

Population genetic dynamics of an invasion reconstructed from the sediment egg bank

MARKUS MÖST,*†‡ SARAH OEXLE,§¶ SILVIA MARKOVÁ,** DALIA AIDUKAITE,* LIVIA BAUMGARTNER,* HANS-BERND STICH,†† MARTIN WESSELS,†† DOMINIK MARTIN-CREUZBURG§ and PIET SPAAK*†

*Eawag, Swiss Federal Institute of Aquatic Science and Technology, CH-8600 Dübendorf, Switzerland, †Institute of Integrative Biology, ETH Zurich, CH-8092 Zurich, Switzerland, ‡Department of Zoology, University of Cambridge, CB2 3EJ Cambridge, UK, §Limnological Institute, University of Konstanz, D-78464 Konstanz, Germany, ¶Laboratory of Aquatic Ecology, Evolution and Conservation, KU Leuven, B-3000 Leuven, Belgium, **Laboratory of Molecular Ecology, Institute of Animal Physiology and Genetics, Academy of Sciences of the Czech Republic, Rumburská 89, 27721 Liběchov, Czech Republic, ††Institute for Lake Research, D-88085 Langenargen, Germany

Abstract

Biological invasions are a global issue with far-reaching consequences for single species, communities and whole ecosystems. Our understanding of modes and mechanisms of biological invasions requires knowledge of the genetic processes associated with successful invasions. In many instances, this information is particularly difficult to obtain as the initial phases of the invasion process often pass unnoticed and we rely on inferences from contemporary population genetic data. Here, we combined historic information with the genetic analysis of resting eggs to reconstruct the invasion of *Daphnia pulicaria* into Lower Lake Constance (LLC) in the 1970s from the resting egg bank in the sediments. We identified the invader as 'European *D. pulicaria*' originating from meso- and eutrophic lowland lakes and ponds in Central Europe. The founding population was characterized by extremely low genetic variation in the resting egg bank that increased considerably over time. Furthermore, strong evidence for selfing and/or biparental inbreeding was found during the initial phase of the invasion, followed by a drop of selfing rate to low levels in subsequent decades. Moreover, the increase in genetic variation was most pronounced during early stages of the invasion, suggesting additional introductions during this period. Our study highlights that genetic data covering the entire invasion process from its beginning can be crucial to accurately reconstruct the invasion history of a species. We show that propagule banks can preserve such information enabling the study of population genetic dynamics and sources of genetic variation in successful invasive populations.

Keywords: ancient DNA, asexual, *Daphnia*, egg bank, genetic variation, invasion

Introduction

Biological invasions, often facilitated by human activities, are a global issue as they can have strong impacts on native species, communities and ecosystems, constitute a major threat to biodiversity and can cause high economic costs (e.g., Sakai *et al.* 2001; Simberloff

et al. 2013). During several decades of biological invasion research, many factors facilitating successful invasions have been identified and considerable scientific effort has been devoted to elucidate the patterns and significance of genetic variation during the invasion process (Sakai *et al.* 2001; Allendorf & Lundquist 2003; Kolbe *et al.* 2004; Dlugosch & Parker 2008; Special Issue: Invasion Genetics: The Baker and Stebbins Legacy 2015).

It is often assumed that introduced populations are undergoing a bottleneck limiting their ability to adapt

Correspondence: Markus Most, Fax: +44 1223 336676;
E mail: mm2083@cam.ac.uk

and increasing the risk of extinction. The apparent success of many invasive species has therefore been described as a 'genetic paradox'; however, this view has been challenged by work suggesting that reduced genetic variation is not a common principle for invasions and can be overcome by several mechanisms (Roman & Darling 2007; Dlugosch & Parker 2008; Bock *et al.* 2015). In many cases, the variation in the introduced range is not or only marginally reduced due to high propagule pressure, a combination of the number of propagules introduced per event and the number of introduction events (Roman & Darling 2007). Multiple introductions from different source populations may result in even greater diversity in the invasive population (Dlugosch & Parker 2008), whereas recurrent introductions from the same source may be more difficult to detect (Benazzo *et al.* 2015). Genetic variation in the introduced range may be limited by the extent of propagule pressure and the overall variation present in the source populations (Roman & Darling 2007), but also by rapid population growth and adaptation of the founding population which may lead to monopolization of resources and competitive exclusion of new immigrant genotypes (De Meester *et al.* 2002).

The bulk of our knowledge on invasion histories and sources of genetic variation is based on inferences from contemporary genetic data, an approach that has widely and often successfully been applied (Cristescu 2015), but has also been shown to have limited power to distinguish between alternative historical invasions scenarios (Barun *et al.* 2013). In contrast, studies measuring genetic variation through time can provide direct information on population histories and genetic processes from invasion to establishment. However, such data are rare as the initial phases of an invasion are inherently difficult to capture and most of the few studies present cover only a short period of the invasion process (e.g. Ortells *et al.* 2014). An alternative approach that allows for high-resolution reconstruction of the entire invasion process is the analysis of archives of resting stages, for example seed or resting egg banks (Fennell *et al.* 2014; Ortells *et al.* 2014). In the present work, we therefore combined available historical information (Einsle 1980, 1987; Stich & Maier 2007) and genetic analyses of a sediment resting egg bank to reconstruct population history and genetic processes during invasion and establishment of a non-native *Daphnia* species in Lower Lake Constance (LLC) from its first appearance during peak eutrophication 40 years ago up to recent years.

Species of the genus *Daphnia* are key organisms in many aquatic food webs and among the most successful aquatic invaders. Several cases of invasions of *Daphnia* spp. into lakes and reservoirs have been reported (Havel *et al.* 1995; Hairston *et al.* 1999a; Duffy

et al. 2000; Mergeay *et al.* 2005, 2006; Brede *et al.* 2009; Spaak *et al.* 2012; Frisch *et al.* 2013). Three essential properties enable them to exert high propagule pressure, mitigate negative effects of diversity loss during the invasion process (Kolbe *et al.* 2004; Frankham 2005a, b; Roman & Darling 2007) and therefore predestine them as successful invaders (Sakai *et al.* 2001; Kolbe *et al.* 2004; Frankham 2005a,b; Lockwood *et al.* 2005; Roman & Darling 2007). First, they are able to reproduce asexually, either by cyclic or obligate parthenogenesis, which provides them with certain demographic advantages, that is avoidance of the 'twofold cost of sex', Allee effects and the ability to found a population starting from a single colonizer (Sakai *et al.* 2001; Mergeay *et al.* 2005, 2006; Roman & Darling 2007). Obligate parthenogenetic populations can furthermore circumvent inbreeding depression and preserve successful genotypes (Vrijenhoek 1998; Sakai *et al.* 2001; Roman & Darling 2007). Second, *Daphnia* spp. produce ephippia, protective chitinous envelopes that enclose up to two resting eggs which are either the product of sexual reproduction in cyclic parthenogenetic populations or asexually produced in the less frequent obligate parthenogenetic populations (Hebert 1981; Zaffagnini 1987; Lampert 2011; Xu *et al.* 2013). Ephippia are highly resistant and can easily be distributed over long distances, for example through migratory water fowl, wind- or human-mediated dispersal, therefore allowing for high propagule pressure (De Meester *et al.* 2002; Lockwood *et al.* 2005; Roman & Darling 2007). Furthermore, a proportion of ephippia is deposited on the lake bottom building up a so-called resting egg bank that reduces the extinction risk in the newly colonized habitat, that is storage effect (Chesson 1983), maintains genetic variation (Hedrick 1995; Brendonck & De Meester 2003) and increases effective population size (Nunney 2002). Third, despite the fact that resting egg banks may also slow the rate of evolution under certain conditions (Hairston & De Stasio 1988), *Daphnia* spp. have been found to adapt quickly to ecological change and there is evidence for rapid local adaptation and monopolization (Cousyn *et al.* 2001; De Meester *et al.* 2002, 2011; Van Doorslaer *et al.* 2010; Colbourne *et al.* 2011; Latta *et al.* 2012; Orsini *et al.* 2013).

A further goal of this study was to identify the most likely origin and phylogenetic position of the invading population in LLC. The species has morphologically been identified as *Daphnia pulicaria* FORBES 1893 (Einsle 1980; Stich & Maier 2007), a member of the *Daphnia pulex* complex. This complex comprises six nominal species and up to ten mitochondrial lineages, and its phylogeny is complicated by hybridization, introgression and polyploidization (Colbourne *et al.* 1998; Weider *et al.* 1999; Vergilino *et al.* 2009, 2011; Marková *et al.*

2013). Although *D. pulicaria* is mainly distributed in the Holarctic, it is also recorded from South America and Cuba (Flößner 2000). The nominal species *D. pulicaria* comprises two separate clades treated as 'European *D. pulicaria*' and 'North American *D. pulicaria*' in the recent literature (Colbourne *et al.* 1998; Vergilino *et al.* 2011; Marková *et al.* 2013). While the 'North American *D. pulicaria*' clade has long been recognized as a sister species of 'North American *D. pulex*', recent work suggests that the European and North American lineages of *D. pulicaria* are sister clades and 'North American *D. pulicaria*' has captured mitochondrial DNA from *D. pulex* in America (Marková *et al.* 2013). 'North American *D. pulicaria*' inhabits deep lakes and reservoirs distributed across the temperate and Arctic region of North America. However, 'European *D. pulicaria*' has been found in a wide range of habitats throughout Europe, ranging from meso- to eutrophic lowland fishponds and reservoirs to oligotrophic alpine lakes (Flößner 2000; Marková *et al.* 2007; Dufresne *et al.* 2011). Recent phylogenetic studies using mitochondrial DNA clearly separated lowland and alpine populations of 'European *D. pulicaria*' into different groups, with exception of populations from the High Tatra Mountains that cluster with both lowland and alpine populations (Dufresne *et al.* 2011; Marková *et al.* 2013). This pattern together with microsatellite results suggests postglacial colonization of the High Tatra mountains from different refugial sources. Hybridization between refugial races has been proposed as cause of obligate parthenogenesis in these populations (Dufresne *et al.* 2011). Interestingly, despite its distribution in the Alps (Flößner 2000; Marková *et al.* 2013), *D. pulicaria* has typically been absent from the large peri-alpine lakes north of the Alps, except for the reported invasion in LLC.

Daphnia egg banks have already successfully been used to reconstruct evolutionary processes and invasions (e.g. Hairston *et al.* 1999a,b; Duffy *et al.* 2000; Cousyn *et al.* 2001; Mergeay *et al.* 2005, 2006; Decaestecker *et al.* 2007; Brede *et al.* 2009; Rellstab *et al.* 2011). However, this is the first time that an invasion comprising more than a single genotype (Mergeay *et al.* 2006) and not being obscured by hybridization events (Brede *et al.* 2009; Rellstab *et al.* 2011) was investigated from its very onset using high-resolution markers to (i) unravel the genetic processes at work and (ii) infer the origin and phylogenetic position of the invading population.

Material and methods

Sampling site and sediment cores

Lower Lake Constance, situated along the border of Switzerland and Germany, represents the smaller and

shallower part of Lake Constance. During the last century, Lake Constance went through a phase of eutrophication, peaking in the 1980s, but has largely recovered since then (Fig. S1, Supporting information). The *Daphnia* community is dominated by the native *D. longispina*, the non-native *D. galeata*, which invaded in the 1950s as a consequence of eutrophication, and their hybrids (Brede *et al.* 2009). *Daphnia pulicaria* has been reported for the first time in 1974 (Einsle 1980) and its occurrence is restricted to LLC (Einsle 1987; Stich & Maier 2007). There, this large-bodied species predominately occurs in deep water layers and its abundance is strongly fluctuating between years, most likely governed by fish predation pressure (Stich & Maier 2007).

LLC can be divided into three sub-basins named Gnadensee (including Markelfinger Winkel), Zeller See and Rheinsee. Sediment cores with a diameter of 105 mm were collected using a gravity corer at two sampling locations in the Gnadensee basin, referred to as 'Markelfinger Winkel' (N47.7265445°/E9.0156945°; 17 m water depth) and 'Gnadensee' (N47.7071778°/E9.0665917°, 21 m water depth) in the literature (Einsle 1980; Stich & Maier 2007), and stored dark at 4 °C until further processing. Three cores from each site were halved and sliced into 0.5-cm-thick slices. To avoid smear contamination, 2–3 mm of the surface and the outer margins of each half core were removed. One half core per sampling site was dated using ¹³⁷Cs and ²¹⁰Pb profiles (Appleby 2002). All cores were aligned using the reconstructed age models and conspicuous sedimentological characteristics.

Ephippia preparation, DNA extraction and age model evaluation

Sediments were sieved through 224-µm and 250-µm mesh size sieves and *Daphnia pulex* species complex ephippia, which could reliably be distinguished from other ephippia produced by the *D. longispina* species complex, were selected under the stereomicroscope, counted and washed with autoclaved ultrapure water. Each ephippium was placed in a single drop of autoclaved water and flipped open with sterilized dissection needles. Eggs were assigned a quality value ranging from 1 (good quality) to 4 (bad quality) based on their colouration and texture and transferred into single 200-µL tube containing 25 µL alkaline lysis buffer. Then, eggs were crushed with a sterile pipette tip and DNA was extracted following the HotSHOT protocol (Montero-Pau *et al.* 2008) using 25 µL of the neutralizing buffer.

For an evaluation of our age models and core alignments, we calculated Spearman's rank correlations of

ephippia fluxes (number of ephippia $\times m^{-2} \times year^{-1}$) between (i) cores and (ii) coring locations between 1974 and 2011, and (iii) ephippia flux and annual means of pelagial abundance at the 'Gnadensee' sampling site reported by Einsle (1987) from 1975 until 1986, using the *cor.test* function in R, version 3.0.3. (R Core Team 2014). To account for the discrete nature of overlapping sediment slices as well as the general assumption that the bulk of ephippia are produced towards the end of the growing season, we predicted *n* equidistant values from *loess*-smoothed (R Core Team 2014) ephippia flux data and allowed for a lag of 1 *n*.

Mitochondrial and nuclear marker sequencing

To resolve the phylogenetic position and identify the most likely origin of *D. pulicaria* found in the LLC sediments, we sequenced mitochondrial and nuclear markers previously used in phylogenetic analyses of the *D. pulex* species complex (Dufresne *et al.* 2011; Vergilino *et al.* 2011; Marková *et al.* 2013).

A fragment (~711 bp) of the mitochondrial NADH dehydrogenase subunit 5 (*ND5*) was amplified for 15 resting eggs collected from different sediment layers in LLC and, additionally, for individuals from seven populations of 'European *D. pulicaria*'. Moreover, *ND5* sequences for four different LLC clones, hatched from different depths from the LLC egg bank in this study, were sequenced in the context of a phylogenetic study of the *D. pulex* complex by Marková *et al.* (2013). We used recently developed primers *ND5newF* 5'-AAA CCT CTA AAB TTC YKA RCT- 3' and *ND5newR* 5'-CAT RTT YAT RTC RGG GGT TGT- 3' and a modified PCR protocol following Dufresne *et al.* (2011). The PCR mix contained 0.625 U/ μ L GoTaq® Flexi Polymerase (Promega AG, Dübendorf, Switzerland), 1X Colorless GoTaq® Flexi Buffer (Promega AG), 2 mM MgCl₂, 0.2 mM dNTP's, 0.4 μ M of each primer (Microsynth AG, Balgach, Switzerland) and 3 μ L of extracted DNA in a total reaction volume of 25 μ L. The PCR conditions were as follows: initial denaturation at 94 °C for 2 min, followed by 38 cycles of denaturation at 94 °C for 40 s, annealing at 53 °C for 60 s and extension at 72 °C for 90 s with a final elongation step of 15 min at 72 °C.

Further, we amplified a fragment (~550 bp) of the nuclear gene coding the small GTPase *Rab 4* for seven LLC resting eggs using primers *F6for* 5'-CGT TTC GAA TTG GCT TAC TGA-3' and *F12rev* 5'-CAT GGT TAT CTG TCT ACG TCT TGA A-3' (Omilian *et al.* 2008). Each 25 μ L PCR contained 0.625 U/ μ L GoTaq® Flexi Polymerase (Promega AG), 1X Colorless GoTaq® Flexi Buffer (Promega AG), 2 mM MgCl₂, 0.2 mM dNTP's, 0.2 μ M of each primer (Microsynth AG) and 3 μ L of extracted DNA. The PCR started with a 2-min initial

denaturation step at 94 °C, followed by 40 cycles of denaturation at 94 °C for 60 s, annealing at 53 °C for 60 s extension at 72 °C for 90 s with a final elongation step of 10 min at 72 °C.

ND5 and *Rab4* PCR products were checked on a 1.2% agarose gel, purified using the Wizard® SV Gel and PCR Clean-Up System (Promega AG) and directly sequenced without cloning in both directions by a commercial sequencing service (Microsynth AG). Chromatograms were inspected visually, and sequences were edited in MEGA 6.0 (Tamura *et al.* 2013). All chromatograms were clear and of high quality, and sequences obtained from both directions were consistent, however, given our sequencing strategy we cannot fully exclude the possibility of allele dropout.

For the phylogenetic tree reconstructions, we compared and complemented *ND5* and *Rab4* data sets with sequences used in a comprehensive study of the *D. pulex* species complex by Marková *et al.* (2013) and additional sequences available on GenBank (Tables S1 and S2, Supporting information). Sequences of 'European *D. pulex*' were used as an outgroup in the *ND5* and the *Rab4* phylogeny, respectively. Alignments, nucleotide substitution model selection based on the Bayesian information criterion and maximum-likelihood (ML) tree reconstruction, were performed in MEGA 6.0 (Tamura *et al.* 2013). The *ND5* phylogeny was inferred using an alignment of 98 sequences with a length of 591 bp, and the Hasegawa Kishino Yano nucleotide substitution model with gamma-distributed rate heterogeneity (HKY + G) was chosen as the best model (Hasegawa *et al.* 1985). For the *Rab4* ML tree, we used an alignment of 82 sequences of 479 bp length and the Tamura 3-parameter nucleotide substitution model with gamma-distributed rate heterogeneity (T92 + G) (Tamura 1992). The node support in each tree was assessed by 1000 bootstrap replicates.

Microsatellite genotyping

To assess genetic variation and population structure of the *D. pulicaria* egg bank in LLC, thirteen microsatellite markers, which had previously been shown to be polymorphic in studies on the *D. pulex* species complex (Colbourne *et al.* 2004; Dufresne *et al.* 2011; Pantel *et al.* 2011; Vergilino *et al.* 2011), were optimized and combined into four multiplex PCRs (MP1-4). All reverse primers were pigtailed (Brownstein *et al.* 1996), and for Dp525altMM and Dp78MM, new primers were designed to improve peak quality. Forward primers were fluorescently labelled (Table S3, Supporting information). All primers were purchased from Microsynth AG. PCRs contained 1.5 μ L of extracted DNA, 5.75 μ L Multiplex PCR Master Mix (Quiagen, Hilden,

Germany), forward and reverse primers at equimolar concentrations (MP1: 0.2 μM Dp183Mark, 0.7 μM Dp502Mark, 0.3 μM Dp512 and 0.35 μM Dp525; MP2: 0.25 μM Dp525altMM, 0.25 μM Dp513, 0.5 μM Dp514 and 0.25 μM Dp196; MP3: 0.3 μM Dp78MM, 0.25 μM Dp519, 0.4 μM Dp514alt; MP4: 0.3 μM Dp433, 0.6 μM Dp461) and were filled up with PCR-grade water to a final volume of 11.5 μL . The PCR amplification protocol consisted of an initial denaturation step at 95 °C for 15 min, followed by 33 (MP1, MP2) or 37 (MP3, MP4) cycles of denaturation at 94 °C for 30 s, annealing at 52 °C (MP1, MP4), 54 °C (MP2) or 49 °C (MP3), respectively, for 90 s and extension at 72 °C for 60 s with a final elongation period of 30 min at 60 °C.

PCR products were diluted 1:10 (MP1, MP2) and 1:15 (MP3, MP4), respectively, in ultrapure water and 0.5 μL of the diluted PCR product was mixed with 9.25 μL of highly deionized formamide and 0.2 μL GeneScan500 LIZ size standard (Life Technologies, Carlsbad, CA, USA) and run on an ABI 3130XL Genetic Analyzer (Life Technologies) for fragment analysis. Binning and scoring of microsatellite alleles were done in GENEMAPPER v.4 (Life Technologies). All samples and alleles were double-checked manually. For rare alleles, PCRs and genotyping were repeated.

Population genetic analysis

We employed two approaches to analyse the development of genetic variation and population structure over time. First, microsatellite data were pooled into eight periods spanning four to five years (Table 1). Periods were assigned according to distinct abundance peaks of *D. pulicaria* in the pelagial separated by times of virtual absence reported by Einsle (1987) (Fig. S2, Supporting information) and Stich & Maier (2007) and under consideration of the dating of the respective sediment slices from six different cores. Second, we used sliding window analyses to control for potential biases of our periods assignment. Several window sizes and step settings were examined, and results for windows of 2.5 years advancing by steps of 1 year are reported.

Standard population genetics parameters and diversity indexes were calculated for each period and sliding window. To assess genetic variation for the defined periods, number of multilocus genotypes (MLGs), genotypic richness (R), number of alleles and unbiased heterozygosity (unbiased H_e) were calculated and tested for significance with the program GENCLONE, version 2.0 (Arnaud-Haond & Belkhir 2007). To compare periods of different sample size, we also computed genotypic richness, unbiased heterozygosity and number of alleles for the minimum sample size ($N = 19$) running 4000 permutations using the same software. Allelic richness

after rarefaction was estimated with the R-package *hierfstat* (Goudet 2005). To investigate mating system and reproductive mode, allele frequencies, linkage disequilibrium (LD; 16 000 permutations) and an exact test of Hardy Weinberg equilibrium (1 000 000 steps in the Markov chain and 100 000 dememorization steps) were calculated with the software ARLEQUIN, version 3.5 (Excoffier & Lischer 2010), and F_{IS} values with 95% confidence intervals (CIs) were estimated in GENETIX, version 4.05 (Belkhir *et al.* 1996–2004), using 10 000 bootstrap replicates. In addition, a robust estimate of selfing rates (s) and 95% CIs were calculated with the software RMES and selfing rate-based F_{IS} values and 95% CIs were calculated according to the formula $F_{IS} = (s/2-s)$ (David *et al.* 2007; Kopp *et al.* 2012). F_{IS} values and selfing rates were also calculated for a separate data set excluding repeated MLGs (unique MLGs data set), to avoid a potential bias introduced by the inclusion of putatively obligate parthenogenetic genotypes. To test whether resting eggs may be produced clonally in LLC, the probability that individuals sharing the same MLG originate from different sexual reproductive events (Psex) was calculated using MLGSIM program, version 2.0 (Ivens *et al.* 2014), an updated version of MLGSIM (Stenberg *et al.* 2003). Statistical significance was assessed by running 1000 simulations. The analysis was repeated for each period and for the full data set under the assumption of random mating (MODEL = HWE) as well as inbreeding (MODEL = FIS). For each analyses, allele frequencies were estimated from either the entire sample (FREQUENCY = SAMPLE) or a subset of individuals including each MLG only once (FREQUENCY = MLG).

To exclude biases due to the grouping into different periods, sliding window analyses were performed using custom scripts in R, version 3.0.3. (R Core Team 2014). Heterozygosities were calculated using the *adegenet* package (Jombart 2008). Numbers of MLGs were obtained with the *poppr* package (Kamvar *et al.* 2014), and for calculating genotypic richness, R custom scripts were used. Allelic richness after rarefaction and selfing rates were estimated as described above.

Furthermore, to assess whether the LLC invasion was a single event or a series of events, discrete or gradual, and whether the same source population or several distinct source populations were involved, population structure was analysed using methods with prior grouping information as well as with individual-based methods without prior information. To assess the presence of distinct clusters and divergence between consecutive periods, a discriminant analysis of principal components (DAPC), a method that does not rely on an underlying population genetics model, was performed using the *adegenet* package (Jombart 2008) in R version 3.0.3. (R Core Team 2014). Prior groups for DAPC were

Table 1 Summary table of population genetics parameters calculated for samples from the *Daphnia pulex* egg bank in LLC pooled into eight consecutive periods

Period	Time span	N	MLGs	R	R*	N alleles	Ar (SD)	unbiased H_e	unbiased H_e (SE) [†]	F_{is} (95% CI)	F_{is} (95% CI) [‡]	s (95% CI)	s (95% CI) [‡]
2007–2011	2006 8–2011 7	43	43	1 00	1 00	29 69 (0 017)	2 28 (0 95)	0 357	0 356 (0 0002)	0 017 (-0 079, 0 09)	0 017 (-0 079, 0 09)	0	0 (0, 0 216)
2002–2006	2001 9–2006 8	42	41	0 98	0 99	28 69 (0 008)	2 21 (0 97)	0 345	0 345 (0 0002)	-0 014 (-0 12, 0 066)	-0 009 (-0 114, 0 072)	0 001	0 044 (0, 0 202)
1997–2001	1997–2001 9	24	24	1 00	1 00	28 75 (0 007)	2 21 (0 98)	0 342	0 341 (0 0001)	0 159 (-0 014, 0 289)	0 159 (-0 016, 0 283)	0 183	0 183 (0, 0 398)
1992–1996	1992 1–1997	87	87	1 00	1 00	28 79 (0 007)	2 22 (0 99)	0 339	0 339 (0 0002)	0 071 (0 003, 0 129)	0 071 (0 003, 0 131)	0 08	0 080 (0, 0 210)
1987–1991	1987 2–1992 1	77	74	0 96	0 99	28 28 (0 012)	2 17 (0 97)	0 309	0 309 (0 0002)	-0 038 (-0 111, 0 023)	-0 033 (-0 110, 0 03)	0	0 (0, 0 133)
1983–1986	1982 6–1987 2	135	107	0 79	0 88	28 84 (0 027)	2 22 (0 92)	0 312	0 311 (0 0003)	0 207 (0 128, 0 28)	0 070 (-0 003, 0 136)	0 501	0 228 (0 113, 0 331)
1979–1982	1978 8–1982 6	92	65	0 70	0 80	28 17 (0 043)	2 16 (0 93)	0 266	0 264 (0 0003)	0 241 (0 147, 0 325)	0 073 (-0 016, 0 152)	0 567	0 240 (0 091, 0 363)
1974–1978	1974–1978 8	19	16	0 83	0 83	19 00 (0 000)	1 46 (0 52)	0 197	0 197 (0 0000)	0 141 (-0 211, 0 41)	0 147 (-0 205, 0 424)	0 52	0 466 (0 107, 0 645)

N, Number of genotyped resting eggs; MLGs, Multilocus genotypes; R, genotypic richness; N alleles, Number of alleles; Ar, allelic richness after rarefaction; unbiased H_e , unbiased heterozygosity; s, selfing rate

*R, genotypic richness, calculated from MLG estimates corrected for minimum sample size (19)

[†]Values corrected for minimum samples size (19) and their standard errors (SE) calculated using 4000 permutations in GENCLONE 2 0

[‡]Values calculated without repeated MLGs

defined based on (i) the eight predefined periods and (ii) an unbiased K-means clustering with the implemented function *find.clusters* evaluating a range of 1–15 clusters. For the latter case, informative numbers of K were inferred by visual inspection of Bayesian information criterion (BIC) values for different values of K and by the *find.clusters* function with the ‘diffNgroup’ criterion. For all analyses, the *optim.a.score* function was used to select the optimum number of PCA axis (Table S4, Supporting information).

In addition, we compared the differentiation between periods by calculating pairwise Jost’s D (Jost 2008) and F_{ST} (Weir & Cockerham 1984) using the *diversity* package (Keenan *et al.* 2013).

Moreover, to avoid potential biases of prior grouping, individual-based analyses of population structure assuming Hardy Weinberg equilibrium and linkage equilibrium were performed with *STRUCTURE*, version 2.3.4. (Pritchard *et al.* 2000; Falush *et al.* 2003) following recommendations of Gilbert *et al.* (2012). A range of clusters $K = 1$ to 10 was tested using a model of admixture with correlated allele frequencies, a burn-in length of 100 000 followed by 250 000 MCMC repetitions. For each value of K , 20 independent replicates were run and summary statistics values were checked for convergence. The program *STRUCTURE HARVESTER* (Earl & vonHoldt 2012) was utilized to summarize likelihood statistics, calculate ΔK and format data for downstream alignment of replicate runs in *CLUMPP*, version 1.1. (Jakobsson & Rosenberg 2007), using the *FULLSEARCH* and *GREEDY* algorithm. Assignment plots for each value of K were produced in *R* version 3.0.3 (R Core Team 2014). To find informative values for K , we combined visual inspection of likelihoods, Pritchard’s (Pritchard *et al.* 2000, 2007), and Evanno’s method (Evanno *et al.* 2005), respectively.

Furthermore, an individual-based factorial correspondence analysis (FCA) implemented in *Genetix* 4.05 (Belkhir *et al.* 1996–2004) was performed to visualize the development of genetic variation since the first invasion of this *D. pulex* in LLC. Six rare genotypes from different populations with a strong impact on the FCA were removed from this analysis as recommended (Jombart *et al.* 2009).

All analyses of population structure were repeated for a data set without repeated MLGs (unique MLGs data set).

Hatching experiment

To test the functionality of the LLC egg bank and assess hatching of resting eggs over time, we exposed a total of 341 ephippia from three Markelfinger Winkel cores to two main hatching cues, light and increased temperature. Ephippia were incubated singly in filtered lake

water (0.45- μ m filter, Sartorius Stedim AG, Switzerland) in 48-well plates and exposed to 20 °C and a light dark cycle of 16:8 h. Hatching was assessed every other day over a minimum period of 2 weeks. Hatchlings were transferred individually to 250-mL jars and kept in culture. Hatching success was calculated as hatchlings per total number of resting eggs and as hatchlings per total number of eggs with quality 1. After the first round of clonal reproduction, hatchlings as well as hatchlings that had died before reproduction were stored in 95% pure ethanol (Merck KGaA, Darmstadt, Germany) until analysis. DNA was isolated from ethanol samples and from all nonhatched eggs and included in the microsatellite analyses (see above). Hatching success was calculated for each period and sliding window and confidence intervals (95%) were estimated using the *binom* package in *R*, version 3.0.3 (R Core Team 2014), using Wilson’s method.

Results

Age model and ephippia distribution

The ^{137}Cs and ^{210}Pb age models for Markelfinger Winkel and Gnadensee sediment cores revealed very similar results. Due to the distinct ^{137}Cs peak in 1986 (Chernobyl nuclear accident), the ^{137}Cs models gave a better resolution for more recent time periods and were therefore used for dating. We extracted a total number of 1407 ephippia from six sediment cores. While the oldest sediment layers examined across cores dated to the beginning of the 20th century, the oldest *Daphnia pulex* ephippia appeared in layers dated between end of 1973 and beginning of 1975 (Fig. 1), which is consistent with

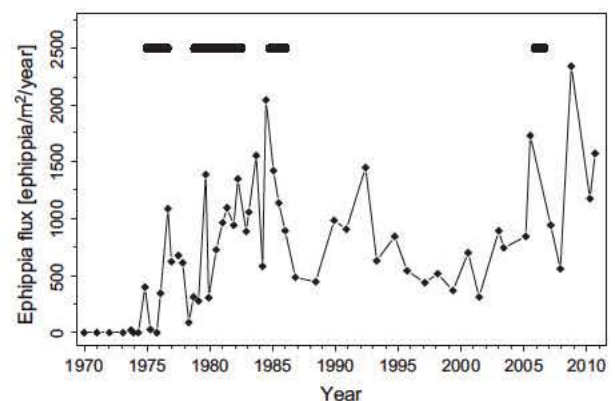


Fig. 1 Ephippia flux in the sediments of ‘Gnadensee’ and ‘Markelfinger Winkel’. Data are pooled for all aligned cores. Each data point represents the mean value of a minimum of 3 sediment slices. Black bars indicate continuous periods of reported pelagic abundance of ‘European *Daphnia pulex*’ (Einsle 1987; Stich & Maier 2007).

the first report of *D. pulicaria* in LLC in the year 1974 (Einsle 1980, 1987). In general, patterns of ehippia flux (number of ehippia $\times m^{-2} \times year^{-1}$) were comparable and significantly positively correlated across cores ('Gnadensee': $n = 25$, Spearman's $\rho = 0.38$ – 0.75 , $P < 0.05$; 'Markelfinger Winkel': $n = 26$, Spearman's $\rho = 0.41$ – 0.52 , $P < 0.05$) and coring locations ($n = 31$, lag = 1, Spearman's $\rho = 0.45$, $P < 0.05$). Moreover, ehippia flux at the 'Gnadensee' sampling site is significantly positively correlated with annual means of *D. pulicaria* pelagial abundance ($n = 12$, lag = 1, Spearman's $\rho = 0.67$, $P < 0.05$) reported from 1975 until 1986 (Fig. S2, Supporting information) from this sampling site (Einsle 1987).

Phylogenetic analysis of mitochondrial and nuclear sequence data

In this study, we obtained 15 sequences of the *ND5* gene (~711 bp) from the LLC resting eggs and combined them with four additional sequences from LLC hatchlings (Marková *et al.* 2013). The phylogenetic analysis revealed three distinct *ND5* haplotypes (LLC1, LLC2 and LLC3) in the LLC egg bank (Fig. 2, Table S1, Supporting information). LLC1 and LLC2 are distinguished from each other by a single nucleotide substitution, whereas both are distinguished from LLC3 by two and one nucleotide substitutions, respectively. LLC1 represents a widely distributed haplotype in European lowland lakes and ponds. In addition, we revealed that during a very recent invasion, LLC1 was introduced also into Lake Greifensee (Switzerland) where *D. pulicaria* was not observed until 2012 (M. Möst, unpublished data). The most common haplotype LLC2 was identified in 11 samples, LLC1 in seven (two of which were siblings) and LLC3 only in one sample, respectively. In isolates from the period 1974 to 1981, only haplotype LLC2 was detected. The phylogenetic reconstruction of a ML tree (log likelihood = 3473.95) (Fig. 2) revealed that all *ND5* haplotypes found in LLC clustered within European *D. pulicaria* haplotypes. Moreover, LLC haplotypes were identical or closely related to European *D. pulicaria* lowland populations from the Czech Republic, Germany, Switzerland, Spain, Albania, Montenegro, Norway, Poland, Germany and High Tatra Mountains (HTM) populations, but clearly distinct from neighbouring alpine populations from Switzerland, Austria, Northern Italy and the populations from the Pyrenees and Sierra Nevada (see also Marková *et al.* 2013).

We identified two *Rab4* haplotypes (LLCA and LLCB) among individuals from LLC that differed only by a single nucleotide substitution. While five individuals were homozygous for the LLCA haplotype, two heterozygous individuals carried LLCA as well as

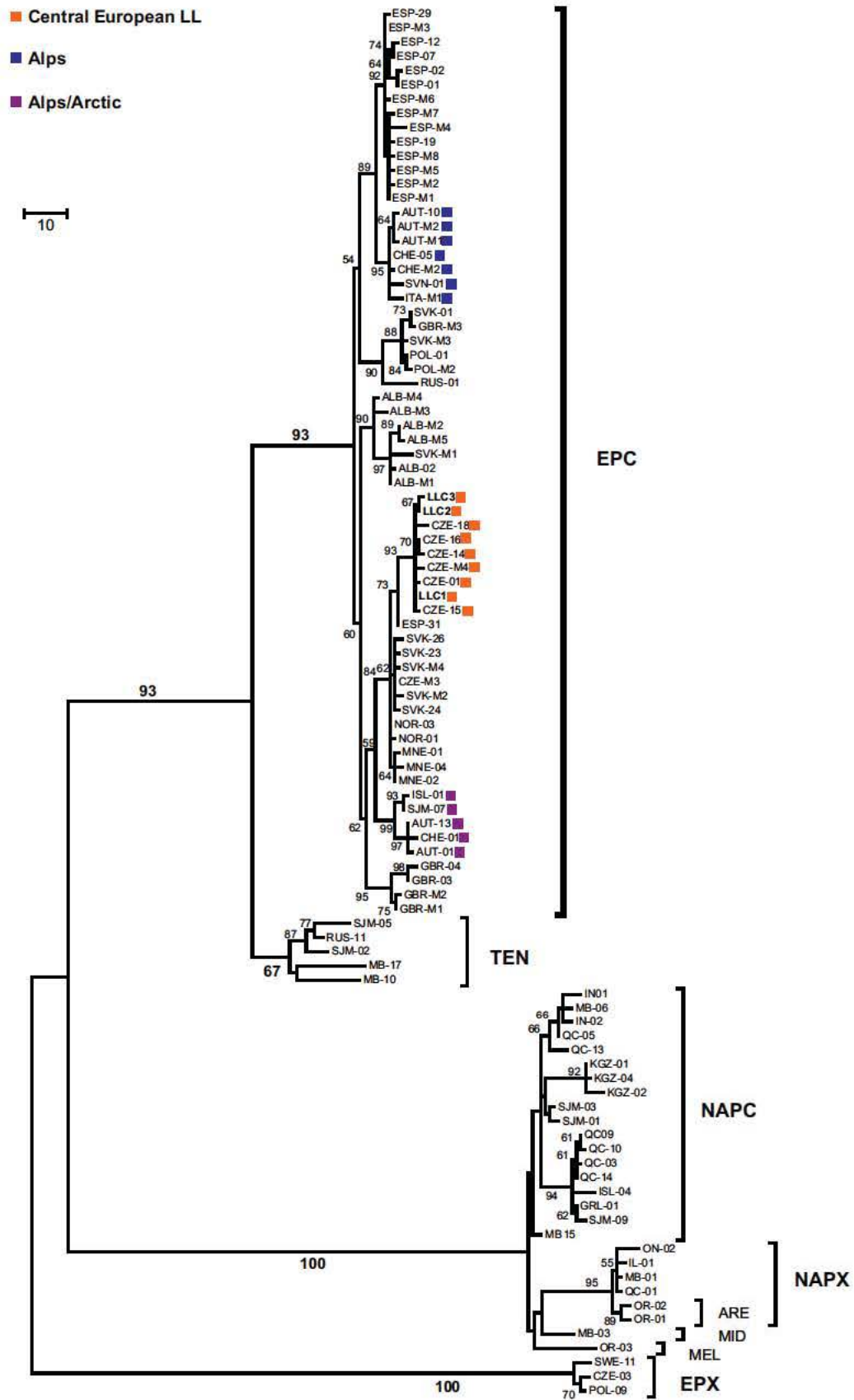
LLCB haplotypes. Maximum-likelihood phylogenetic analysis (log likelihood = 1684.37) placed the LLC *Rab4* haplotypes into a clade formed by individuals identified as 'European *D. pulicaria*' (Fig. S3, Table S2, Supporting information, see also Fig. 2 in Marková *et al.* 2013). The LLCA haplotype is identical to a haplotype reported from lowland populations in the Czech Republic, Poland, and the United Kingdom (King George Reservoir), while the LLCB sequence is one of the most widely distributed *Rab4* haplotypes in Europe (Table S2, Supporting information) (see also Fig. 2 in Marková *et al.* 2013).

Microsatellite genotyping and population genetic analysis

After quality filtering and double-checking, we retained 519 highly reliable genotypes, which were used in subsequent statistical analyses. In total, 37 alleles were detected at thirteen microsatellite loci. The average number of alleles per locus was 2.85 with a maximum of eight alleles at locus Dp514alt (Fig. 3). Eleven loci were polymorphic; however, loci Dp513 and Dp514 were monomorphic in the LLC egg bank, although they are known to be polymorphic in other 'European *D. pulicaria*' populations (Dufresne *et al.* 2011; Vergilino *et al.* 2011, S. Marková, unpublished data).

Overall, both the analyses based on eight predefined populations and on sliding windows revealed an increase in genetic and genotypic diversity over time (Table 1, Figs 3 and 4 and S4, Supporting information). We found a sharp increase in number of alleles and allelic richness after the first appearance of *D. pulicaria* that subsequently started to approach a plateau from the mid-1980s (Table 1, Figs 3 and 4, and S4, Supporting information), corresponding to period 1983–1986 when a third phase of high pelagial abundances in Gnadensee was recorded (Einsle 1987) (Fig. S2, Supporting information). Strikingly, we never observed more than two alleles per locus during the first appearance of *D. pulicaria* in period 1974–1978. A number of 17 new alleles appeared at low frequencies in periods 1979–1982 and 1983–1986. While six of these alleles were not recovered in the following periods, 10 alleles were established in the egg bank and were found in each subsequent period, and one allele was only reencountered during period 2007–2011 (Fig. 3). In the most recent period 2007–2011, one additional new allele was observed (Dp78MM 201, Fig. 3).

Genetic diversity measured as unbiased expected heterozygosity was found to increase in all analyses by approximately 45% (Table 1, Figs 4 and S4, Supporting information), with a steep increase during the initial phase, and reached a level of 0.36 in period 2007–2011.



Considerable departures from Hardy Weinberg equilibrium and linkage disequilibrium, both measures that are affected by inbreeding and selfing, were detected in periods 1979–1982 and 1983–1986. This was also reflected by significantly increased F_{IS} values in these periods and marginally but significantly increased F_{IS} values in periods 1992–1996 for the complete data set. However, F_{IS} values for the unique MLGs data set (457 genotypes) were not significantly different from zero for periods 1979–1982 and 1983–1986 (Table 1). Moreover, considerably high and significant selfing rates were observed in the three earliest periods 1974–1978, 1979–1982 and 1983–1986. Selfing rates calculated for the unique MLGs data set were lower but remained significant for the same three periods (Table 1). A comparison of F_{IS} values with F_{IS} values calculated from selfing rates revealed consistent results and did not suggest the presence of null alleles (data not shown). Sliding window analyses showed a consistent pattern with elevated selfing rates until the mid-1980s (Figs 4G and S4G, Supporting information).

Altogether, 424 multilocus genotypes (MLGs) were resolved for the full data set with all samples pooled, and a total of 457 MLGs were detected when predefined periods were analysed separately and MLGs were summed up, ranging from 16 MLGs in period 1974–1978 to 107 MLGs in period 1983–1986 (Table 1). Identical MLGs and consequently reduced genotypic richness were found during five of the eight periods, predominantly during early stages of the invasion with a peak around 1980 (Table 1, Figs 4D and S4D, Supporting information). Notably, these observations could mainly be attributed to the multiple occurrence of a single MLG that was completely homozygous across all loci, occurred already in the founding period 1974–1978 (2 eggs) and reached abundance peaks in consecutive periods 1979–1982 (20 eggs) and 1983–1986 (22 eggs) with very low and highly significant P_{sex} values ($P < 0.001$) (Table S5, Supporting information). Only a few additional identical MLGs were found, always at low copy numbers (two to five copies), often representing siblings, and in most cases, P_{sex} values were significant only under the more conservative assumption of departures from Hardy Weinberg equilibrium (Table S5, Supporting information). Moreover, most MLGs were

unique suggesting that ephippia are generally the result of sexual reproduction in LLC.

The results of population structure analyses indicated a gradual change of population structure through time rather than a sequence of clearly distinguishable populations. A discriminant analysis of principal components (DAPC) based on the eight predefined periods reflected genetic change through time, but no clear separation into distinct consecutive clusters was evident (Figs 5, S5 and S6, Supporting information). Successive periods overlap, which was best illustrated by their density distributions on the first discriminant function axis (Fig. 5) and the admixture of individuals (Figs S5 and S6, Supporting information). This finding is supported by overall low values of pairwise Jost's D and F_{ST} that increase with temporal distance between pairs of periods (Fig. S12, Table S6, Supporting information).

For the DAPC analysis based on unbiased K-means clusters, the inspection of the BIC values did not indicate strong population structure (Figs S7A and S8A, Supporting information). The *find.clusters* function suggested $K = 7$ for data sets with and without repeated MLGs. Also for the STRUCTURE analyses, the interpretation of different values of K was not straightforward with putatively informative values for K ranging from $K = 2$ to $K = 8$ depending on the method and data set used (Figs S9A,B and S10A,B, Supporting information). A combined consideration of likelihood plots and Evanno's and Pritchard's method suggested that a maximum of $K = 6$ for the full data set and $K = 5$ for the unique MLGs data set capture the major structure in the data. Assignment plots for a range of $K = 2$ to $K = 7$ for DAPC and STRUCTURE methods are shown (Figs S7B G, S8B G, S9C H, and S10C H, Supporting information). In agreement with a lack of strong population structure, individuals were often admixed and no distinct consecutive clusters were observed. Such a pattern is typically found in the case of 'isolation by distance', or, as in our case, overlap of generations as a consequence of the egg bank rendering the interpretation of such results challenging (see STRUCTURE 2.3 Documentation). The contribution of different inferred clusters, however, was changing over time and new clusters emerged, often corresponding to the appearance of new alleles. The only clear pattern was the consistent discrimination of

Fig. 2 Maximum likelihood phylogenetic tree of the mitochondrial *ND5* gene for the *Daphnia pulex* species complex. Numbers on major branches represent percentage bootstrap values larger than 50% (1000 replications). The scale bar indicates number of substitutions. Haplotypes found in LLC (LLC1, LLC2, and LLC3) are shown in bold, and sample information is summarized in Table S1. Orange squares indicate samples forming a well supported cluster of Central European haplotypes including all three LLC haplotypes. Blue and purple squares indicate clusters comprising haplotypes found in the neighbouring populations in the Alps. (EPC 'European *Daphnia pulicaria*', TEN *D. tenebrosa*, NAPC 'North American *D. pulicaria*', NAPX 'North American *D. pulex*', ARE *D. arenata*, MID *D. middendorffiana*, MEL *D. melanica*, EPX 'European *D. pulex*').

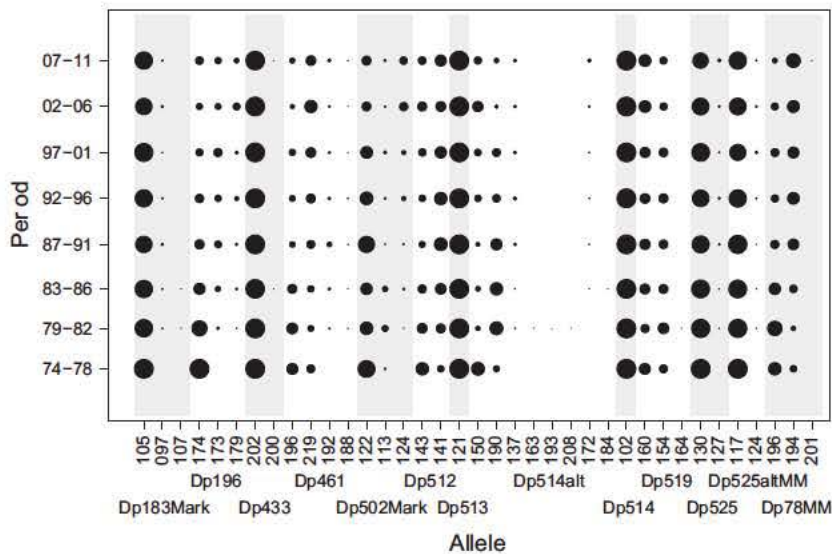


Fig. 3 Allele frequencies at thirteen microsatellite loci in the LLC egg bank for eight consecutive periods (1974–2011). Numbers on the x axis indicate allele length and labels indicate loci. Circles are drawn at sizes relative to allele frequencies.

one cluster driven by the repeated fully homozygous MLG in DAPC and STRUCTURE analyses of the full data set (Figs S7B G and S9C H, Supporting information; dark blue cluster).

The representation of the data by a factorial correspondence analysis (FCA), in which the first two axes explained 31.7% of the genetic variation, reflected an increase in genetic variation, in particular during the early stages, and a partial overlap of different time periods (Fig. S11, Supporting information).

Hatching experiment

Throughout the hatching experiment, we observed a total of 86 hatchlings from 372 exposed eggs. Hatching success was higher (>60%) for ephippia from more recent layers (Figs 4H and 6), in particular for period 1997–2001 (>80%). We also found a clear decrease in hatching success with age (Figs 4H and 6). No hatching was observed in periods 1974–1978, and the single hatchling observed in period 1979–1982 was infertile. Interestingly, opposing trends were observed for hatching success and selfing rates (Figs 4G, H and 6).

Discussion

In the present study, we aimed to reconstruct the history and population genetics of an aquatic invasion from resting eggs deposited in the sediments of LLC. Our analysis of the egg bank clearly revealed that a member of the *Daphnia pulex* species complex has invaded LLC around 1974 during the peak of eutrophication and subsequently established a population with a functional resting egg bank. Patterns of ephippia abundance and flux are highly consistent with reports

on pelagial densities (Einsle 1980, 1987; Stich & Maier 2007), suggesting that our reconstruction of the population history of *Daphnia pulex* in LLC is accurate. This combination of historical records and the dating of the egg bank allows us to determine the timing of the invasion with high certainty. Obtaining such information independently from genetic data is crucial for study systems with germ banks because resting stage banks may impede the inference of past demographic events from recent patterns of genetic diversity (Živković & Tellier 2012).

Our mitochondrial and nuclear phylogenies together with findings from Marková *et al.* (2013) using more recent LLC hatchlings confirmed that LLC has been colonized by the 'European *D. pulex*' lineage rather than other invasive members of the *D. pulex* species complex recently reported in Europe (Marková *et al.* 2007; Fadda *et al.* 2011). 'European *D. pulex*' inhabits strongly contrasting habitats, that is meso- to eutrophic lowland lakes and ponds as well as oligotrophic high-altitude lakes (Hrbáček 1959, 1977; Flößner 2000; Marková *et al.* 2007; Dufresne *et al.* 2011). The clade comprising all LLC mitochondrial haplotypes and haplotypes from Central European lowland populations (Czech Republic, Poland, and Germany) is clearly distinct from other lineages, in particular the high-altitude populations found in the surrounding Alps or in the Pyrenees. The fact that haplotype LLC1 is also found in one alpine lake in the HTM is remarkable. However, Dufresne *et al.* (2011) have shown already that the HTM have recently been colonized from different glacial refugia and therefore share some haplotypes with lowland populations (see Fig. 4 in Dufresne *et al.* 2011). Based on the currently available data, one mitochondrial haplotype (LLC3) is unique to LLC and was only found in one sample from the

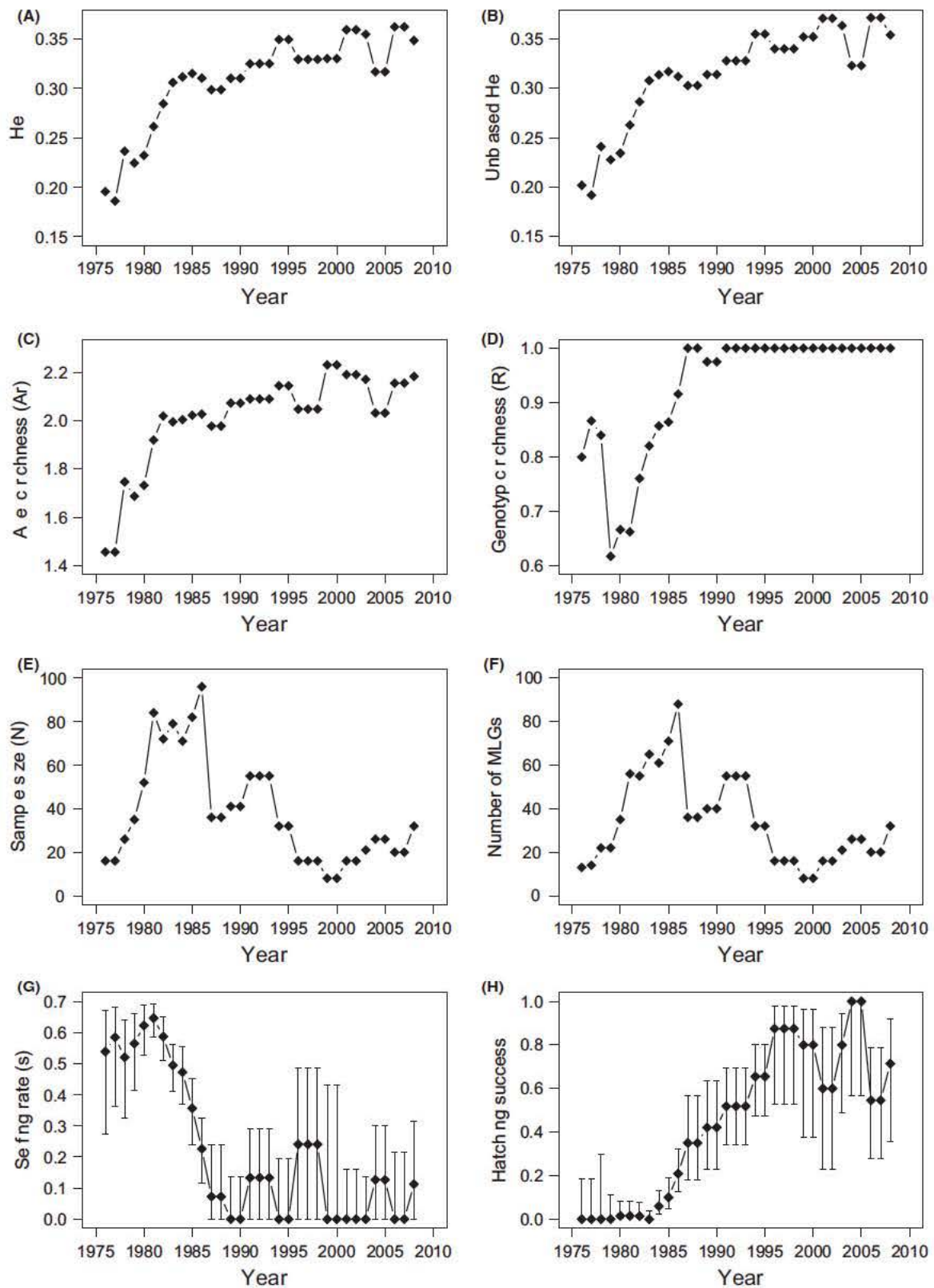


Fig. 4 Sliding window analyses of all samples from the LLC egg bank. For each window heterozygosity (A), unbiased heterozygosity (B), allelic richness (C), genotypic richness (D), sample size (E), number of MLGs (F), selfing rate (G) and hatching success (H) are shown (window width: 2.5 years).

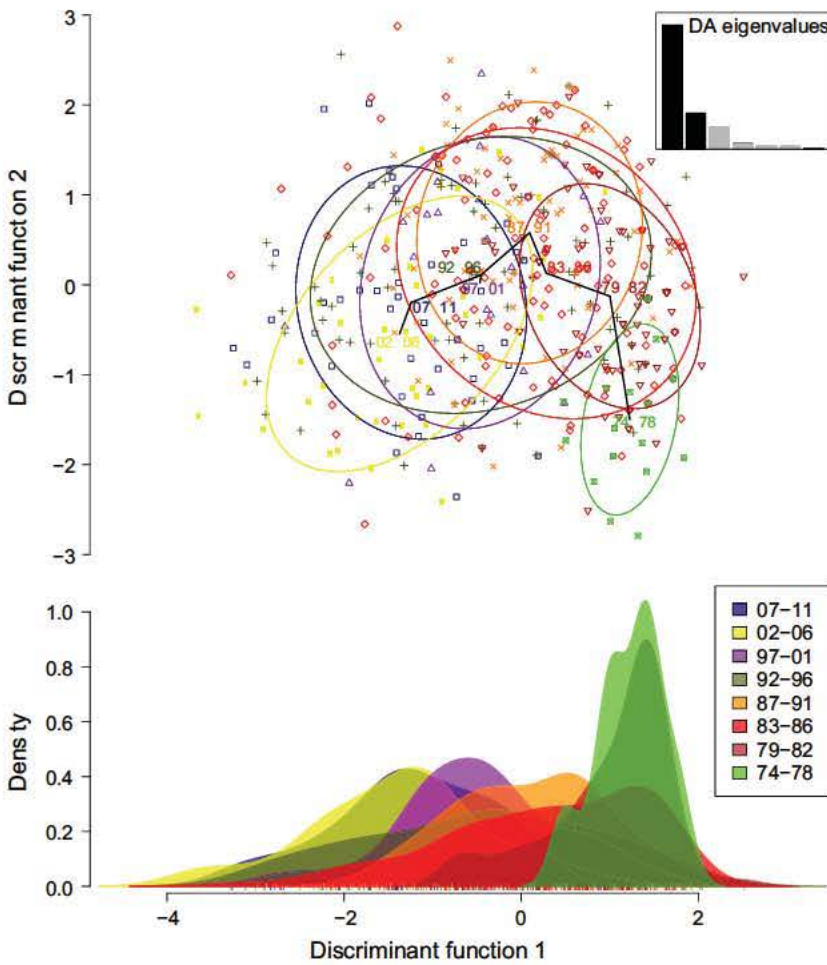


Fig. 5 Results of a discriminant analysis of principal components (DAPC) for samples from the *Daphnia pulicaria* egg bank in LLC pooled into eight consecutive periods. The upper panel depicts a scatter plot using discriminant function 1 and discriminant function 2. Periods are labelled and depicted with different colours and symbols. Inertia ellipses (67%) and a minimum spanning tree are also shown. The lower panel presents the density distribution on discriminant function 1 for each population.

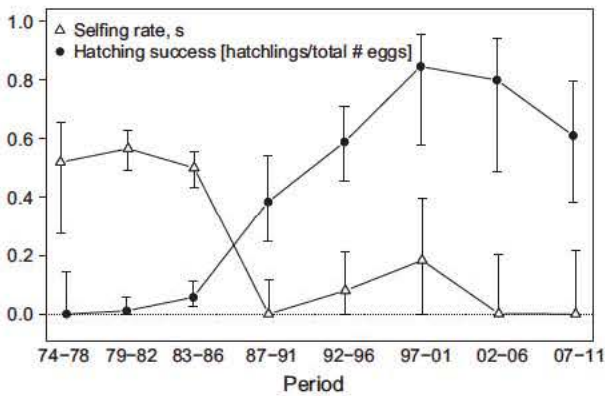


Fig. 6 Hatching success and selfing rate (corrected for repeated MLGs) and their 95% confidence intervals for *Daphnia pulicaria* in LLC for eight consecutive periods.

late 1980s, and the fact that it differs by only one nucleotide substitution from LLC2 indicates that it may have evolved locally. The overall pattern therefore strongly suggests that the LLC population originated from meso- to eutrophic lowland ponds in Central

Europe. However, a more specific determination of the location of origin is limited by the resolution of the phylogenetic trees. The suggested origin of the LLC population from lakes with high trophic state is in accordance with earlier ideas that this invasion has been facilitated by the eutrophication of Lake Constance (Flößner 1972; Einsle 1987). However, a better characterization of the ecological preferences and nutrient requirements for the different lineages is needed to support this conclusion.

The initial phases of invasions and the temporal patterns and sources of genetic variation, for example a single introduction or multiple introductions from either the same or distinct populations or mutations, are of particular interest for understanding the invasion process and the conditions resulting in successful establishment (Roman & Darling 2007; Dlugosch & Parker 2008; Bock *et al.* 2015). For many organisms and in particular for aquatic species in large water bodies, it is likely that these initial stages are overlooked and therefore not completely captured. In addition, many invasions have started at a time when appropriate genetic methods have not yet been available. Here, we

attempted to overcome these limitations using the resting egg bank of 'European *D. pulicaria*' as a proxy to reconstruct the invasion of LLC from the very beginning. The temporal patterns of genetic variation and LLC population structure suggest a scenario of multiple introductions from genetically similar source populations, possibly the same source. This is clearly different from invasions of a single asexual genotype described for another member of the *D. pulex* species complex in African lakes (Mergeay *et al.* 2006). While the diversity of the founding population is extremely low, with no more than two alleles per locus, and could therefore in principal represent an invasion by a single clone, all measures of genetic diversity and richness show a steep increase over the first 12 years and genetic variation reaches levels that lie within a range that is commonly encountered in populations of 'European *D. pulicaria*' (Dufresne *et al.* 2011). Moreover, the concerted occurrence of new microsatellite alleles and the successive appearance of haplotypes are indicative for additional introductions during that time. The low frequencies of most of the new microsatellite alleles are in accordance with the expectation that the population in place, already recruiting from an established egg bank, is outnumbering the newly arriving invaders in the pelagial during sexual reproduction. Indeed, mutations may cause similar patterns. However, the fact that new alleles often occurred simultaneously and even within the same sample (e.g. new alleles at 4 different loci in sample MW 127a) and the virtual lack of new alleles in more recent periods despite comparable population sizes (Einsle 1987; Stich & Maier 2007) both argue against a strong contribution of mutation to the observed pattern. A single new allele in the most recent period may indicate that propagules are still transported to LLC, although we cannot exclude the possibility of a single mutation in this case.

An alternative explanation could be a single introduction event already comprising all the genetic diversity. Such a scenario would, however, need to invoke that several clones persisted clonally without resting egg production in the pelagial for more than a decade until they finally produced ephippia or males. Moreover, studies on the abundance of 'European *D. pulicaria*' in LLC revealed high peaks followed by long periods of virtual absence from the pelagial (Einsle 1987; Stich & Maier 2007), a pattern ascribed to high fish predation pressure on this large-bodied species (Stich & Maier 2007). Consequently, the 'European *D. pulicaria*' population in LLC seems to rely heavily on its egg bank rendering long-term existence of clones that do not invest in resting eggs in the pelagial unlikely. In support of this, Ortells *et al.* (2014) reported that a newly founded *Daphnia magna* population was well

reflected by its egg bank. From 20 alleles present in the pelagial population, 19 alleles were detectable in the egg bank after only one growing season. A scenario of additional introductions appears thus to be the most likely explanation for our data. The lower number of new alleles in the more recent periods may be caused by (i) a decline in propagule pressure, (ii) reaching equilibrium with the source population(s), (iii) monopolization (De Meester *et al.* 2002), that is rapid local adaptation of founders decreasing chances of establishment for later immigrants, or combinations of these processes.

Population structure analyses revealed evidence neither for strong population structure nor for distinct successive invasions replacing each other, as one may expect considering the strong fluctuations in pelagial abundance (Einsle 1987; Stich & Maier 2007). However, an increase in the number and contribution of different clusters is observed with time. Analysing and visualizing the data according to a prior assignment to different time periods revealed a pattern of gradual change over time, a strong overlap of consecutive periods, and an increase in genetic variation during the initial periods. These results are overall consistent with the idea of additional introductions from a genetically similar source introducing new alleles without a strong effect on population structure.

A further characteristic that one would expect for an invasion founded by a small number of genotypes of a cyclical parthenogenetic species followed by additional introductions is reflected by significant selfing rates during the initial stages followed by a striking decline over time. As long as the number of genotypes in the lake and in the egg bank is low, it is likely that the pelagial population is dominated by few clones, possibly even a single clone, increasing the chance of selfing (i.e. intracolonial mating) and biparental inbreeding during sexual reproduction. With the arrival of new genotypes, the number of different genotypes available for mating in the pelagial is increasing, the chance of reproduction among identical or closely related clones ceases, and consequently selfing and inbreeding become less likely. This hypothesis is consistent with the observed decrease in selfing rate with increasing genetic diversity and the accumulation of resting eggs in LLC.

Intriguingly, we observed dominance of a single homozygous MLG from the late 1970s until mid-1980s, resulting in a drop of genotypic richness, significant F_{IS} values and strong linkage disequilibrium during this period. One explanation for such an observation could be obligate parthenogenesis, known for members of the *D. pulex* species complex (e.g. Colbourne *et al.* 1998). A mixed breeding system was recorded among populations of 'European *D. pulicaria*'. While populations from

lowland lakes and ponds reproduce primarily by cyclic parthenogenesis, alpine populations from HTM reproduce asexually by obligate parthenogenesis (Dufresne *et al.* 2011). Assuming obligate apomictic parthenogenesis, we would expect a high level of heterozygosity at different loci due to mutation accumulation (e.g. Hebert 1981). Recent studies revealed that, mechanistically, parthenogenetic reproduction rather represents an atypical automixis via abortive meiosis in *D. pulex*. However, no evidence for recombination was found, and therefore, this mechanism is also likely increasing heterozygosity (Hiruta *et al.* 2010). Automixis with terminal fusion or post-meiotic doubling is a mechanisms resulting in highly homozygous offspring, but to our best knowledge, there is currently no evidence for these processes in *Daphnia*. Substantial loss of heterozygosity can, however, occur via hemizygous deletions and gene conversion during asexual reproduction in the *D. pulex* species complex (Omilian *et al.* 2006; Xu *et al.* 2011; Tucker *et al.* 2013). Nevertheless, transitions to obligate parthenogenesis, most probably originating from hybridization, as reported for 'European *D. pulicaria*' (Marková *et al.* 2007, 2013; Dufresne *et al.* 2011) and 'North American *D. pulex*' (Tucker *et al.* 2013; Xu *et al.* 2013), were associated with increased heterozygosity.

An alternative explanation for the presence of repeated homozygous MLGs in LLC is strong inbreeding (i.e. biparental inbreeding and selfing). The highly homozygous individuals are supposedly best explained by local inbreeding or colonization by already inbred individuals during the initial phase of the invasion followed by a dominance of these inbred clone(s) in the pelagial population with sustained inbreeding thus maintaining the production of highly homozygous resting eggs. Indeed, *Daphnia* populations are often founded by only a few clones favouring inbreeding and selfing (Hairston *et al.* 1999a; Ebert *et al.* 2002; Haag *et al.* 2002; Ortells *et al.* 2014). The high selfing rate and low number of alleles found during the initial periods of the invasion in LLC suggest that the number of colonizers was low, possibly only a single clone or few closely related clones. The fact that the disappearance of the homozygous genotype from the egg bank coincides with an increase in genetic diversity and a decrease in selfing rates supports this hypothesis. Moreover, homozygote clones of *D. pulicaria* reproducing via cyclical parthenogenesis have also been found in the Pyrenees and southern Quebec (Dufresne *et al.* 2011; F. Dufresne, unpublished data). Additional work to reveal the reproductive mode of the dominant homozygous genotype in LLC is still needed. As yet, however, our attempts to hatch this specific genotype have failed.

The LLC egg bank is functional as we observed fertile hatchlings up to 29 years back in time. The observed

age for successful hatching is not exceptional for *Daphnia* resting egg banks and at the lower limit compared to other studies (Weider *et al.* 1997; Cáceres 1998; Hairston *et al.* 1999a,b; Keller & Spaak 2004; Hembre & Peterson 2013; Frisch *et al.* 2014). Despite increased effort, we failed to hatch individuals from earlier periods even from eggs classified as quality 1, except for a single viable but infertile individual. We observed an intriguing coincidence of decreasing selfing rates with increasing hatching success. Inbreeding depression is well known for *Daphnia* spp. (De Meester 1993; Deng 1997; Haag *et al.* 2002; Cáceres *et al.* 2009) and could potentially have contributed to the reduced hatching success (Frommen *et al.* 2008) in the earliest populations. However, our data do not allow to disentangle effects of inbreeding and senescence, and environmental effects on hatching success.

In summary, by analysing the LLC resting egg bank, we were able to reconstruct the population genetic patterns associated with the invasion of a key planktonic grazer. Our data suggest the following scenario: 'European *D. pulicaria*', most likely originating from meso- to eutrophic lowland lakes or ponds in Central Europe, invaded LLC around 1974 and successfully established a population and resting egg bank. The founding population comprised few genotypes, possibly only one clone, resulting in inbreeding. Subsequent introductions from the same or very similar source populations contributed to an increase in genetic variation which may have facilitated the successful establishment of the population (Sakai *et al.* 2001; Kolbe *et al.* 2004; Frankham 2005b; Roman & Darling 2007; Dlugosch & Parker 2008; Forsman 2014; Rius & Darling 2014; Bock *et al.* 2015). Interestingly, 'European *D. pulicaria*' with the same mitochondrial haplotype as found in LLC has been detected for the first time in the pelagial of the nearby Lake Greifensee in winter 2012. Additional work is now required to monitor whether this invader is able to successfully establish a population in Lake Greifensee and whether it expands into other surrounding lakes.

We also show that the initial phase of the invasion is highly dynamic and the information on genetic patterns during this phase was crucial to make conclusion about the invasion history of 'European *D. pulicaria*' in LLC. This is an important consideration for studies aiming to reconstruct histories and sources of genetic variation of invasions. For example, in their influential review, Dlugosch & Parker (2008) show differences in genetic variation between source and invasive populations as a function of time since first introduction for 27 and 29 species. They report a quadratic relationship between years since introduction and allelic richness with introductions of intermediate age showing the lowest levels (see Fig. 3 in Dlugosch & Parker 2008). However,

the lag between the known date of first introduction and the assessment of genetic variation is often several decades, and none of the data points falls into the initial period of the introduction during which we observed the most dramatic changes and a very different temporal pattern of genetic variation in our study. Our findings therefore highlight the importance of the very early phases of an invasion and contribute to our knowledge on genetic processes associated with successful invasions.

The system we describe here and comparable systems with resting stages, for example seed banks in plants (Fennell *et al.* 2014), have the potential to serve as excellent models to reconstruct the histories of invading populations. For instance, the inference of multiple introductions from the same or closely related source populations from contemporary genetic data is inherently difficult, and little is known about the importance and frequency of this scenario. However, these cryptic additional invasion waves may strongly affect invasion success and consequently management decisions. Such considerations also motivated the development of model-based approaches to detect such signals (Benazzo *et al.* 2015). As our results show, genetic analyses of resting stage banks offer a more direct route to study such questions. They could also help to evaluate and improve modelling tools which will unquestionably be required to study systems lacking resting stage banks. Moreover, resting stage banks also preserve genetic evidence of failed invasions (Hairston *et al.* 1999a) and analysing and comparing these cases with successful invasions as reported here offers a promising route to identify crucial factors affecting invasion success.

Finally, this approach may also help to elucidate patterns of adaptive variation and the role of adaptive processes during the invasion process (Orsini *et al.* 2013). The egg bank of 'European *D. pulicaria*' in LLC seems to be particularly suited to address such questions because the lake has been subjected to well-documented environmental changes during the course of the invasion, that is eutrophication and climate change, genomic resources are available (Colbourne *et al.* 2011) and the population history has now been characterized.

Acknowledgements

We thank Esther Keller, Irene Brunner, Alfred Luck and Marian Fujak for help in the molecular and sedimentology laboratory and Christoph Tellenbach for statistical advice. We thank Nora Brede for initiating this project collaboration. We are grateful to Robert Obad for providing data on total phosphorus (© BOWIS Daten aus dem Bodensee Wasserinformationssystem der Internationalen Gewässerschutzkommission für den Bodensee, IGKB). We also thank France Dufresne, Larry Weider, four

anonymous reviewers and the editor for helpful comments. Markus Most and Piet Spaak were supported by the Swiss Science Foundation (CR3213 125211). Additional support was provided to Markus Most by a SNF Early PostDoc.Mobility fellowship (P2EZP3 148773) and to Silvia Marková from the from the institutional grant RVO: 67985904.

References

- Allendorf FW, Lundquist LL (2003) Introduction: population biology, evolution, and control of invasive species. *Conservation Biology*, **17**, 24–30.
- Appleby PG (2002) Chronostratigraphic techniques in recent sediments. In: *Tracking Environmental Change Using Lake Sediments* (eds Last W, Smol J), pp. 171–203. Springer, Dordrecht, the Netherlands.
- Arnaud Haond S, Belkhir K (2007) GENCLONE: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Molecular Ecology Notes*, **7**, 15–17.
- Barun A, Niemiller ML, Fitzpatrick BM, Fordyce JA, Simberloff D (2013) Can genetic data confirm or refute historical records? The island invasion of the small Indian mongoose (*Herpestes auro-punctatus*). *Biological Invasions*, **15**, 2243–2251.
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996–2004) GENETIX 4.05, logiciel sur Windows™ pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).
- Benazzo A, Ghirrotto S, Vilaca ST, Hoban S (2015) Using ABC and microsatellite data to detect multiple introductions of invasive species from a single source. *Heredity*. doi:10.1038/hdy.2015.38 [Epub ahead of print].
- Bock DG, Caseys C, Cousens RD *et al.* (2015) What we still don't know about invasion genetics. *Molecular Ecology*, **24**, 2277–2297.
- Brede N, Sandrock C, Straile D *et al.* (2009) The impact of human made ecological changes on the genetic architecture of *Daphnia* species. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 4758–4763.
- Brendonck L, De Meester L (2003) Egg banks in freshwater zooplankton: evolutionary and ecological archives in the sediment. *Hydrobiologia*, **491**, 65–84.
- Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non templated nucleotide addition by tag DNA polymerase: Primer modifications that facilitate genotyping. *BioTechniques*, **20**, 1004–1006.
- Cáceres CE (1998) Interspecific variation in the abundance, production, and emergence of *Daphnia* diapausing eggs. *Ecology*, **79**, 1699–1710.
- Cáceres CE, Hartway C, Paczolt KA (2009) Inbreeding depression varies with investment in sex in a facultative parthenogen. *Evolution*, **63**, 2474–2480.
- Chesson PL (1983) Coexistence of competitors in a stochastic environment: The storage effect. In: *Population Biology* (eds Freedman H, Strobeck C), pp. 188–198. Springer, Berlin Heidelberg.
- Colbourne JK, Crease TJ, Weider LJ *et al.* (1998) Phylogenetics and evolution of a circumpolar species complex (Cladocera: *Daphnia pulex*). *Biological Journal of the Linnean Society*, **65**, 347–365.

- Colbourne J, Robison B, Bogart K, Lynch M (2004) Five hundred and twenty eight microsatellite markers for ecological genomic investigations using *Daphnia*. *Molecular Ecology Notes*, **4**, 485–490.
- Colbourne JK, Pfrender ME, Gilbert D *et al.* (2011) The ecoreponsive genome of *Daphnia pulex*. *Science*, **331**, 555–561.
- Cousyn C, De Meester L, Colbourne JK *et al.* (2001) Rapid local adaptation of zooplankton behavior to changes in predation pressure in absence of neutral genetic changes. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 6256–6260.
- Cristescu ME (2015) Genetic reconstructions of invasion history. *Molecular Ecology*, **24**, 2212–2225.
- David P, Pujol B, Viard F, Castella V, Goudet J (2007) Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology*, **16**, 2474–2487.
- De Meester L (1993) Inbreeding and outbreeding depression in *Daphnia*. *Oecologia*, **96**, 80–84.
- De Meester L, Gomez A, Okamura B, Schwenk K (2002) The monopolisation hypothesis and the dispersal gene flow paradox in aquatic organisms. *Acta Oecologica*, **23**, 121–135.
- De Meester L, Van Doorslaer W, Geerts A, Orsini L, Stoks R (2011) Thermal genetic adaptation in the water flea *Daphnia* and its impact: An evolving metacommunity approach. *Integrative and Comparative Biology*, **51**, 703–718.
- Decaestecker E, Gaba S, Raeymaekers JAM *et al.* (2007) Host parasite “Red Queen” dynamics archived in pond sediment. *Nature*, **450**, 870–873.
- Deng H W (1997) Increase in developmental instability upon inbreeding in *Daphnia*. *Heredity*, **78**, 182–189.
- Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431–449.
- Duffy MA, Perry LJ, Kearns CM, Weider LJ, Hairston NG (2000) Paleogenetic evidence for a past invasion of Onondaga Lake, New York, by exotic *Daphnia curvirostris* using mtDNA from dormant eggs. *Limnology and Oceanography*, **45**, 1409–1414.
- Dufresne F, Marková S, Vergilino R, Ventura M, Kotlík P (2011) Diversity in the reproductive modes of European *Daphnia pulex* deviates from the geographical parthenogenesis. *PLoS ONE*, **6**, e20049.
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.
- Ebert D, Haag C, Kirkpatrick M *et al.* (2002) A selective advantage to immigrant genes in a *Daphnia* metapopulation. *Science*, **295**, 485–488.
- Einsle U (1980) Weitere Untersuchungen (1976/1977) über das Crustaceen Plankton des Gnadensees (Bodensee Untersee). *Schriften des Vereins für die Geschichte des Bodensees und seiner Umgebung*, **98**, 195–218.
- Einsle U (1987) Die Entwicklung des Crustaceen Planktons im Bodensee Obersee (1972–1985) und Untersee Gnadensee und Rheinsee (1974–1985). *Internationale Gewässerschutzkommission für den Bodensee: Bericht*, **37**, 1–103.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.
- Fadda A, Marková S, Kotlík P *et al.* (2011) First record of planktonic crustaceans in Sardinian reservoirs. *Biologia*, **66**, 856–865.
- Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.
- Fennell M, Gallagher T, Vintro LL, Osborne B (2014) Using soil seed banks to assess temporal patterns of genetic variation in invasive plant populations. *Ecology and Evolution*, **4**, 1648–1658.
- Floßner D (1972) Kiemen und Blattfüßer, Branchiopoda Fischlaube, Branchiura. In: *Die Tierwelt Deutschlands* (ed. Dahl F), pp. 499. Gustav Fischer Verlag, Jena.
- Floßner D (2000) *Die Haplopoda und Cladocera (Ohne Bosminidae) Mitteleuropas*. Backhuys, Leiden.
- Forsman A (2014) Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 302–307.
- Frankham R (2005a) Genetics and extinction. *Biological Conservation*, **126**, 131–140.
- Frankham R (2005b) Invasion biology Resolving the genetic paradox in invasive species. *Heredity*, **94**, 385–385.
- Frisch D, Havel JE, Weider LJ (2013) The invasion history of the exotic freshwater zooplankton *Daphnia lumholtzi* (Cladocera, Crustacea) in North America: a genetic analysis. *Biological Invasions*, **15**, 817–828.
- Frisch D, Morton PK, Chowdhury PR *et al.* (2014) A millennial scale chronicle of evolutionary responses to cultural eutrophication in *Daphnia*. *Ecology Letters*, **17**, 360–368.
- Frommen JG, Luz C, Mazzi D, Bakker TCM (2008) Inbreeding depression affects fertilization success and survival but not breeding coloration in threespine sticklebacks. *Behaviour*, **145**, 425–441.
- Gilbert KJ, Andrew RL, Bock DG *et al.* (2012) Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program STRUCTURE. *Molecular Ecology*, **21**, 4925–4930.
- Goudet J (2005) HIERFSTAT, a package for R to compute and test hierarchical F statistics. *Molecular Ecology Notes*, **5**, 184–186.
- Haag CR, Hotinger JW, Riex M, Ebert D (2002) Strong inbreeding depression in a *Daphnia* metapopulation. *Evolution*, **56**, 518–526.
- Hairston NG, De Stasio BT (1988) Rate of evolution slowed by a dormant propagule pool. *Nature*, **336**, 239–242.
- Hairston NG, Perry LJ, Bohonak AJ *et al.* (1999a) Population biology of a failed invasion: Paleolimnology of *Daphnia exilis* in upstate New York. *Limnology and Oceanography*, **44**, 477–486.
- Hairston NG, Lampert W, Cáceres CE *et al.* (1999b) Lake ecosystems Rapid evolution revealed by dormant eggs. *Nature*, **401**, 446–446.
- Hasegawa M, Kishino H, Yano TA (1985) Dating of the human ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.
- Havel JE, Mabee WR, Jones JR (1995) Invasion of the exotic cladoceran *Daphnia lumholtzi* into North American reser-

- voirs. *Canadian Journal of Fisheries and Aquatic Sciences*, **52**, 151–160.
- Hebert PDN (1981) Obligate asexuality in *Daphnia*. *The American Naturalist*, **117**, 784–789.
- Hedrick PW (1995) Genetic polymorphism in a temporally varying environment—effects of delayed germination or diapause. *Heredity*, **75**, 164–170.
- Hembre LK, Peterson LA (2013) Evolution of predator avoidance in a *Daphnia* population: evidence from the egg bank. *Hydrobiologia*, **700**, 245–255.
- Hiruta C, Nishida C, Tochinai S (2010) Abortive meiosis in the oogenesis of parthenogenetic *Daphnia pulex*. *Chromosome Research*, **18**, 833–840.
- Hrbáček J (1959) Über die angebliche Variabilität von *Daphnia pulex* Leydig. *Zoologischer Anzeiger*, **162**, 116–126.
- Hrbáček J (1977) Competition and predation in relation to species composition of freshwater zooplankton, mainly Cladocera. In: *Aquatic Microbial Communities* (ed. Cairns JJ), pp. 305–353. Garland Reference Library of Science and Technology, New York City, New York.
- Ivens ABF, deVan Sanden M, Bakker J (2014) MLGsim2.0 software. <http://www.rug.nl/research/theoreticalbiology/downloads>
- Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.
- Jombart T, Pontier D, Dufour AB (2009) Genetic markers in the playground of multivariate analysis. *Heredity*, **102**, 330–341.
- Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.
- Kamvar ZN, Tabima JF, Grunwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, **2**, e281.
- Keenan K, McGinnity P, Cross TF, Crozier WW, Prodohl PA (2013) diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, **4**, 782–788.
- Keller B, Spaak P (2004) Nonrandom sexual reproduction and diapausing egg production in a *Daphnia* hybrid species complex. *Limnology and Oceanography*, **49**, 1393–1400.
- Kolbe JJ, Glor RE, Schettino LRG *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177–181.
- Kopp K, Wolff K, Jokela J (2012) Natural range expansion and human assisted introduction leave different genetic signatures in a hermaphroditic freshwater snail. *Evolutionary Ecology*, **26**, 483–498.
- Lampert W (2011) *Daphnia: Development of a Model Organism in Ecology and Evolution*. International Ecology Institute, Oldendorf/Luhe, Germany.
- Latta LC, Weider LJ, Colbourne JK, Pfrender ME (2012) The evolution of salinity tolerance in *Daphnia*: a functional genomics approach. *Ecology Letters*, **15**, 794–802.
- Lockwood JL, Cassey P, Blackburn T (2005) The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution*, **20**, 223–228.
- Marková S, Dufresne F, Rees DJ, Cerny M, Kotlík P (2007) Cryptic intercontinental colonization in water fleas *Daphnia pulex* inferred from phylogenetic analysis of mitochondrial DNA variation. *Molecular Phylogenetics and Evolution*, **44**, 42–52.
- Marková S, Dufresne F, Manca M, Kotlík P (2013) Mitochondrial capture misleads about ecological speciation in the *Daphnia pulex* complex. *PLoS ONE*, **8**, e69497.
- Mergeay J, Verschuren D, Meester LD (2005) Cryptic invasion and dispersal of an American *Daphnia* in East Africa. *Limnology and Oceanography*, **50**, 1278–1283.
- Mergeay J, Verschuren D, De Meester L (2006) Invasion of an asexual American water flea clone throughout Africa and rapid displacement of a native sibling species. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2839–2844.
- Montero Pau J, Gomez A, Munoz J (2008) Application of an inexpensive and high throughput genomic DNA extraction method for the molecular ecology of zooplanktonic diapausing eggs. *Limnology and Oceanography: Methods*, **6**, 218–222.
- Most M, Oexle S, Marková S *et al.* (2015) Data from: Population genetic dynamics of an invasion reconstructed from the sediment egg bank. Dryad Digital Repository. <http://dx.doi.org/10.5061/dryad.0rh0t>.
- Nunney L (2002) The effective size of annual plant populations: The interaction of a seed bank with fluctuating population size in maintaining genetic variation. *The American Naturalist*, **160**, 195–204.
- Omilian AR, Cristescu MEA, Dudycha JL, Lynch M (2006) Asexual recombination in asexual lineages of *Daphnia*. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 18638–18643.
- Omilian AR, Scofield DG, Lynch M (2008) Intron presence absence polymorphisms in *Daphnia*. *Molecular Biology and Evolution*, **25**, 2129–2139.
- Orsini L, Schwenk K, De Meester L *et al.* (2013) The evolutionary time machine: using dormant propagules to forecast how populations can adapt to changing environments. *Trends in Ecology & Evolution*, **28**, 274–282.
- Ortells R, Vanoverbeke J, Louette G, De Meester L (2014) Colonization of *Daphnia magna* in a newly created pond: founder effects and secondary immigrants. *Hydrobiologia*, **723**, 167–179.
- Pantel JH, Leibold MA, Juenger TE (2011) Population differentiation in *Daphnia* alters community assembly in experimental ponds. *American Naturalist*, **177**, 314–322.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- Pritchard JK, Wen X, Falush D (2007) Documentation for structure software: Version 2.3. Available from <http://pritch.bsd.uchicago.edu/software>.
- R Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rellstab C, Keller B, Girardclos S, Anselmetti F, Spaak P (2011) The limits of invasion—Trophic state shapes the past and present taxonomic composition of *Daphnia* populations in ultra oligotrophic lakes. *Limnology and Oceanography*, **56**, 292–302.
- Rius M, Darling JA (2014) How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology & Evolution*, **29**, 233–242.

- Roman J, Darling JA (2007) Paradox lost: Genetic diversity and the success of aquatic invasions. *Trends in Ecology & Evolution*, **22**, 454–464.
- Sakai AK, Allendorf FW, Holt JS *et al.* (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics*, **32**, 305–332.
- Simberloff D, Martin JL, Genovesi P *et al.* (2013) Impacts of biological invasions: What's what and the way forward. *Trends in Ecology & Evolution*, **28**, 58–66.
- Spaak P, Fox J, Hairston NG (2012) Modes and mechanisms of a *Daphnia* invasion. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 2936–2944.
- Special Issue: Invasion Genetics: The Baker and Stebbins Legacy (2015) *Molecular Ecology*, **24**, 1927–2297.
- Stenberg P, Lundmark M, Saura A (2003) mlgSim: a program for detecting clones using a simulation approach. *Molecular Ecology Notes*, **3**, 329–331.
- Stich HB, Maier GM (2007) Distribution and abundance of *Daphnia pulicaria*, a large *Daphnia* of the "pulex group", in Lake Constance (Lower Lake). *Limnologia*, **37**, 303–310.
- Tamura K (1992) Estimation of the number of nucleotide substitutions when there are strong transition transversion and G+C content biases. *Molecular Biology and Evolution*, **9**, 678–687.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, **30**, 2725–2729.
- Tucker AE, Ackerman MS, Eads BD, Xu S, Lynch M (2013) Population genomic insights into the evolutionary origin and fate of obligately asexual *Daphnia pulex*. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 15740–15745.
- Van Doorslaer W, Stoks R, Swillen I *et al.* (2010) Experimental thermal microevolution in community embedded *Daphnia* populations. *Climate Research*, **43**, 81–89.
- Vergilino R, Belzile C, Dufresne F (2009) Genome size evolution and polyploidy in the *Daphnia pulex* complex (Cladocera: Daphniidae). *Biological Journal of the Linnean Society*, **97**, 68–79.
- Vergilino R, Marková S, Ventura M, Manca M, Dufresne F (2011) Reticulate evolution of the *Daphnia pulex* complex as revealed by nuclear markers. *Molecular Ecology*, **20**, 1191–1207.
- Vrijenhoek RC (1998) Animal clones and diversity. *BioScience*, **48**, 617–628.
- Weider LJ, Lampert W, Wessels M, Colbourne JK, Limburg P (1997) Long term genetic shifts in a microcrustacean egg bank associated with anthropogenic changes in the Lake Constance ecosystem. *Proceedings of the Royal Society of London B: Biological Sciences*, **264**, 1613–1618.
- Weider LJ, Hobaek A, Colbourne JK *et al.* (1999) Holarctic phylogeography of an asexual species complex I. Mitochondrial DNA variation in arctic *Daphnia*. *Evolution*, **53**, 777–792.
- Weir BS, Cockerham CC (1984) Estimating F Statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- Xu S, Omilian AR, Cristescu ME (2011) High rate of large scale hemizygous deletions in asexually propagating *Daphnia*: Implications for the evolution of sex. *Molecular Biology and Evolution*, **28**, 335–342.
- Xu S, Innes DJ, Lynch M, Cristescu ME (2013) The role of hybridization in the origin and spread of asexuality in *Daphnia*. *Molecular Ecology*, **22**, 4549–4561.
- Zaffagnini F (1987) Reproduction in *Daphnia*. In: *Daphnia* (eds Peters RH, de Bernardi R), pp. 245–284. Memorie dell'Istituto Italiano di Idrobiologia, Verbania Pallanza, Italy.
- Živković D, Tellier A (2012) Germ banks affect the inference of past demographic events. *Molecular Ecology*, **21**, 5434–5446.

M.M., P.S., D.M.-C. and S.M. conceived the study. M.M. collected the data together with S.O., D.A. and L.B. M.M. analysed the data and interpreted them with input from S.M., L.B. and H.-B.S. MW provided sediment cores and helped to analyse and interpret the sediment record. M.M. wrote the manuscript with contributions from P.S., S.M., D.M.-C. and H.-B.S.

Data accessibility

DNA sequences have been uploaded to GenBank; Accession nos: KR809887–KR809891. Sample information, microsatellite marker information and settings for DAPC analyses have been uploaded as Supporting Information. Microsatellite genotypes, sample IDs, dating information, sequence alignments, tree files, hatching data, STRUCTURE input files and R codes for sliding window analyses and pairwise D_{ST} and F_{ST} calculations are available from the Dryad Digital Repository doi:10.5061/dryad.0rh0t (Möst *et al.* 2015).

Supporting information

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

Fig. S1 Patterns of eutrophication and re oligotrophication in Lower Lake Constance over the last four decades.

Fig. S2 Abundance patterns of *D. pulicaria* in the pelagial of 'Gnadensee' in LLC from 1975 to beginning of 1986 (modified from Einsle 1987).

Fig. S3 Maximum likelihood phylogenetic tree of the nuclear Rab4 gene for the *D. pulex* species complex.

Fig. S4 Sliding window analyses of all unique multi locus genotypes (MLGs) from the LLC egg bank.

Fig. S5 Assignment plots for all LLC samples for a Discriminant Analysis of Principal Components shown in Fig. 5.

Fig. S6 Assignment plots for unique LLC multi locus genotypes for a Discriminant Analysis of Principal Components.

Fig. S7 Discriminant Analysis of Principal Components for all LLC samples.

Fig. S8 Discriminant Analysis of Principal Components for unique multi locus genotypes.

Fig. S9 Results of STRUCTURE analyses for all LLC samples.

Fig. S10 Results of STRUCTURE analyses for unique multi locus genotype samples from LLC.

Fig. S11 The first two axes of a factorial correspondence analysis (FCA) reflecting the increase of genotypic diversity of 'European *D. pulicaria*' in LLC over time.

Fig. S12 Pairwise F_{ST} and Jost's D (D_{ST}) for all samples (top row) and all unique multilocus genotypes (bottom row).

Table S1 Sample information for ND5 sequences used for reconstructing a maximum likelihood tree (Fig. 2).

Table S2 Sample information for Rab4 sequences used for reconstructing a maximum likelihood tree (Fig. S3).

Table S3 Summary information on microsatellite markers, primers, dyes and multiplexing.

Table S4 Settings used for Discriminant Analyses of Principal Components for an analysis with eight pre defined groups (Figs 5, S5 and S6) and analyses for K means clusters for $K = 2$ to $K = 7$ (Figs S7 and S8).

Table S5 Summary statistics for MLGsim 2.0 (http://www.rug.nl/research/theoretical_biology/downloads) analyses calculating the probability of finding at least as many identical MLGs as observed in a population.

Table S6 Jost's D (A) and F_{ST} (B) for pairwise comparisons of periods and 95% confidence intervals based on 10 000 bootstrap replicates.