

# Supporting Information

Brede et al. 10.1073/pnas.0807187106

## SI Text

**Statistical Analysis: Factorial Correspondence Analysis (Fig. 3).** We used a correspondence analysis (CA; also called reciprocal averaging), which is based on a unimodal model. Factorial correspondence analysis (FCA) is adapted for use with diploid genetic data following She et al. (1) and is implemented in GENETIX (2). FCA is an exploratory technique, suitable for categorical data, which allows investigation of correspondence between rows (e.g., individuals) and columns (e.g., alleles) in a two-way table. It enables visualization of individuals in multidimensional space, with no a priori assumptions about grouping, using each allele as an independent variable. Axes are generated from combinations of alleles that explain portions of the total observed “inertia” of the table. Hence, those alleles exhibiting the strongest nonrandom association with groups of individuals will contribute most to the axes. We have analyzed the global variability of the whole collection of *Daphnia* individuals (per lake) and projected the “centers of gravity” for each sample (sample = all individuals of a particular time period). CA axis 1 has been used as an indicator of the genetic changes and associated with P concentration on the y axis.

**Statistical Analysis: Hardy-Weinberg Equilibrium (Table S1 and Fig. S1).** We tested the hypothesis that the relative abundance of interspecific hybrids is determined by the relative abundance of

parental species. This test is based on the assumption that both species (*D. hyalina* and *D. galeata*) mate randomly in each lake during each time period. All available genetic information (microsatellites and ITS RFLP patterns) were used to group all individuals into 3 hybrid categories (*D. hyalina* and *D. galeata* or hybrid). If both *Daphnia* species (alternative homozygotes) mate randomly (e.g., no mating preferences, no selection), the frequency of interspecific hybrids (heterozygotes) should be proportional to the relative frequency of parental species according to the Hardy-Weinberg equilibrium (HWE). In the majority of cases (67%), Hardy-Weinberg assumptions were not violated, suggesting that in most time periods, mating of parental taxa was random without any prevalent pre- or postzygotic barrier. However, deviations from the HWE (e.g., in Greifensee) suggest that either selection or assortative mating caused a higher abundance of interspecific hybrids than expected from random mating. The slope of regression was significantly higher than 1 ( $P < 0.0001$ ) in Greifensee but not different from 1 in Lake Constance ( $P > 0.15$ ). In summary, although the assumptions of the HWE are not met in all cases, our data analysis does not reject the hypothesis that abundances of parental taxa are responsible for the occurrence and relative abundances of interspecific hybrids.

1. She JX, Autem M, Kotulas G, Pasteur N, Bonhomme F (1987) Multivariate analysis of genetic exchanges between *Solea aegyptiaca* and *Solea senegalensis* (Teleosts, Soleidae). *Biol J Linn Soc* 32:357–371.

2. Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996–2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations (Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France).



**Table S1. Hardy-Weinberg expectations of *Daphnia* taxa**

Lake	Time	$\chi^2$	<i>P</i>
Lake Constance	1945–1953	0.031	0.860
	1954–1959	0.270	0.603
	1960–1963	0.068	0.794
	1965–1970	10.101	0.001
	1971–1975	45.265	<0.001
	1980–1987	0.003	0.955
	1990–1995	0.122	0.727
	1999–2004	4.967	0.026
Greifensee	1927–1936	0.050	0.823
	1942–1945	0.082	0.775
	1948–1950	0.919	0.338
	1951–1954	33.199	<0.001
	1958–1963	0.502	0.479
	1966–1970	1.823	0.177
	1971–1974	0.166	0.684
	1977–1980	2.323	0.127
	1988–1992	38.122	<0.001
	1999–2004	21.748	<0.001

We used the relative abundances of species, based on ITS and microsatellite DNA markers, to estimate expected taxa proportions under random mating in both lakes. We found 3 of 8 (Lake Constance) and 3 of 10 (Greifensee) deviations from Hardy-Weinberg expectations ( $\chi^2$  test).

**Table S2. Number of *Daphnia* resting eggs subjected to DNA analyses**

Lake	Time	ITS	16S	$\mu$ sat
Lake Constance	1908–1926	2	0	2
	1935–1939	3	0	1
	1940–1944	1	0	0
	1945–1953	9	0	0
	1954–1959	27	27	27
	1960–1963	30	0	38
	1965–1970	71	0	0
	1971–1975	64	37	38
	1980–1987	81	0	0
	1990–1995	56	57	56
	1999–2004	100	0	0
	Total	444	121	156
	Greifensee	1927–1936	30	30
1937–1939		15	14	14
1942–1945		15	15	15
1948–1950		38	34	34
1951–1954		66	61	61
1958–1963		44	25	25
1966–1970		97	76	76
1971–1974		29	18	18
1977–1980		100	49	49
1988–1992		66	39	39
1999–2004		56	52	52
Total		556	413	413

ITS, RFLP analysis of the internal ITS region; 16S, 16S rDNA mitochondrial DNA RFLP;  $\mu$ sat, microsatellite analysis of 6 (Lake Constance) and 8 (Greifensee) polymorphic loci.

**Table S3. Species-specific mitochondrial DNA markers in *Daphnia***

Species	Haplotype	<i>Rsa</i> I	<i>Dde</i> I	<i>Mnl</i> I
<i>D. cucullata</i>	<i>c1</i>	560	290–180–100	250–230–100
	<i>c3</i>	560	450–100	250–260–100
<i>D. galeata</i>	<i>g1</i>	510–50	290–180–100	250–210–100
<i>D. hyalina</i>	<i>h1</i>	560	390–100–80	230–180–100–90
	<i>x7</i>	560	450–100	250–260–100

RFLP analysis of 16S rDNA fragments using restriction enzymes *Rsa* I, *Dde* I, and *Mnl* I. Haplotypes *c1*, *g1*, and *h1* were reported in previous studies; *x7* and *c3* exhibit identical RFLP patterns, but DNA sequencing of 12S rDNA revealed large sequence differences (11.1% sequence divergence). Each resting egg of *D. cucullata* and a subsample of *D. hyalina* resting eggs were subjected to PCR and sequencing of mtDNA. Haplotype *c3* was found in only 12 *D. cucullata* resting eggs during the time period of 1958–1963, indicating a short but unsuccessful invasion of Greifensee.