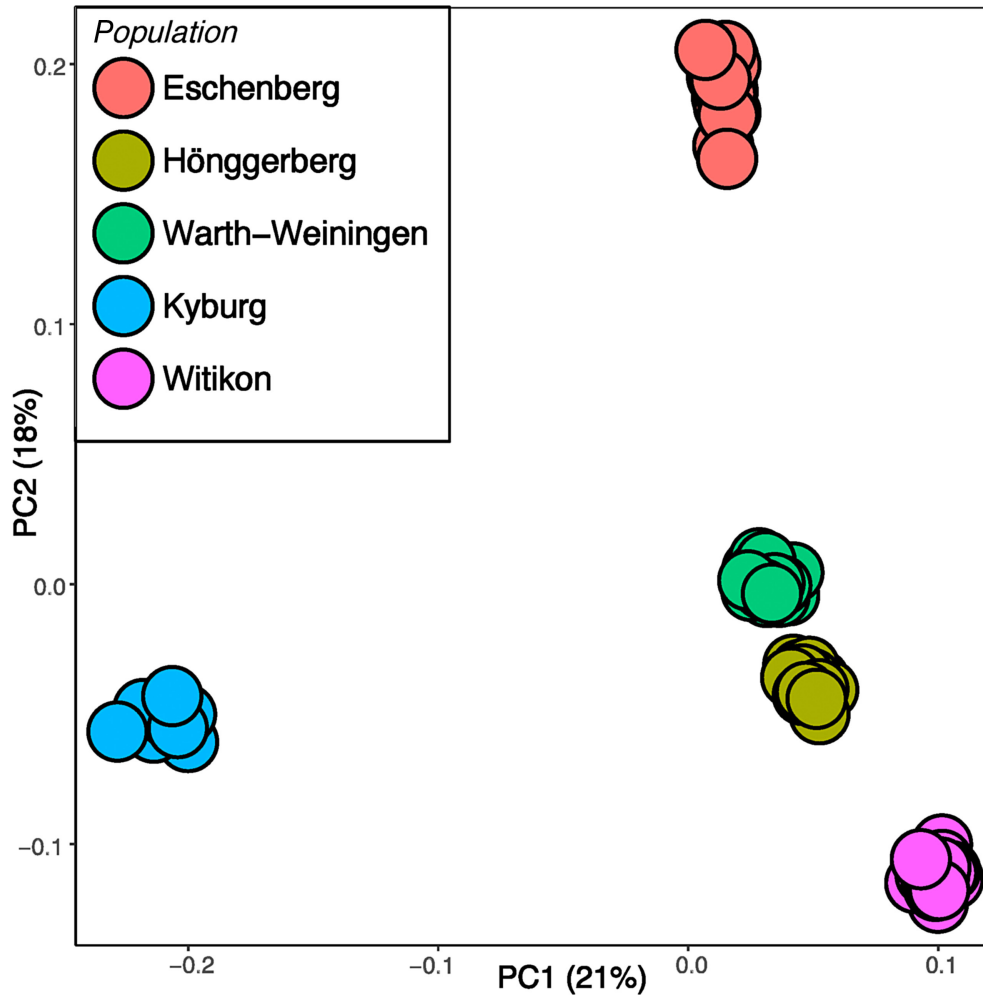
## 2.5 | Mating system estimates

To evaluate the importance of the contemporary mating system on the patterns of genetic variation, we estimated rates of inbreeding,

**TABLE 2** Pairwise  $F_{ST}$  estimates between all five populations. The degree of color intensity, as indicated by the scale bar, corresponds to increasing  $F_{ST}$  values.

Sites	Eschenberg	Hönggerberg	Warth-Weiningen	Kyburg
Hönggerberg	0.19			
Warth-Weiningen	0.19	0.14		
Kyburg	0.28	0.23	0.24	
Witikon	0.26	0.18	0.21	0.32

0.14  0.32



**FIGURE 2** A scatterplot showing individual variation in principal component (PC) scores derived from principal component analysis (PCA) of the genomic data (2263 SNPs). The amounts of variation explained by each PC axis are given as percentages.

relatedness, and self-fertilization in each population using the thinned dataset. We first estimated the inbreeding rate in five populations using the fixation index  $F_{IS}$ , which is the proportional reduction in heterozygosity due to inbreeding. We calculated the  $F_{IS}$  using the “populations” module of the STACKS software v.2.41.

We estimated selfing rates,  $s$ , from  $F_{IS}$  using the formula  $s = 2 F_{IS} / (1 + F_{IS})$  (Wright, 1969). We also estimated selfing rates,  $s(g_2)$ , using a maximum-likelihood approach based on the identity disequilibrium ( $g_2$ ) with the software Robust Multilocus Estimation of Selfing (RMES) (David et al., 2007). RMES uses the distribution of multilocus heterozygosity, which is considered more reliable compared to

estimates derived from  $F_{IS}$  (Bürkli et al., 2017; David et al., 2007). Because RMES cannot calculate  $s(g_2)$  using more than 1000 SNPs, we took five subsamples of 1000 SNPs from the data for each population and calculated  $s(g_2)$  for each of them. We report the average  $s(g_2)$  of each population. We used 1000 iterations of RMES to generate  $p$ -values for  $s(g_2)$  (Miller et al., 2014).

To evaluate the rates of bi- and uniparental inbreeding, we estimated the relatedness among all pairwise genotypes and calculated individual inbreeding coefficients,  $f$ . Relatedness is defined as the probability of identity by descent of two alleles at a given locus that two individuals share (Keller & Waller, 2002; Wright, 1922).



Population	$N_e$ (95% CI)	Tajima's $D$	$I_A$ ( $p$ value)	$r_d$ ( $p$ value)
Eschenberg	91 (65–149)	0.99	<b>3.37 (.001)</b>	<b>0.0031 (.001)</b>
Hönggerberg	195 (97–4763)	0.63	<b>2.06 (.001)</b>	<b>0.0015 (.001)</b>
Warth-Weiningen	231 (105–infinite)	0.81	1.94 (.098)	0.0015 (.081)
Kyburg	90 (57–195)	0.96	<b>1.69 (.002)</b>	<b>0.0021 (.002)</b>
Witikon	41 (29–65)	1.19	<b>5.66 (.001)</b>	<b>0.0064 (.001)</b>

Note: Effective population sizes ( $N_e$ ) and their associated 95% confidence intervals (CI), Tajima's  $D$ , and estimates of linkage disequilibrium (the index of association,  $I_A$ , based on pairwise allele correlations, and the index of association standardized for the number of loci,  $r_d$ ) are shown. Significant estimates ( $p < .05$ ) for  $I_A$ , and  $r_d$  are shown in bold.

We estimated relatedness using the triadic likelihood estimator (TrioML) of Wang (2007) with the R package “related” v.1 (Pew et al., 2015). This estimator is considered more robust toward inbreeding and bias due to the small sample size and genotyping errors compared to the five commonly used moment estimators and the dyadic likelihood method (Wang, 2007). We used the settings “unknown allele frequencies” (i.e., allele frequencies calculated directly from the genetic data) and “inbreeding.” We also tested whether these estimates differed among populations. We used individual TrioML inbreeding and pairwise TrioML relatedness estimates as response variables and population as a factor in a linear model in SPSS 26.

## 2.6 | Relationships between genetic diversity, genetic drift, and mating system estimates

We estimated the determinants of genetic diversity in the examined populations by calculating Pearson correlations among nucleotide diversity ( $\pi$ ) and environmental characteristics of ponds (elevation and area of the ponds) and the most relevant measures for drift ( $N_e$  and Tajima's  $D$ ) and mating type (selfing rates) using SPSS v. 26. Because relatedness ( $r$ ) and  $F_{IS}$  can be influenced by other factors such as biparental inbreeding and population size, we did not use them in the correlations with genetic diversity (Bierne et al., 2000; Szulkin et al., 2010). Additionally, we examined possible correlations between the explanatory variables before examining their correlations with genetic diversity. The variables representing drift and mating type did not covary (Table S2).

## 3 | RESULTS

### 3.1 | Population genetic diversity and differentiation

Populations differed in their observed heterozygosity (range: 0.09–0.15), the number of polymorphic sites (range: 1454–2685 out of 4054 SNPs), nucleotide diversity ( $\pi$ ) (range: 0.0024–0–0038), and the number of private alleles (range: 90–229) (Table 1).

TABLE 3 Estimates for genetic drift based on the thinned dataset.

All populations had lower observed than expected heterozygosity (Table 1). Populations of Hönggerberg and Warth-Weiningen were the most genetically diverse, having almost twice as many polymorphic sites as the least diverse population, Kyburg. Although Kyburg and Witikon were the least diverse populations, the number of private alleles in them differed more than twofold from each other (Table 1). The elevation and area of the ponds did not covary with their genetic diversity (Table 1,  $r = -.86$ ;  $p = .14$  and  $r = -.44$ ;  $p = .56$ , respectively).

Examined populations were clearly differentiated from each other based on pairwise  $F_{ST}$  estimates (overall  $F_{ST} = 0.17$ , Table 2). PCA resolved five distinct and nonoverlapping clusters separating the populations, the first two principal components explaining 21% and 18% of the total variation (Figure 2). The genetic differentiation of populations showed no obvious spatial pattern; even the closest populations, Kyburg and Eschenberg, exhibited large genetic differences in allele frequencies (pairwise  $F_{ST} = 0.28$ , Table 2, Figure 2). The most genetically diverse populations, Hönggerberg and Warth-Weiningen, were genetically most similar (pairwise  $F_{ST} = 0.14$ , Table 2, Figure 2). In contrast, the genetic differentiation between the two least diverse populations, Kyburg and Witikon, was the highest (pairwise  $F_{ST} = 0.32$ , Table 2, Figure 2).

### 3.2 | Estimates of genetic drift and historical inbreeding

Effective population size ( $N_e$ ) estimates of the examined populations were significantly different from each other, ranging from 40 to 231 (Table 3). The highest  $N_e$  estimates were obtained for Warth-Weiningen and Hönggerberg, while the lowest was for Witikon (Table 3). The effective population sizes of Kyburg and Eschenberg were intermediate (Table 3). All populations had a positive Tajima's  $D$  value, even without applying any MAF (Table 3, Table S3). For all populations except Warth-Weiningen, linkage disequilibrium ( $r_d$ ) differed significantly from zero (Table 3). Genetically most diverse populations, Warth-Weiningen and Hönggerberg, had lower  $r_d$  compared to other populations (Table 3). Populations'  $N_e$  estimates and Tajima's  $D$  values covaried with their genetic diversity ( $r = .98$ ;  $p = .02$  and  $r = -.95$ ;  $p = .05$ , respectively, Table S2).

### 3.3 | Mating system estimates

The two most diverse populations, Hönningerberg and Warth-Weiningen, also had the highest population level inbreeding estimates ( $F_{IS}$ ) (Table 4). In comparison, the two least diverse populations, Kyburg and Witikon, had the lowest  $F_{IS}$  values (Table 4). Selfing rate estimates based on  $F_{IS}$  values,  $s(F_{IS})$ , ranged between 0.18 and 0.31 (Table 4). However, selfing rate estimates based on identity disequilibrium,  $s(g_2)$ , are more suited for multilocus data ( $F_{IS}$  averages across the markers and could be inflated by linkage disequilibrium caused by other factors than selfing (Bürkli et al., 2017; David et al., 2007)). The obtained  $s(g_2)$  estimates were lower than  $s(F_{IS})$  values but significantly higher than zero for all populations except Hönningerberg (Table 4), ranging between 0.04 and 0.10 (Table 4). There was no significant correlation between selfing rates and population genetic diversity ( $r = -.72$ ;  $p = .28$ , Table S2) nor among variables representing drift and nonrandom mating (Table S2).

Using polymorphic sites in each population as a reference, mean relatedness values,  $r$ , were low, ranging from 0.04 to 0.05 (Table 4), but still significantly different among populations ( $p < .01$ ). More than half (55%) of all pairwise  $r$  values within the populations exhibited a level of relatedness characteristic of first cousins (0.125) to second cousins (0.03125) (Figure 3). Mean individual level inbreeding coefficients,  $f$ , of populations were high overall (mean value of 0.26) and did not significantly differ among them ( $p = .26$ ). The distributions of relatedness and inbreeding estimates for individuals of each population are given in Figure 3 and Figure S1, respectively.

## 4 | DISCUSSION

We examined the importance of genetic drift, including historical inbreeding and the contemporary mating system, in determining the genetic diversity and structure in populations of the freshwater snail, *L. stagnalis*, inhabiting ponds. We found a strong population genetic structure among five geographically closely located populations. The genetic differentiation of populations showed no obvious spatial pattern; even nearby populations (a few km apart) exhibited large genetic differences in allele frequencies. Patterns in genetic variation suggested genetic drift and historical inbreeding (e.g., due

to founder events and/or small population size) in the examined populations observed as (1) large variation in  $N_e$ , (2) a positive correlation between  $N_e$  and genetic variation, (3) the absence of rare alleles indicated by positive Tajima's  $D$  estimates, and (4) a negative correlation between Tajima's  $D$  estimates and genetic variation of the populations. In contrast, we did not find evidence for a strong effect of the contemporary mating system on the genetic diversity and structure because of (1) low selfing rates and low relatedness estimates in all populations and (2) the lack of a correlation between selfing rates and genetic variation of the populations. Our analyses suggest that populations of *L. stagnalis* inhabiting ponds likely exhibit extinction/(re)colonization dynamics and that their populations are shaped by founder events with limited influence of their contemporary mating system.

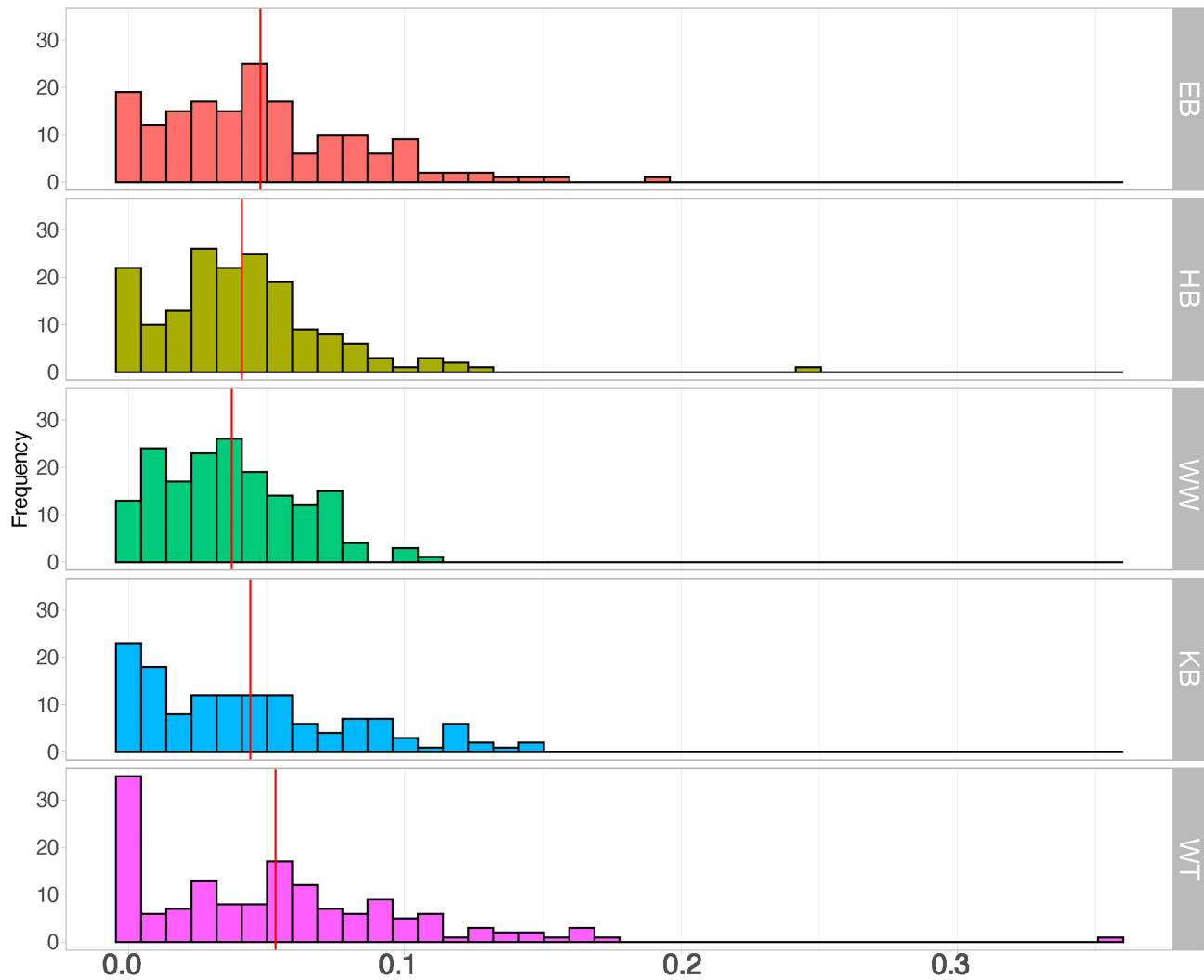
Strong population genetic structure is frequently reported in previous empirical studies on species with high rates of selfing or biparental inbreeding (Duminil et al., 2007; Jarne, 1995; Jullien et al., 2019). The observed large variation in genetic diversity and strong differentiation among populations align with earlier studies examining *L. stagnalis* populations using microsatellite markers in Switzerland, Finland, and New Zealand (Kopp et al., 2012; Puurtinen et al., 2004). Furthermore, covariation between the strength of signals related to genetic drift (e.g.,  $N_e$  estimates and Tajima's  $D$  values) and genetic variation estimates of populations suggests that larger populations are less likely to experience local extinctions than small populations (Hill, 1981; Laurie-Ahlberg & Weir, 1979). On the other hand, we did not find evidence for contemporary self-fertilization driving the differences in diversity among the studied populations because selfing rates were low and did not covary with the genetic diversity estimates of the populations. For example, the most diverse populations (Hönningerberg and Warth-Weiningen) had similar selfing rates to the least diverse ones (Kyburg and Witikon). However, self-fertilization during the founding of populations might have strengthened genetic differentiation among populations (Ingvarsson, 1998, 2002). This is because, in self-compatible hermaphrodites, self-fertilization may further reduce the  $N_e$  by decreasing the number of independent gametes (Burgess et al., 2019; Jarne, 1995).

Although we found signs of nonrandom mating at the population level, our analysis did not reveal evidence for strong biparental inbreeding. This is because of low and similar mean values

TABLE 4 Estimates for nonrandom mating based on the thinned dataset.

Population	$F_{IS}$	$s(F_{IS})$	$g_2$	$s(g_2)$	$f$	$r$
Eschenberg	0.14	0.25	0.02	<b>0.07</b>	0.25	.05
Hönningerberg	0.17	0.29	0.01	0.04	0.26	.04
Warth-Weiningen	0.18	0.31	0.02	<b>0.08</b>	0.28	.04
Kyburg	0.11	0.20	0.02	<b>0.08</b>	0.28	.04
Witikon	0.10	0.18	0.03	<b>0.10</b>	0.23	.05

Note: The table presents the inbreeding coefficient ( $F_{IS}$ ) and selfing rate for each population, derived from  $s(F_{IS})$ . It also shows the selfing rate,  $s$ , which is calculated using the identity disequilibrium ( $g_2$ , which is used to estimate  $s(g_2)$ ). Mean inbreeding ( $f$ ) and relatedness ( $r$ ) are also provided for each population, calculated using within-population allele frequencies. Significant estimates ( $p < .05$ ) for  $s(g_2)$  are shown in bold.



**FIGURE 3** Pairwise relatedness estimates from the probability-based estimator TrioML (Wang, 2007) performed on each of the five populations separately, using within-population allele frequencies in corresponding populations. The red lines indicate the mean values for each population.

of pairwise relatedness estimates among populations. However, occasional self-fertilization most probably leads to deviations from Hardy–Weinberg equilibrium in all populations, observed as excess homozygosity in this study. An alternative explanation for excess homozygosity could be within-population structure (i.e., Wahlund effect). This effect arises when the sampled population is composed of subpopulations with varying allele frequencies. However, because there is no visible substructure within populations based on pairwise relatedness distributions and the PCA results, the Wahlund effect is unlikely to influence our study populations. Previous studies on *L. stagnalis* populations mostly found limited inbreeding depression (Coutellec & Lagadic, 2006; Koene et al., 2008; Puurtinen et al., 2007), but see Leicht et al. (2019) and Seppälä and Langeloh (2016), suggesting that recessive deleterious mutations may have been purged by population bottlenecks and founder events. However, these predictions have not previously been validated by genome-wide markers. We suggest that further research employing genome-wide analyses investigate

the interplay between inbreeding, genetic diversity, demographic history, and life-history traits in small populations experiencing extinction-recolonization dynamics.

However, because we only examined five populations, the lack of significant correlations between mating system estimates and genetic diversity could be due to low statistical power. Furthermore, there are potential weaknesses in the SNP data when estimating genetic diversity and Tajima's *D*. This is because of systematic biases such as allelic dropouts, which might lead to decreased diversity and increased population differentiation estimates (Andrews et al., 2016; Arnold et al., 2013). We minimized these issues using a very high sequencing depth (mean site depth: 253), low mean missing data (1%), and the same sequencing and filtering criteria for all populations (Arnold et al., 2013). It is also important to note that rare alleles can introduce some bias in LD estimates (Do et al., 2014; Wang et al., 2016). However, we expect there to be little bias in  $LDN_e$  estimates due to rare alleles. This is because (1) our populations already exhibit a general deficit of rare alleles even before applying any minor



allele filtering; (2) we used the same filtering thresholds for all of our populations; and (3) we tested the effect of minor allele filtering on  $LDN_e$  estimates and there was no difference in the results.

Limited gene flow in *L. stagnalis* populations, potentially caused by extinction-recolonization cycles, likely explains the observed deficit of rare alleles (Wakeley, 2004). Self-fertilization during founder events and/or occasional biparental inbreeding and self-fertilization might further accentuate this effect (Bousset et al., 2004; Kopp et al., 2012; Puurtinen et al., 2004). While genetic drift may lead to lower evolutionary potential to respond to environmental change (Hoffmann et al., 2017), it can also transiently increase adaptive potential by epistatic or dominance variance to additive genetic variance by fixing common alleles in populations (Barton & Turelli, 2004). However, in the long term, repeated founder events and inbreeding are expected to slow down responses to directional selection due to decreased genetic variation. This aligns with findings from other freshwater snails (Escobar et al., 2011; Jarne et al., 2021; Lamy et al., 2012) and even organisms with larger populations (Sin et al., 2021), suggesting that demographic history can be a major driver of genetic diversity and inbreeding across species.

Understanding the timescales of genetic signals related to mating systems (recent) and genetic drift (historical) is crucial for interpreting their ecological significance. Some metrics, like identity disequilibrium, relatedness, and  $F_{IS}$ , reflect recent mating patterns, typically corresponding to the last few generations. Conversely, others, such as  $\pi$  or Tajima's  $D$ , provide insights into longer-term trends as they capture genetic variation accumulated over longer timescales, potentially including past inbreeding events. For example, selfing estimates derived from identity disequilibrium (ID) offer a snapshot of contemporary breeding patterns. This is because ID focuses on the probability of encountering identical genotypes at two loci due to recent ancestry (David et al., 2007; Ritland et al., 1990). On the other hand, linkage disequilibrium decays at longer timescales through recombination; therefore, linkage disequilibrium provides information about historical processes affecting population size, including bottlenecks (Hill & Robertson, 1968; Slatkin, 2008). Tajima's  $D$  estimates also potentially reflect more historical events, which lead to differences in the frequency of the rare alleles (Tajima, 1989). Thus, while our study found no strong evidence for recent selfing or biparental inbreeding, we cannot definitely rule out their historical influence on genetic diversity. Demographic simulations that incorporate various mating system parameters and extinction/recolonization events would be very useful for gaining a more comprehensive understanding of the demographic processes shaping these populations and also potential biases arising from different timescales of genetic signals.

In conclusion, our results suggest that founder effects and genetic drift play a larger role in defining the population genetic structure of *L. stagnalis* in the studied region than the contemporary mating system. Previous studies analyzing the role of genetic drift and the mating system using genomic techniques compared different

taxa (Diaz-Martin et al., 2023; Latron et al., 2020) or only one population (Sin et al., 2021). Our results illustrate how the investigation of population genetic diversity and structure in small populations characterized by a mixed-mating system benefits from genome-wide multilocus analyses. The differences in genetic drift and diversity among populations that we report might reflect the degree to which genetic drift and/or founder events, including historical inbreeding, erode the adaptive potential of small populations (Lynch et al., 1995). However, if those populations undergo extinction/(re)colonization dynamics frequently, genetic diversity and adaptive potential would be largely determined by the source populations. Thus, investigating small ponds and more permanent habitats to identify potential source and sink populations would be necessary to understand the relevant spatial units for biodiversity conservation.

## AUTHOR CONTRIBUTIONS

**Cansu Çetin:** Formal analysis (lead); investigation (equal); methodology (equal); writing – original draft (lead); writing – review and editing (equal). **Jukka Jokela:** Resources (equal); supervision (equal); validation (equal); writing – original draft (supporting); writing – review and editing (equal). **Philine G. D. Feulner:** Formal analysis (supporting); methodology (equal); validation (equal); writing – review and editing (supporting). **Tamara Schlegel:** Methodology (supporting); writing – review and editing (supporting). **Nadine Tardent:** Methodology (supporting); writing – review and editing (supporting). **Otto Seppälä:** Conceptualization (lead); funding acquisition (lead); investigation (equal); methodology (lead); project administration (equal); resources (equal); supervision (equal); validation (equal); writing – review and editing (supporting).

## ACKNOWLEDGMENTS

We thank Rishi De-Kayne, Niklaus Zemp, and Linda Haltiner for their help in providing scripts for data analyses and discussions. We thank Natalia Zajac, Fatma Gözde Cilingir, Piotr Zielinski, and Krystyna Nadachowska-Brzyska for discussions and their feedback. We thank the editor and reviewers for helpful comments on earlier versions of this manuscript. We also thank the Genetic Diversity Centre (GDC) of ETH Zürich for their support during data analyses.

## FUNDING INFORMATION

The study was funded by the ETH research commission (grant ETH-20 17-1) and SNF grant (31003A 169531) to OS.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## OPEN RESEARCH BADGES



This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results.

## DATA AVAILABILITY STATEMENT

Raw genetic data (fastq files for all 92 individuals) are deposited on short read archive, SRA under accession numbers SAMN36306817-SAMN36306908. Genotypes (vcf format) are available at the dryad repository <https://datadryad.org/stash/share/Of1xul7HzOpe2yJ3il1o5GuvEzGVhLr9HwIMBNx5g>.

## ORCID

Cansu Çetin  <https://orcid.org/0000-0001-6514-5858>

## REFERENCES

- Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews. Genetics*, 17(2), 81–92.
- Arnold, B., Corbett-Detig, R. B., Hartl, D., & Bomblies, K. (2013). RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology*, 22(11), 3179–3190.
- Attard, C. R. M., Beheregaray, L. B., & Möller, L. M. (2018). Genotyping-by-sequencing for estimating relatedness in nonmodel organisms: Avoiding the trap of precise bias. *Molecular Ecology Resources*, 18(3), 381–390.
- Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, 3(10), e3376.
- Barton, N. H., & Turelli, M. (2004). Effects of genetic drift on variance components under a general model of epistasis. *Evolution; International Journal of Organic Evolution*, 58(10), 2111–2132.
- Bierne, N., Tsitroni, A., & David, P. (2000). An inbreeding model of associative overdominance during a population bottleneck. *Genetics*, 155(4), 1981–1990.
- Bousset, L., Henry, P.-Y., Sourrouille, P., & Jarne, P. (2004). Population biology of the invasive freshwater snail *Physa acuta* approached through genetic markers, ecological characterization and demography. *Molecular Ecology*, 13(7), 2023–2036.
- Brown, A. H., Feldman, M. W., & Nevo, E. (1980). Multilocus structure of natural populations of *Hordeum spontaneum*. *Genetics*, 96(2), 523–536.
- Burgess, S. C., Sander, L., & Bueno, M. (2019). How relatedness between mates influences reproductive success: An experimental analysis of self-fertilization and biparental inbreeding in a marine bryozoan. *Ecology and Evolution*, 9(19), 11353–11366.
- 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), 525–533.
- Byrne, K., & Nichols, R. A. (1999). *Culex pipiens* in London underground tunnels: Differentiation between surface and subterranean populations. *Heredity*, 82(Pt 1), 7–15.
- Cain, G. L. (1956). Studies on cross-fertilization and self-fertilization in *Lymnaea stagnalis appressa* say. *The Biological Bulletin*, 111(1), 45–52.
- 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., & Charlesworth, D. (2010). *Elements of evolutionary genetics*. Roberts & Company.
- Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics*, 134(4), 1289–1303.
- Coutellec, M.-A., & Lagadic, L. (2006). Effects of self-fertilization, environmental stress and exposure to xenobiotics on fitness-related traits of the freshwater snail *Lymnaea stagnalis*. *Ecotoxicology (London, England)*, 15(2), 199–213.
- Coutellec-Vreto, M. A., Guiller, A., & Daguzan, J. (1994). Allozyme variation in some populations of the freshwater snails *Lymnaea peregra*, *L. auricularia* and *L. stagnalis* (Gastropoda: Pulmonata). *The Journal of Molluscan Studies*, 60(4), 393–403.
- Crnokrak, P., & Roff, D. A. (1999). Inbreeding depression in the wild. *Heredity*, 83(Pt 3), 260–270.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158.
- David, P., Pujol, B., Viard, F., Castella, V., & Goudet, J. (2007). Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology*, 16(12), 2474–2487.
- DeRose, M. A., & Roff, D. A. (1999). A comparison of inbreeding depression in life-history and morphological traits in animals. *Evolution; International Journal of Organic Evolution*, 53(4), 1288–1292.
- Diaz-Martin, Z., Cisternas-Fuentes, A., Kay, K. M., Raguso, R. A., Skogen, K., & Fant, J. (2023). Reproductive strategies and their consequences for divergence, gene flow, and genetic diversity in three taxa of clarkia. *Heredity*, 131(5–6), 338–349.
- 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 procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19, 11–15.
- 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.
- Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews. Genetics*, 17(7), 422–433.
- Escobar, J. S., Auld, J. R., Correa, A. C., Alonso, J. M., Bony, Y. K., Coutellec, M. A., Koene, J. M., Pointier, J. P., Jarne, P., & David, P. (2011). Patterns of mating-system evolution in hermaphroditic animals: Correlations among selfing rate, inbreeding depression, and the timing of reproduction. *Evolution; International Journal of Organic Evolution*, 65(5), 1233–1253.
- Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv [q-bio.GN]*. <http://arxiv.org/abs/1207.3907>
- Henry, P. Y., Pradel, R., & Jarne, P. (2003). Environment-dependent inbreeding depression in a hermaphroditic freshwater snail. *Journal of Evolutionary Biology*, 16(6), 1211–1222.
- Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetics Research*, 38(3), 209–216.
- Hill, W. G., & Robertson, A. (1968). Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics*, 38(6), 226–231.
- Hoffmann, A. A., Sgrò, C. M., & Kristensen, T. N. (2017). Revisiting adaptive potential, population size, and conservation. *Trends in Ecology & Evolution*, 32(7), 506–517.
- Hohenlohe, P. A., Phillips, P. C., & Cresko, W. A. (2010). Using population genomics to detect selection in natural populations: Key concepts and methodological considerations. *International Journal of Plant Sciences*, 171(9), 1059–1071.
- Ingvarsson, P. K. (1998). Kin-structured colonization in *Phalacrus substriatus*. *Heredity*, 80(4), 456–463.
- Ingvarsson, P. K. (2002). A metapopulation perspective on genetic diversity and differentiation in partially self-fertilizing plants. *Evolution; International Journal of Organic Evolution*, 56(12), 2368–2373.

- Jarne, P. (1995). Mating system, bottlenecks and genetic polymorphism in hermaphroditic animals. *Genetics Research*, 65(3), 193–207.
- Jarne, P., Lozano del Campo, A., Lamy, T., Chapuis, E., Dubart, M., Segard, A., Canard, E., Pointier, J.-P., & David, P. (2021). Connectivity and selfing drives population genetic structure in a patchy landscape: A comparative approach of four co-occurring freshwater snail species. *Peer Community Journal*, 1, e21.
- Jullien, M., Navascués, M., Ronfort, J., Loridon, K., & Gay, L. (2019). Structure of multilocus genetic diversity in predominantly selfing populations. *Heredity*, 123(2), 176–191.
- 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.
- Kardos, M., Luikart, G., & Allendorf, F. W. (2015). Measuring individual inbreeding in the age of genomics: Marker-based measures are better than pedigrees. *Heredity*, 115(1), 63–72.
- Keller, L., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, 17(5), 230–241.
- Koene, J. M., Loose, M. J., & Wolters, L. (2008). Mate choice is not affected by mating history in the simultaneously hermaphroditic snail *Lymnaea stagnalis*. *The Journal of Molluscan Studies*, 74(4), 331–335.
- Kopp, K. C., Wolff, K., & Jokela, J. (2012). Natural range expansion and human-assisted introduction leave different genetic signatures in a hermaphroditic freshwater snail. *Evolutionary Ecology*, 26(3), 483–498.
- Lamy, T., Lévy, L., Pointier, J. P., Jarne, P., & David, P. (2012). Does life in unstable environments favour facultative selfing? A case study in the freshwater snail *Drepanotrema depressissimum* (Basommatophora: Planorbidae). *Evolutionary Ecology*, 26(3), 639–655.
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nature Methods*, 9(4), 357–359.
- Latron, M., Arnaud, J.-F., Ferla, H., Godé, C., & Duputié, A. (2020). Effects of contemporary shifts of range margins on patterns of genetic structure and mating system in two coastal plant species. *Heredity*, 124(2), 336–350.
- Laurie-Ahlberg, C. C., & Weir, B. S. (1979). Allozymic variation and linkage disequilibrium in some laboratory populations of *Drosophila melanogaster*. *Genetics*, 92(4), 1295–1314.
- Leicht, K., Jokela, J., & Seppälä, O. (2019). Inbreeding does not alter the response to an experimental heat wave in a freshwater snail. *PLoS One*, 14(8), e0220669.
- Lemopoulos, A., Prokkola, J. M., Uusi-Heikkilä, S., Vasemägi, A., Huusko, A., Hyvärinen, P., Koljonen, M. L., Koskiniemi, J., & Vainikka, A. (2019). Comparing RADseq and microsatellites for estimating genetic diversity and relatedness – Implications for brown trout conservation. *Ecology and Evolution*, 9(4), 2106–2120.
- Lynch, M., Conery, J., & Burger, R. (1995). Mutation accumulation and the extinction of small populations. *The American Naturalist*, 146(4), 489–518.
- Miller, J. M., Malenfant, R. M., David, P., Davis, C. S., Poissant, J., Hogg, J. T., Festa-Bianchet, M., & Coltman, D. W. (2014). Estimating genome-wide heterozygosity: Effects of demographic history and marker type. *Heredity*, 112(3), 240–247.
- Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*, 76(10), 5269–5273.
- Noël, E., Jarne, P., Glémin, S., MacKenzie, A., Segard, A., Sarda, V., & David, P. (2017). Experimental evidence for the negative effects of self-fertilization on the adaptive potential of populations. *Current Biology: CB*, 27(2), 237–242.
- Pannell, J. R., & Charlesworth, B. (2000). Effects of metapopulation processes on measures of genetic diversity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 355(1404), 1851–1864.
- Pew, J., Muir, P. H., Wang, J., & Frasier, T. R. (2015). Related: An R package for analysing pairwise relatedness from codominant molecular markers. *Molecular Ecology Resources*, 15(3), 557–561.
- Puritz, J. B., Hollenbeck, C. M., & Gold, J. R. (2014). dDocent: A RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ*, 2, e431.
- Puurtinen, M., Knott, K. E., Suonpää, S., Nissinen, K., & Kaitala, V. (2007). Predominance of outcrossing in *Lymnaea stagnalis* despite low apparent fitness costs of self-fertilization. *Journal of Evolutionary Biology*, 20(3), 901–912.
- Puurtinen, M., Knott, K. E., Suonpää, S., van Ooik, T., & Kaitala, V. (2004). Genetic variability and drift load in populations of an aquatic snail. *Evolution; International Journal of Organic Evolution*, 58(4), 749–756.
- Ritland, K., & Jain, S. (1984). A comparative study of floral and electrophoretic variation with life history variation in *Limnanthes alba* (Limnanthaceae). *Oecologia*, 63(2), 243–251.
- Ritland, K., Soltis, D. E., & Soltis, P. S. (1990). A two-locus model for the joint estimation of intergametophytic and intragametophytic selfing rates. *Heredity*, 65(3), 289–296.
- 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.
- Seppälä, O., & Langeloh, L. (2016). Estimating genetic and maternal effects determining variation in immune function of a mixed-mating snail. *PLoS One*, 11(8), 1–14.
- Sin, S. Y. W., Hoover, B. A., Nevitt, G. A., & Edwards, S. V. (2021). Demographic history, not mating system, explains signatures of inbreeding and inbreeding depression in a large outbred population. *The American Naturalist*, 197(6), 658–676.
- Slatkin, M. (2008). Linkage disequilibrium – Understanding the evolutionary past and mapping the medical future. *Nature Reviews. Genetics*, 9(6), 477–485.
- Szulkin, M., Bierne, N., & David, P. (2010). Heterozygosity-fitness correlations: A time for reappraisal. *Evolution; International Journal of Organic Evolution*, 64(5), 1202–1217.
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123, 585–595.
- Wakeley, J. (2004). Metapopulations and coalescent theory. In I. Hanski & O. Gaggiotti (Eds.), *Ecology, genetics and evolution of metapopulations* (pp. 175–198). Elsevier.
- Wang, J. (2007). Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetical Research*, 89(3), 135–153.
- Wang, J. (2014). Marker-based estimates of relatedness and inbreeding coefficients: An assessment of current methods. *Journal of Evolutionary Biology*, 27(3), 518–530.
- Wang, J., Santiago, E., & Caballero, A. (2016). Prediction and estimation of effective population size. *Heredity*, 117(4), 193–206.
- Waples, R. S., & Do, C. (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. (2010). Linkage disequilibrium estimates of contemporary  $N_e$  using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3(3), 244–262.
- Weir, B. S. (1996). *Genetic data analysis II: Methods for discrete population genetic data*. Sinauer.
- Wright, S. (1922). Coefficients of inbreeding and relationship. *The American Naturalist*, 56(645), 330–338.
- Wright, S. (1931). Evolution in mendelian populations. *Genetics*, 16(3), 97–159.

- Wright, S. (1969). *Evolution and the genetics of populations: The theory of gene frequency*, Vol. 2. University of Chicago Press.
- Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, 28, 3326–3328.

#### 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:** Çetin, C., Jokela, J., Feulner, P. G. D., Schlegel, T., Tardent, N., & Seppälä, O. (2024). Population genetic structure in a self-compatible hermaphroditic snail is driven by drift independently of its contemporary mating system. *Ecology and Evolution*, 14, e70162. <https://doi.org/10.1002/ece3.70162>