

## RESEARCH ARTICLE

## Open Access

# Populations of a cyprinid fish are self-sustaining despite widespread feminization of males

Patrick B Hamilton<sup>1\*</sup>, Elizabeth Nicol<sup>2</sup>, Eliane SR De-Bastos<sup>1</sup>, Richard J Williams<sup>3</sup>, John P Sumpter<sup>2</sup>, Susan Jobling<sup>2</sup>, Jamie R Stevens<sup>1</sup> and Charles R Tyler<sup>1</sup>

## Abstract

**Background:** Treated effluents from wastewater treatment works can comprise a large proportion of the flow of rivers in the developed world. Exposure to these effluents, or the steroidal estrogens they contain, feminizes wild male fish and can reduce their reproductive fitness. Long-term experimental exposures have resulted in skewed sex ratios, reproductive failures in breeding colonies, and population collapse. This suggests that environmental estrogens could threaten the sustainability of wild fish populations.

**Results:** Here we tested this hypothesis by examining population genetic structures and effective population sizes ( $N_e$ ) of wild roach (*Rutilus rutilus* L.) living in English rivers contaminated with estrogenic effluents.  $N_e$  was estimated from DNA microsatellite genotypes using approximate Bayesian computation and sibling assignment methods. We found no significant negative correlation between  $N_e$  and the predicted estrogen exposure at 28 sample sites. Furthermore, examination of the population genetic structure of roach in the region showed that some populations have been confined to stretches of river with a high proportion of estrogenic effluent for multiple generations and have survived, apparently without reliance on immigration of fish from less polluted sites.

**Conclusions:** These results demonstrate that roach populations living in some effluent-contaminated river stretches, where feminization is widespread, are self-sustaining. Although we found no evidence to suggest that exposure to estrogenic effluents is a significant driving factor in determining the size of roach breeding populations, a reduction in  $N_e$  of up to 65% is still possible for the most contaminated sites because of the wide confidence intervals associated with the statistical model.

**Keywords:** Fishery, Genetic diversity, DNA microsatellites, Waste water treatment work, Ecotoxicology

## Background

Approximately two-thirds of the world's freshwater is used to dilute wastewater discharges. The demand for freshwater is expected to rise by 70% by 2050 [1] driving an urgent need to understand the impacts of treated waste effluent discharges on aquatic ecosystems. Wastewater treatment works (WWTW) effluents contain tens of thousands of chemicals, including pharmaceuticals and natural steroid estrogens that are biologically active at low (ng/L) exposure concentrations [2]. However, the long-term consequences of exposure to most of these chemicals on fish health and population sustainability are not known.

There is substantial evidence showing that experimental exposure of fish to WWTW effluents and the estrogens they contain can result in adverse health effects, including effects on reproductive development and breeding output. This has led to concerns that freshwater fish populations might also be affected with cascading consequences for freshwater ecosystems. Feminization of male fish is widespread in stretches of rivers downstream of WWTW outfalls as demonstrated in studies using wild [3,4] and caged [5,6] fish. Feminized phenotypes include the presence of vitellogenin, a female-specific protein in the blood of male fish [7] and the intersex condition: the presence of oocytes and/or female reproductive ducts in otherwise male gonads [3]. Feminization has been attributed to the presence of estrogens in effluents: estradiol ( $E_2$ ) and estrone ( $E_1$ ) from human excretion; 17 alpha-ethinylestradiol ( $EE_2$ ), a component of the female

\* Correspondence: [p.b.hamilton@exeter.ac.uk](mailto:p.b.hamilton@exeter.ac.uk)

<sup>1</sup>Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK

Full list of author information is available at the end of the article

contraceptive pill [8]; and a large number of other estrogenic chemicals from industrial and domestic effluents. WWTW effluents can also induce genotoxic effects [9], alterations in immune function [10], decreased reproductive output [11], altered stress response [12] and changes in reproductive behavior [13].

Concern about estrogens in rivers in the United Kingdom drove a £40 M programme to evaluate the efficacy of various tertiary treatment processes in the removal of estrogens [14,15]. Implementation of such processes will, however, incur considerable costs and a greater carbon footprint for WWTW [14,16], emphasising the need to understand better the population-level consequences for exposure to estrogenic and other so-called endocrine disrupting chemicals (EDCs).

A critical question is whether chronic exposure to estrogenic effluents negatively impacts the viability of wild fish populations, but this has been difficult to address experimentally as it requires controlled experiments extending over periods of several years. Limited studies suggest that high concentrations of EE<sub>2</sub> (between 3 to 6 ng/L) in the aquatic environment could be a threat to the sustainability of fish populations. For example, a controlled exposure of an entire lake to EE<sub>2</sub> in Canada resulted in the collapse of the fathead minnow (*Pimephales promelas*) population within three years [17]. Likewise, long-term (>204 days) laboratory exposures of a range of fish species have resulted in the absence of breeding males [18-20] and a three-year exposure of roach (*Rutilus rutilus* L.) to an undiluted WWTW effluent in large tanks resulted in an all-female population [21]. It is not known, however, if this occurs in rivers contaminated by effluents. Female fecundity can also be reduced through estrogen exposure, which can potentially reduce population growth rates [22]. Although the exposure concentrations in these studies were high compared to those typically experienced by wild fish populations [23], exposures to EE<sub>2</sub> at concentrations below 1 ng/L during the period of sexual development, have been shown to result in feminized gonads in roach [19] and decreased egg fertilization and female-skewed sex ratios in fathead minnows [24]. Evidence from wild roach living in UK rivers has similarly shown that feminized fish (generally less than 10% of males) with large numbers of eggs in their gonads have impaired semen quality [25] and severely (up to 76%) reduced reproductive success [26].

While these studies suggest that exposure to high concentrations of effluent could threaten the viability of fish populations, aggregates of cyprinid fish, including roach, are often found in effluent contaminated rivers. However, numbers alone may provide a misleading assessment of population sustainability as these could be sink populations maintained by substantial immigration from less contaminated locations where successful reproduction

still occurs. Likewise, effective population sizes ( $N_e$ ) – related to the breeding population of fish – may be decreased without necessarily impacting on population sizes [27], as effluent exposure can affect the number of reproducing individuals and can skew reproductive success [21,26]. Density-dependent growth and survival can also play an important role [28], so a few reproducing individuals can potentially maintain large adult population sizes. Indeed, studies in several species of marine fish with high fecundity have shown that  $N_e$  can be several orders of magnitude smaller than census population sizes. One study found two populations of the exploited New Zealand snapper (*Pagrus auratus*) to have values of  $N_e$  less than 1,000 despite adult census population sizes in the millions [29]. Similarly, a study of striped bass (*Morone saxatilis*), a freshwater fish species, found cohorts to consist of a few, full sib families, despite an adult census size of over 300,000 [30]. Critically,  $N_e$  influences long-term sustainability as it determines the rate at which genetic diversity is lost from a population through genetic drift [31]. High genetic diversity increases the long-term potential for populations to adapt to changes in the environment and also acts to reduce the risk of inbreeding [32]. Small  $N_e$ , however, may act to increase the chances of losing some lethal or sub-lethal mutations through genetic purging.

Understanding the impact of estrogenic effluents on the sustainability of fish populations is, therefore, paramount, but has been limited to date by the logistical challenges involved in undertaking long-term exposures to realistic effluent concentrations, and understanding the demographic history of wild fish populations at highly contaminated sites. In this study, we examine evidence for population impacts on wild roach (*R. rutilus*), a fish species in which feminization is widespread, in southern England. Southern England has some of the highest proportions of WWTW effluent in rivers known globally, and numerous weirs and locks which potentially confine fishes to heavily polluted stretches of river. We have used this system to evaluate whether stretches of river highly contaminated with estrogenic effluents have impaired breeding populations of roach. To do this we undertook analysis of population genetic structures of roach in the region using DNA microsatellite analysis. Microsatellite data were also used to calculate  $N_e$  and estimate levels of gene flow to determine the extent to which these populations are maintained through immigration of fish from less contaminated stretches of river.

## Results

### Genetic diversity and genetic bottlenecks

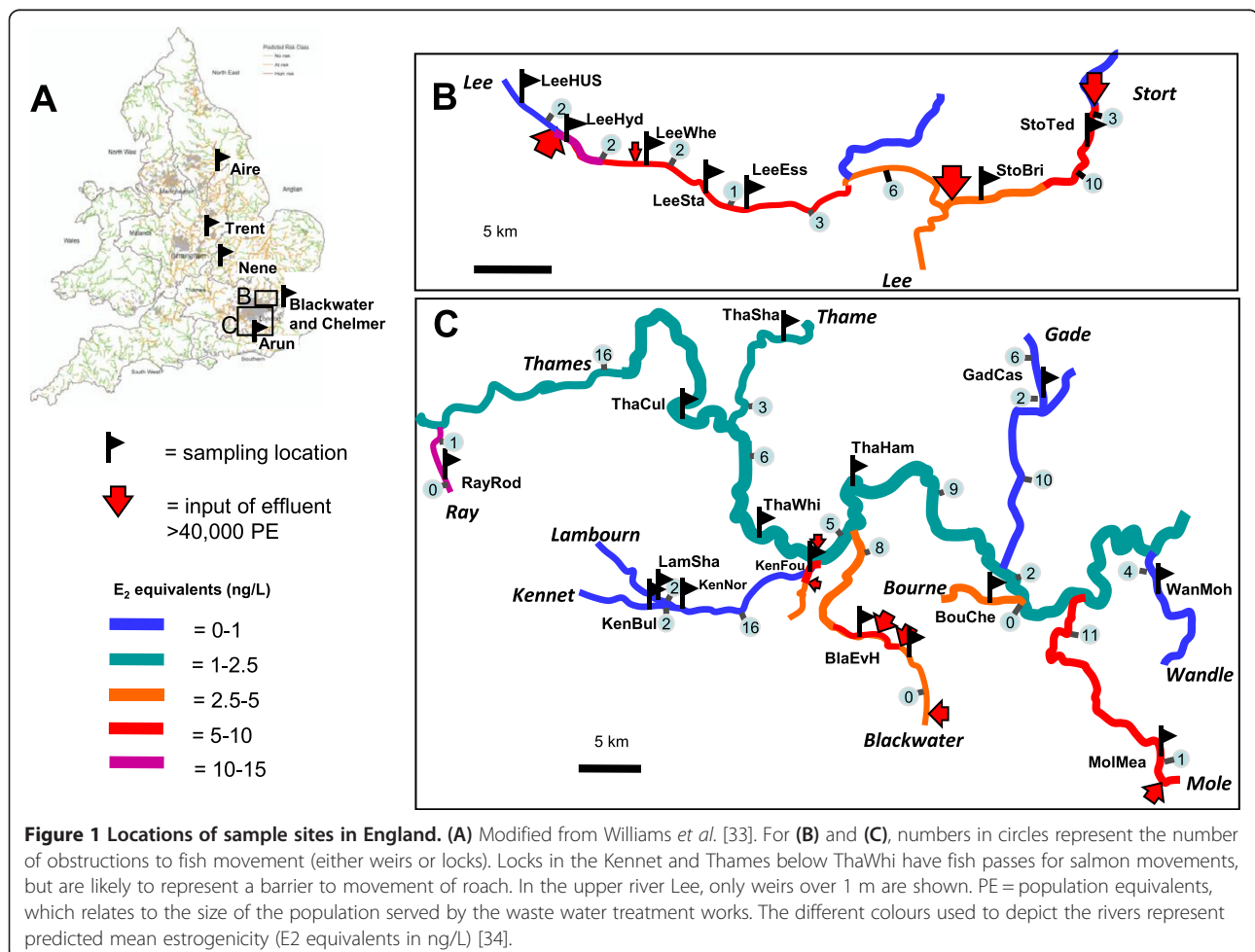
A total of 1,769 roach, constituting 39 samples (roach sampled from one location within a river – typically a stretch of approximately 100 m – at one time point) from

32 different geographic locations in England were genotyped (Figure 1). Data for 14 microsatellite loci [see Additional file 1 for details] revealed high genetic diversity in all 39 samples (Table 1). Allelic richness ( $AR$ ) ranged from 6.8 to 8.9, and expected heterozygosity ( $H_e$ ) ranged from 0.69 to 0.75 (Table 1, see Additional file 2 for diversity statistics for each locus). Nevertheless, significant differences in  $AR$  (analysis of variance (ANOVA),  $F_{(38,532)} = 2.1398$ ,  $P = 0.00014$ ) and observed heterozygosity ( $H_o$ ) (ANOVA,  $F_{(38,532)} = 1.8677$ ,  $P = 0.0017$ ) among samples were detected. Roach sampled at two relatively unpolluted sites (LamSha and LeeHUS'95) exhibited comparatively low  $AR$  and LeeHyd'95 exhibited comparatively relatively low  $H_o$  [see Additional file 3]. No significant differences in  $H_e$  were found. Additionally, there was evidence for genetic bottlenecks at two relatively unpolluted sites sampled within the rivers Arun (AruHUS) and Lee (LeeHUS), and at a polluted site in the Lee (LeeWhe) (Table 1).

#### Population genetic structure of roach in English rivers

We undertook an analysis of population genetic structure in order to examine the genetic similarity of roach

between and within catchments. Analysis of molecular variance (AMOVA) indicated the majority of variation was partitioned among individuals within river locations, with river location accounting for a small 2.27%, but highly significant proportion of the genetic variation (Table 2). Average pairwise  $F_{ST}$  between roach samples from different catchments was 0.028 and comparisons were consistently highly significant (Table 3, Additional file 4). The population tree (Figure 2) shows distinct clusters of samples in different catchments: the Arun, the Nene, the Anglian Blackwater and the Trent, supported with moderate–high bootstrap values (>64%). Samples from the Arun and the Nene also group in the principal component (PCA) and the STRUCTURE analyses [see Additional file 5 and Additional file 6] demonstrating a distinct genetic identity of fish at these sites. Utilising the method of Evanno *et al.* [35], the inferred most likely number of genetically distinct clusters in the STRUCTURE analysis was three, comprising: the Arun, the two most upstream sample sites in the river Lee, and all remaining sites genotyped. However, from visual examination of STRUCTURE plots run with higher levels of  $K$



**Table 1 Sampling locations, genetic diversity statistics (allelic richness (AR) and expected heterozygosity ( $H_e$ )) for each population sample**

Sample code <sup>a</sup>	River	Year	Lat.	Long.	E <sub>2</sub> Eq (ng/L) <sup>b</sup>	No. fish genotyped	Bottleneck (TPM) <sup>c</sup>	AR	H <sub>e</sub>	Ref. <sup>d</sup>
<b>Blackwater</b>										
BlaBlM	Blackwater	2010	51.78529	0.651013	4.1	55	0.40	7.9	0.71	
BlaSti	Blackwater	2010	51.88792	0.606184	7.1	55	0.60	8.2	0.72	
CheAbB	Chelmer	2010	51.84596	0.423716	1.2	50	0.06	7.7	0.75	
<b>Nene</b>										
NenBro'95	Nene	1995	52.24627	-0.77987	1.2	47	0.69	8.3	0.72	[3]
NenBro'99	Nene	1999	"	"	1.2	47	0.80	8.3	0.73	[25]
NenEct	Nene	2007	52.24234	-0.80281	4.2	51	0.79	8.6	0.73	
<b>Aire</b>										
AirDar	Aire	2011	53.79336	-1.54911	2.7	43	0.73	8.6	0.75	
<b>Arun</b>										
AruHor'95	Arun	1995	51.05516	-0.36197	4.1	54	0.90	7.8	0.73	
AruHor'00	Arun	2000	51.0556	-0.36124	4.1	34	0.29	7.6	0.73	[25]
AruHor'08	Arun	2008	"	"	4.1	69	0.73	7.9	0.74	[26]
AruHUS	Arun	1995	51.05953	-0.35326	0.2	48	<b>0.01</b>	7.4	0.73	[3]
<b>Thames catchment</b>										
BlaEvH'10	Blackwater	2010	e		4.2	41	0.98	8.5	0.72	
BlaEvH'00	Blackwater	2000	51.327864	-0.769635	8.8	47	0.98	8.3	0.71	[25]
BouChe'11	Bourne	2011	51.38086	-0.47711	4.8	56	0.62	8.3	0.74	
BouChe'02	Bourne	2002	51.40286	-0.54223	5.8	31	0.55	8.8	0.75	
BouChe'06	Bourne	2006	51.40330	-0.54150	5.8	48	0.25	8.3	0.75	[26]
GadCas	Gade	2010	51.65893	-0.42559	NM <sup>f</sup>	56	0.38	8.1	0.74	
KenBul	Kennet	2010	51.39707	-1.28485	0.6	51	0.82	8.4	0.74	
KenFou	Kennet	2010	51.43564	-0.97664	8.1	32	0.64	8.7	0.75	
KenNor	Kennet	2010	51.40165	-1.33725	0.2	52	0.62	8.5	0.73	
LamSha	Lambourn	2011	51.40816	-1.30843	0.03	41	0.10	6.8	0.72	
LeeEss	Lee	2010	51.77279	-0.18818	6.6	56	0.60	8.3	0.73	
LeeHyd	Lee	2010	51.83958	-0.35825	10.3	28	0.36	8.2	0.73	
LeeHyd	Lee	1995	51.84751	-0.37111	11.6	44	0.62	7.7	0.70	[3]
LeeHUS	Lee	1995	51.84935	-0.37395	NM	37	<b>0.001</b>	7.0	0.73	[3]
LeeSta	Lee	2010	51.7894	-0.22496	6.6	31	0.85	8.4	0.71	
Lee'00	Lee	2000	g		NM	41	0.90	8.1	0.70	[25]
LeeWhe	Lee	2010	51.81424	-0.28903	6.6	55	<b>0.05</b>	8.5	0.75	
MolMea	Mole	2010	51.19028	-0.18581	5.8	42	0.84	8.4	0.73	
RayRod	Ray	2003	51.57093	1.81815	10.9	30	0.40	7.7	0.72	
StoBri	Stort	2010	51.77989	0.05024	4.1	52	0.92	8.1	0.71	
StoTed	Stort	2010	51.83115	0.16892	6.0	30	0.97	8.2	0.69	
ThaCul	Thames	2010	51.65046	-1.26739	1.6	44	0.45	8.8	0.74	
ThaHam	Thames	2010	51.55989	-0.87347	1.8	44	0.45	8.2	0.74	
ThaWhi	Thames	2010	51.48662	1.08974	1.5	60	0.88	8.2	0.74	
ThaSha	Thame	2010	51.75281	-1.03319	1.9	50	0.38	7.9	0.74	
WanMoh	Wandle	2011	51.40329	0.18821	3.3	48	0.33	8.2	0.75	

**Table 1 Sampling locations, genetic diversity statistics (allelic richness (AR) and expected heterozygosity ( $H_e$ )) for each population sample (Continued)**

Trent									
TreWol	Trent	1995	52.781178	-1.971789	3.7	45	0.55	8.0	0.72
TreNot	Trent	2007	<sup>g</sup>			24	0.50	8.9	0.75

Full diversity statistics for each locus are given in Additional file 8 and statistical differences in genetic diversity are given in Additional file 3. <sup>a</sup>Headings in bold indicate different catchments; <sup>b</sup>Estradiol equivalents, the predicted average estrogenicity at the sample site, calculated from the predicted concentrations of  $E_1$ ,  $E_2$  and  $EE_2$ ; <sup>c</sup>TPM = two phase model for microsatellite evolution used for this test; <sup>d</sup>Source of samples; <sup>e</sup>A composite sample of fish caught at BlaEve (51.354119, -0.8584853) and BlaHaw (51.324119, -0.7665606). Effluent concentrations are an average between the two sites for statistical analysis; <sup>f</sup>NM = not modelled. For this site there are no major upstream discharges. For statistical analysis, and average  $E_2Eq$  value for sites with no major discharges was used. <sup>g</sup>The exact locations for these sample sites are unknown. Numbers in bold indicate significance ( $p \leq 0.05$ ).

[see Additional file 6] other possible groups are apparent. We found no evidence that roach in the Thames catchment constitute a distinct genetic group, as samples failed to group together in any analysis (Figure 2, Additional file 5 and Additional file 6). This may reflect a true lack of genetic distinctiveness of roach in this catchment, but may also result from the limited ability of the microsatellite markers used to resolve population genetic structure at this level.

#### Population genetic structure of roach within rivers and catchments

Despite Thames catchment roach appearing not to constitute a single genetic unit, distinct from roach in other regions, the study did find evidence for significant genetic structuring in roach populations within the Thames catchment. This suggests the existence of local subpopulations exchanging a limited number of effective migrants (breeding individuals) rather than panmixia (where all individuals are potential partners). For example, average  $F_{ST}$  between samples in the Thames catchment was 0.022, only slightly lower than the average for between-catchment comparisons (0.028) for the study as a whole, while 262 of the 325 pairwise  $F_{ST}$  comparisons in the catchment were highly significant. There was a

weak, but significant, relationship between genetic and geographic distance ( $r^2 = 0.1089$ ,  $P = 0.010$ ) within the catchment, indicating a tendency for individuals to produce offspring with fish from nearby populations rather than distant populations [see Additional file 7]. Additionally, the population tree (Figure 2) and PCA analyses [see Additional file 5] showed groups comprising samples from neighbouring Thames sites: three from the main Thames; four from the Kennet and its tributary (Lambourn); samples from the Stort and the Lee; and samples from the Wandle and Mole. Samples from the Thames Blackwater collected in the years 2000 and 2010 (approximately two to three generations) clustered with very high bootstrap support (98%) in the population tree (Figure 2). This indicates that this roach population is largely restricted to this stretch of river, which includes both moderately and highly polluted sites, and has no substantial uncontaminated upstream stretch (Figure 1).

Despite the proximity of some populations in the PCA and tree, we also found significant genetic differentiation between samples from some neighbouring stretches of the same river, sometimes occurring over small distances of separation (<10 km), for example, within the upper Lee (see below), between the Lee and the Stort, between the Blackwater and main Thames, between the Lambourn

**Table 2 Analysis of molecular variance (AMOVA) testing for partitioning of genetic variation among roach samples, grouped according to geography**

Source of variation	d.f.	Sum of squares	Variance	% Total	P-value
<b>Geographical partition, location<sup>a</sup></b>					
Among locations	32	584.063	0.11631	2.27	<0.00001
Among samples within locations	6	35.197	0.00953	0.19	0.03226
Within samples	3,499	17,489.296	4.99837	97.54	<0.00001
Total	3,537	18,108.556	5.12421		
<b>Geographical partition, catchment<sup>b</sup></b>					
Among groups	5	189.529	0.06309	1.22	<0.00001
Among samples within groups	33	429.731	0.08909	1.73	<0.00001
Within samples	3,499	17,489.296	4.99837	97.05	<0.00001
Total	3,537	18,108.556	5.15055		

<sup>a</sup>Samples from same location caught in different years are grouped; <sup>b</sup>Samples grouped by catchment (Aire, Arun, 'Blackwater and Chelmer', Nene, Thames and Trent).



**Table 3 Summary of pairwise  $F_{ST}$  and  $D_{est}$  among roach samples (see Additional file 4 for full table of values)**

	$F_{ST}$	$D_{est}$	P-value <sup>a</sup>
<b>Between catchments</b>			
Thames/Arun	0.033, (0.002-0.074) <sup>b</sup>	0.0617, (0.0124-0.1612) <sup>b</sup>	
Thames/Nene	0.026, (0.007-0.076)	0.0453, (0.0126-0.1458)	
Thames/Trent	0.026, (0.008-0.069)	0.0447, (0.0059-0.1394)	
Thames/Chelmer, Blackwater	0.022, (0.005-0.057)	0.0444, (0.0035-0.1191)	
<b>Within catchments</b>			
Thames	0.022, (-0.004-0.090)	0.0376, (-0.0058-0.1914)	
Arun	0.007, (0.002-0.014)	0.0064, (0.0015-0.0104)	
Nene	0.002, (0.000-0.003)	0.0010, (-0.0001-0.0020)	
Blackwater/Chelmer	0.022, (0.005-0.057)	0.0454, (0.0058-0.0652)	
<b>Neighbouring samples in the river Lee/Stort</b>			
LeeHUS/LeeHyd'95	-0.002	0.0017	0.78
LeeHyd'10/LeeWhe	0.020	0.0352	>0.00001
LeeWhe/LeeSta	0.009	0.0189	>0.00001
LeeSta/LeeEss	0.009	0.0127	>0.00001
LeeEss/StoBri	0.015	0.0349	>0.00001
<b>Other neighbouring stretches</b>			
BlaBIM/BlaSti	0.007	0.0058	>0.00001
AruHUS/AruHor'95	0.002	0.0043	0.22

<sup>a</sup>For  $F_{ST}$  estimate; <sup>b</sup>average, (range). Pairwise values for neighbouring stretches in the Lee/Stort, the Anglian Blackwater and the Arun are also given.

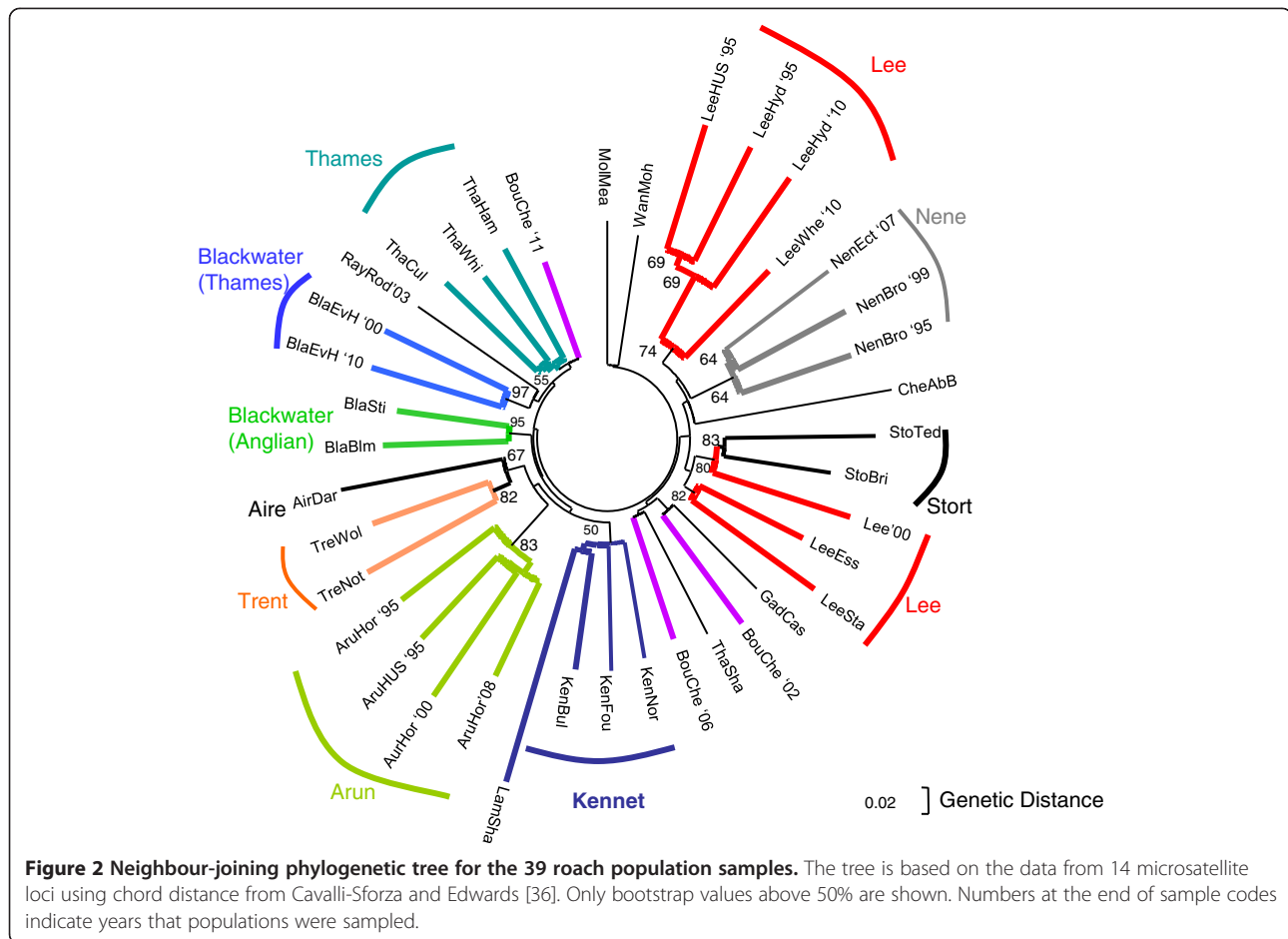
and the Kennet and within the Anglian Blackwater. In other cases, despite the separation of sampling locations by in-river impoundments such as weirs, we found no significant genetic differentiation between sites, for example, within the Stort, the main Thames, the Kennet, the Arun, the Nene and the Trent. Thus, patterns of within-river genetic structure differed between river stretches. For some other fish species analysis of genome-wide SNP data has provided greater resolution in population structure than that achievable using microsatellite data [37], and it is possible that some fine-scale genetic structure in the roach populations has not been detected with the microsatellites used in the current study.

#### Relationship between exposure to estrogenic effluents and effective population size

Estimates of  $N_e$  calculated from the microsatellite data using the approximate Bayesian computation method ( $N_{e(ABC)}$ ), ranged from 54 to 301 for each sample, with higher precision for small  $N_e$  estimates (Figure 3A). We found no evidence for a correlation between  $N_{e(ABC)}$  and predicted  $E_2$  equivalents ( $E_2Eq$ ), a measure of total estrogenicity of the river water due to contamination by sewage effluent (generalized linear models (GLM),  $F_{(1,20)} = 0.7468$ ,  $P = 0.40$ ) or for an interaction between sample site and estrogen exposure (GLM,  $F_{(6,19)} = 1.9954$ ,  $P = 0.14$ ) across the 28 sample sites where no recent restocking had occurred and had sufficient sample sizes for robust  $N_e$

calculation. However, the 95% confidence intervals (CI) for the model coefficient indicated  $N_{e(ABC)}$  could decrease by a maximum of 5.6% for each incremental increase in exposure of 1 ng/L  $E_2Eq$ , or 65% at 11.6 ng/L  $E_2Eq$ , equivalent to the most polluted river stretch included in this study. The inclusion of roach density as an additional covariate within the model also produced a non-significant result (GLM,  $F_{(1,16)} = 1.3966$ ,  $P = 0.26$ ), albeit for a reduced number of sites (19). Similarly, there was no significant correlation between the other variables included in the statistical analyses (average flow rate, geographic/phylogenetic group and roach density) with  $N_{e(ABC)}$ . This analysis makes the assumption that immigration of fish from remote sites is limited and this is discussed below.

There was limited evidence for reduced variation in  $N_e$  in roach sampled from more contaminated stretches of river compared to those sampled from less contaminated sites; all estimates above 6 ng/L  $E_2Eq$  were below 100, whereas there was greater variation (54 to 301) where  $E_2Eq$  was below 6 ng/L (Figure 3B). Estimates of  $N_e$  using other methods were of the same order of magnitude but had wider confidence intervals for each estimate;  $N_{e(SA)}$  ranged from 36 to 145 and also showed no relationship with  $E_2Eq$  (Figure 3C). Temporal estimates of  $N_e$ , calculated from allele frequency changes over several generations using the Jorde and Ryman method [40], varied from 14 to 265, but were available for too



few locations to make meaningful comparisons (Table 4). Overall, the relatively small variations in  $N_e$  observed in this study could not be explained by any of the environmental or other variables measured in this study.

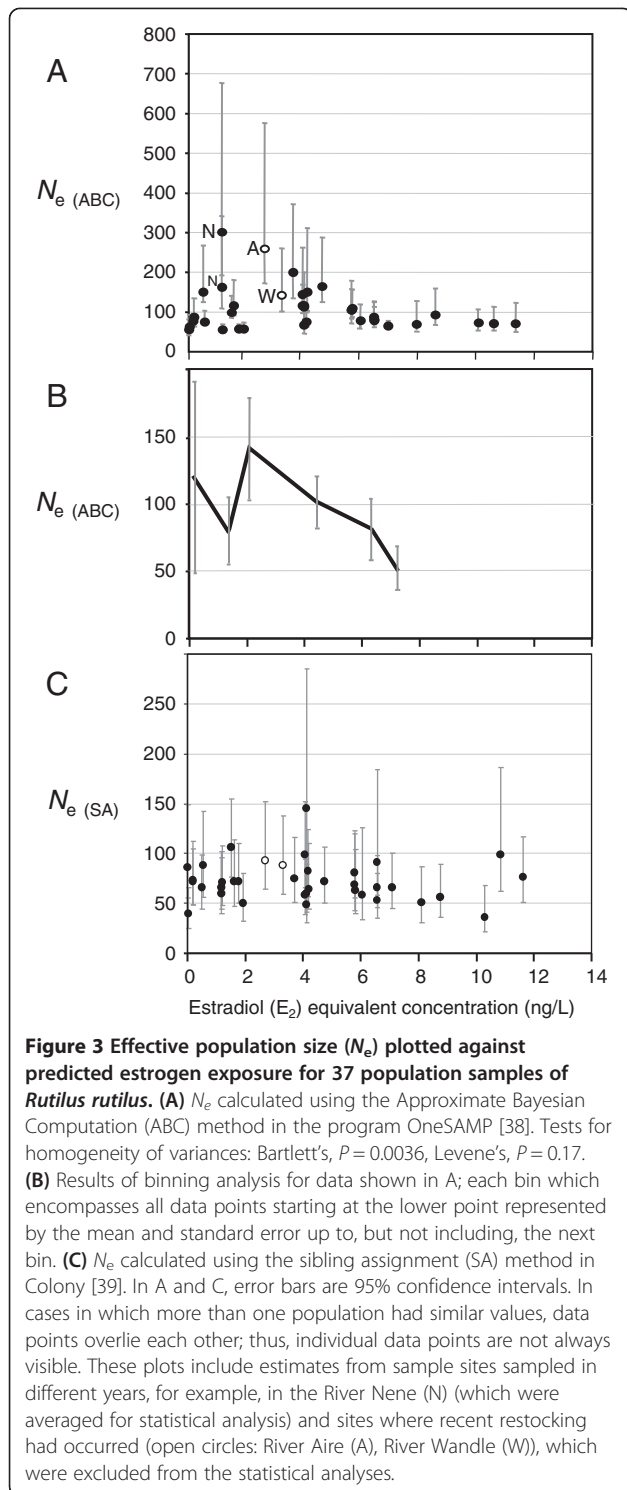
#### Population genetic structure within the River Lee, a high effluent river

The average proportion of effluent in the upper Lee downstream from Harpenden and East Hyde WWTWs ranges from 28% to 70% in different stretches. Significant genetic differentiation was detected between fish sampled from four of the five locations in this stretch of river ( $F_{ST}$  values  $\geq 0.009$ ), and between these samples and two from its tributary, the Stort ( $F_{ST} \geq 0.015$ ) shown in Table 3. The presence of numerous large weirs (Figure 1) likely confines fish to particular areas of this river; nonetheless, samples from the Lee and the Stort did cluster together in some analyses (Figure 2, Additional file 5). A sample from the upstream, unpolluted sample site (LeeHUS) grouped with two samples (collected in 1995 and 2010) from the most polluted river stretch immediately downstream (LeeHyd). The next sample site downstream, LeeWhe, was distinct from these (Figure 2, Additional

file 5 and Additional file 6), indicating restricted movement of fish between LeeHyd and LeeWhe over at least three to five generations (Figure 2). Analysis using the program  $IM_2$  [42] suggested that there was less than one effective (breeding) migrant per generation between LeeHyd and LeeWhe in either direction, and about one migrant per generation from LeeWhe downstream to LeeSta (Figure 4). Collectively, these data suggest that roach populations at LeeWhe and those downstream do not rely on migration from the uncontaminated stretch of this river. Despite this,  $N_{e(ABC)}$  estimates for these polluted sites in the upper Lee ranged from 70 to 84 (95% CI: 50 to 127) compared to only 54 (95% CI: 42 to 82) for the upstream uncontaminated location, suggesting no substantial impact of the effluent on the effective population size of these roach.

#### Discussion

In this study, analyses of population genetic structure via the analysis of DNA microsatellite loci identified distinct subpopulations of roach in two tributaries of the Thames, the rivers Lee and the Blackwater, that were largely restricted to high-effluent stretches of the rivers



over multiple generations. This is despite evidence for widespread feminization of male fish in the studied rivers and previous evidence that feminization alters breeding capabilities [21,26]. Both of these tributaries contain feminized fish [25]), with predicted average exposure of between 4 and 9 ng/L  $E_2Eq$ . We also found no statistically

**Table 4 Temporal estimates of effective population size ( $N_e$ ) among roach samples**

Site	Time interval	$G^a$	$N_e$ (95% CI) <sup>b</sup>	$N_e$ (95% CI) <sup>c</sup>
Nene	1995 to 1999	1	265 (66-∞)	619 (82-∞)
AruHor	1995 to 2000	1	14 (8-79)	48 (26-137)
	2000 to 2008	2	60 (31-1733)	321 (106-∞)
	1995 to 2008	3	73 (43-247)	232 (123-669)
BouChe	2002 to 2006	1	32 (17-219)	51 (25-151)
	2006 to 2011	1	45 (23-9572)	63 (35-162)
	2002 to 2011	2	87 (44-3897)	495 (295-1405)
LeeHyd	1995 to 2010	3	137 (81-∞)	346 (145-35518)
BlaEvH	2000 to 2010	2	141 (68-∞)	206 (83-∞)

<sup>a</sup> $G$  = assumed number of generations between sampling points; <sup>b</sup>Calculated using the Jorde and Ryman method [40]; <sup>c</sup>Calculated using the classical moment-based method of Waples [41]. CI, confidence interval.

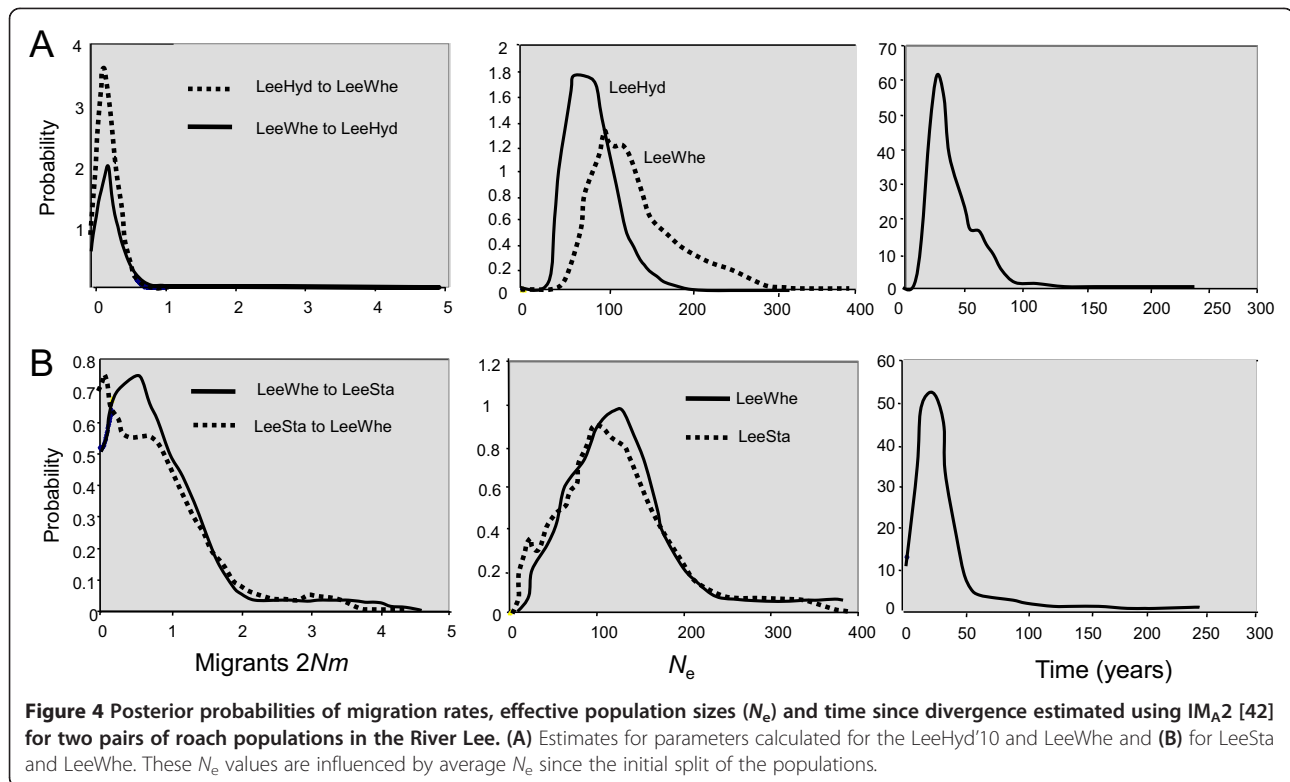
robust evidence for a substantial impact of estrogenic sewage effluents on  $N_e$  of roach. The possibility of a reduction in  $N_e$  of up to 65% for roach living in the most polluted river stretches ( $E_2Eq$  of 11.6 ng/L) could not be ruled out, due to the wide 95% confidence intervals associated with the statistical model. Moreover, our analysis included relatively few samples from rivers in the highest risk category, largely because these sites are rare.

#### Caveats

As with any modelling exercise, this analysis makes assumptions that may affect the interpretation of the results. One of these assumptions is that migration between sites with different pollution profiles is limited over two to three generations, the time frame likely to have the greatest influence on  $N_{e(ABC)}$  [43]. This was ensured by selecting sites with physical obstructions between them. However, quantifying migration rates over this timescale was not always possible because all potential source populations could not be sampled and, in some cases, we found no significant genetic differentiation between roach at sites distant from one another. Genetic differentiation can take many generations to manifest with low levels of migration [44]. Histological data from the Arun and the Lee show that feminized gonads in roach were approximately 6-fold (Lee) and approximately 2.5-fold (Arun) more prevalent in populations living in the stretches downstream of major WWTW inputs compared with those living upstream [3]. This demonstrates that migration in these rivers was indeed restricted to stretches delimited by physical barriers, despite no significant genetic differentiation observed between river stretches ( $F_{ST} < 0.002$ , Table 3).

A second assumption is that no restocking of the rivers sampled had occurred or that the effect of restocking activities on  $N_{e(ABC)}$  was relatively minor. Approximately 500,000 hatchery-reared roach just over one-year-old





(so called '1+' fish) have been introduced into the Thames catchment since 2000; broodstock for hatchery fish originate from the river Trent. The influence of restocking activities on  $N_{e(ABC)}$ , however, is likely to be relatively minor, as the sites sampled in this study were separated by major physical barriers from sites where these introductions had occurred. Moreover, introductions prior to 2000 are unlikely to have had a large influence on  $N_{e(ABC)}$  as this is primarily affected by the size and variance in reproductive success of the parental generation, which would have spawned between 2004 and 2007 for most of the samples in this study. However, we cannot exclude some influence of introductions prior to 2000 as some of the summary statistics used to calculate  $N_{e(ABC)}$  are known to be affected by demographic processes over a longer time period [38]. The effects of introductions on genetic diversity, the detection of bottlenecks and population structure are likely to be greater, as these factors are affected by demography over many generations. However, neither the success of the reintroduced fish nor the size of the roach population in the Thames is currently known. In salmonids, restocking success is highly variable and has been attributed to local adaptation [45]. Using our microsatellite dataset, 73% of 48 individual roach from a stretch of the Wandle (restocked in 2007, 2009, 2010) assigned to the Thames reporting regions. Only 5% (two fish) assigned to the Trent (the source of the parents of introduced fish), which may be mis-assignments, as 5% also assigned to the

Arun, 10% to the Anglian Blackwater and 10% to the Nene, from where no restocking had taken place. Thus, the success of the re-introduced fish may be low, but this requires further investigation.

#### Evidence for self-sustaining populations in effluent contaminated rivers

While this study does not exclude the possibility that estrogenic effluents reduce  $N_e$  of fish populations, it suggests that roach populations can be self-sustaining despite exposure to estrogens over several generations. These findings are consistent with the fact that the prevalence of male fish with moderate to severely feminized gonads (that have been shown to have substantially reduced reproductive competitiveness in controlled breeding studies) is generally less than 10% in English rivers [3,25,26,46]. The reproductive competitiveness of fish with the more common mild-intersex condition is similar to those of fish without gonadal feminization [26]. In roach, the gonads of male fish exposed to estrogens become progressively feminized with age [46], so gonadal feminization could theoretically increase  $N_e$  by reducing the reproductive dominance of large older males in estrogen-contaminated rivers. While the effects on females are less well studied, female roach exposed to an undiluted effluent for three years in large tanks were able to breed, despite the fact that this exposure caused complete gonadal feminization of males [21]; similarly, the majority of females collected from two

effluent-polluted rivers examined in this study were also able to breed [26].

#### **Population risks of long-term exposures to estrogenic effluents**

The results of this study on wild roach populations seem to contrast with studies that have assessed population risk through long-term exposures to estrogens, where exposure to concentrations between 3 and 6 ng EE<sub>2</sub>/L [17-20] or to a full-strength effluent [21] resulted in all-female populations and/or reproductive failure. The apparent difference between the wild populations and those experiments designed to simulate 'real world' exposure, however, may be because the fish living in the effluent-contaminated rivers examined in this study have been exposed to a lower level of estrogen or because all of the estrogen is not bioavailable; organic pollutants can bind to particulates and dissolved organic matter [47]. The most contaminated river in this study has a mean proportion of effluent of approximately 70%, although the majority of contaminated English rivers average approximately 10% to 30% [34]. While EE<sub>2</sub> has been measured up to approximately 4 to 8 ng/L in English WWTW effluents [33,48], for the most part, they are lower [23,49], and estrogen concentrations vary greatly over short periods. For instance, EE<sub>2</sub> was detected in only 21 of 135 water samples from the Lee, although occasionally reaching 4 ng/L [33]. Considering the totality of estrogen content, the predicted average estrogenicity of the most contaminated site in this study is 12 ng/L E<sub>2</sub>Eq and would be below 21 ng/L for 90% of the time. Only 1% to 3% of 10,313 individual river reaches in the UK receiving WWTW effluent were predicted to have average E<sub>2</sub>Eq >10 ng/L, and, of these, many are ditches composed almost entirely of sewage effluent [34]. As E<sub>2</sub> is approximately 10 times less potent than EE<sub>2</sub> in inducing gonadal feminization in fish [34,50], it is probable that average life-time exposure to estrogens in the wild does not currently reach the concentrations shown to cause sex-reversal and population collapse in controlled experimental exposures. Green *et al.* [51] recently predicted a doubling of estrogen exposure concentrations in some rivers with population growth and climate change by 2050 suggesting an increased likelihood of population level effects of estrogenic effluents in the future, unless mitigated by substantial improvements in sewage treatment processes.

#### **Influences on the population genetic structure of roach**

The population genetic structure of roach in southern England observed in this study may have been influenced by historical biogeography, migratory behaviour, human translocations and in-river barriers. Roach can be highly mobile and can migrate over 10 km, particularly in the spawning period April to June, if migration is not obstructed [52];

additionally, there is some evidence that roach show fidelity in migration and return to spawning sites they have used previously [53]. Within the Thames catchment, the observed population genetic structure likely results, at least in part, from the large number of physical barriers, such as weirs and locks (Figure 1); these have been recognized as major factors restricting movement (including downstream) of roach [54]. Similarly, the importance of barriers in driving intra-catchment genetic variation is well documented in other fish species, for example, brown trout [55]. Only obstructions in the main River Thames and the Kennet are equipped with fish passes and, although some passes can be used by roach [56], the effectiveness of these passes in allowing fish movement has not been studied. As we identified significant genetic differentiation between roach from the Kennet and the Thames, despite being connected by fish passes, these passes may represent major physical separation barriers to this fish species.

#### **Conclusions**

Despite the widespread feminization of male roach in effluent-contaminated rivers of southern England, using nuclear DNA microsatellites we were able to identify some populations that have been confined to stretches of river with moderate to high exposure to estrogenic effluents over multiple generations. We also found no evidence of a correlation between the  $N_e$  of roach populations and predicted exposure to estrogens, although because of the wide confidence intervals, a reduction in  $N_e$  of up to 65% is still possible at the most contaminated sites.

#### **Methods**

##### **Study location**

Southern England, particularly the region within the Thames catchment, was chosen for this study for four reasons. Firstly, it is a densely populated region with relatively low rainfall and, therefore, includes some river stretches with some of the highest concentrations of WWTW effluents in the United Kingdom [34]. Secondly, feminization of roach has been widely reported in the region [3,46]. Thirdly, many rivers in the region have locks, dams or weirs which are likely to limit movement of fish species between stretches of river with different pollution profiles. Fourthly, the effluent concentrations and risk of estrogenic endocrine disruption have been modelled [34]. Sample sites are shown in Figure 1 and Table 1 and were selected to span the full range of predicted estrogen concentrations in English rivers and where obstructions are likely to restrict fish movements [3,46].

##### **Roach study species**

Roach was selected as the study species because it is native and widely distributed in the United Kingdom, including

in rivers polluted with WWTW effluents. Additionally, widespread feminization has been reported in wild populations and with a proven association with exposure to estrogenic effluents [14,19,21,57]. Roach generally reach sexual maturity between two and three years and spawn annually in the spring. Adult roach can migrate considerable distances, but where weirs obstruct upstream and downstream movement they are able to complete their lifecycles in a single stretch of river [54].

### Population-genetic analyses

To understand the extent to which roach populations are restricted to various stretches of river, several approaches were used to investigate population genetic structure. We analysed microsatellite loci variation in 1,769 fish sampled between 1995 and 2011. Each fish was genotyped at between 14 to 19 microsatellite loci. Microsatellite genotypes are provided in Additional file 8. Protocols for DNA extraction and details of amplification of the microsatellite loci are illustrated in Additional file 1. Data for 14 microsatellite loci were used to calculate three measures of genetic diversity: observed heterozygosity ( $H_O$ ) and expected heterozygosity ( $H_e$ ) using GenAEx 6 [58]; allelic richness ( $AR$ ) was calculated using Fstat v2.9.3 [59] – see Additional file 1 for full details. The programme BOTTLENECK [60,61] was used to test for recent genetic bottlenecks. This programme tests for a relative excess in heterozygosity that is apparent for a few generations after a bottleneck and develops because allelic diversity declines faster than heterozygosity, due to loss of rare alleles. Pairwise genetic differentiation between the sampled sites was estimated using  $F_{ST}$ , calculated using Arlequin 3.5 [62] and Jost's  $D$ ,  $D_{est}$  [63], calculated using SMOGD [64]. The significance of the  $F_{ST}$  estimates was assessed based on 10,000 permutations. AMOVA was performed using Arlequin. In order to test whether fish are more likely to produce offspring with local mates, compared to mates in geographically distant locations within the Thames catchment, isolation by distance analysis was performed using the Mantel test [65] in GenAEx 6 [58]. Genetic similarity between populations was investigated using population based trees, calculated in POPULATIONS, v1.2.30beta [66], PCA in GenAEx 6 [58] and a Bayesian clustering approach in STRUCTURE [67]. Finally, the program  $IM_A2$  [42,68] was used to estimate migration rates between adjacent populations within high effluent stretches of the Lee, giving relatively high pairwise  $F_{ST}$  values (LeeHyd, LeeWhe, and LeeSta). See Additional file 1 for further details. To investigate the influence of restocking, genetic assignment of fish from the Wandle was undertaken using the 'leave one out test' in the computer program, ONCOR [69], based on their microsatellite genotypes. The reporting regions comprised: (1) the Wandle, (2) Lee/Stort, (3) rest of the Thames,

(4) Trent, (5) Nene, (6) Arun, (7) Chelmer and (8) Anglian Blackwater. All animals used in this research were treated humanely and with regard for the alleviation of suffering; all procedures were subject to approval by the local ethical review process as required under the U.K. Animals (Scientific Procedures) Act (1986).

### Effective population size

To test whether WWTW effluents substantially reduce the size of breeding populations, effective population sizes ( $N_e$ ), which relate to the number of breeding fish and skews in breeding success, were estimated using the microsatellite genotypes. We compared  $N_e$  from sites ranging from little/no upstream WWTW effluent inputs to those where the majority of the flow can comprise WWTW effluent. Two single sample (generation) methods, that use different aspects of the microsatellite data, were used to estimate  $N_e$  for each population; the Approximate Bayesian Computation (ABC) method using ONEsAMP 1.2 [38], hereafter referred to as  $N_{e(ABC)}$ ; and the sibling assignment method (SA),  $N_{e(SA)}$  [39]. Temporal estimates for  $N_e$ , which are calculated from the change in allele frequencies between generations, were also estimated for sites where fish had been sampled more than once using TempoFs [40] and NeEstimator [70]. For further details see Additional file 1.

### Statistical analysis

GLM were used to examine the relationship between predicted exposure to estrogenic effluents and  $N_{e(ABC)}$ . For further details see Additional file 1. Differences in genetic diversity among sampled populations were tested using ANOVA. All statistical analysis was performed using the software R 2.13.0 [71].

### Additional files

**Additional file 1:** Details of microsatellite genotyping methods, population-genetic analysis and statistical analysis.

**Additional file 2:** Genetic diversity estimates for each locus in each population.

**Additional file 3:** Graph showing statistical differences in genetic diversity among roach *Rutilus rutilus* populations.

**Additional file 4:** Pairwise  $F_{ST}$  and  $D_{est}$  among roach population samples.

**Additional file 5:** Multidimensional scaling plots of pairwise  $D_A$  distances [72].

**Additional file 6:** Structure analyses plots using the *locprior* model and analysis of optimum number of genetic units in Structure Harvester [73].

**Additional file 7:** Correlation between genetic distance and geographic distance (km) between pairs of sites for 24 population samples from the Thames catchment.

**Additional file 8:** Microsatellite genotypes for 1,769 roach genotyped in this study.

## Abbreviations

AMOVA: analysis of molecular variance; ANOVA: analysis of variance; AR: allelic richness; CI: confidence interval; E<sub>1</sub>: Estrone; E<sub>2</sub>: Estradiol; EE<sub>2</sub>: 17 alpha-ethinylestradiol; GLM: Generalized Linear Models; H<sub>e</sub>: expected heterozygosity; HWE: Hardy-Weinberg equilibrium; MCMC: Markov chain Monte Carlo; MEGA: Molecular Evolutionary Genetic Analysis; N<sub>b</sub>: effective number of breeders; N<sub>e</sub>: effective population size; N<sub>e(SA)</sub>: effective population size estimate calculated using the sibling assignment method; N<sub>e(ABC)</sub>: effective population size estimate calculated using the approximate Bayesian computation method; PCA: principal component analysis; SNP: single nucleotide polymorphism; WWTW: Waste Water Treatment Work (WWTW).

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

PH, EN, JS, SJ, JS, JS and CT participated in the design of the research. RW modeled estrogen exposure in rivers. PH, EB and EN participated in data collection, analysis of microsatellite data and statistical analysis. PH, EN, EB, RW, JS, SJ, JS and CT wrote the paper. All authors read and approved the final manuscript.

## Authors' information

PH, the corresponding author, is a molecular ecologist and evolutionary biologist, and is a Research Fellow at the Biosciences Department at the University of Exeter.

## Acknowledgements

This work was funded by Natural Environmental Research Council (NERC) grant numbers NE/G019355/1 and NE/K004263/1 and the Environment Agency (England and Wales). Contains Environment Agency information © Environment Agency and database right. We thank the Environment Agency for support from their Fisheries Teams and databases, Jinliang Wang and David Tallmon for their advice on using and interpreting the output of Colony and OneSAMP, and Tiago Antao for advice on calculating and interpreting N<sub>e</sub>. We are grateful to Anke Lange for providing EST sequences, Andrew Griffiths and Andy King for advice on population-genetic analysis, and Veronica Vinciotti, Marta Söfker, Venura Perera and James Creswell for support on statistical aspects of this work.

## Author details

<sup>1</sup>Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK. <sup>2</sup>Institute for the Environment, Brunel University, Uxbridge, Middlesex UB8 3PH, UK. <sup>3</sup>Centre for Ecology and Hydrology, Oxfordshire OX10 8BB, UK.

Received: 31 October 2013 Accepted: 20 December 2013

Published: 13 January 2014

## References

1. Gilbert N: Water under pressure. *Nat Biotechnol* 2012, **483**:256–257.
2. Heberer T: Occurrence, fate, and removal of pharmaceutical residues in the aquatic environment: a review of recent research data. *Toxicol Lett* 2002, **131**:5–17.
3. Jobling S, Nolan M, Tyler CR, Brighty G, Sumpter JP: Widespread sexual disruption in wild fish. *Environ Sci Technol* 1998, **32**:2498–2506.
4. Hinck JE, Blazer VS, Schmitt CJ, Papoulias DM, Tillitt DE: Widespread occurrence of intersex in black basses (*Micropterus* spp.) from US rivers, 1995–2004. *Aquat Toxicol* 2009, **95**:60–70.
5. Harries JE, Sheahan DA, Jobling S, Matthiessen P, Neall M, Sumpter JP, Taylor T, Zaman N: Estrogenic activity in five United Kingdom rivers detected by measurement of vitellogenesis in caged male trout. *Environ Toxicol Chem* 1997, **16**:534–542.
6. Lu GH, Song WT, Wang C, Yan ZH: Assessment of *in vivo* estrogenic response and the identification of environmental estrogens in the Yangtze River (Nanjing section). *Chemosphere* 2010, **80**:982–990.
7. Purdom CE, Hardiman PA, Bye WJ, Eno NC, Tyler CR, Sumpter JP: Estrogenic effects of effluents from sewage treatment works. *Chem Ecol* 1994, **8**:275–285.
8. Desbrow C, Routledge EJ, Brighty GC, Sumpter JP, Waldock M: Identification of estrogenic chemicals in STW effluent. 1. Chemical fractionation and *in vitro* biological screening. *Environ Sci Technol* 1998, **32**:1549–1558.
9. Gravato C, Santos MA: *Dicentrarchus labrax* biotransformation and genotoxicity responses after exposure to a secondary treated industrial/urban effluent. *Ecotoxicol Environ Saf* 2003, **55**:300–306.
10. Filby AL, Neuparth T, Thorpe KL, Owen R, Galloway TS, Tyler CR: Health impacts of estrogens in the environment, considering complex mixture effects. *Environ Health Perspect* 2007, **115**:1704–1710.
11. Thorpe KL, Maack G, Benstead R, Tyler CR: Estrogenic wastewater treatment works effluents reduce egg production in fish. *Environ Sci Technol* 2009, **43**:2976–2982.
12. Pottinger TG, Henrys PA, Williams RJ, Matthiessen P: The stress response of three-spined sticklebacks is modified in proportion to effluent exposure downstream of wastewater treatment works. *Aquat Toxicol* 2013, **126**:382–392.
13. Sebire M, Katsiadaki I, Taylor NGH, Maack G, Tyler CR: Short-term exposure to a treated sewage effluent alters reproductive behaviour in the three-spined stickleback (*Gasterosteus aculeatus*). *Aquat Toxicol* 2011, **105**:78–88.
14. Baynes A, Green C, Nicol E, Beresford N, Kanda R, Henshaw A, Churchley J, Jobling S: Additional treatment of wastewater reduces endocrine disruption in wild fish—a comparative study of tertiary and advanced treatments. *Environ Sci Technol* 2012, **46**:5565–5573.
15. Filby AL, Shears JA, Drage BE, Churchley JH, Tyler CR: Effects of advanced treatments of wastewater effluents on estrogenic and reproductive health impacts in fish. *Environ Sci Technol* 2010, **44**:4348–4354.
16. Owen R, Jobling S: Environmental science: The hidden costs of flexible fertility. *Nature* 2012, **485**:441.
17. Kidd KA, Blanchfield PJ, Mills KH, Palace VP, Evans RE, Lazorchak JM, Flick RW: Collapse of a fish population after exposure to a synthetic estrogen. *Proc Natl Acad Sci U S A* 2007, **104**:8897–8901.
18. Lange R, Hutchinson TH, Croudace CP, Siegmund F, Schweinfurth H, Hampe P, Panter GH, Sumpter JP: Effects of the synthetic estrogen 17 alpha-ethinylestradiol on the life-cycle of the fathead minnow (*Pimephales promelas*). *Environ Toxicol Chem* 2001, **20**:1216–1227.
19. Lange A, Paull GC, Coe TS, Katsu Y, Urushitani H, Iguchi T, Tyler CR: Sexual reprogramming and estrogenic sensitization in wild fish exposed to ethinylestradiol. *Environ Sci Technol* 2009, **43**:1219–1225.
20. Nash JP, Kime DE, Van der Ven LT, Wester PW, Brion F, Maack G, Stahlschmidt-Allner P, Tyler CR: Long-term exposure to environmental concentrations of the pharmaceutical ethinylestradiol causes reproductive failure in fish. *Environ Health Perspect* 2004, **112**:1725–1733.
21. Lange A, Paull GC, Hamilton PB, Iguchi T, Tyler CR: Implications of persistent exposure to treated wastewater effluent for breeding in wild roach (*Rutilus rutilus*) populations. *Environ Sci Technol* 2011, **45**:1673–1679.
22. Grist EP, Wells NC, Whitehouse P, Brighty G, Crane M: Estimating the effects of 17 alpha-ethinylestradiol on populations of the fathead minnow *Pimephales promelas*: Are conventional toxicological endpoints adequate? *Environ Sci Technol* 2003, **37**:1609–1616.
23. Hannah R, D'Aco VJ, Anderson PD, Buzby ME, Caldwell DJ, Cunningham VL, Ericson JF, Johnson AC, Parke NJ, Samuelian JH, Sumpter JP: Exposure assessment of 17 alpha-ethinylestradiol in surface waters of the United States and Europe. *Environ Toxicol Chem* 2009, **28**:2725–2732.
24. Parrott JL, Blunt BR: Life-cycle exposure of fathead minnows (*Pimephales promelas*) to an ethinylestradiol concentration below 1 ng/L reduces egg fertilization success and demasculinizes males. *Environ Toxicol* 2005, **20**:131–141.
25. Jobling S, Coey S, Whitmore JG, Kime DE, Van Look KJ, McAllister BG, Beresford N, Henshaw AC, Brighty G, Tyler CR, Sumpter JP: Wild intersex roach (*Rutilus rutilus*) have reduced fertility. *Biol Reprod* 2002, **67**:515–524.
26. Harris CA, Hamilton PB, Runnalls TJ, Vinciotti V, Henshaw A, Hodgson D, Coe TS, Jobling S, Tyler CR, Sumpter JP: The consequences of feminization in breeding groups of wild fish. *Environ Health Perspect* 2011, **119**:306–311.
27. Cotton S, Wedekind C: Population consequences of environmental sex reversal. *Conserv Biol* 2009, **23**:196–206.
28. Lorenzen K, Enberg K: Density-dependent growth as a key mechanism in the regulation of fish populations: evidence from among-population comparisons. *Proc Biol Sci* 2002, **269**:49–54.
29. Hauser L, Adcock GJ, Smith PJ, Ramirez JH, Carvalho GR: Loss of microsatellite diversity and low effective population size in an



- overexploited population of New Zealand snapper (*Pagrus auratus*). *Proc Natl Acad Sci U S A* 2002, **99**:11742–11747.
30. Liu JX, Ely B: Sibship reconstruction demonstrates the extremely low effective population size of striped bass *Morone saxatilis* in the Santee-Cooper system, South Carolina, USA. *Mol Ecol* 2009, **18**:4112–4120.
31. Charlesworth B: Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* 2009, **10**:195–205.
32. Reed DH, Frankham R: Correlation between fitness and genetic diversity. *Conserv Biol* 2003, **17**:230–237.
33. Williams RJ, Johnson AC, Smith JJ, Kanda R: Steroid estrogens profiles along river stretches arising from sewage treatment works discharges. *Environ Sci Technol* 2003, **37**:1744–1750.
34. Williams RJ, Keller VD, Johnson AC, Young AR, Holmes MG, Wells C, Gross-Sorokin M, Benstead R: A national risk assessment for intersex in fish arising from steroid estrogens. *Environ Toxicol Chem* 2009, **28**:220–230.
35. Evanno G, Regnaut S, Goudet J: Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 2005, **14**:2611–2620.
36. Cavalli-Sforza L, Edwards A: Phylogenetic analysis: models and estimation procedures. *Am J Hum Genet* 1967, **19**:233–257.
37. Corander J, Majander KK, Cheng L, Merila J: High degree of cryptic population differentiation in the Baltic Sea herring *Clupea harengus*. *Mol Ecol* 2013, **22**:2931–2940.
38. Tallmon DA, Koyuk A, Luikart G, Beaumont MA: ONeSAMP: a program to estimate effective population size using approximate Bayesian computation. *Mol Ecol Resour* 2008, **8**:299–301.
39. Jones OR, Wang J: COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol Ecol Resour* 2009, **10**:551–555.
40. Jorde PE, Ryman N: Unbiased estimator for genetic drift and effective population size. *Genetics* 2007, **177**:927–935.
41. Waples RS: A generalized approach for estimating effective population-size from temporal changes in allele frequency. *Genetics* 1989, **121**:379–391.
42. Hey J: Isolation with migration models for more than two populations. *Mol Biol Evol* 2010, **27**:905–920.
43. Neel MC, McKelvey K, Ryman N, Lloyd MW, Bull RS, Allendorf FW, Schwartz MK, Waples RS: Estimation of effective population size in continuously distributed populations: there goes the neighborhood. *Heredity* 2013, **111**:189–199.
44. Mills LS, Allendorf FW: The one-migrant-per-generation rule in conservation and management. *Conserv Biol* 1996, **10**:1509–1518.
45. Griffiths AM, Ellis JS, Clifton-Dey D, Machado-Schiaffino G, Bright D, Garcia-Vazquez E, Stevens JR: Restoration versus recolonisation: the origin of Atlantic salmon (*Salmo solar* L.) currently in the River Thames. *Biol Conserv* 2011, **144**:2733–2738.
46. Jobling S, Williams R, Johnson A, Taylor A, Gross-Sorokin M, Nolan M, Tyler CR, van Aerle R, Santos E, Bright G: Predicted exposures to steroid estrogens in UK rivers correlate with widespread sexual disruption in wild fish populations. *Environ Health Perspect* 2006, **114**:32–39.
47. Gong J, Ran Y, Chen DY, Yang Y, Zeng EY: Association of endocrine-disrupting chemicals with total organic carbon in riverine water and suspended particulate matter from the Pearl River, China. *Environ Toxicol Chem* 2012, **31**:2456–2464.
48. Rodgers-Gray TP, Jobling S, Morris S, Kelly C, Kirby S, Janbaksh A, Harries JE, Waldock MJ, Sumpter JP, Tyler CR: Long-term temporal changes in the estrogenic composition of treated sewage effluent and its biological effects on fish. *Environ Sci Technol* 2000, **34**:1521–1528.
49. Williams RJ, Churchley JH, Kanda R, Johnson AC: Comparing predicted against measured steroid estrogen concentrations and the associated risk in two United Kingdom river catchments. *Environ Toxicol Chem* 2012, **31**:892–898.
50. Caldwell DJ, Mastrocco F, Anderson PD, Länge R, Sumpter JP: Predicted-no-effect concentrations for the steroid estrogens estrone, 17 $\beta$ -estradiol, estril, and 17 $\alpha$ -ethinylestradiol. *Environ Toxicol Chem* 2012, **31**:1396–1406.
51. Green C, Williams R, Kanda R, Churchley J, He Y, Thomas S, Goonan P, Kumar A, Jobling S: Modeling of steroid estrogen contamination in UK and South Australian Rivers predicts modest increases in concentrations in the future. *Environ Sci Technol* 2013, **47**:7224–7232.
52. Baade U, Fredrich F: Movement and pattern of activity of the roach in the River Spree, Germany. *J Fish Biol* 1998, **52**:1165–1174.
53. Vollestad LA, Labeelund JH: Reproductive biology of stream-spawning roach, *Rutilus rutilus*. *Environ Biol Fishes* 1987, **18**:219–227.
54. Geeraerts C, Ovidio M, Verbiest H, Buysse D, Coeck J, Belpaire C, Philippart JC: Mobility of individual roach *Rutilus rutilus* (L.) in three weir-fragmented Belgian rivers. *Hydrobiologia* 2007, **582**:143–153.
55. Griffiths AM, Koizumi I, Bright D, Stevens JR: A case of isolation by distance and short-term temporal stability of population structure in brown trout (*Salmo trutta*) within the River Dart, southwest England. *Evol Appl* 2009, **2**:537–554.
56. Knaepkens G, Baekelandt K, Eens M: Fish pass effectiveness for bullhead (*Cottus gobio*), perch (*Perca fluviatilis*) and roach (*Rutilus rutilus*) in a regulated lowland river. *Ecol Freshw Fish* 2006, **15**:20–29.
57. Lange A, Katsu Y, Ichikawa R, Paull GC, Chidgey LL, Coe TS, Iguchi T, Tyler CR: Altered sexual development in roach (*Rutilus rutilus*) exposed to environmental concentrations of the pharmaceutical 17 alpha-ethinylestradiol and associated expression dynamics of aromatases and estrogen receptors. *Toxicol Sci* 2008, **106**:113–123.
58. Peakall R, Smouse PE: GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 2006, **6**:288–295.
59. Goudet J: FSTAT, a program to estimate and test gene diversities and fixation indices. 2.9.3 edition. 2002. Available at: <http://www2.unil.ch/popgen/softwares/fstat.htm>.
60. Cornuet JM, Luikart G: Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 1996, **144**:2001–2014.
61. Piry S, Luikart G, Cornuet JM: BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J Hered* 1999, **90**:502–503.
62. Excoffier L, Laval G, Schneider S: Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* 2005, **1**:47–50.
63. Jost L:  $G_{ST}$  and its relatives do not measure differentiation. *Mol Ecol* 2008, **17**:4015–4026.
64. Crawford NG: SMOGD: software for the measurement of genetic diversity. *Mol Ecol Resour* 2010, **10**:556–557.
65. Mantel NA: The detection of disease clustering and a generalized regression approach. *Cancer Res* 1967, **27**:209–220.
66. Langella O: Populations 1.2.30; Population genetic software: individuals or populations distances based on allelic frequencies, phylogenetic trees, file conversions. 1999. Available at: [http://www.bioinformatics.org/project/?group\\_id=84](http://www.bioinformatics.org/project/?group_id=84).
67. Pritchard JK, Stephens M, Donnelly P: Inference of population structure using multilocus genotype data. *Genetics* 2000, **155**:945–959.
68. Hey J, Nielsen R: Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc Natl Acad Sci USA* 2007, **104**:2785–2790.
69. Kalinowski ST, Manlove KR, Taper ML: *ONCOR: a Computer Program for Genetic Stock Identification*, Volume 2. Bozeman, MT: Department of Ecology, Montana State University; 2008.
70. Ovenden JR, Peel D, Street R, Courtney AJ, Hoyle SD, Peel SL, Podlich H: The genetic effective and adult census size of an Australian population of tiger prawns (*Penaeus esculentus*). *Mol Ecol* 2007, **16**:127–138.
71. R Core Team: *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2012.
72. Nei M, Tajima F, Tateno Y: Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *J Mol Evol* 1983, **19**:153–170.
73. Earl DA: Structure Harvester v0.3. 2009. Available at <http://taylor0.biology.ucla.edu/structureHarvester/>.

doi:10.1186/1741-7007-12-1

Cite this article as: Hamilton et al.: Populations of a cyprinid fish are self-sustaining despite widespread feminization of males. *BMC Biology* 2014 **12**:1.