MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 1394–1410 doi: 10.1111/j.1365-294X.2012.05478.x

Testing metapopulation dynamics using genetic, demographic and ecological data

T. LAMY,* J. P. POINTIER,† P. JARNE* and P. DAVID*

**UMR 5175 CEFE, campus CNRS, 1919 route de Mende, 34293 Montpellier cedex 5, France, †USR 3278 CNRS-EPHE, 52 avenue Paul Alduy, 66860 Perpignan cedex, France*

Abstract

The metapopulation concept is a cornerstone in the recent history of ecology and evolution. However, determining whether a natural system fits a metapopulation model is a complex issue. Extinction-colonization dynamics are indeed often difficult to quantify because species detectability is not always 100%, resulting in an imperfect record of extinctions. Here, we explore whether combining population genetics with demographic and ecological surveys can yield more realistic estimates of metapopulation dynamics. We apply this approach to the freshwater snail *Drepanotrema depressissimum* in a fragmented landscape of tropical ponds. In addition to studying correlations between genetic diversity and demographical or ecological characteristics, we undertake, for the first time, a detailed search for genetic signatures of extinction–recolonization events using temporal changes in allele frequencies within sites. Surprisingly, genetic data indicate that extinction is much rarer than suggested by demographic surveys. Consequently, this system is better described as a set of populations with different sizes and immigration rates than as a true metapopulation. We identify several cases of apparent extinction owing to nondetection of low-density populations, and of aestivating individuals in desiccated ponds. More generally, we observed a frequent mismatch between genetic and demographical/ecological information at small spatial and temporal scales. We discuss the causes of these discrepancies and show how these two types of data provide complementary information on population dynamics and history, especially when temporal genetic samples are available.

Keywords: colonization, dispersal, effective size, extinction, genetic structure, metapopulation, seed bank, snails

Received 22 July 2011; revision revised 21 December 2011; accepted 30 December 2011

Introduction

The metapopulation concept is central in modern ecological and evolutionary literature (Hanski & Gaggiotti 2004). According to the original definition (Levins 1969), a metapopulation is composed of separate subpopulations that have limited lifespans, and its dynamics depends on a balance between extinction and colonization. This has deep influences on demography, evolution and community dynamics. For instance, it affects the degree of genetic variation (Whitlock & Barton 1997), it increases stochasticity in invasion dynamics (Facon & David 2006), it influences the evolution of dispersal (Olivieri *et al.* 1995), dormancy (Rajon *et al.* 2009) and self-fertilization (Pannell & Barrett 1998) and it can allow species coexistence through colonization/competition tradeoffs (Calcagno *et al.* 2006).

Yet, we have surprisingly few indubitable examples of natural metapopulations. Although spatial fragmentation of habitats is extremely common (e.g. forest patches, butterflies living on patchily distributed plants, frogs inhabiting ponds, fishes inhabiting coral reefs, etc.), extinction-colonization cycles are less well documented. Observations, when available, consist of demographic surveys in which population presence/absence is recorded in a set of habitat patches along a time series (e.g. Hanski 1994, 1999). However, seed banks in plants

Correspondence: Thomas Lamy, Fax: +33 (0)467613336;
E-mail: thomas.lamy27@gmail.com

(e.g. Bekker *et al.* 1998) and resting stages in animals (e.g. De Stasio 1989) remain undetected in most demographic censuses, and even adult individuals may go unnoticed when their density is very low, resulting in erroneous records of extinction and recolonization. Symmetrically, true extinction events can be overlooked when recolonization occurs between surveys (i.e. the 'rescue-effect' *sensu* Hanski 1994). Intensifying demographic surveys is certainly an option to improve reliability, but systematic biases seem unavoidable (e.g. noncapturable life stages). In this study, we explore another possibility, that is, to use genetic data as an independent validation of observed extinction and colonization events. Indeed, recolonization by immigrants after a true extinction and population reconstitution from local individuals after an apparent extinction are expected to leave different genetic signatures: in the first case, the new population is derived from an external gene pool, while in the second case, it comes from the local gene pool. In principle, temporal changes in genetic structure should therefore allow to distinguish true extinctions from apparent ones. However, although many genetic studies have investigated fine-scale spatial structure (Guillot *et al.* 2009), relatively few have explored temporal changes in allele frequencies (see Luikart *et al.* 2010). Moreover, the latter focused on effective population size within supposedly isolated populations rather than on metapopulation structure. So far, the only genetic test of the metapopulation concept has been to correlate estimates of apparent population age (APA) (e.g. date of most recent apparent extinction) or habitat age (most recent perturbation) with diversity and differentiation (Whitlock 1992; McCauley *et al.* 1995; Giles & Goudet 1997; Ingvarsson *et al.* 1997; Haag *et al.* 2005): younger populations of a metapopulation tend to exhibit less diversity and more differentiation than older ones. Nevertheless, we are not aware of any study that directly tested the validity of demographic extinction records using temporal genetic analysis.

Our aim here is to show how spatial and temporal genetic data can be combined with demographic and ecological observations to test the metapopulation concept. We illustrate this approach in a snail living in fragmented and unstable aquatic habitats, and exhibiting high apparent rates of local extinction and recolonization. On this basis, this system could qualify as a textbook example of the metapopulation concept. We confront demographic and ecological observations with an analysis of spatial and temporal genetic structure to test whether apparent population extinctions or perturbations of habitat patches indeed correspond to extinctions of local gene pools. More generally, we evaluate whether metapopulation dynamics inferred from demography remains valid after genetic evaluation.

Materials and methods

Alternative models of population structure and predictions

To test the metapopulation concept, we must consider alternative, competing models of structured population, and associated predictions (Table 1). Although no natural system perfectly matches a given model, such predictions should allow identifying which model best describes a natural situation. The simplest alternative to the metapopulation model is the island model (Wright 1931), in which extinction-colonization cycles do not occur. Yet, the original island model is highly idealized because of the symmetry assumption (all demes have equal size and immigration rate). Therefore, we also consider a more realistic version that relaxes this assumption, the 'asymmetric island model'. In this model, the source of asymmetry is not the same as in a metapopulation: demes differ in size and connectivity rather than in population age. For each model (symmetric island, asymmetric island and metapopulation model), we listed predictions that can be tested by comparing genetic, ecological and demographic data. They are summarized in Table 1 and explained in more detail in the Supporting information (Appendix S1). Importantly, spatial genetic structure is not sufficient to distinguish metapopulation dynamics from alternative models. This requires studying correlations between genetic changes and ecological/demographic ones in space/time, and especially the analysis of temporal changes in allele frequencies. To summarize, stronger inferences can be drawn by observing genetic structure before and after observed extinction–recolonization events than by looking for the phantoms of past extinction and recolonization.

Species, habitats and sampling

Our work was conducted in the Grande-Terre of Guadeloupe, an Island of the Lesser Antilles (French West Indies). Grande-Terre is a plateau of about 800 km^2 harbouring many small ponds (*c.* 2000), and a few small rivers and swamp grasslands connected to mangroves. The 29 recorded species of molluscs (Pointier 2008) constitute a major fraction of the macrobenthos in these environments. Here, we focus on *Drepanotrema depressissimum* (Gastropoda: Basommatophora: Planorbidae), a very common hermaphroditic snail found mostly in small ponds with abundant aquatic vegetation. A fraction of these ponds completely dry out either yearly, or more irregularly. Sites can stay dry for up to several months, especially in the northern and eastern parts of Grande-Terre. An extensive survey was initiated in

Table 1 Expectations on spatial and temporal distribution of genetic variation and on demography in three models of subdivided populations

	Symmetric patch model	Asymmetric patch model	Metapopulation model
Symmetric	Yes	No	No
Extinction/recolonization	No	No	Yes
Spatial distribution of genetic diversity			
Genetic diversity and allelic richness	Identical in all patches	Variable	Variable
Spatial F_{ST}	Only dependent on distance	Dependent on distance and H_e	Dependent on distance and H_e
Effect of patch characteristics on F_{ST} and genetic diversity			
Connectivity	None	+ on diversity, − on F_{ST}	+ on diversity, − on F_{ST}
Size	None	+ on diversity, − on F_{ST}	+ on diversity, − on F_{ST}
Stability	None	None	+ on diversity, − on F_{ST}
Apparent population age	None	None	+ on diversity − on F_{ST}
Temporal distribution of genetic diversity			
Genetic diversity and allelic richness	No significant change between successive generations	No significant change between successive generations	No significant change or slight increase in diversity in some patches, important loss of diversity in others
Temporal F_{ST}	Equal and small; a fraction of genes in the $(t + 1)$ sample comes from the same patch at t	Variable among sites; a fraction of genes in the $(t + 1)$ sample comes from the same patch at t	A few very large values; in some patches, all genes in the $(t + 1)$ sample come from different patches at t
Effect of patch history			
Apparent extinction	None	None	+ on temporal F_{ST}; in sites with apparent extinction, genes at time $t + 1$ come from other sources than the same patch at t

The models considered are the symmetric patch model (island or stepping stone model with identical patches) and two asymmetric models, namely the asymmetric patch model (island model with differences in size or connectivity among patches and without extinction/recolonization) and the metapopulation model (with extinction/recolonization). We consider how variation (mean genetic diversity and F_{ST} over loci) is distributed and how it is affected by demographic characteristics.
Apparent extinction refers to the nondetection of species in a patch between t and $t + 1$ (null apparent density or disturbed site).

2001 with yearly visits in 244 sites distributed over the whole Grande-Terre; *D. depressissimum* was observed at least once in 214 of these sites. Each year, we repeated some visits in approximately 30 randomly chosen sites, from which we derived an estimate of species detectability [77.5%; 95% CI (73.9–81.0)]. The apparent yearly extinction rate (presence at year t, absence at $t + 1$) was 21.7% over the 2001–2011 period, and the apparent colonization rate (absence at t, presence at $t + 1$) was 19.6%. The yearly fraction of dry sites was on average 4.9%. This underestimates the true frequency as sampling is performed at the beginning of the dry season (January–February), when water level is high and molluscs are very abundant. Many sites may dry out later in the season. In short, population dynamics and habitat perturbation (drought) strongly suggest a metapopulation structure, with yearly extinction and recolonization of a significant fraction of populations. However, (i) populations at very low density can be overlooked; (ii) previous field observations suggested that this species can survive for several weeks under rocks or vegetation when sites are dry ['aestivation'; Pointier & Combes (1976)]; and (iii) populations can go extinct and be recolonized during the time interval between two visits.

We defined a spatially and temporally stratified sampling scheme for population genetic analysis (Fig. 1). (i) Twenty-five populations were sampled all over the island (mean pairwise Euclidean distance = 16.3 km); (ii) Twelve of these formed four well-separated clusters of three neighbouring sites (mean Euclidean distance within clusters = 3.1 km) allowing to explore fine-grained differentiation; and (iii) Twelve sites were repeatedly sampled in different years (two to four times). These sites were chosen to cover most of Grande-Terre and to be representative of habitat types. However, sampling was constrained by the possibility

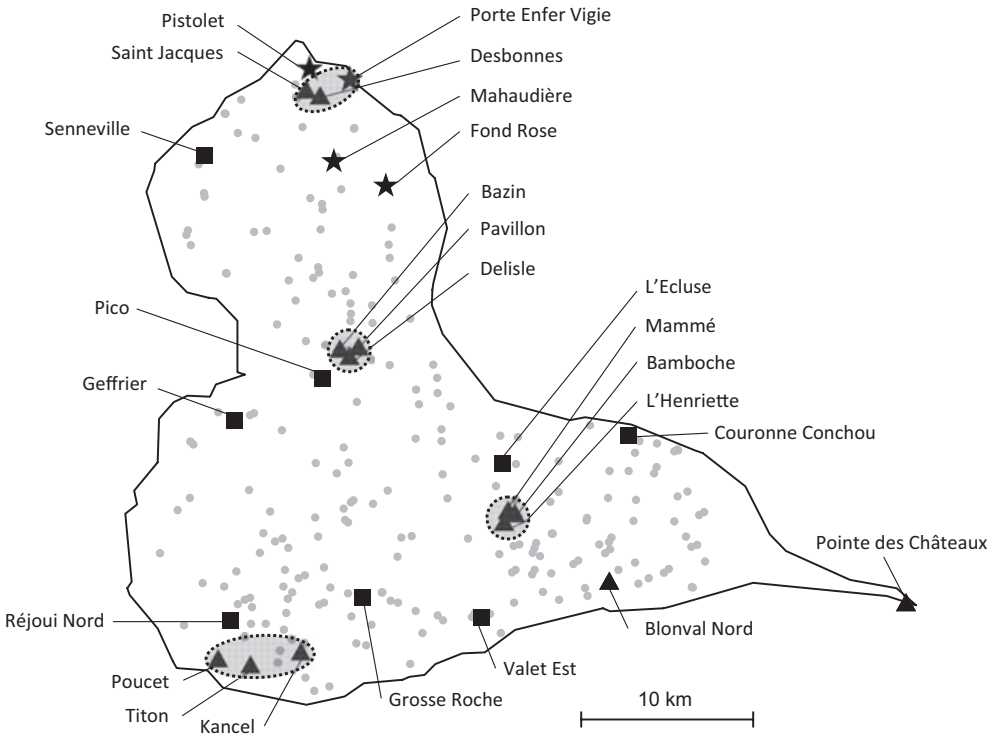

Fig. 1 Distribution of the 25 sites from Grande-Terre Island (Guadeloupe, Lesser Antilles) studied in the genetic analysis. Sites sampled several times (see Table 2) are indicated by either squares or stars (the latter are sites that were dry in 2010 and in which samples were collected in 2009 and 2011). Shaded areas correspond to clusters of three neighbouring populations. Grey dots are sites that have been surveyed annually since 2001.

of collecting large enough numbers of individuals ($N > 22$). On the whole, 42 samples (i.e. site by year; 1270 individuals) were collected (Table 2). Snails were killed in 80 °C water for 1 min and preserved in 95° ethanol prior to genetic analysis.

Demographical and ecological information on sampled sites

Yearly field surveys including (among others) all sites sampled for the genetic analysis have been performed since 2001. Each site was explored by three people for at least 15 min each (total searching time 45 min). We tried to explore all the favourable subhabitats (mostly near the shore) and consequently increased searching time in larger sites in order to walk around the site margin until every part of the shore had been explored. However, we stopped after a maximum of 30 min (total searching time 90 min), and therefore, very large sites are not as intensively explored as small ones. Snails were caught using a scoop (0.5 m) allowing to forage both the sediment and plant strata. We also visually surveyed rock surfaces or floating debris when present. A set of environmental variables was characterized at each site. For the present analysis, we retained size (pond diameter or river width) and vegetation cover (in %) (Table 3). We also assessed two permanent characteristics of sites: overall hydrological regime (five levels; from fully permanent to frequently dry) and water connectivity to neighbouring watersheds (four levels; 0: always completely isolated; 1: connected to the local watershed through water flow less than once a year; 2: connected to the local watershed on most years during the rainy season; and 3: nearly always connected to the local watershed). Hydrological regime and connectivity were estimated based on field experience, the visual aspect of water margins and topographical considerations (outlets, slopes, etc.), independently at each visit, and then averaged over years. We also defined site stability as the first axis (56% of total variance) of a principal component analysis (PCA) including hydrological regime, the proportion of visits during which the site was dry over the 2001–2011 period, the temporal variance in size of the water body, and the temporal variance in the percentage of vegetation cover during the same period. We also used detailed maps (Bruyere & Questel 2001) to estimate the number of water bodies (including those not surveyed) within a radius of 2000 m around each site. This estimated the local 'density' of favourable habitats (Table 3). It should be noted that metapopulation connectivity is often defined

Table 2 Characteristics of the 42 populations (25 sites) of *Drepanotrema depressissimum* sampled in Grande-Terre

Site	Year	n	R_A	H_e	f	$\hat{s}$
Pico†	2006	34	11.21	0.897	**0.038**	0
	2007	24	11.36	0.911	0.022	0
	2009‡	32	11.4	0.896	0.024	0.03
Grosse Roche†	2006	32	11.21	0.831	0.005	0
	2007‡	32	11.04	0.82	0.018	0
Senneville†	2006	32	11.2	0.876	−0.008	0
	2007‡	30	11.59	0.88	**0.046**	0.026
Valet Est†	2006	31	10.54	0.864	0.034	0.03
	2007	32	11.27	0.854	**0.044**	0.033
	2008‡	31	11.18	0.871	0.014	0
Geffrier†	2006	29	11.27	0.89	−0.038	0
	2007	29	12.33	0.905	**0.039**	0
	2008	32	12.07	0.899	**−0.001**	0
	2009‡	30	11.92	0.901	**0.012**	0
Réjoui Nord†	2006	31	13.82	0.897	0.031	0
	2007	31	13.67	0.908	0.034	0
	2008‡	31	14	0.919	**0.017**	0
L'Ecluse†	2006	31	12.4	0.898	0.024	**0.068**
	2009‡	30	11.75	0.884	**0.064**	0
Couronne Conchou†	2006	31	10.77	0.878	−0.022	**0.036**
	2009‡	22	10.74	0.875	**0.025**	0
Fond Rose§	2009‡	32	10.56	0.861	−0.003	0
	2011	32	12.5	0.895	0.015	**0.039**
Mahaudière§	2009‡	31	11.62	0.878	**0.087**	0
	2011	29	9.05	0.823	**0.087**	0
Pistolet§	2009‡	30	10	0.82	**0.054**	0
	2011	31	11.31	0.839	0.056	0
Porte Enfer Vigie§	2009‡	32	9.84	0.848	0.005	0
	2011	30	9.71	0.848	−0.014	0
Blonval Nord	2009‡	29	10.82	0.887	0.04	0.043
Poucet	2009‡	30	12.21	0.904	0.013	0
Desbonnes	2009‡	31	6.26	0.736	**0.075**	0
Bamboche	2009‡	32	9.65	0.83	**0.068**	0
Bazin	2009‡	32	12.66	0.892	**0.052**	0.025
L'Henriette	2009‡	24	8.01	0.786	0.02	0
Mammé	2009‡	31	12.04	0.901	**0.021**	0
Titon	2009‡	28	12.57	0.904	0.019	0
Delisle	2010‡	32	10.49	0.863	**0.096**	0
Pavillon	2009‡	31	14.16	0.921	**0.02**	0
Saint Jacques	2010‡	30	12.77	0.904	0.041	0
Kancel	2010‡	32	12.41	0.895	0.03	0
Pointe des Châteaux	2009‡	24	6.62	0.761	**0.049**	**0.168**
Mean	—	30	11.24	0.870	0.030	0.006¶

Year is the sampling year and n the sample size. Populations considered in the spatial analysis are indicated by‡, and those in the temporal analysis by † and §. †Indicates populations that have not gone extinct (non-APE) and §those that have gone extinct (APE) between temporal samples. R_A is the allelic richness based on a sample size of 14 individuals, H_e the gene diversity and f the inbreeding coefficient. Estimates of the selfing rate ($\hat{s}$) are based on a multilocus maximum-likelihood method. Values in bold characters are significantly higher than 0 at $P < 0.05$. Underlined values are significant after Bonferroni correction ($P < 0.0012$). ¶Estimate using the ML method assuming that all populations have the same selfing rate.

as a function of patch area and a dispersal kernel that scales the effect of distance on migration rate (Moilanen & Hanski 2001). This is not the definition we retained here: connectivity does not vary from year to year and characterizes the overall probability that the focal population has to be connected to surrounding freshwater habitats (i.e. the closest watershed) during the rainy season when flood occurs.

Table 3 Variables used to characterize patches and the demography of *Drepanotrema depressissimum* populations in the 25 sites studied based on yearly surveys (2001–2011)

Variable	Symbol	Description	Transformation
Patch characteristics			
Size	Size	Largest diameter in meter	$\log(1 + X)$
Vegetation cover	V	Fraction of site covered by aquatic vegetation (macrophytes and algae)	$\log(1 + X)$
Connectivity	C	Connectivity to neighbouring sites	—
Density of favourable habitats	D	Number of water bodies within a radius of 2 km	—
Stability	Stab	Temporal stability (first axis of a principal component analysis including four variables) Hydrological regime Fraction of years at which the site was dry Variance of site size over years Variance of vegetation cover over years	—
Population demography			
Apparent population age	APA	Time in years since the most recent potential founder effect	—
Long-term population size	N_{LT}	Combination of observed density and site size averaged over years	—

Symbols used as well as transformation for statistical analyses are also indicated. For more details, see text.

The density of *D. depressissimum* was estimated visually on a semi-log scale [10 levels: 0 (species not detected), 1 (<1 ind/m^2), 2 (1–5 ind/m^2), 3 (5–10 ind/m^2), up to 9 (5000–10 000 ind/m^2)]. When the pond included subhabitats with contrasted densities, we estimate the average density over all subhabitats weighted by their area. The APA was defined per year and site as the number of years since the last record of null density (including when the site was dry). APA got the maximum score (i.e. the total number of years in the survey) when densities were always non-null.

Population size can be approximated by the product of pond perimeter (favourable habitat is usually only at pond margins) by snail density. Long-term population size (over years) was computed on a log scale as:

$$N_{LT} = \text{Mean}(1/2d_t + \log_{10}(\text{size}_t)) \quad \text{(eqn 1)}$$

where d_t is the semi-log density index, size_t is site diameter and t refers to years. The scaling factor 1/2 reflects the fact that d_t increases by two units (not one) when actual densities are multiplied by 10. The arithmetic mean of logarithms gives weight to years with very low population size, which have a stronger influence on long-term diversity (Wright 1938).

Microsatellite amplification

DNA was extracted using a Chelex® method (Bio-Rad). Genotypes were obtained at ten polymorphic microsatellite loci (Table S3, Supporting information; Nicot *et al.* 2009). PCRs were conducted in 10 μL final volume including 1 μL of primers (2 μM), 5 μL of Qiagen multiplex PCR kit (Qiagen, Inc.), 3 μL of water and 1 μL of genomic DNA (1/10 dilution). PCR conditions were as in Nicot *et al.* (2009). Three microlitres of diluted amplicon was pooled with 15 μL of deionized formamide and 0.2 μL GeneScan-500 LIZ Size Standard and analysed on an ABI PRISM 3100 Genetic Analyser.

Statistical analyses

Genetic diversity. The number of polymorphic loci, allelic richness (R_A, Petit *et al.* 1998; see Appendix S2, Supporting information) and gene diversity (H_e, unbiased estimator; Nei 1987; see Appendix S2, Supporting information) were computed for all populations and loci. Deviations from Hardy–Weinberg equilibrium (HWE) were tested at each locus using exact tests (Raymond & Rousset 1995a), and a global P-value for all loci was obtained using Fisher's method (Sokal & Rohlf 1995). All calculations were made using GENEPOP 4.0.9 (Raymond & Rousset 1995b), GENETIX 4.05.2 (Belkhir *et al.* 2000) and FSTAT 2.9.3 (Goudet 2001). The estimator of Wright's inbreeding coefficient F_{IS}, f, was calculated following Weir & Cockerham (1984). Its significance was assessed using 10 000 permutations using GENETIX 4.05.2 (Belkhir *et al.* 2000). The selfing rate ($\hat{s}$) was estimated using the maximum-likelihood multilocus method implemented in RMES (David *et al.* 2007).

Spatial genetic structure. The estimator θ of F_{ST} between population pairs was calculated following Weir & Cockerham (1984), and their significance was assessed by exact tests, using genepop 4.0.9 (Raymond & Rousset 1995a). Hereafter, we refer to these estimates as pairwise F_{ST} (between sites) or temporal F_{ST} (between samples from the same site). Isolation by distance (IBD) was tested using Euclidian distance between populations with GENETIX 4.05.2 (Belkhir *et al.* 2000), based on 10 000 permutations. An analysis of molecular variance (AMOVA) was performed to quantify variance within and between the four clusters of three populations using ARLEQUIN 3.5.1.2 (Excoffier *et al.* 2005), with significance tests based on 10 000 permutations. The effect of patch characteristics (size, vegetation cover, connectivity, stability and density of favourable habitat), long-term population size and APA on population genetic structure was investigated using the Bayesian method GESTE 2.0 (Foll & Gaggiotti 2006). This method estimates the genetic differentiation between each local population and the overall metapopulation (hereafter, site-specific F_{ST}) and relates it to environmental factors using a generalized linear model. We used 10 pilot runs of 5000 iterations to obtain proposal distributions. Posterior probabilities were obtained by a MCMC with 5×10^4 burn-in iterations, a thinning interval of 20 and a sample size of 10 000.

Estimates of effective size and migration rate from time series of genetic data. We used temporal samples to jointly estimate the effective population size (N_e) and immigration rate (m) using the likelihood method of Wang & Whitlock (2003). The method assumes an infinitely large source population providing immigrants to the focal population in which N_e and m are estimated. Generation time was fixed at two months (based on laboratory cultures, Lamy *et al.* 2012), and the maximum N_e was set at 4000. In our data set, $N_e = 4000$ is practically indistinguishable from infinity (both result in insignificant changes in allele frequencies) and allowing higher values would only waste computing time. The runs often failed to converge owing to a very high polymorphism. We therefore binned alleles into eight size categories at each locus. This potentially entails a loss of precision because size homoplasy increases. However, microsatellite alleles always display homoplasy (Estoup *et al.* 2002), and homoplasy cannot bias estimates of N_e and m under the assumption of selective neutrality because drift and immigration are independent of allelic states. Indeed, changes in the frequency of a composite allele C (representing alleles A and B pooled together) over time reflect the same drift and immigration processes as changes in the frequencies of A and B separately, although the precision of the estimations of N_e and m may be lower because the information on the change in the relative frequencies of A and B within class C is lost. In addition, pooling alleles into size categories minors the potential impacts of misreading and of some of the mutations occurring between sampling dates ('stepwise' mutations), as the resulting erroneous or mutant allele sizes are likely to be binned with the original allele. The source population was composed of all other populations except the focal one. We also tested more specific sources composed of neighbouring populations only, and estimates of N_e and m were not affected (results not shown).

Clustering and assignment. We performed a Bayesian clustering of all temporal samples (12 sites and 29 samples) with STRUCTURE version 2.3.1 (Pritchard *et al.* 2000). We completed 20 runs under the constraint that the number of clusters (K) equals the number of sites ($K = 12$). Our goal was not to assess the number of groups that best fits our data but rather to test whether genotypes collected at different times in the same site cluster together, without prior information. We applied the admixture model with a burn-in of 10^6 and a run length of 10^6. Summary output was displayed graphically using the software DISTRUCT (Rosenberg 2004).

We also used an assignment method (Rannala & Mountain 1997) implemented in GENECLASS 2 (Piry *et al.* 2004) to identify whether individuals from the 12 sites are related to their local gene pool, or to external ones, at the previous time step. In each site sampled at year t, we estimated the likelihood that multilocus genotypes are drawn from the same site at year $t - 1$ and from each of the other 24 sites using Rannala & Mountain's method (1997). In addition, we computed the probability that each individual sampled at year t belongs to the same site at year $t - 1$ using a Monte-Carlo resampling method (Paetkau *et al.* 2004, 1000 simulations). Type I error was set to 0.01 (other thresholds gave the same qualitative results). We then computed the number of individuals, $N_{P < 0.01}$, for which this probability was lower than the type I error threshold. Clustering and assignment were performed on raw data without binning alleles.

When significant, migration rates derived from Wang & Whitlock's (2003) method were used to estimate total genetic turnover between two temporal samples (t_1, t_2) taking into account the time interval (in generations) between t_1 and t_2. Indeed, even in the absence of extinction and recolonization, repeated immigration during several generations may dilute local gene pools in such a way that most genes sampled at t_2 do not originate

from the local gene pool at t_1. The genetic turnover coefficient is simply the probability that the nth-generation ancestor of a gene resided outside the local population, computed as:

$$R = 1 - (1 - m)^n$$

with n being the number of generations elapsed between the two samples. Populations that went extinct and were refounded by immigrants during the time interval studied should have $R = 1$, whereas $R = 0$ in the absence of migration.

Results

Genetic diversity within populations

All loci were polymorphic in all populations. Allelic richness (standardized for a sample size of $N = 14$) was high but variable among sites, ranging from 6.26 to 14.16 alleles per locus, while gene diversity, H_e, ranged from 0.736 to 0.921. Moderate heterozygote deficiencies (average $f = 0.030$, maximum = 0.096) were observed in 19 of the 42 samples, and in 48 of 418 sample-locus combinations (Table S5, Supporting information). Selfing rates were not significantly different from zero in most populations (Table 2). Three populations exhibited very low, though significant selfing rates (4–7%, or ~1–2 individuals of ~30), and one had a higher value (Pointe des Châteaux, 17%, or ~4 selfed individuals of 24).

Patterns of spatial differentiation

All pairwise F_{ST} were significant, ranging from 0.006 to 0.191 (Table S6, Supporting information). The overall F_{ST} was 0.060 [95% CI (0.054–0.065)]. Mantel test of IBD was significant ($r = 0.318$; $P = 0.020$), although the influence of distance on pairwise F_{ST} was moderate (Fig. 2). For example, the mean pairwise F_{ST} within clusters of neighbouring sites was 0.059 ± 0.036 SD, not much lower than that among pairs from different clusters (0.062 ± 0.030 SD). The amount of molecular variance among clusters was indeed low (AMOVA, 0.40%, $P = 0.032$) in comparison with the variance among populations within clusters (5.74%; $P < 0.001$). On the other hand, the site-specific F_{ST} was predicted quite accurately by the average genetic diversity per site (Pearson $r^2 = -0.953$; Fig. 2).

We explored the influence of five patch characteristics on site-specific F_{ST} using GESTE: size, vegetation cover, connectivity, density of favourable habitats and stability (Table 4). Two models, out of the 32 explored, exhibited higher posterior probabilities (PP) than the null model (which has PP = 0.07; PP < 0.05 in all other models). The best model (PP = 0.37) included the effects of connectivity and size, while the second-best model (PP = 0.36) included connectivity only. Connectivity

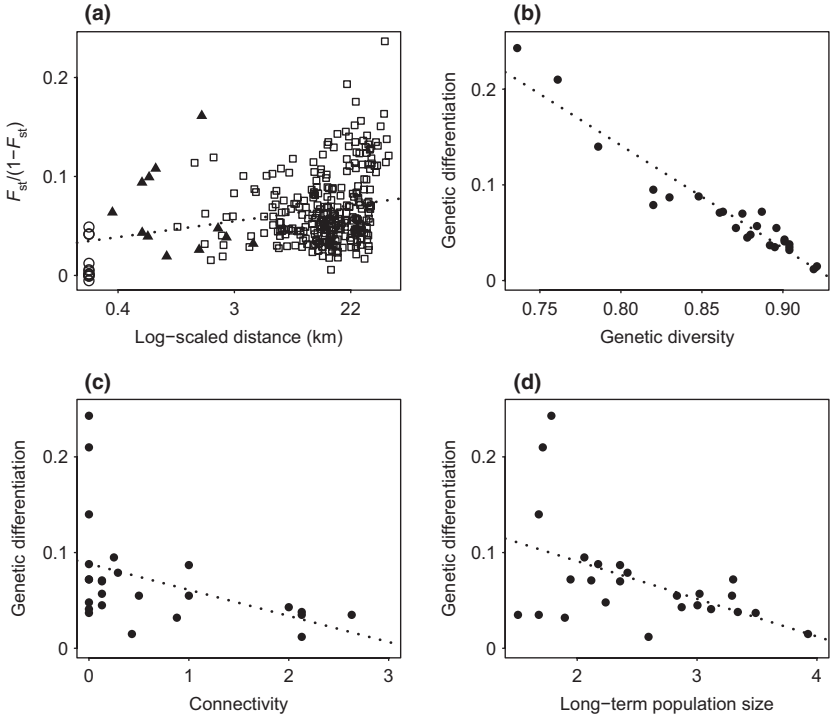

Fig. 2 Relationships between genetic, geographic, demographic and ecological variables. (a) Genetic distance [$F_{ST}/(1 - F_{ST})$] between population pairs as a function of log-transformed geographic distance in km. Black triangles represent populations within clusters of neighbouring populations and black circles samples within the same sites. Genetic differentiation (site-specific F_{ST} from GESTE) as a function of: (b) genetic diversity (H_e), (c) site connectivity and (d) long-term population size. Dashed lines represent linear regressions.

Table 4 Characteristics of the 25 sites sampled in the analysis of spatial genetic structure

Site	Type	F_{ST} GESTE	Size	V	C	Stab	D	APA	N_{LT}
Pico	m	0.055	1.61	1.89	0.50	0.14	35	6	3.29
Grosse Roche	m	0.095	1.41	0.99	0.25	−0.09	37	4	2.06
Senneville	m	0.048	1.21	0.98	0	−0.47	40	3	2.24
Valet Est	m	0.055	1.36	1.90	1.00	−0.20	36	5	2.83
Geffrier	mg	0.043	1.19	1.34	2.00	−1.78	9	9	2.87
Réjoui Nord	m	0.012	1.55	1.48	2.13	0.93	34	5	2.60
L'Ecluse	m	0.057	1.93	1.16	0.13	−0.93	56	1	3.02
Couronne Conchou	m	0.07	1.29	1.40	0.13	−2.39	29	6	2.36
Fond Rose	m	0.071	0.84	1.02	0.13	−3.93	36	1	2.12
Mahaudière	m	0.045	1.64	1.73	0.13	−1.41	32	3	3.00
Pistolet	m	0.079	1.42	1.23	0.29	−3.31	27	1	2.42
Porte Enfer Vigie	m	0.088	1.00	1.34	0	−5.56	14	1	2.18
Blonval Nord	m	0.072	1.17	0.62	0	−2.27	103	6	1.95
Poucet	r	0.035	0.87	0.86	2.63	0.16	7	2	1.51
Desbonnes	m	0.243	1.06	0.45	0	−1.81	35	3	1.79
Bamboche	m	0.087	1.33	0.76	1.00	1.09	172	2	2.36
Bazin	m	0.037	1.58	1.62	0	0.78	75	9	3.49
L'Henriette	m	0.14	1.13	1.50	0	1.29	189	1	1.68
Mammé	m	0.041	2.39	1.82	0	1.02	164	2	3.12
Titon	mg	0.038	1.84	1.67	2.13	−1.09	13	9	3.34
Delisle	m	0.072	1.52	1.72	0	0.95	72	7	3.30
Pavillon	m	0.015	1.88	1.32	0.43	0.43	88	6	3.93
Saint Jacques	m	0.032	1.17	0.97	0.88	−0.21	38	1	1.90
Kancel	m	0.035	0.93	1.96	2.13	−0.30	26	2	1.68
Pointe des Châteaux	m	0.21	1.38	1.26	0	−1.20	0	NA	1.71

Type is habitat type (p = pond, r = small river and s = swamp grasslands). F_{ST} GESTE is population-specific F_{ST} as returned by GESTE. Size refers to site size, V to vegetation cover, C to connectivity, Stab to stability, D to density of favourable habitats, APA to apparent population age and N_{LT} to long-term population size (see Table 3).

and size had negative effects on F_{ST} [mode of slope estimate: connectivity = −0.442, 95% CI (−0.699; −0.168); size = −0.334, 95% CI (−0.609; −0.0644)]. Similar results were obtained using multivariate linear regressions with site-specific F_{ST} or genetic diversity (H_e) as the dependent variable: connectivity and size were the only significant variables and had positive effects on diversity (size slope = 0.052, $P = 0.03$; connectivity slope = 0.029, $P = 0.006$).

The same analyses were run using long-term population size and APA in addition to habitat connectivity (Table 4). Habitat size was discarded because it is included in the computation of long-term population size. Using GESTE, the best model (PP = 0.86) did not include APA as a predictor but included connectivity and long-term population size, both having a negative effect on F_{ST} [mode of slope estimate: connectivity mode = −0.394, 95% CI (−0.655; −0.173); long-term population size mode = −0.408, 95% CI (−0.648; −0.154); Fig. 2]. Similarly, using multivariate regressions, APA influenced neither F_{ST}, nor H_e, while both connectivity and long-term population size had a significantly positive effect on H_e (long-term population size slope = 0.039, $P = 0.002$; connectivity slope = 0.028, $P = 0.003$).

Temporal genetic turnover

Seven populations, out of 12, did not undergo apparent population extinction (hereafter, non-APE sites) during the sampling period. Temporal F_{ST} were extremely low in these populations (<0.006), well below those estimated among populations (0.060). An exception is the Couronne Conchou population (0.049). Five populations underwent apparent population extinction (APE sites), of which four were found dry in 2010 while sampled in 2009 and 2011. In the last one (L'Ecluse), *D. depressissimum* was absent between the two samples but the pond was not dry. Among these sites, temporal F_{ST} were not significant in Pistolet and Ecluse (−0.0007 and 0.001, respectively), low but significant in Porte Enfer Vigie (0.012, $P < 0.001$), and much larger in Fond Rose and Mahaudière (0.042 and 0.042, $P < 0.001$ in both cases). In no site did we observe any detectable drop in H_e or allelic richness between temporal samples, all values being consistently high (>0.8 and >9, respectively; Table 2).

Table 5 Estimates of effective size (N_e), migration rate (m), genetic turnover (R) and temporal F_{ST} in the 12 populations considered in the temporal analysis

Site	n	N_e	m	R	F_{ST}
No apparent population extinction					
Pico	3	448.7 (185.4–1691.9)	0.009 (0.003–0.02)	0.15	**0.006**
Grosse Roche	2	∞ (240.1–∞)	0 (0–0.001)	0	−0.005
Senneville	2	512.2 (132.7–∞)	0 (0–0.12)	0	0.001
Valet Est	3	∞ (958.62–∞)	0 (0–0.005)	0	−0.001
Geffrier	4	∞ (458.3–∞)	0 (0–0.011)	0	**0.005**
Réjoui Nord	3	∞ (958.62–∞)	0 (0–0.027)	0	0.001
Couronne Conchou	2	49.5 (25.1–96.1)	0.07 (0.03–0.15)	0.73	**0.049**
Apparent population extinction					
Ecluse	2	∞ (603.3–∞)	0 (0–0.12)	0	0.001
Fond Rose	2	51.1 (25.4–98.0)	0.1 (0.05–0.23)	0.72	**0.042**
Mahaudière	2	52.1 (29.03–95.6)	0.05 (0.026–0.1)	0.46	**0.042**
Pistolet	2	∞ (469.7–∞)	0 (0–0.023)	0	−0.001
Porte Enfer Vigie	2	175.9 (80.2–532.9)	0.016 (0.005–0.035)	0.18	**0.012**

The first seven sites correspond to populations that apparently did not go extinct. The five other sites experienced an apparent population extinction because either the site dried out in 2010, or *D. depressissimum* was not detected between two samples (Ecluse). n is the number of temporal samples. N_e and m were derived according to Wang & Whitlock (2003) assuming six generations per year and are reported together with their 95 % confidence intervals. F_{ST} is the mean pairwise F_{ST} between temporal samples within site. Significant values ($P < 0.05$) are in bold characters.

In sites with temporal $F_{ST} < 0.01$ (six non-APE and two APE), Wang & Whitlock's (2003) method either yielded very high values of N_e (>400) and low values of m (<0.009) or failed to converge (Table 5 and Fig. 3). In the latter case, the maximum-likelihood estimates for m were zero. That of N_e was beyond the maximum authorized value (4000), i.e. indistinguishable from infinity. In the four populations with temporal $F_{ST} > 0.01$, effective sizes were around 50 (except in Porte Enfer Vigie, $N_e = 176$) and immigration rates ranged between 1.6% and 10.2% per generation (all significantly different from 0). The genetic turnover R could be as high as 70% in some sites (Mahaudière and Fond Rose; Table 5).

STRUCTURE was run to produce 12 genetic clusters using the 884 individuals from the 29 temporal samples, without any prior information on site or sample. A striking match was detected between clusters and site boundaries. Indeed, temporal samples from a given site were consistently classified in the same cluster (Fig. 3) with few exceptions. In Réjoui Nord (a population with very high genetic diversity), the three temporal samples seemed to be of mixed origin (i.e. several clusters) but remained very similar to each other. Differences in cluster distribution among temporal samples were only observed in the two sites with the highest turnover rates, Mahaudière and Fond Rose (Fig. 3). In Mahaudière, the 2009 sample contains mostly two types of genotypes with distinct cluster memberships (see Fig. 3); the 2011 sample is composed mostly of one of these types. In Fond Rose, the 2009 sample is entirely assigned to a single genetic cluster, and very few individuals in the 2011 sample have the same profile; nearly all of them are of a different, mixed genetic make-up (Fig. 3).

Results of assignment tests are reported in Table 6. In all samples except Fond Rose and Couronne Conchou, the most likely population of origin of the sample at year (t) was the same site at ($t - 1$), and the number of individuals whose genotype was unlikely to be of local origin ($N_{P < 0.01}$) was low (0–5 out of a total of around 30). For both Couronne Conchou (non-APE site) and Fond Rose (APE site), the most likely site of origin was Réjoui Nord. However, the likelihood of being the potential origin of its own population, respectively three and two years later, was much higher in Couronne Conchou ($\Delta_{-\log(L)} = 51$, ranks 6th among all 24 possible candidate populations; $N_{P < 0.01} = 13$ of 22) than in Fond Rose ($\Delta_{-\log(L)} = 198$, ranks 16th; $N_{P < 0.01} = 31$ of 32).

Discussion

Asymmetric island model vs. metapopulation

Drepanotrema depressissimum showed virtually no selfing, confirming a previous study (Lamy *et al.* 2012), and exhibited a higher polymorphism than any other freshwater snail we know of (Jarne 1995; Escobar *et al.* 2011). However, genetic diversity was highly

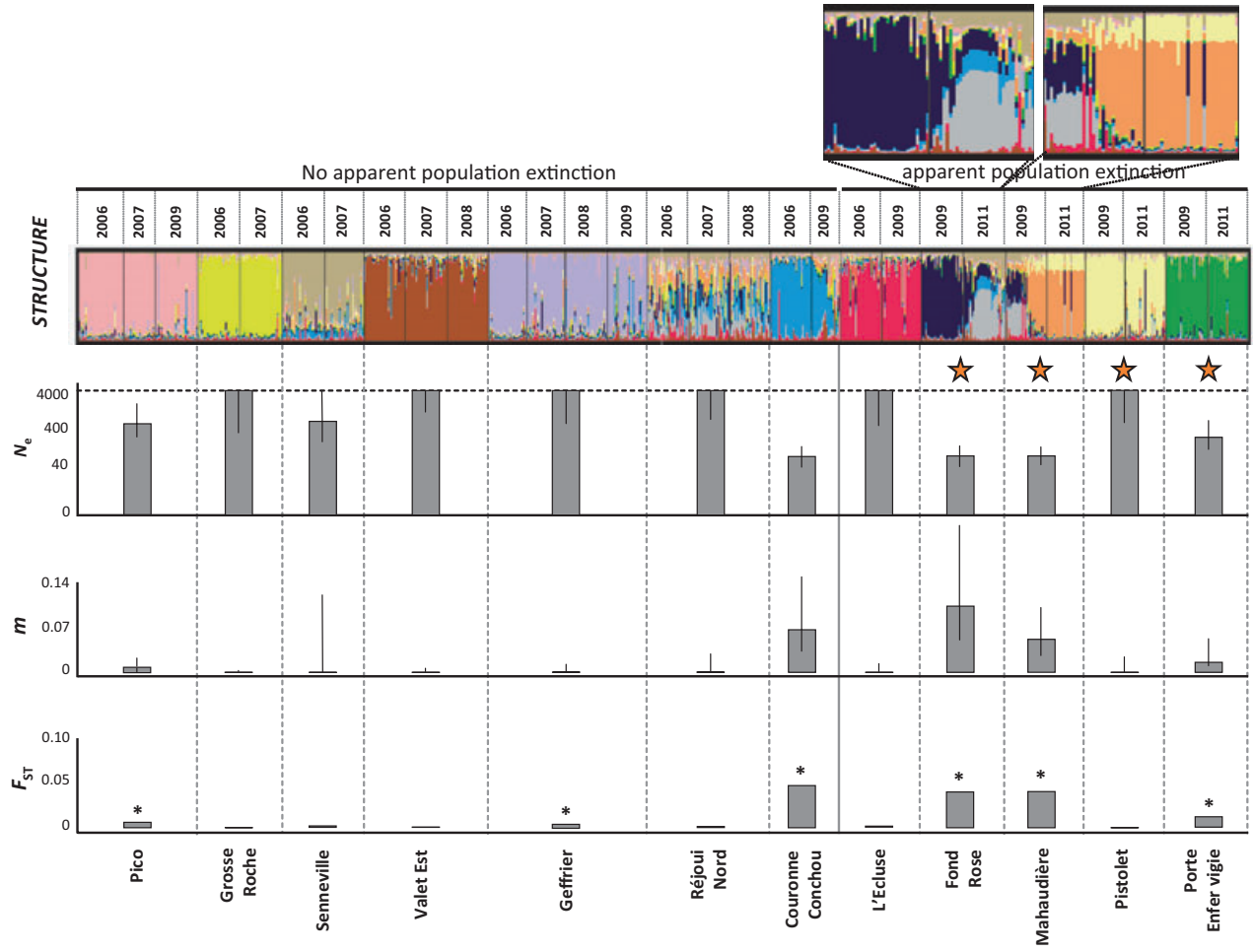

Fig. 3 Results of STRUCTURE analysis, estimates of effective size (N_e in log scale) and immigration rate (m) from Wang & Whitlock's (2003) method, and mean pairwise F_{ST} between temporal samples in the 12 sites (names are given at the bottom of the figure) that were sampled several times (29 samples in total). STRUCTURE analysis: each colour refers to one of the 12 clusters considered in the analysis, and bar plots (one per individual) to cluster membership. Plots were produced with Distruct based on 20 replicate runs, with a close-up on two populations Fond Rose and Mahaudière (see Results). N_e and m are reported with their 95% confidence interval, and dotted line (N_e row) represents maximum N_e value. Black stars (F_{ST} row) refer to significant F_{ST} value ($P < 0.01$) and orange stars to sites that dried out in 2010. See Table 5 for estimate values.

variable among sites, and so was genetic differentiation, the two variables being strongly correlated. This variation indicates that the symmetry assumption is not met in our system. This feature is probably frequent in natural fragmented populations and is shared by asymmetric island and metapopulation models (Table 1). In the former, population sizes and immigration rates remain constant in time but differ among subpopulations, while in the latter, subpopulations differ with respect to age, recently recolonized ones being less diverse and more differentiated. A given system may combine both sources of asymmetry, the important question for metapopulation studies being to assess their relative importance (i.e. a sensitivity analysis).

To address this question, we have to study the effect of ecological variables and temporal genetic structure (Table 1). Among ecological variables, patch connectivity and size had a negative effect on genetic differentiation and a positive effect on genetic diversity. However, neither habitat stability (an indicator of both perturbation frequency and potential extinction) nor APA had any detectable effect. Thus, our system seems to be better described by an asymmetrical island model than by a metapopulation model.

Temporal changes in allele frequencies in 12 sites with (5) and without (7) apparent extinction further confirmed this conclusion. In general, temporal samples from the same site were very weakly differentiated, had the same genetic diversity and clustered together using

Table 6 Results of assignment tests in the 12 populations considered in the temporal analysis

Site	Year	Rank	Best	$\Delta_{-\log(L)}$	Mean($\Delta_{-\log(L)}$)	$N_{P < 0.01}$
Non-apparent population extinction						
Pico	2007	1	—	0	293 ± 104	0/24
	2009	1	—	0	279 ± 126	2/32
Grosse Roche	2007	1	—	0	387 ± 145	0/32
Senneville	2007	1	—	0	262 ± 117	2/30
Valet Est	2007	1	—	0	337 ± 140	0/32
	2008	1	—	0	335 ± 135	1/31
Geffrier	2007	1	—	0	248 ± 109	4/29
	2008	1	—	0	329 ± 133	0/32
	2009	1	—	0	306 ± 131	3/30
Réjoui Nord	2007	1	—	0	210 ± 106	0/31
	2008	1	—	0	195 ± 113	3/31
Couronne Conchou	2009	6	REJ2008	51	109 ± 84	13/22
Apparent population extinction						
L'Ecluse	2009	1	—	0	257 ± 103	0/30
Fond Rose	2011	16	REJ2008	198	175 ± 106	31/32
Mahaudière	2011	1	—	0	220 ± 95	2/29
Pistolet	2011	1	—	0	324 ± 122	5/31
Porte Enfer Vigie	2011	1	—	0	321 ± 146	2/30

Year is the sampling year. L is the likelihood that the set of individuals sampled at year t entirely comes from the same site at year $t-1$ or from one of the 24 other samples of the metapopulation according to Rannala & Mountain (1997). *Rank* refers to the rank of the likelihood of the focal site. When rank $\neq 1$, *Best* indicates the sample displaying the highest likelihood, and $\Delta_{-\log(L)}$ the difference in likelihood between the two sites. Mean($\Delta_{-\log(L)}$) represent the average pairwise difference in likelihood between the focal sample and the 24 other samples. $N_{P < 0.01}$ is the number of individuals for which the probability of being assigned to the same site was lower than 0.01.

STRUCTURE. Assignment tests also showed that, irrespective of apparent extinctions, genes sampled in a site at t mostly came from the local gene pool at $t-1$ (to a few exceptions discussed below). Importantly, in three cases (Ecluse, Pistolet and Porte d'Enfer Vigie), apparent population extinction was not associated with any genetic change. This does not mean that extinction never happens, rather that its rate is much lower than estimated based on field surveys (21.7%), as most of, if not all, apparent extinctions turned out not to be true extinctions. Our study therefore illustrates that demographical surveys can suggest a metapopulation structure that is later invalidated by genetic data.

Apparent vs. real extinction–recolonization cycles

Several studies have previously found an increase in genetic diversity with population age (Whitlock 1992; McCauley *et al.* 1995; Giles & Goudet 1997; Ingvarsson *et al.* 1997; Haag *et al.* 2005). Only two focused on a metapopulation (Giles & Goudet 1997; Haag *et al.* 2005), and they probably estimated population age with good accuracy. In the *Daphnia* metapopulation of Haag *et al.* (2005), habitats were small rock pools, and the temporal frequency of visits was high enough to state extinction with some certainty. In Giles & Goudet (1997), population age was based on the date of emergence of new islands, before which the absence of population was certain. However, population age might not in general be known so precisely when extinction records are imperfect, as in our system. It should be noted that the detectability of *D. depressissimum* (0.78) is not particularly low; for example, detectability of Glanville fritillary butterfly is around 0.5 (Hanski 2011). More generally, field ecologists can never be certain that presence/absence data are 100% reliable. A species can go undetected due to dilution in large sites (Kéry *et al.* 2006), or because it is present as a cryptic form (such as seed banks, Honnay *et al.* 2008).

Detection probability depends on many factors including density, season, individual size and sampling effort (Royle & Nichols 2003; Bailey *et al.* 2004; Chen *et al.* 2009). Sampling effort per unit area often decreases with patch size (Altermatt & Ebert 2010), especially when the latter is very variable. Indeed, maintaining a constant sampling effort would require unrealistic amounts of time in the largest sites (e.g. a 100-fold longer time in a 150-m-diameter pond than in a 15-m pond). Many 'false extinctions' may thus occur in large sites, although true extinctions are probably less frequent there. This primarily concerns invertebrate or plant species inhabiting habitat fragments with

considerable size variation, spanning several orders of magnitude [e.g. butterflies (Hanski 1999), *Daphnia* (Haag *et al.* 2005)]. Although one may want to target landscapes with small habitat fragments in which extinction/colonization is both more frequent and easier to characterize, real metapopulations often combine very large and very small sites (Massol *et al.* 2011).

In addition, cryptic forms, such as banks of seeds or resting stages, can buffer local populations against demographic stochasticity (Kalisz & McPeek 1992) and increase population effective size (Vitalis *et al.* 2004). Seed bank effects (*sensu lato*) have an important impact in our system as illustrated by the Pistolet and Porte d'Enfer Vigie ponds. These ponds completely dried out in 2010, without any genetic signature of extinction or of bottleneck. The question arises whether the genetic similarity of samples taken in 2009 and 2011 in these two sites is due to the persistence of aestivating individuals in the soil or to extinction followed by recolonization from very close populations, genetically indistinguishable from the local population. However, this second hypothesis is highly improbable because IBD is weak in our system and geographically close sites are not genetically identical. Indeed, differentiation at a local scale (clusters of three neighbouring sites) is in general as high as at the regional scale, and this is true for pairs of sites which are true nearest neighbours to each other and still show high F_{ST} (Desbonnes/Porte Enfer Vigie = 0.139; Desbonnes/Saint Jacques = 0.098; L'Henriette/Mammé = 0.086; L'Henriette/Bamboche = 0.09). Therefore, substantial numbers of individuals must have persisted in the ground [aestivation; Pointier & Combes (1976)], preventing extinction.

Although aestivation has previously been observed in freshwater snails (Brown 1994), our study is the first demonstration that population resurrection after a long drought can rely on local recruitment and does not require recolonization. More generally, genetic evidence for seed bank effects is scarce and indirect in all organisms, including plants, where it relies on comparisons of spatial genetic structure between species with long- and short-lived seeds (Honnay *et al.* 2008). Our study provides a more direct approach to the contribution of resting stages to persistence during unfavourable periods. Direct quantification of seed banks (*sensu lato*) is often challenging, and temporal genetic studies represent an interesting, and accessible, alternative.

Genetic bottlenecks are not always associated with detected demographic accidents

Even if apparent extinction does not result in true extinction, it might be associated with a demographic bottleneck, and therefore a large genetic turnover. Such events were detected here in three populations, resulting in significant temporal F_{ST} and changes in cluster membership (STRUCTURE). Two of them (Mahaudière and Fond Rose) underwent desiccation during the time interval considered, and one did not undergo apparent extinction (Couronne Conchou). For Mahaudière and Couronne Conchou, genetic data are incompatible with total genetic resetting and rather reflect the combined action of drift (bottleneck) and immigration. The case of Fond Rose is more ambiguous because local genotypes in 2011 were not more assigned to Fond Rose in 2009 than to any other population. Although this is compatible with extinction and recolonization, Wang & Whitlock's (2003) method suggests that a moderate bottleneck (N_e = 50) together with a large immigration rate (m = 0.1) over 12 generations is also compatible with the data. On the whole, large genetic turnovers sometimes occur in our system, but it seems hard to predict where and when, as they are not systematically associated with detected demographic accidents or apparent extinction. We know of no other study that tried to match small-scale genetic changes to demographic estimates. However, differences between genetical N_e and demographical N are frequently detected (Luikart *et al.* 2010). Although previous studies focused on large populations and long temporal scales, they put forward two explanations that also apply to the smaller spatio-temporal scale of our study: (i) genetic and demographical surveys are not sensitive to the same range of variation in population size and (ii) dispersal affects the two types of data in a different way. We consider them in turn.

Changes in effective size affect genetic diversity as a saturating function. In most situations, bottlenecks become detectable when N_e remains below a few tens of individuals for several generations. A true extinction–recolonization, a bottleneck of a few individuals and a bottleneck of a few tens of individuals leave different genetic signatures (Cornuet & Luikart 1996). However, in small animals or plants, all these situations may correspond to relatively low densities, difficult to distinguish using demographic surveys. Reciprocally, demographic changes easily detected by observation (for instance, from 1000 ind/m^2 to a few ind/m^2 in our data) leave no significant trace on genetic diversity, as total population size remains large and genetic drift in a few generations remains negligible. For example, *D. depressissimum* was not found in 2008 in the Ecluse pond, but allele frequencies remained stable from 2006 to 2009. Indeed, given the large diameter (150 m) of the pond, even densities that we can hardly detect ($\ll$1 ind/m^2) can represent a local population of a few hundreds, not small enough to

leave a detectable genetic signal of bottleneck. Thus, demography and genetics are sensitive to different ranges of variation in population size. This may not hold in large species, such as mammals or birds, in which demographic observations are more efficient at detecting changes in small populations. Note however, that long-term population size is clearly related to genetic diversity and differentiation in our system. Therefore, although genetic data do not fit year-to-year variation in demography, demographic data reflect true long-term properties of populations with significant impact on genetic diversity at a larger temporal scale. This illustrates that genetics and demography are complementary because they document both different ranges of population size variation and different timescales.

Propagule movement is very difficult to observe (Clobert *et al.* 2001), which is why many studies must rely on genetic data (Bohonak 1999). In our data set, Wang & Whitlock's (2003) method revealed an unsuspected intensity of migration among ponds in *D. depressissimum*. In some populations, the effective size was so high that even a substantial number of migrants per generation would not change allele frequencies (dilution effect). However, the genetic impact of immigrants became very important in small populations, where they resulted in large genetic turnover rates (20–70%). The effective number of immigrants per generation (N_e m) ranged between 2 and 5 (which can represent as much as 10% of the effective size). There is little reason to believe that N_e m is any smaller in larger populations, given the confidence intervals, although the actual estimates may be nonsignificant because of dilution. Thus, most sites are probably submitted to constant immigration, maintaining high genetic diversity despite demographic fluctuations.

Gene flow among sites can occur through passive transportation in the rainy season, when neighbouring sites can be connected through water flow. The importance of connexions is revealed here by the positive effect of connectivity on genetic diversity. However, distance is only weakly correlated with genetic differentiation in *D. depressissimum*, suggesting that dispersal also occurs over long distance. Possible vectors are water birds (Figuerola & Green 2002), although estimating their role directly is difficult (Bohonak 1999; Bilton *et al.* 2001). Frequent long-distance dispersal should allow empty sites to be quickly colonized and to rapidly reach high genetic diversity. Importantly, this inference can be drawn only when gene flow and colonization rely on the same process (here, snail movement). Similar reasoning may not apply to other organisms. For example, gene flow in plants results from both pollen and seed dispersal, but colonization can be initiated by seeds only.

Scope and limitations

This study raises an important general question: which fraction of extinction-colonization cycles observed in natural systems is real? While the extinction-colonization assumption lies at the heart of the metapopulation theory, solid evidence for metapopulation dynamics in nature boils down to very few examples (Hanski 1999; Haag *et al.* 2005), and too few systems have been intensively studied, especially from both a demographic and genetic point of view. Our study shows that temporal genetic studies would be a useful complement to demographic surveys to gain insight into metapopulation dynamics. Small invertebrates and plants seem particularly appropriate for this approach, owing to the difficulty of exhaustive samples and to the presence of seed banks and resting stages.

Can our methodology be generalized to species that do not share the particulars of our system? Clearly, the high levels of genetic diversity and spatial differentiation found in our study create favourable conditions to detect genetic signatures of extinction/recolonization. However, very high diversities can limit the range of F_{ST}. There is an ongoing debate on the possible use of other metrics such as G'_{ST} (Hedrick 2005) or D (Jost 2008) that do not share this property. High diversity is derived from high mutation rate, which tends to obliterate the effects of migration and/or recolonization. In our case, replacing F_{ST} by other measures of differentiation does not alter our conclusions (data not shown). In principle, if a genetic signal is weak and overwhelmed by mutational noise, no change in the differentiation metric (G'_{ST}, D or F_{ST}) is expected to cure the problem (Whitlock 2011), and the only solution would be to find markers with lower mutation rates. Fortunately, our study shows that even with genetic diversity as high as H_e = 0.8–0.9, F_{ST} carries significant information on migration and extinction-colonization dynamics. Future studies could benefit from analysing the effects of choosing a given differentiation metric.

A further possible limitation to the approach developed here is limited genetic differentiation, reducing the performances of F_{ST}, STRUCTURE and genetic assignment. However, multilocus methods such as STRUCTURE behave surprisingly well even with F_{ST} of a few per cents (Coulon *et al.* 2008; Schwartz & McKelvey 2009), as they use linkage disequilibrium information in addition to within-locus differentiation. An important requirement of our methods, however, is the existence of clearly delimited populations. Indeed, continuous

landscapes complicate the interpretation of assignment and STRUCTURE outputs (Schwartz & McKelvey 2009). They also represent situations under which local extinction and recolonization can hardly be defined (a reason not to test metapopulation dynamics in such systems).

In conclusion, our methodology has a large range of applicability; as soon as populations are clearly delimited, genetic diversity is not extremely high and differentiation not too low. Although no population ever fits perfectly any simple model, the importance of the metapopulation concept in ecological and evolutionary sciences is such that it is worth knowing to how many natural systems it can, even approximately, be applied. The increasing availability of molecular markers opens the way to routinely checking the validity of metapopulation models in systems hitherto studied only demographically, and potentially to develop future methods combining genetics and demography to precisely quantify extinction and recolonization rates.

Acknowledgements

The authors thank Christoph R Haag for fruitful comments on the manuscript, Arnaud Estoup for discussion on STRUCTURE and for access to CBGP bioinformatics cluster, the SMGE at CEFE and the environmental genomics platform of the 'Montpellier Environnement Biodiversité' Research Federation for help with genotyping. T. Lamy was supported by a fellowship from the French Ministry of Research. This work was funded by the CNRS, ANR JCJC-0202 to P. David, the Ecosystèmes Tropicaux program and a 'Chercheurs d'Avenir' Grant to P. David from the Région Languedoc-Roussillon.

Conflict of interest

The authors declare no conflict of interest.

References

Altermatt F, Ebert D (2010) Populations in small, ephemeral habitat patches may drive dynamics in a *Daphnia magna* metapopulation. *Ecology*, **91**, 2975–2982.

Bailey LL, Simons TR, Pollock KH (2004) Estimating site occupancy and species detection Probability parameters for terrestrial salamanders. *Ecological Applications*, **14**, 692–702.

Bekker RM, Bakker JP, Grandin U *et al.* (1998) Seed size, shape and vertical distribution in the soil: indicators of seed longevity. *Functional Ecology*, **12**, 834–842.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2000) GENETIX, logiciel sous Windows TM pour la génétique des populations. Université de Montpellier II.

Bilton DT, Freeland JR, Okamura B (2001) Dispersal in freshwater invertebrates. *Annual Review of Ecology and Systematics*, **32**, 159–181.

Bohonak AJ (1999) Dispersal, gene flow, and population structure. *Quarterly Review of Biology*, **74**, 21–45.

Brown DS (1994) *Freshwater Snails of Africa and Their Medical Importance*, 2nd edn. Taylor and Francis Ltd, London.

Bruyere F, Questel Y (2001) *Etude de recensement des mares et des canaux en Guadeloupe*. DIREN Guadeloupe, Basse-Terre.

Calcagno V, Mouquet N, Jarne P, David P (2006) Coexistence in a metacommunity: the competition-colonization trade-off is not dead. *Ecology Letters*, **9**, 897–907.

Chen G, Kéry M, Zhang J, Ma K (2009) Factors affecting detection probability in plant distribution studies. *Journal of Ecology*, **97**, 1383–1389.

Clobert J, Danchin E, Dhondt AA, Nichols JD (2001) *Dispersal*. Oxford University Press, Oxford.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Coulon A, Fitzpatrick JW, Bowman R *et al.* (2008) Congruent population structure inferred from dispersal behaviour and intensive genetic surveys of the threatened Florida scrub-jay (*Aphelocoma coerulescens*). *Molecular Ecology*, **17**, 1685–1701.

David P, Pujol B, Viard F, Castella V (2007) Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology*, **16**, 2474–2487.

De Stasio BT (1989) The seed bank of a freshwater crustacean: copepodology for the plant ecologist. *Ecology*, **70**, 1377–1389.

Escobar JS, Auld JR, Correa AC *et al.* (2011) Patterns of mating-system evolution in hermaphroditic animals: correlations among selfing rate, inbreeding depression and the timing of reproduction. *Evolution*, **65**, 1233–1253.

Estoup A, Jarne P, Cornuet J-M (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, **11**, 1591–1604.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Facon B, David P (2006) Metapopulation dynamics and biological invasions: a spatially explicit model applied to a freshwater snail. *The American Naturalist*, **168**, 769–783.

Figuerola J, Green AJ (2002) Dispersal of aquatic organisms by waterbirds: a review of past research and priorities for future studies. *Freshwater Biology*, **47**, 483–494.

Foll M, Gaggiotti O (2006) Identifying the environmental factors that determine the genetic structure of populations. *Genetics*, **174**, 875–891.

Giles BE, Goudet J (1997) Genetic differentiation in *Silene dioica* metapopulations: estimation of spatiotemporal effects in a successional plant species. *The American Naturalist*, **149**, 507–526.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3).

Guillot G, Leblois R, Coulon A, Frantz A (2009) Statistical methods in spatial genetics. *Molecular Ecology*, **18**, 4734–4756.

Haag CR, Riek M, Hottinger JW, Pajunen VI, Ebert D (2005) Genetic diversity and genetic differentiation in *Daphnia* metapopulations with subpopulations of known age. *Genetics*, **170**, 1809–1820.

Hanski I (1994) A practical model of metapopulation dynamics. *Journal of Animal Ecology*, **63**, 151–162.

Hanski I (1999) *Metapopulation Ecology*. Oxford University Press, Oxford.

Hanski I (2011) Eco-evolutionary spatial dynamics in the Glanville fritillary butterfly. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 14397–14404.

Hanski I, Gaggiotti OE (2004) *Ecology, Genetics, and Evolution of Metapopulations*. Academic Press, San Diego.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633–1638.

Honnay O, Bossuyt B, Jacquemyn H, Shimono A, Uchiyama K (2008) Can a seed bank maintain the genetic variation in the above ground plant population? *Oikos*, **117**, 1–5.

Ingvarsson PK, Olsson K, Ericson L (1997) Extinction-recolonization dynamics in the mycophagous beetle *Phalacrus substriatus*. *Evolution*, **51**, 187–195.

Jarne P (1995) Mating system, bottlenecks and genetic polymorphism in hermaphroditic animals. *Genetics Research*, **65**, 193–207.

Jost L (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

Kalisz S, McPeek MA (1992) Demography of an age-structured annual: resampled projection matrices, elasticity analyses, and seed bank effects. *Ecology*, **73**, 1082–1093.

Kéry M, Spillmann JH, Truong C, Holderegger R (2006) How biased are estimates of extinction probability in revisitation studies? *Journal of Ecology*, **94**, 980–986.

Lamy T, Lévy L, Pointier J-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*, in press, doi: 10.1007/s10682-011-9520-8.

Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomology Society of America*, **71**, 237–240.

Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW (2010) Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conservation Genetics*, **11**, 355–373.

Massol F, Duputié A, David P, Jarne P (2011) Asymmetric patch size distribution leads to disruptive selection on dispersal. *Evolution*, **65**, 490–500.

McCauley DE, Raveill J, Antonovics J (1995) Local founding events as determinants of genetic structure in a plant metapopulation. *Heredity*, **75**, 630–636.

Moilanen A, Hanski I (2001) On the use of connectivity measures in spatial ecology. *Oikos*, **95**, 147–151.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nicot A, Jarne P, David P (2009) Development of polymorphic microsatellite loci in the hermaphroditic freshwater snails *Drepanotrema surinamense* and *Drepanotrema depressissimum*. *Molecular Ecology Resources*, **9**, 897–902.

Olivieri I, Michalakis Y, Gouyon P-H (1995) Metapopulation genetics and the evolution of dispersal. *The American Naturalist*, **146**, 202–228.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, **13**, 55–65.

Pannell JR, Barrett SCH (1998) Baker's law revisited: reproductive assurance in a metapopulation. *Evolution*, **52**, 657–668.

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, **95**, 536–539.

Pointier J-P (2008) *Guide to the Freshwater Molluscs of the Lesser Antilles*. Conch Books, Hackenheim.

Pointier J-P, Combes C (1976) La saison sèche en Guadeloupe et ses conséquences sur la démographie des mollusques dans les biotopes à *Biomphalaria glabrata* (Say, 1818), vecteur de la bilharziose intestinale. *La Terre et la Vie*, **30**, 121–147.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Rajon E, Venner S, Menu F (2009) Spatially heterogeneous stochasticity and the adaptive diversification of dormancy. *Journal of Evolutionary Biology*, **22**, 2094–2103.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences USA*, **94**, 9197–9201.

Raymond M, Rousset F (1995a) An exact test for population differentiation. *Evolution*, **49**, 1280–1283.

Raymond M, Rousset F (1995b) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Royle JA, Nichols JD (2003) Estimating abundance from repeated presence-absence data or point counts. *Conservation Genetics*, **10**, 441–452.

Schwartz MK, McKelvey KS (2009) Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conservation Genetics*, **10**, 441–452.

Sokal RR, Rohlf FJ (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd edn. Freeman and Co, New York.

Vitalis R, Glémin S, Olivieri I (2004) When genes go to sleep: the population genetic consequences of seed dormancy and monocarpic perenniality. *The American Naturalist*, **163**, 295–311.

Wang J, Whitlock MC (2003) Estimating effective population size and migration rates from genetic samples over space and time. *Genetics*, **163**, 429–446.

Weir B, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Whitlock MC (1992) Nonequilibrium population-structure in forked fungus beetles: extinction, colonization, and the genetic variance among populations. *The American Naturalist*, **139**, 952–970.

Whitlock MC (2011) G'_{ST} and D do not replace F_{ST}. *Molecular Ecology*, **20**, 1083–1091.

Whitlock MC, Barton NH (1997) The effective size of a subdivided population. *Genetics*, **146**, 427–441.

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

Wright S (1938) Size of population and breeding structure in relation to evolution. *Science*, **87**, 430–431.

This research was part of T.L.'s PhD research which combines demographical and genetic studies to study metacommunity and metapopulation dynamics in freshwater habitats. J.-P.P. is interested in freshwater snail ecology and systematics, and has started studying snail communities in the West Indies and South America 40 years ago. P.D. and P.J. are population geneticists and evolutionary biologists interested in the evolution of mating systems, life-history traits and how they influence community dynamics in fragmented habitats. Together with J.-P.P. they have initiated a long-term project on freshwater snail communities in the French Antilles which has been running for 11 years.

Data accessibility

Microsatellite data uploaded as online Supporting information.

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Types of population structure and associated predictions.

Appendix S2 Different measures and associated equations used in the main text.

Table S1 Polymorphism at the ten microsatellite loci studied in *Drepanotrema depressissimum*.

Table S2 Population codes used in Tables S3 and S4.

Table S3 Wright's inbreeding coefficient F_{IS}, f, estimated following Weir & Cockerham (1984), per sampling year and locus.

Table S4 Estimates of pairwise F_{ST}, θ, calculated following Weir & Cockerham (1984) for the 25 populations of *Drepanotrema depressissimum* included in the spatial analysis.

Table S5 Estimates of pairwise F_{ST}, θ, calculated following Weir & Cockerham (1984) for the 12 populations (29 samples) of *Drepanotrema depressissimum* included in the temporal analysis.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.