Received: 23 October 2018 | Revised: 19 March 2020 | Accepted: 8 April 2020

DOI: 10.1111/1755-0998.13166

RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES WILEY

A newly developed genetic sex marker and its application to understanding chemically induced feminisation in roach (*Rutilus rutilus*)

Anke Lange[1] | Josephine R. Paris[1] | Karim Gharbi[2,3] | Timothée Cézard[2] | Shinichi Miyagawa[4] | Taisen Iguchi[5] | David J. Studholme[1] | Charles R. Tyler[1]

[1]Biosciences, College of Life & Environmental Sciences, University of Exeter, Exeter, UK

[2]Edinburgh Genomics, School of Biological Sciences, University of Edinburgh, Edinburgh, UK

[3]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

[4]Department of Biological Science and Technology, Faculty of Industrial Science and Technology, Tokyo University of Science, Tokyo, Japan

[5]Graduate School of Nanobioscience, Yokohama City University, Yokohama, Japan

Correspondence
Anke Lange, Biosciences, College of Life & Environmental Sciences, University of Exeter, Exeter EX4 4QD, UK.
Email: A.Lange@exeter.ac.uk

Present address
Karim Gharbi, Earlham Institute, Norwich, UK
Timothée Cézard, The European Bioinformatics Institute (EMBL-EBI), Hinxton, UK

Funding information
Department for Environment, Food and Rural Affairs; Natural Environment Research Council, Grant/Award Number: NBAF608 and NE/K004263/1; University of Exeter

Abstract

Oestrogenic wastewater treatment works (WwTW) effluents discharged into UK rivers have been shown to affect sexual development, including inducing intersex, in wild roach (*Rutilus rutilus*). This can result in a reduced breeding capability with potential population level impacts. In the absence of a sex probe for roach it has not been possible to confirm whether intersex fish in the wild arise from genetic males or females, or whether sex reversal occurs in the wild, as this condition can be induced experimentally in controlled exposures to WwTW effluents and a steroidal oestrogen. Using restriction site-associated DNA sequencing (RAD-seq), we identified a candidate for a genetic sex marker and validated this marker as a sex probe through PCR analyses of samples from wild roach populations from nonpolluted rivers. We also applied the sex marker to samples from roach exposed experimentally to oestrogen and oestrogenic effluents to confirm suspected phenotypic sex reversal from males to females in some treatments, and also that sex-reversed males are able to breed as females. We then show, unequivocally, that intersex in wild roach populations results from feminisation of males, but find no strong evidence for complete sex reversal in wild roach at river sites contaminated with oestrogens. The discovered marker has utility for studies in roach on chemical effects, wild stock assessments, and reducing the number of fish used where only one sex is required for experimentation. Furthermore, we show that the marker can be applied nondestructively using a fin clip or skin swab, with animal welfare benefits.

KEYWORDS

endocrine disruption, intersex, RAD-seq, Rutilus rutilus, sex marker, sex reversal

1 | INTRODUCTION

In mammals, sexual determination is under genetic control whereas in fish, sex determination and sexual differentiation are the result of a delicate interplay of genetic and environmental factors (reviewed in Baroiller, Guiguen, & Fostier, 1999; Baroiller & D'Cotta, 2001; Heule, Salzburger, & Böhne, 2014). Any imbalance caused by exogenous substances, including endocrine disrupting compounds

(EDCs), can ultimately impact sex assignment, even in gonochoristic (single-sexed) species (Jalabert et al., 2000; Nakamura, Kobayashi, Chang, & Nagahama, 1998; Patino, 1997; Strüssmann & Nakamura, 2002). Oestrogens play key roles in sexual differentiation and gametogenesis, and exposure to oestrogens or oestrogen-mimicking chemicals during critical periods of differentiation has been shown to disrupt sexual development in fish (reviewed in Scholz & Klüver, 2009). Exposure to very low (ng/L) concentrations of the synthetic steroidal oestrogen 17α-ethinyloestradiol (EE2) has been shown to result in the feminisation of male fish (reviewed in Scholz & Klüver, 2009). These feminised phenotypes include intersex, characterised by the presence of developing oocytes and/or female reproductive ducts (oviducts) in the testes of otherwise male fish (Nolan, Jobling, Brighty, Sumpter, & Tyler, 2001). Globally, feminised responses, including intersex, have been reported for freshwater fish living in rivers polluted with endocrine disrupting chemicals and in more heavily contaminated rivers there have even been reports of populations with a sex ratio skewed towards females (Bengu, du Plessis, Modley, & van Dyk, 2017; Bernet et al., 2008; Bjerregaard, Korsgaard, & Bjerregaard, 2006; Hashimoto et al., 2000; Jobling, Nolan, Tyler, Brighty, & Sumpter, 1998; van Aerle et al., 2001; Vajda et al., 2008; Zheng, Liu, Liu, Jin, & An, 2015).

Over the past two decades, the roach (*Rutilus rutilus*; Figure 1) has been widely adopted in central and Northern Europe as a species for studies into endocrine disruption (Burkhardt-Holm, Peter, & Segner, 2002; Gerbron et al., 2014; Jobling et al., 1998; McGee, Brougham, Roche, & Fogarty, 2012; Minier, Caltot, Leboulanger, & Hill, 2000; Noaksson, Tjärnlund, Bosveld, & Balk, 2001; Trubiroha et al., 2010; Wiklund, Lounasheimo, Lom, & Bylund, 1996). These studies have been aided by the fact that the roach is a gonochorist (developing as either a male or female, thus avoiding complications where sex changes occur as part of the natural process of sexual development), a good understanding of the normal reproductive development (Paull, Lange, Henshaw, & Tyler, 2008), and the availability of analytical tools for assessing oestrogenic effects in this species (Hamilton & Tyler, 2008; Harris et al., 2011; Lange et al., 2008, 2009; Tyler, van der Eerden, Jobling, Panter, & Sumpter, 1996; Tyler et al., 2009). Surveys of wild roach populations in UK rivers contaminated with wastewater treatment works (WwTW) effluent have shown high incidences of intersex (Enivornment Agency, 2004; Jobling et al., 1998) and moderately to severely intersex fish have reduced milt volume and sperm density, and reduced fertility (Jobling, Beresford, et al., 2002; Jobling, Coey, et al., 2002) which might have population level consequences for roach in some UK rivers (Harris et al., 2011). These feminised phenotypes observed in wild roach populations can be induced through controlled chronic exposures to both EE2 or to a treated WwTW effluent. Furthermore, exposure to both 4 ng EE2/L or to a full strength treated WwTW effluent have been shown to result in phenotypically all-female populations (Lange et al., 2009; Lange, Paull, Hamilton, Iguchi, & Tyler, 2011), although that sex reversal could not be confirmed in the absence of a genetic sex probe. The ability to determine the fish's genetic sex is crucial for assessing whether fish in wild populations undergo sex reversal as a consequence of pollution exposures and understanding whether sex reversed fish reproduce in this ecologically important cyprinid species.

Despite the absence of heteromorphic sex chromosomes in many fish species, genetic sex determination has been found in a wide variety of fish species (Devlin & Nagahama, 2002). The best known example is DMY, a Y chromosome-specific duplicate of an autosomal gene called dmrt1 in medaka (*Oryzias latipes*) identified through two independent approaches, namely a positional cloning strategy and a

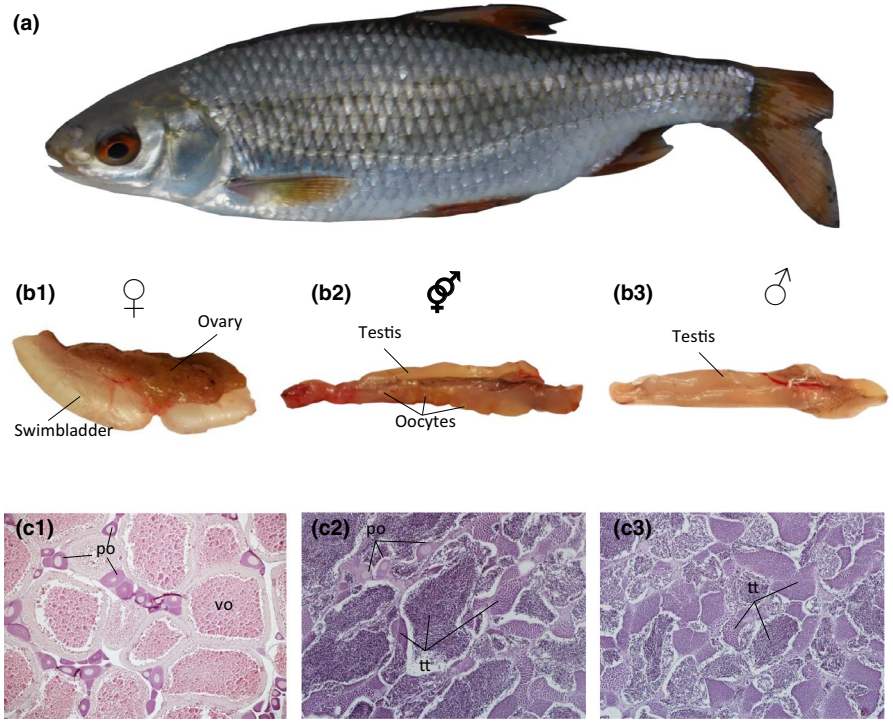

FIGURE 1 Image of an adult roach and macroscopic and histological images of roach gonads. Adult roach (a), macroscopic images of a roach ovary (b1), an intersex gonad (b2),a testis (b3) and histological sections of a roach ovary (c1), an intersex gonad (c2) and a testis (c3). po, primary oocyte; tt, testicular tissue; vo, vitellogenic. Photograph credit: Dr Gregory Paull (adult roach; a) and Dr Alice Baynes (macroscopic gonad images (b1–3)

candidate gene approach (Matsuda et al., 2002; Nanda et al., 2002). In the past, various molecular methods have led to the isolation of sex-specific markers in various fish species, including three-spined stickleback (*Gasterosteus aculeatus*), fathead minnow (*Pimephales promelas*), half smooth tongue sole (*Cynoglossus semilaevis*), Nile tilapia (*Oreochromis niloticus*), African catfish (*Clarias gariepinus*) and common carp (*Cyprinus carpio*) (Chen, Du, Yue, Dang, & Chang, 2010; Chen et al., 2009; Chen et al., 2007; Ezaz et al., 2004; Griffiths, Orr, Adam, & Barber, 2000; Kovács, Egedi, Bartfai, & Orban, 2000; Olmstead et al., 2011). In other fish species, including the three-spined stickleback, sex-linked allozyme markers have been identified (Allendorf, Gellman, & Thorgaard, 1994; Liu, Goudie, Simco, & Davis, 1996; Volff et al., 2013; Withler, McPhail, & Devlin, 1986), but in these cases, sex linkage could be a consequence of sex-specific gene expression of autosomal genes (Devlin & Nagahama, 2002).

The development of restriction-site-associated DNA sequencing (RAD-seq) (Miller, Dunham, Amores, Cresko, & Johnson, 2007) has made the discovery of genetic markers in wild populations of non-model species more readily attainable. Application of RAD-seq has resulted in the discovery of informative polymorphic markers and in the construction of linkage maps for a variety of fish species (Houston et al., 2014; Leitwein et al., 2017; Manousaki et al., 2016; Willing, Hoffmann, Klein, Weigel, & Dreyer, 2011). RAD-seq has also proven powerful in the identification of single nucleotide polymorphisms (SNPs) and genetic sex markers (Brown et al., 2016; Carmichael et al., 2013; Etter, Preston, Bassham, Cresko, & Johnson, 2011; Fowler & Buonaccorsi, 2016; Gamble et al., 2015, 2017; Gamble & Zarkower, 2014; Hohenlohe et al., 2010; Mathers et al., 2015; Palaiokostas, Bekaert, Davie, et al., 2013; Palaiokostas, Bekaert, Khan, et al., 2013; Palaiokostas et al., 2015).

Here, we used RAD-seq to isolate a male-specific genetic sex marker in roach and verified this using phenotypically assigned males and females derived from clean water study sites. The sex marker was then applied to a series of samples firstly, to confirm male to female sex reversal in roach for controlled chronic exposures to EE2 and treated WwTW effluents. We then applied the sex probe to confirm that induced intersex in wild roach populations arises due to feminisation of genetic males, and to assess evidence for complete sex reversal in wild roach populations of roach inhabiting rivers heavily contaminated with oestrogenic WwTW effluents. We also sought to investigate the use of the genetic marker to sex roach prior to the appearance of a histologically sex differentiated gonad and to sex roach with non-destructive tissue (fin clip and skin swab) sampling methods. Our results illustrate the wide-ranging applications of the sex probe for studies on roach sex and the effects of chemicals (here oestrogens) on roach populations.

2 | MATERIALS AND METHODS

2.1 | Fish populations and sampling

For RAD-seq and subsequent analysis for sex marker isolation, fish were sampled from two independent sites (Table 1) in the United Kingdom with no known pollution inputs. The first population was collected from a lake in Yorkshire (site 1), in northern England and

TABLE 1 Summary of roach populations from which samples were derived for this study

Population	Purpose	Origin	Pollution history	Feminisation/Intersex[a]	Reference
1	RAD-seq	Field	None known	None	Sampled for this study
2	RAD-seq	Lab[b]	None known	None	Hamilton et al. (2015) – controls only
A	Validation	Field	None known	None	P. B. Hamilton, A. L. Baynes, J. R. Stevens, S. Jobling, & C. R.Tyler (personal communication, October 29, 2016)
B	Validation	Field	None known	None	Hamilton et al. (personal communication)
C	Validation	Field	None known	None	David et al. (2017)
D	Validation	Field	None known	None	Unpublished
E	Validation	Field	None known	None	Defra (2009), Harris et al. (2011)
F	Application	Field	Effluent impacted	41% intersex	Defra (2009), Harris et al. (2011)
G	Application	Field	Effluent impacted	13% intersex	Defra (2009)
H	Application	Field	Effluent impacted	3% intersex	Defra (2009)
I	Application	Field	Effluent impacted	39% intersex	Defra (2009), Harris et al. (2011)
EE2	Application	Lab[b]	Controlled exposure	100% phenotypic females	Lange et al. (2009)
Effluent	Application	Lab[b]	Controlled exposure	100% phenotypic females[c]	Lange et al. (2011)
Early life	Application	Lab[b]	None	None	Sampled for this study

[a]Defined as having oocytes in their testes.

[b]Fish were only one generation away from the wild and parental fish were all caught from the same location which was also the same as for population E.

[c]Some putative sex-reversed males appeared to have bred as females.

the second population were the offspring of parental fish that originated from the River Trent, in the Midlands (site 2) and that had been bred in the laboratory. One hundred males and 100 female roach were sampled from each population.

For validation and application of the sex marker, samples from a total of 429 roach were analysed: (i) Eighty for the initial validation of the isolated sex marker; (ii) Eighty-two fish to verify complete male to female sex reversal can occur for exposure to environmental oestrogens and; (iii) two hundred and sixty-seven fish to test that intersex roach in the wild derive from genetic males. These samples were derived from previous studies (David, Lange, Abdul-Sada, Tyler, & Hill, 2017; Defra, 2009; ; Harris et al., 2011; Lange et al., 2009, 2011) and details on the different roach populations used are provided below and in Table 1.

The fish used for validation of the sex marker (n = 80, 42 males and 38 females, based on gonadal phenotype) originated from roach populations collected at five independent field sites across England with no known pollution inputs (populations A–E, Table 1). Application of the sex probe to verify complete male to female sex reversal can occur after exposure to environmental oestrogens, was carried out on samples derived from 82 roach that had been chronically exposed to EE2 or a treated WwTW effluent and resulted in phenotypic all-female populations (Lange et al., 2009, 2011). In the exposure study to a WwTW effluent by Lange et al. (2011), a subsequent competitive breeding experiment that included parentage analysis on the offspring suggested that presumptive males in the effluent-exposed treatment reproduced as females, but this could not be confirmed at the time due to the lack of a genetic sex marker.

A further application of the sex probe involved testing the hypothesis that intersex roach in the wild derive from genetic males and to investigate for evidence for possible complete reversal in roach. To this end, the sex marker was applied to 267 (177 males and 90 females) samples collected at sites heavily contaminated with oestrogenic WwTW effluents. For this, samples were used from four independent roach populations that varied in the proportions and degrees of intersex (populations F–I, Table 1).

For each individual fish used in the study, a fin clip was collected from the caudal fin and preserved in ethanol until DNA extraction. All procedures employed were in accordance with UK Home Office regulations (Animals (Scientific Procedures) Act 1986).

2.2 | DNA extraction and RAD-seq library preparation

For RAD-seq library preparation, total DNA was extracted from individual fish using a small fraction of the preserved fin tissue using Qiagen DNeasy Blood & Tissue columns (Qiagen, Germany), with a few modifications of the manufacturer's protocol: Briefly, tissue lysis was carried out with proteinase K (20 µl of a 20 mg/ml stock) under gentle rotation at 37°C for 1 hr, followed by incubation with RNase A (4 µl of a 100 mg/ml stock) prior to loading onto the extraction column. DNA was eluted from the extraction columns with 100 µl of AE buffer. After DNA extraction, DNA and RNA concentrations were determined by fluorometry using a Qubit fluorometer. DNA quality was assessed spectrometrically with a NanoDrop 1,000 Spectrophotometer to measure the ratio of absorbance at 260 nm and 280 nm. DNA integrity was assessed by gel electrophoresis.

Four different DNA pools containing equimolar proportions of samples from each site and sex were prepared: pool 1 contained 42 females from site 1; pool 2 contained 42 males from site 1; pool 3 contained 45 females from site 2 and pool 4 contained 45 males from site 2. *Pst*I-digested RAD libraries were prepared for each pool following the study by Baird et al. (2008), using 20 units of *Pst*I-HF (New England BioLabs, UK) (recognition cut site 5'-CTGCA/G-3' and 3'-G/ACGTC-5') to digest 1 µg of DNA per pool. Digested DNA from each pool was ligated to a different P1-barcoded adapter (total 4 µg DNA per library) and then sheared to a target peak of 400 bp using a Covaris S2 sonicator (Covaris Ltd, UK). The sheared DNA was cleaned using the Nucleospin kit (Macherey-Nagel, Germany) and run on a 1.25% agarose gel. DNA fragments between 250 and 700 bp were excised from the gel. The libraries were blunt ended and A-tailed before purification with Agencourt AMPure XP magnetic beads (Beckman Coulter, UK) at a volume DNA:beads ratio of 1:0.8. P2 adapters were then ligated followed by a further bead clean-up at a volume DNA:beads ratio of 1:0.7. The four libraries were PCR amplified using 12 cycles. PCR-enriched libraries were purified with AMPure XP beads, normalised to 8 nM and pooled together for sequencing on one lane of an Illumina HiSeq 2000 flowcell (101 bp paired-end reads).

Validation and application of the sex marker were performed on DNA extracted from individual fin clips using the HotSHOT method (Truett et al., 2000). Briefly, a small section of fin tissue was incubated in 75 µl alkaline lysis reagent (25 mM NaOH, 0.2 mM Na_2EDTA) at 95°C for 45 min. The samples were placed on ice for 5 min before the adding 75 µl neutralising reagent (40 mM Tris-HCl, pH 5.0).

2.3 | RAD-seq identification of a male-specific marker and locus extension

The software Stacks version 0.99 (Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011) was used to assemble markers from the RAD-seq data. Raw reads were cleaned and demultiplexed using process radtags, with options to clean (-c), quality filter (-q) and rescue barcodes (-r). Demultiplexed clean reads were formed into polymorphic loci using denovo_map.pl with the minimum number of reads set to 2 (-m), number of mismatches set to 2 (-M) and the number of differences between the four DNA pools set to 4 (-n), using parameter optimisation as outlined elsewhere (Paris, Stevens, & Catchen, 2017). Catalogue consensus sequences built from the males and females from both sites were used for downstream coverage analysis.

In order to identify sex-specific markers, we undertook a read mapping coverage analysis of the RAD-seq catalogue consensus sequences. Clean, demultiplexed forward reads from the males and

females of both sites were mapped to the Stacks' catalogue consensus sequences (687,048 sequences, see Results) using bwa mem (Li, 2013). Resulting bam files were used as input for read coverage analysis using a custom script (RAD_coverage_read1.py; available at https://github.com/tcezard/RADmapper). The script takes each aligned RAD read and calculates per consensus coverage. As well as calculating raw read coverages, samtools (Li et al., 2009) was used to report read coverages after removal of potential duplicates. Output of the coverage script is a tab-separated file (coverage.tsv) detailing for each consensus tag: the consensus ID; total coverage across all samples; total coverage after duplicate filtering across all samples; the number of samples with read coverage > 2; and two columns per sample detailing the total coverage and the coverage without duplicate reads. The output coverage.tsv file was used to search for sex-specific markers. Male-specific markers were where females from both sites showed no mapped reads (site1 and site2 female columns = 0), and males showed greater or equal to five mapped reads (site1 and site2 columns => 5 reads). A read coverage threshold for males was applied in order to provide confidence in male-specificity. A read coverage cutoff of five as chosen specifically, as it represented more than double the number of reads for the average of each male population (site1 = 1.9× and site2 = 1.8×) after filtering for 0 read mappings for females. The analysis resulted in three markers being identified as male-specific (see Results). No female-specific markers were uncovered.

Given that Stacks version 0.99 was used, the ability to assemble paired-end read contigs within the Stacks software was not possible at the time of analysis. Therefore, we used a different, but widely adopted approach for assembling contigs from the paired-end read data. We did this in the following way: The reverse reads (read 2) were aligned to the catalogue consensus sequences, and were grouped using the read 1 alignment bam files. In order to assemble long and contiguous contigs that could be used to design primers, several different assembler software were used (https://github.com/tcezard/RADmapper/blob/master/bin/RAD_assemble_read2.py) and assessed (https://github.com/tcezard/RADmapper/blob/master/bin/RAD_assess_read2_contig.py). The software IDBA-UD version 1.0.9 (Peng, Leung, Yiu, & Chin, 2012), assembled the longest contigs from the paired-end reads, using the following parameters: min_contig = 200, --mink = 40, -- min_count = 8 and --min_support = 4).

The paired-end contig from each of the three identified markers were used downstream for PCR probe design. All markers were tested as candidates for sex-specific genetic markers using PCR (see Section 2.4.1., Figures S1 and S2). The paired-end contig of Marker_780797 successfully distinguished between male and female roach (see Results, Figure 2 and Figure S2).

As the read-mapping method relies on the identification of restriction site-associated presence-absence polymorphisms, we explored whether the identified sex marker was more likely to be the result of high sequence divergence between males and females, or a male-specific insertion-deletion polymorphism. If the presence-absence site was the result of the former, we might expect other contigs in the data set to show high sequence divergence. We assessed this by performing an allele frequency divergence analysis. The allele frequencies of all contigs were calculated (RAD_allele_frequency_count.py), which reports the number of ATCG nucleotides if coverage Q20 is > 5. The allele frequencies were filtered so that major alleles were reported as different between males and females at a major allele frequency greater than 0.8. Analysis of potential high sequence divergence between male and female haplotypes was also assessed by *k*mersing the male marker contig and aligning the resulting *k*mers to the female-derived catalogue consensus tags. The contig was *k*merised using jellyfish count version 1.1.4 (Marçais & Kingsford, 2011), by *k*merising it to word lengths of 11, 15, 21, 31. Resulting *k*mers were aligned to female consensus tags using bwa mem (Li, 2013). Any identified alignments were further assessed using mafft online version 7 (Katoh & Standley, 2013).

Finally, in order to identify a putative function of the male-specific region, all reads with a match against the putative sex marker sequence were extracted from the RAD-seq data set, including both members of the paired-end read (even if only one member of the pair actually matched the target sequence), and pairs of reads were then assembled using SPAdes version 3.6.2 (Bankevich et al., 2012). The resulting scaffolds were analysed for the presence of the original target sequence. A BLASTN search against the NCBI Nucleotide database was used to map scaffolds containing the original sequence to published sequences. Primers were designed at the outer ends of a newly assembled scaffold to experimentally confirm the sequences (Table 2; Figure S3).

TABLE 2 PCR primers used to amplify sex-specific marker in roach

		Amplicon size with			
Primer	**Sequence (5'–3')**	**Rr_780797_r1**	**Rr_780797_r2**	**Rr_780797_r3**	**Rr_sml_r1**
Rr_780797_f1	AGGGGCACCATGTGAAAATCC	247 bp	381 bp	–	–
Rr_780797_f2	AGAGATGTCTGGAGTTATATAGGGG	–	–	400 bp	–
Rr_780797_r1	TATGCCTCCTCCCAGCACAA	–	–	–	–
Rr_780797_r2	ACAGCCTTATAGTTGCTTGCTC	–	–	–	–
Rr_780797_r3	CAGCCTTATAGTTGCTTGCTCC	–	–	–	–
Rr_sml_f1	TGACGAACCATAACCCATTGTG	–	–	–	1,457 bp
Rr_sml_r1	GCTTTCCATCTTGCTTTCTTGC	–	–	–	–

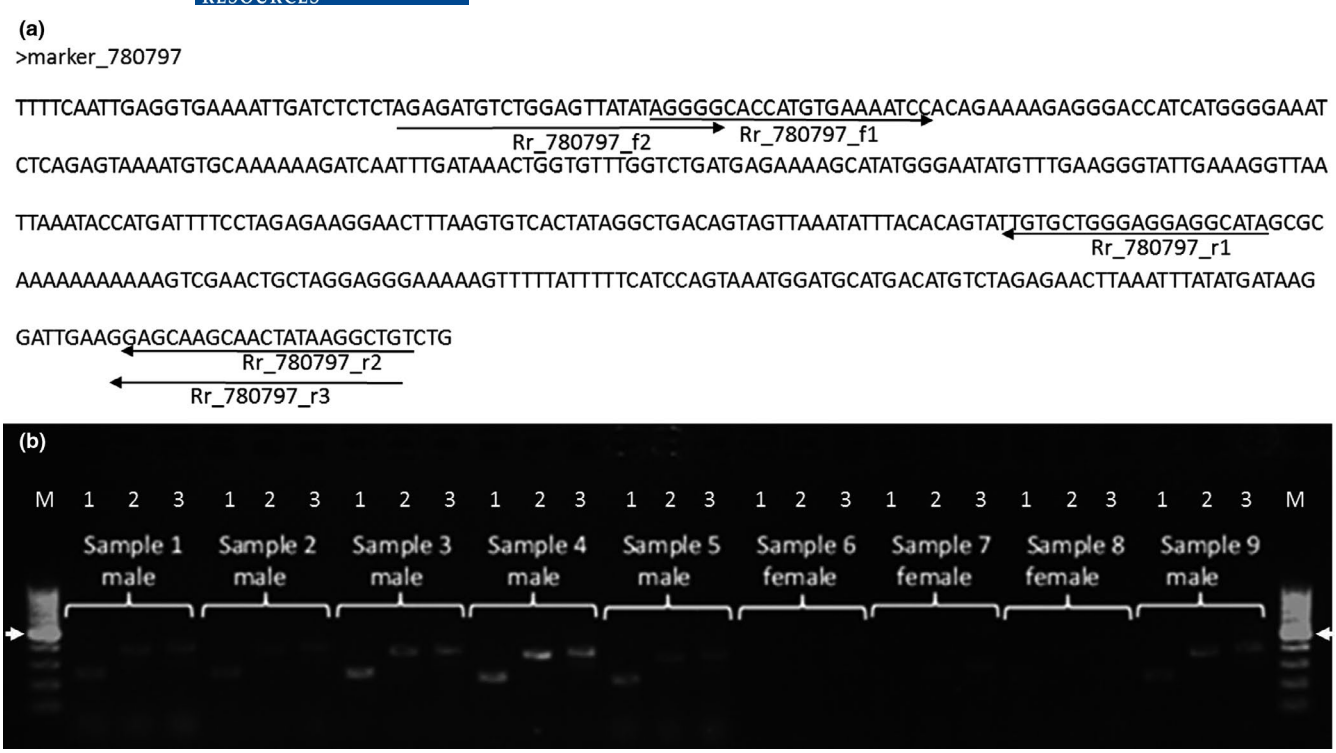

FIGURE 2 Putative sex-specific marker identified by RAD-seq and its validation by PCR. (a) Sequence (434 bp) of the male-specific sex marker (marker_780797) distinguishing between male and female roach. Arrows show the annealing sites of the used PCR primers. (b) PCR validation of the sex-specific marker, present in males and absent in females. For each fish, DNA was analysed in three separate PCR reactions using three different primer combinations (see methods and Table 2) in order to avoid erroneous assignment of genetic sex determination due to failed PCR reactions in case of polymorphisms. Expected fragment sizes were 247 bp (Rr_780797_f1/r1; lane 1 for each fish), 381 bp (Rr_780797_f1/r2, lane 2 for each fish) and 400 bp (Rr_780797_f2/r3; lane 3 for each fish). A 100 bp ladder was loaded into the outer lanes (M) with the arrow indicating the 500 bp band of the size marker

2.4 | Polymerase chain reactions

2.4.1 | Roach sex marker

In the first instance, all three markers were tested as candidates for sex-specific genetic markers using polymerase chain reaction (PCR) (see Supporting Information Methods and Figure S2). Using Primer-BLAST (Ye et al., 2012), primers were designed to the sequences identified as present in males and absent in females (see Section 2.3., Supporting Information Methods and Figures S1 and S2). Marker_780797 successfully distinguished between male and female roach (see Results, Figure 2, Figures S1 and S2) and this was applied to all subsequent analyses/studies. Five different primers (two sense and three antisense primers, Table 2) were then designed for the marker_780797 which successfully distinguished between male and female roach (see Results and Figure 2 and Figure S2).

Samples from each fish were analysed in three separate PCR reactions using three different primer combinations of two sense and three antisense primers (Table 2). If amplification was successful for at least two out of three PCRs reactions, the individual was assigned as a genetic male. This approach was chosen to avoid erroneous assignment of genetic (female) sex due to failed PCR reactions in case of polymorphisms. PCR reactions were carried out using GoTaq Flexi DNA Polymerase (Promega, UK), 1.5 mM $MgCl_2$, 0.2 mM dNTP mix (Thermo Scientific, UK), 0.2 μM of each forward and reverse primer (Eurofins Genomics, Germany) and 2 μl DNA in a total volume of 20 μl. An initial denaturing step at 95°C for 5 min was followed by 30 cycles of denaturation (1 min at 95°C), annealing (30 s at 56°C) and extension (45 s at 72°C), followed by a final extension of 5 min at 72°C. Amplicons were resolved on 1.5% agarose gels. PCR bands were scored blindly of the donor fish and only after scoring was completed were the phenotypic and genetic sex results combined and compared.

2.4.2 | Internal transcribed spacer

Roach can breed (thus hybridise) with other cyprinid fish species including bream (*Abramis brama*) and rudd (*Scardinius erytrophthalmus*) (Pitts, 1994). Therefore, all individuals used in this study (those used for the RAD-library pools and those for the validation and application approach) were verified as genetically pure-bred roach using species-specific forward primers and an universal cyprinid reverse primer designed to the ITS1 nuclear ribosomal DNA region according to Wyatt, Pitts and Butlin (2006). In addition, this amplification also served as positive control for successful DNA extraction.

2.4.3 | Locus extension

Primers were designed at the outer ends of the newly assembled scaffold to experimentally confirm the sequences (Table 2; Figure S3). PCR reactions were carried out using Q5 High-Fidelity 2X Master Mix (New England BioLabs), 0.5 μM of each forward (Rr_sml_f1, Table 2) and reverse (Rr_sml_r1, Table 2) primer (Eurofins Genomics) and 2.5 μl DNA in a total volume of 50 μl. An initial denaturing step at 98°C for 30s was followed by 35 cycles of denaturation (5 s at 98°C), annealing (20 s at 66°C) and extension (45 s at 72°C), followed by a final extension of 2 min at 72°C. Amplicons were resolved on 1.5% agarose gels. PCR products were subsequently purified using the Nucleospin kit (Macherey-Nagel) according to the manufacturer's instructions. Purified PCR products were sequenced (Eurofins Genomics), the resulting sequences cleaned of low quality ends and aligned to the assembled scaffold.

2.5 | Genetic sex identification prior to histological gonadal differentiation

Roach fry at 54 days post hatch (dph) were obtained from the Environment Agency Fish Farm Calverton (UK) and individually kept in 11 L tanks, maintained under flow through conditions at 18 ± 1°C with a photoperiod regime of 16 hr:8 hr light:dark. At 70 dph, fry were anesthetised with 50 mg/L Ethyl-p-aminobenzoate, a small section of the caudal fin taken and preserved in ethanol for identifying genetic sex. The fish were then allowed to recover from the anaesthesia before being placed back into their growth tanks where they were kept individually until termination. Fish were provisioned with dietary requirements according to their age (Paull et al., 2008). Briefly, roach were fed three times daily until satiation with freshly hatched Artemia nauplii (ZM Premium Grade Artemia; ZM Ltd., UK) and at all life stages, the diet was supplemented with Cyprico Crumble EX dry food (Coppens International bv, The Netherlands).

After five months, fish were terminated by lethal anaesthesia using ethyl-p-aminobenzoate in accordance with UK Home Office regulations (Animals (Scientific Procedures) Act 1986). Fork length and wet weight were recorded to the nearest 1 mm and 0.01 g, respectively. For identifying genetic sex, both a fin clip and skin swab (taken with sterile cotton tips; Technical Service Consultants Ltd, UK) were collected from each fish and preserved in ethanol or stored at −20°C, respectively. The remainder of each body was preserved in Bouin's fixative for subsequent determination of phenotypic sex via gonad histopathology. Histological processing and determination of phenotypic sex was carried out as described previously (Paull et al., 2008).

DNA was extracted from fin clips and mucus swabs using the HotSHOT method or Chelex extraction protocol (Estoup, Largiadèr, Perrot, & Chourrout, 1996), respectively. All samples were genetically identified as pure roach origin using the ITS1 nuclear ribosomal DNA region according to Wyatt et al. (2006). For the early life stage fish, genetic sex of each fish was determined by PCR using the Multiplex PCR Kit (Qiagen) with individual reactions for each primer combination described above for marker_780797. The PCR reaction mix included the kit's Q-solution, 0.2 μM of each primer and 2 μl DNA in a total reaction volume of 20 μl. An initial activation step of 95°C for 15 min was followed by 45 cycles of denaturation (30 s at 94°C), annealing (90 s at 60°C) and extension (1 min at 72°C), followed by a final extension of 10 min at 72°C. Amplicons were resolved on 1.5% agarose gels. For the later roach stage fin clips and mucus swabs, PCRs were performed as described above (Section 2.4.1).

3 | RESULTS

3.1 | Restriction-site associated DNA sequencing (RAD-seq) and bioinformatical identification of putative sex markers

Sequencing of the *Pst*I RAD libraries produced between 30,435,402 and 37,086,239 raw reads, of which between 24,950,577 and 31,846,048 were retained after cleaning (Table S1). De novo assembly of the first read was achieved using Stacks (Catchen et al., 2011), resulting in 737,393 consensus RAD-tags, for which coverage varied from 14x to 17x (Table S1).

To detect presence-absence sex markers, we undertook a read mapping analysis and selected sites where five or more reads mapped in the male sex for both sites, and no reads mapped in either of the female pools. This identified three male-specific contigs: marker_780797 (site 1 coverage: 16; site 2 coverage: 15; contig length 434 bp); marker_808205 (site 1 coverage: 14; site 2 coverage: 9; contig length 160 bp); marker_815983 (site 1 coverage: 7; site 2 coverage: 15; contig length 400 bp) (Table S2). The associated paired-end reads were assembled into contigs using IDBA_UD, which successfully assembled 687,048 paired-end contigs. The paired-end contigs resulting from all three markers were tested as candidates for sex-specific genetic markers using PCR (Figure S2) and marker_780797 successfully distinguished between male and female roach (Figure 2 and Figure S2).

In an attempt to disentangle the origin of the presence-absence polymorphism, two additional analyses were performed. Analysis of allele-frequency divergence between males and females on other contigs (using a major allele frequency cutoff of 0.8), did not result in the recovery of any other sex-linked contigs. We also *k*merised the male marker and aligned it to the female consensus catalogue tags. *K*mers of length 11,15 and 21 did not align to any female consensus tags, and although word length 31 did show alignment to two female consensus tags, further inspection of this alignment using mafft, suggested that the alignment does not represent true sequence homology (Table S3). We therefore hypothesise that the identified sex marker is probably the result of a male-specific insertion-deletion polymorphism, rather than the result of high haplotype divergence between the sexes (although further research would be required to confirm this; see Discussion).

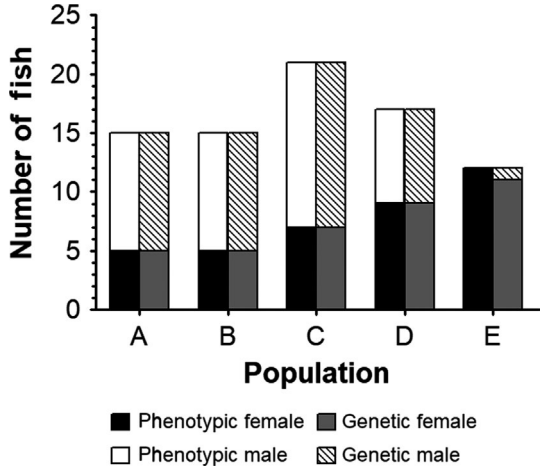

FIGURE 3 Validation of the sex-specific marker. The genetic sex of 80 roach collected from five independent, unpolluted river sites across England (populations A–E; see Table 1) identified by PCR. The phenotypic sex of these fish had been determined previously via gonadal histopathology, or based on the release of gametes in mature fish. Each group of stacked bars shows phenotypic sex (left hand bars) and the identified genetic sex (right hand bars) for each population; the lower segment of each bar corresponds to female fish and the upper segments male fish

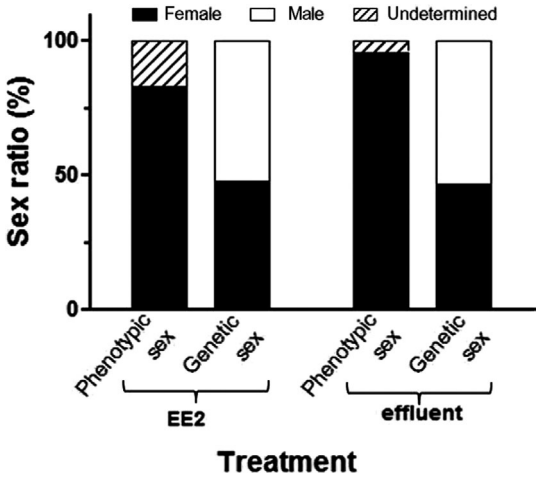

FIGURE 4 Identification of genetic males in phenotypic all-female roach populations. Genetic sex of fish derived from two controlled long-term laboratory exposures to EE2 or effluent (Lange et al., 2009, 2011). The majority of fish were phenotypic females with a small proportion of fish where sex could not be assigned by gonadal histopathology

3.2 | Validation of the sex-specific marker

Seventy-nine of the 80 roach (corresponding to 99%) used to validate the sex probe were identified correctly according to their gonadal sex, suggesting that the sex marker distinguishes genetic sex in these roach populations with high fidelity (Figure 3).

3.3 | Verification of complete male to female sex reversal for oestrogen exposed fish

The marker proved complete male-to-female sex reversal occurred in samples analysed from experiments for controlled long-term exposures to EE2 (4 ng/L) and to a full strength WwTW effluent. These studies used roach from independent populations and both oestrogenic exposures resulted in phenotypic all-female populations (based on gonadal histopathology), whereas the sex marker identified genetic male:female ratios of 52%:48% and 53%:47% for the EE2 and effluent experiment, respectively (Figure 4).

3.4 | Identifying genetic sex in populations with intersex fish

The sex marker was subsequently applied to identify the genetic sex of 267 (177 males and 90 females based on the histological gonadal phenotype) fish collected from four wild roach populations (with varying proportions and degrees of intersex) and assigned 172 of the 177 males (based on gonadal histopathology) as genetic males, a 97% agreement (Figure 5). These phenotypic males included 51 individuals with ovotestis and of these fish, 50 were confirmed as genetic males (Figure 5), proving that intersex fish in the wild derive from genetic males. There was no evidence for complete sex reversal in these fish populations. Of the 90 phenotypic females analysed from the wild populations, 85 were identified as genetic females (Figure 5). Overall, the genetic sex of 257 of these 267 fish (corresponding to 96%) aligned with the phenotypic (histologically derived) sex.

3.5 | Determining genetic sex prior to gonadal sex differentiation

Of the 44 roach that were sexed via histopathology at seven months old, 19 were phenotypic males and 25 phenotypic females and the genetic sex determined agreed with these phenotypic sexes for all females and for 14 of the 19 of the males. Thus, five phenotypic males were identified as genetic females. All fish were kept individually to allow us to trace each fish. For the same individuals at 70 dph, prior to gonadal differentiation, all 25 phenotypic females were correctly assigned with the genetic sex probe, but six of the 19 males were identified as genetic females (Figure 6).

3.6 | Extending the probe sequence

Extracting all reads with a match against the sex marker sequence from the full data set yielded 2,308 pairs of reads. The subsequent assembly resulted in 53 scaffolds, one of which contained the original target sequence. The length of this scaffold was 1,514 bp, around three times the length of the original sequence. A BLASTN

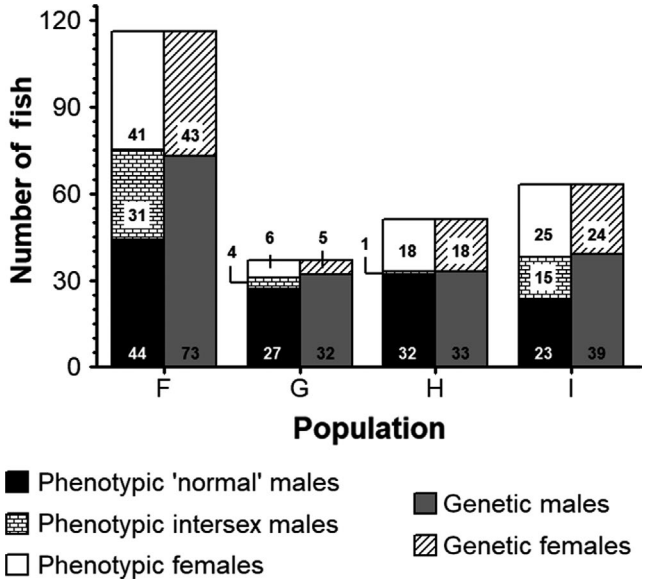

FIGURE 5 Application of the genetic sex marker to wild roach populations with varying levels of intersex condition. The genetic sex of 267 fish from four independent roach populations (populations F–I; see Table 1) with varying proportions of intersex males was identified by PCR. For these fish, gonadal phenotypic sex had been determined previously via histopathology, or via release of their gametes in male fish (Defra, 2009; ; Harris et al., 2011). Each group of stacked bars shows phenotypic sex (left hand bars) and the identified genetic sex (right hand bars) for each population; the lower segment of each bar (in case of phenotypic sex also the middle segment) represents male fish and the upper segments female fish. Numbers indicate the number of fish represented in each segment

search against the NCBI Nucleotide database mapped this scaffold to GenBank accession EU621898.1, representing a 300 kb region of the Atlantic salmon (*Salmo salar*) genome that contains the gene for growth hormone 1 (GenBank: ACD37713.1) among other protein-coding genes (von Schalburg et al., 2008). Primers designed to the outer ends of this newly assembled scaffold amplified a PCR product of the expected size in males, but did not yield a PCR product in females. The PCR products were subsequently sequenced. After trimming low quality ends, 1,332 bp of the experimental sequence aligned to the assembled scaffold with 100% identity (Figure S3).

4 | DISCUSSION

Globally, feminised responses, including intersex, have been reported for freshwater fish living in rivers containing EDCs with reports of sex ratio skewed towards females for populations in some of the more polluted rivers. Extensive research conducted on roach in UK rivers has established this species as a sentinel for endocrine disruption in wild freshwater fish populations and it has since been adopted widely for this purpose across central and Northern Europe. To date, however, the lack of a genetic sex marker has prevented answering the question on whether wild intersex roach arise as a consequence of the demasculisation (oestrogenisation) of males or the masculinisation (androgenisation) of females and whether female-biases observed in some wild roach populations are a consequence of sex-reversal in males. This lack of a genetic probe has thus been a limiting factor for studies into the population level consequences of sexual disruption in roach. Using RAD-seq, we isolated and subsequently validated a genetic sex-specific marker for roach, and applied it using PCR to samples derived from previous studies on chemical and effluent exposures and wild populations to confirm genetic sex in sexually disrupted fish and to sex fish prior to gonadal sex differentiation.

RAD-seq has been applied recently in the discovery of sex markers in various species of fish, reptiles and crustaceans (Brown et al., 2016; Carmichael et al., 2013; Fowler & Buonaccorsi, 2016; Gamble et al., 2015, 2017; Gamble & Zarkower, 2014; Mathers et al., 2015; Palaiokostas, Bekaert, Davie, et al., 2013; Palaiokostas, Bekaert, Khan, et al., 2013; Palaiokostas et al., 2015) and in some cases, these studies have been able to construct linkage maps based on single nucleotide polymorphisms (SNPs) and map a major sex determining locus to certain linkage groups. In our study, we have been able to identify presence-absence markers pertaining to sex only, and so more extensive sequencing of the roach genome would be required to construct linkage maps and further characterise this marker.

We performed two additional analyses in order to provide insight into the potential origin of the RAD-seq derived male marker, i.e. whether the presence-absence marker is the result of high sequence divergence between the sexes, or is missing in females due to it being a male-specific insertion-polymorphism. As the coverage analysis was performed on read 1 (where the *Pst*I cut-site is), it is possible that a female-specific mutation in the cut-site, due to high sequence divergence between male and female haplotypes at this region, resulted in male-read recovery, and no female read mappings. As the PCR region sits in the paired-end contig, and the size-selection step targeted regions of 250–700 bp, the paired-end contig could reside some distance from the read 1. This means a divergent sex-specific haplotype of between 250–700 bp (or even larger) could indeed produce such a signal. An alternative hypothesis is the presence of an insertion-deletion polymorphism in which the female haplotype lacks both the restriction enzyme cut site and a proportion of the PCR region in the paired-end contig. The assessment of allele frequency divergence did not identify any sex-linked divergence in other contigs, and the *k*mer analysis provided further support that the identified sex-linked marker is truly male-specific (and is missing in females). Due to the lack of a reference genome for roach, a more comprehensive analysis of these two hypotheses is not possible, yet this represents an interesting area of future research.

Sex markers alone, however, can be indicative of the underlying sex determining system with male sex markers indicative of a XY system, whereas female sex markers are indicative of a ZW system (Fowler & Buonaccorsi, 2016). It is not known which sex chromosome system roach have, if indeed they have one, but the isolation of a sex marker present in males and absent in females suggests male

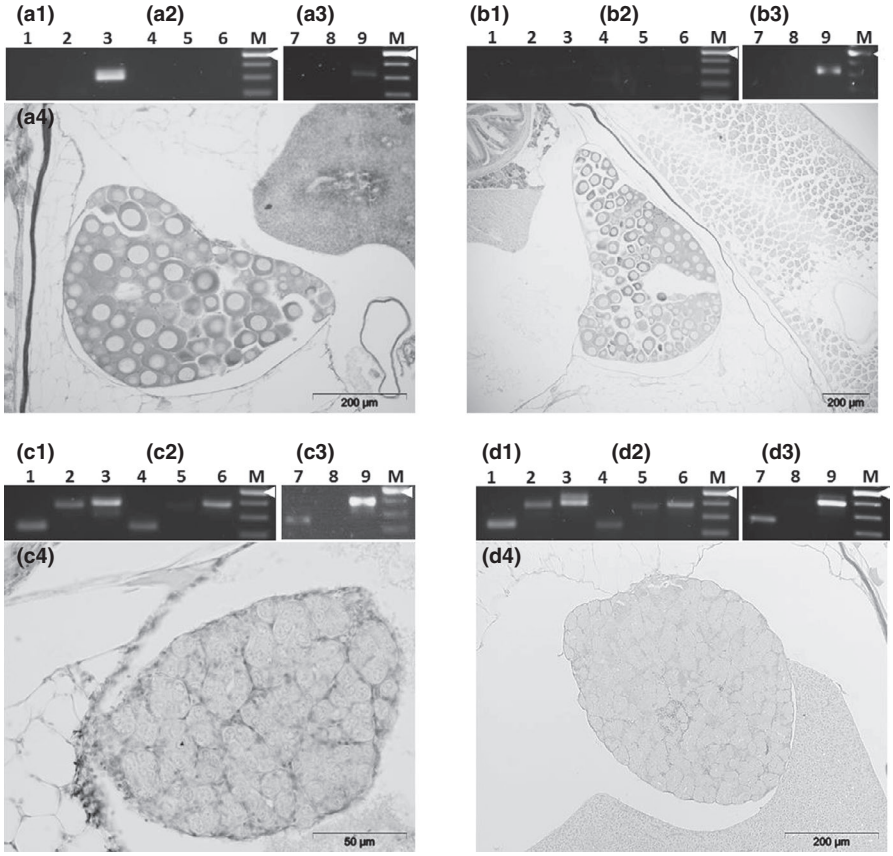

FIGURE 6 Identifying genetic sex in the same individual roach prior to, and after gonadal differentiation. Genetic sex of female (A & B) and male (C & D) roach was determined by PCR prior to (A.1–D.1, lanes 1–3) and after gonadal sex differentiation (fin clips: A.2–D.2, lanes 4–6 and skin swabs: A.3–D.3, lanes 7–9). After gonadal sex differentiation, gonadal phenotypic sex was determined by histology (A.4–D.4). Lanes 1, 4 and 7: PCR product Rr_780797_f1/r1 (247 bp); 2, 5 and 8: PCR product Rr_780797_f1/r2 (381 bp); 3, 6 and9: PCR product Rr_780797_f2/r3 (400 bp) and M: 100 bp ladder with arrow indicating the 500 bp band of the size marker

heterogamety and as such an XX/XY as opposed to the ZW/ZZ sex determination system. To date, only a few sex-determining genes have been identified in fish but male heterogamety seems to predominate, including in medaka (Matsuda et al., 2002), three-spined stickleback (Peichel et al., 2004), the cichlids *Pseudocrenilabrus philander* and *Astatotilapia burtoni* (Böhne et al., 2019; Böhne, Wilson, Postlethwait, & Salzburger, 2016) and salmonids (Woram et al., 2003). There is little evidence of a dominant female-determining sex-determination gene, i.e., female heterogamety (Pan et al., 2016). The exception to this is an indication that in zebrafish (wild populations), the closest species studied evolutionary to the roach, where there is a female heterogamety sex determination system (Tong, Hsu, & Chung, 2010; Wilson et al., 2014).

Our approach to obtain a population-independent sex marker for the roach was to sequence four genomic DNA pools created from equal numbers of individuals from two independent roach populations. Adopting this approach, one of the three markers identified proved to be specific for sex. Previous approaches to obtain sex markers have sequenced individuals rather than pool individuals (e.g. Fowler & Buonaccorsi, 2016; Gamble et al., 2015; Gamble & Zarkower, 2014) which then requires extensive bioinformatics to confirm the identification of many putative of sex-loci/specific markers, most of which turn out to be false-positives in the subsequent sex validation (Gamble & Zarkower, 2014). We acknowledge that potential genotyping and sex-identification error can also be introduced by pooling males and females, and the biases in pooled approaches, especially in allele frequency determination, is well recognised (Gautier et al., 2013; Rellstab, Zoller, Tedder, Gugerli, & Fischer, 2013). However, pooled-based approaches have shown great promise for delineating diversity in species lacking reference genomes (Kurland et al., 2019; Neethiraj, Hornett, Hill, & Wheat, 2017), and, moreover, pool-RAD-seq approaches have been used to successfully identify species-specific SNPs across a range of species (Delord et al., 2018).

The PCR approach developed to determine genetic sex of individuals is rapid and easily applied, but the reliance on a presence/absence marker means false assignment is possible due to a failure in the DNA extraction or PCR. In the roach this could lead to assigning genetic males as females. We circumvented this potential problem through the inclusion of a positive control (ITS1 nuclear ribosomal DNA region) to verify successful DNA extraction and through routinely running PCR assays with three different primer combinations for each sample. Yields on two PCRs in bands of the correct size were taken as a confident call on a genetic male.

Sex specificity of the RAD marker was validated based on the analysis of 80 individual roach (42 males and 38 females) from five different river sites for which there was 99% match between phenotypic (gonadal) and genetic sex. One phenotypic female was identified as being genetic male. As all of these roach were derived from the wild, sex-reversal could be a possible explanation for the observed mismatches between phenotypic and genetic sex, although this seems unlikely given that rivers sites from where they were derived are reported to have low levels of pollution. In the case for the phenotypic female, that was a genetic

male, such a sex-reversal is possible as this has been shown to occur in fish chronically exposed to low oestrogen levels including during the period of sexual differentiation (Kidd et al., 2007; Lange et al., 2009). Female-to-male sex-reversal on the other hand would imply complete masculinisation. This could result from exposure to androgens, but unlike that for specific cases in the USA for exposure to cattle feedlots containing high concentrations of the androgen trenbolone (Ankley et al., 2003), androgens are not present at concentrations in English rivers to cause such a phenotypic change in roach.

The isolated sex marker shows high fidelity for identifying sex-reversed individuals also based on our analysis of roach derived from two controlled, life-long exposures to oestrogens (a full strength WwTW effluent and 4 ng EE2/L) both of which resulted in all-female populations as determined by gonad histopathology (Lange et al., 2009, 2011). In both of these experiments, the genetic sex ratios of exposed fish were near 50:50 (52:48 and 53:47) male:female, respectively. The phenotypic sex ratios of the controls in these experiments were 69:31 (EE2) and 56:44 (effluent) (Lange et al., 2009, 2011). Previous studies on roach maintained in the laboratory have reported sex ratios of 50:50, but also male and female biased populations (Lange et al., 2009, 2011, 2015; Paull, Filby, & Tyler, 2009; Rodgers-Gray et al., 2001) which may reflect the fact that growth rates and/or social factors potentially affect sex assignment in this species (Paull et al., 2009). In wild roach populations evenly balanced as well as both male- or female-biased populations have been reported (Geraudie, Gerbron, & Minier, 2017; Jobling, Beresford, et al., 2002) and outside the spawning season, even the existence of monosex populations has been suggested (Tyler, Lange, Paull, Katsu, & Iguchi, 2007).

The use of the sex probe in these instances has proven that full male-to-female sex reversal can occur for exposure to a full strength WwTw effluent and to EE2 (4 ng/L) when they are exposed during the periods of sex differentiation and gonadal development supporting the possibility for this to occur in wild populations (but see later in the discussion). Dilution of WwTW effluents in UK rivers, however, averages 10%–30% and in only a few exceptional circumstances and during periods of low flow do they exceed 50% (Hamilton et al., 2014), albeit in exceptional circumstances the full flow of the river can be made up of treated wastewater effluent (Jobling et al., 1998). In the case of EE2, concentrations in the river reaches where roach live are most likely to be in the range of 0.15–1.14 ng/L (Williams, Churchley, Kanda, & Johnson, 2012), although concentrations up to 3.4 ng/L have been reported previously (Williams, Johnson, Smith, & Kanda, 2003). Application of the genetic sex marker to phenotypic female roach derived from the controlled WwTW effluent exposure and then subsequently subjected to a competitive breeding experiment furthermore identified that sex reversed genetic males spawned as females, producing viable offspring. These findings add a new and fascinating dimension for studies investigating potential impacts of exposure to WwTW effluents on roach (and potentially other fish) populations. Based on modelling approaches applied to investigate for the effect of compromised reproductive fitness, as a result of endocrine disruption, on other fish populations (An, Hu, Giesy, & Yang, 2009; Cotton & Wedekind, 2009) the levels of effect seen in roach would indicate this is unlikely for wild populations living in UK rivers. This should, however, not be discounted, as predictive modelling for EDCs has shown that it may be many generations of chemical exposure before fish populations are reduced substantially and such changes are difficult to monitor in the wild, particularly for long-lived species such as the roach (Hamilton et al., 2016).

In the studies on wild roach populations exposed to WwTW effluents with varying proportions and degrees of intersex males, we confirmed that the intersex condition occurs due to feminisation of males. For these study populations, the sex marker assigned ≥ 97% of phenotypic males (including intersex males) as genetic males and ≥ 94% of phenotypic females as genetic females. Of the intersex fish, 98% were assigned as genetic males. As mentioned above, some studies have reported female-biased populations (Geraudie et al., 2017; Jobling, Beresford, et al., 2002), leading to the question of possible complete male to female sex reversal. In our analyses of 128 phenotypic females, 90 of which were derived from river sites with oestrogenic contamination, only six registered as genetic males, strongly indicating that complete sex reversal in roach does not occur commonly, if at all, in roach in UK rivers as a consequence of pollution exposure.

By extending the initial sequence of the sex marker to around three times the original sequence length, the subsequent BLASTN search mapped the sequence to a 300 kb region of the Atlantic salmon (*Salmo salar*) genome containing a gene encoding growth hormone (GH) (von Schalburg et al., 2008). In fish, GH is involved in a variety of physiological processes, including the regulation of ionic and osmotic balance, lipid, protein and carbohydrate metabolism, skeletal and soft tissue growth, reproduction and immune function (Reinecke et al., 2005). Interestingly, salmonids are considered to have a male heterogamety sex determination system (Yano et al., 2013) and amongst several Y chromosome-specific gene markers that are located near the sex-determining gene, one is a GH pseudogene (GHp). This has been successfully applied to determine males of various Pacific salmon species including chum salmon (*Oncorhynchus keta*), pink salmon (*Oncorhynchus gorbuscha*), coho salmon (*Oncorhynchus kisutch*), Chinook salmon (*Oncorhynchus tshawytscha*) and masu salmon (*Oncorhynchus masou*) (Devlin, Biagi, & Smailus, 2001; Du, Devlin, & Hew, 1993; Micheletti & Narum, 2018; Zhang et al., 2001), whereas it did not produce sex-specific patterns in other *Oncorhynchus* salmonid species including sockeye salmon (*Oncorhynchus nerka*), rainbow trout (*Oncorhynchus mykiss*), or in Atlantic salmon (Devlin et al., 2001). Given that the GHp appears sex-specific for some Pacific salmon species only, the BLAST hit does not align with the coding region for Atlantic salmon GH, and salmonids are distant in evolutionary terms to cyprinids it is unlikely that there is a relationship between our isolated genetic sex marker and GH.

The finding that the sex marker can also be applied to fish at a life stage prior to gonadal sex differentiation and the analysis can be undertaken nondestructively from fin clips and even skin swabs gives the marker further considerable utility. For example, where

monosex populations are required for chemical exposures or for other laboratory biological investigations the sex probe can be applied to preselect sexed individuals, as immature or mature fish, prior to experimentation and thus halving the number of experimental animals used and thus addressing the principle of 3Rs. The same principle applies for surveys of wild roach populations (even as juveniles) where only one sex is required for subsequent termination and analysis. The ability to identify the genetic sex of roach non-destructively also has utility for studies on the wider ecology of the roach, including understanding of sex related behaviours that include fish movement and feeding patterns, and for appropriate sex balancing in the re-stocking of rivers and other water bodies for recreational fishing.

ACKNOWLEDGMENTS

DNA sequencing and bioinformatics analyses were performed by the GenePool genomics facility in the University of Edinburgh under a NERC Biomolecular Analysis Facility grant (NBAF608). We thank staff of the GenePool genomics facility at the University of Edinburgh for assistance with RAD library preparation and sequencing, Alan Henshaw and his team at Calverton Fish Farm (UK Environment Agency) for their support, Eliane Bastos for assistance with the DNA extractions for the libraries, Gregory Paull and the Aquatic Resources Centre technical team for support with the fish husbandry. We are grateful to Patrick Hamilton for advice throughout the study and Ronny van Aerle for initial bioinformatics suggestions. This work was funded on grants to CRT and DJS from the Natural Environmental Research Council (NE/K004263/1), the Department for the Environmental Food and Rural Affairs and the University of Exeter to CRT.

AUTHOR CONTRIBUTIONS

A.L., D.J.S. and C.R.T. conceived the project. A.L. performed the DNA extractions, PCRs and analysed the data. J.R.P. led the bioinformatics analyses with input from T.C., K.G. and D.J.S. A.L. and C.R.T. wrote the paper with contributions from J.R.P., K.G., T.C. and D.J.S. All authors contributed to the final version of the manuscript and gave their final approval.

DATA AVAILABILITY STATEMENT

RAD-seq reads are available from the European Nucleotide Archive (ENA): ena-STUDY-ED-09–05–2018–12:07:59:395–193. Primary accession: PRJEB26657. A github repository providing all scripts used the in analysis can be found here: https://github.com/tcezard/RADmapper.

ORCID

Anke Lange https://orcid.org/0000-0003-0665-8404
Josephine R. Paris https://orcid.org/0000-0001-6868-3416
Karim Gharbi https://orcid.org/0000-0003-1092-4488
Shinichi Miyagawa https://orcid.org/0000-0003-1360-6215
David J. Studholme https://orcid.org/0000-0002-3010-6637
Charles R. Tyler https://orcid.org/0000-0003-2353-5748

REFERENCES

Allendorf, F. W., Gellman, W. A., & Thorgaard, G. H. (1994). Sex-linkage of 2 enzyme loci in *Oncorhynchus mykiss* (Rainbow trout). *Heredity*, *72*, 498–507. https://doi.org/10.1038/hdy.1994.67

An, W., Hu, J. Y., Giesy, J. P., & Yang, M. (2009). Extinction risk of exploited wild roach (*Rutilus rutilus*) populations due to chemical feminization. *Environmental Science & Technology*, *43*(20), 7895–7901. https://doi.org/10.1021/es900857u

Ankley, G. T., Jensen, K. M., Makynen, E. A., Kahl, M. D., Korte, J. J., Hornung, M. W., … Gray, L. E. (2003). Effects of the androgenic growth promoter 17-β-trenbolone on fecundity and reproductive endocrinology of the fathead minnow. *Environmental Toxicology and Chemistry*, *22*(6), 1350–1360. https://doi.org/10.1002/etc.5620220623

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., … Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLOS One*, *3*(10), e3376. https://doi.org/10.1371/journal.pone.0003376

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., … Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Baroiller, J. F., & D'Cotta, H. (2001). Environment and sex determination in farmed fish. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, *130*(4), 399–409. https://doi.org/10.1016/S1532-0456(01)00267-8

Baroiller, J.-F., Guiguen, Y., & Fostier, A. (1999). Endocrine and environmental aspects of sex differentiation in fish. *Cellular and Molecular Life Sciences CMLS*, *55*(6), 910–931. https://doi.org/10.1007/s000180050344

Bengu, T. S., du Plessis, J., Modley, L. S., & van Dyk, J. C. (2017). Health effects in fish from the polluted Orlando Dam and Klipspruit wetland system, Soweto, South Africa. *African Journal of Aquatic Science*, *42*(2), 131–141. https://doi.org/10.2989/16085914.2017.1347083

Bernet, D., Liedtke, A., Bittner, D., Eggen, R. I. L., Kipfer, S., Küng, C., … Segner, H. (2008). Gonadal malformations in Whitefish from Lake Thun: Defining the case and evaluating the role of EDCs. *CHIMIA International Journal for Chemistry*, *62*(5), 383–388. https://doi.org/10.2533/chimia.2008.383

Bjerregaard, L. B., Korsgaard, B., & Bjerregaard, P. (2006). Intersex in wild roach (*Rutilus rutilus*) from Danish sewage effluent-receiving streams. *Ecotoxicology and Environmental Safety*, *64*(3), 321–328. https://doi.org/10.1016/j.ecoenv.2005.05.018

Böhne, A., Weber, A.-A.-T., Rajkov, J., Rechsteiner, M., Riss, A., Egger, B., & Salzburger, W. (2019). Repeated evolution versus common ancestry: Sex chromosome evolution in the haplochromine cichlid *Pseudocrenilabrus philander*. *Genome Biology and Evolution*, *11*(2), 439–458. https://doi.org/10.1093/gbe/evz003

Böhne, A., Wilson, C. A., Postlethwait, J. H., & Salzburger, W. (2016). Variations on a theme: Genomics of sex determination in the cichlid fish *Astatotilapia burtoni*. *BMC Genomics*, *17*(1), 883–883. https://doi.org/10.1186/s12864-016-3178-0

Brown, J. K., Taggart, J. B., Bekaert, M., Wehner, S., Palaiokostas, C., Setiawan, A. N., … Penman, D. J. (2016). Mapping the sex determination locus in the hāpuku (*Polyprion oxygeneios*) using ddRAD sequencing. *BMC Genomics*, *17*(1), 448. https://doi.org/10.1186/s12864-016-2773-4

Burkhardt-Holm, P., Peter, A., & Segner, H. (2002). Decline of fish catch in Switzerland - Project Fishnet: A balance between analysis and synthesis. *Aquatic Sciences*, *64*(1), 36–54. https://doi.org/10.1007/s00027-002-8053-1

Carmichael, S. N., Bekaert, M., Taggart, J. B., Christie, H. R. L., Bassett, D. I., Bron, J. E., … Sturm, A. (2013). Identification of a sex-linked SNP marker in the salmon louse (*Lepeophtheirus salmonis*) using RAD sequencing. *PLOS One*, *8*(10), e77832. https://doi.org/10.1371/journal.pone.0077832

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci *de novo* from short-read sequences. *G3-Genes Genomes Genetics*, *1*(3), 171–182. https://doi.org/10.1534/g3.111.000240

Chen, J. J., Du, Q. Y., Yue, Y. Y., Dang, B. J., & Chang, Z. J. (2010). Screening and identification of male-specific DNA fragments in common carps *Cyprinus carpio* using suppression subtractive hybridization. *Journal of Fish Biology*, *77*(2), 403–413. https://doi.org/10.1111/j.1095-8649.2010.02700.x

Chen, J. J., Wang, Y. L., Yue, Y. Y., Xia, X. H., Du, Q. Y., & Chang, Z. J. (2009). A novel male-specific DNA sequence in the common carp, Cyprinus carpio. *Molecular and Cellular Probes*, *23*(5), 235–239. https://doi.org/10.1016/j.mcp.2009.04.004

Chen, S.-L., Li, J., Deng, S.-P., Tian, Y.-S., Wang, Q.-Y., Zhuang, Z.-M., … Xu, J.-Y. (2007). Isolation of female-specific AFLP markers and molecular identification of genetic sex in half-smooth tongue sole (*Cynoglossus semilaevis*). *Marine Biotechnology*, *9*(2), 273–280. https://doi.org/10.1007/s10126-006-6081-x

Cotton, S., & Wedekind, C. (2009). Population consequences of environmental sex reversal. *Conservation Biology*, *23*(1), 196–206. https://doi.org/10.1111/j.1523-1739.2008.01053.x

David, A., Lange, A., Abdul-Sada, A., Tyler, C. R., & Hill, E. M. (2017). Disruption of the prostaglandin metabolome and characterization of the pharmaceutical exposome in fish exposed to wastewater treatment works effluent as revealed by nanoflow-nanospray mass spectrometry-based metabolomics. *Environmental Science & Technology*, *51*(1), 616–624. https://doi.org/10.1021/acs.est.6b04365

Defra (2009). *Final report of the EDCAT project – Endocrine Disruption in Catchments*. Retrieved from http://webarchive.nationalarchives.gov.uk/20130301194201/http://www.ceh.ac.uk//science/EDCATProjectDocuments.html.

Delord, C., Lassalle, G., Oger, A., Barloy, D., Coutellec, M.-A., Delcamp, A., … Petit, E. J. (2018). A cost-and-time effective procedure to develop SNP markers for multiple species: A support for community genetics. *Methods in Ecology and Evolution*, *9*(9), 1959–1974. https://doi.org/10.1111/2041-210x.13034

Devlin, R. H., Biagi, C. A., & Smailus, D. E. (2001). Genetic mapping of Y-chromosomal DNA markers in Pacific salmon. *Genetica*, *111*(1), 43–58. https://doi.org/10.1023/a:1013759802604

Devlin, R. H., & Nagahama, Y. (2002). Sex determination and sex differentiation in fish: An overview of genetic, physiological, and environmental influences. *Aquaculture*, *208*, 191–364. https://doi.org/10.1016/S0044-8486(02)00057-1

Du, S. J., Devlin, R. H., & Hew, C. L. (1993). Genomic structure of growth hormone genes in Chinook salmon (*Oncorhynchus tshawytscha*): Presence of two functional genes, GH-I and GH-II, and a male-specific pseudogene, GH-ψ. *DNA and Cell Biology*, *12*(8), 739–751. https://doi.org/10.1089/dna.1993.12.739

Enivornment Agency (2004). *Causes and consequences of feminisation of male fish in English rivers*. Science Report SC030285/SR, Bristol, UK. Retrieved from http://webarchive.nationalarchives.gov.uk/20150516033153/https://www.gov.uk/government/publications/causes-and-consequences-of-feminisation-of-male-fish-in-english-rivers

Estoup, A., Largiadèr, C. R., Perrot, E., & Chourrout, D. (1996). Rapid one-tube DNA extraction for reliable PCR detection of fish polymorphic markers and transgenes. *Molecular Marine Biology and Biotechnology*, *5*(4), 295–298.

Etter, P. D., Preston, J. L., Bassham, S., Cresko, W. A., & Johnson, E. A. (2011). Local de novo assembly of RAD paired-end contigs using short sequencing reads. *PLOS One*, *6*(4), e18561. https://doi.org/10.1371/journal.pone.0018561

Ezaz, M. T., Harvey, S. C., Boonphakdee, C., Teale, A. J., McAndrew, B. J., & Penman, D. J. (2004). Isolation and physical mapping of sex-linked AFLP markers in nile tilapia (*Oreochromis niloticus* L.). *Marine Biotechnology*, *6*(5), 435–445. https://doi.org/10.1007/s10126-004-3004-6

Fowler, B. L. S., & Buonaccorsi, V. P. (2016). Genomic characterization of sex-identification markers in *Sebastes carnatus* and *Sebastes chrysomelas* rockfishes. *Molecular Ecology*, *25*(10), 2165–2175. https://doi.org/10.1111/mec.13594

Gamble, T., Castoe, T. A., Nielsen, S. V., Banks, J. L., Card, D. C., Schield, D. R., … Booth, W. (2017). The discovery of XY sex chromosomes in a *Boa* and *Python*. *Current Biology*, *27*(14), 2148–2153. https://doi.org/10.1016/j.cub.2017.06.010

Gamble, T., Coryell, J., Ezaz, T., Lynch, J., Scantlebury, D. P., & Zarkower, D. (2015). Restriction site-associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions among gecko sex-determining systems. *Molecular Biology and Evolution*, *32*(5), 1296–1309. https://doi.org/10.1093/molbev/msv023

Gamble, T., & Zarkower, D. (2014). Identification of sex-specific molecular markers using restriction site-associated DNA sequencing. *Molecular Ecology Resources*, *14*(5), 902–913. https://doi.org/10.1111/1755-0998.12237

Gautier, M., Foucaud, J., Gharbi, K., Cézard, T., Galan, M., Loiseau, A., … Estoup, A. (2013). Estimation of population allele frequencies from next-generation sequencing data: Pool-versus individual-based genotyping. *Molecular Ecology*, *22*(14), 3766–3779. https://doi.org/10.1111/mec.12360

Geraudie, P., Gerbron, M., & Minier, C. (2017). Endocrine disruption effects in male and intersex roach (*Rutilus rutilus*, L.) from French rivers: An integrative approach based on subcellular to individual responses. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *211*, 29–36. https://doi.org/10.1016/j.cbpb.2017.05.006

Gerbron, M., Geraudie, P., Fernandes, D., Rotchell, J. M., Porte, C., & Minier, C. (2014). Evidence of altered fertility in female roach (*Rutilus rutilus*) from the River Seine (France). *Environmental Pollution*, *191*, 58–62. https://doi.org/10.1016/j.envpol.2014.04.015

Griffiths, R., Orr, K. J., Adam, A., & Barber, I. (2000). DNA sex identification in the three-spined stickleback. *Journal of Fish Biology*, *57*(5), 1331–1334. https://doi.org/10.1006/jfbi.2000.1386

Hamilton, P. B., Cowx, I. G., Oleksiak, M. F., Griffiths, A. M., Grahn, M., Stevens, J. R., … Tyler, C. R. (2016). Population-level consequences for wild fish exposed to sublethal concentrations of chemicals – a critical review. *Fish and Fisheries*, *17*(3), 545–566. https://doi.org/10.1111/faf.12125

Hamilton, P. B., Lange, A., Nicol, E., Bickley, L. K., De-Bastos, E. S. R., Jobling, S., & Tyler, C. R. (2015). Effects of exposure to WwTW effluents over two Generations on sexual development and breeding in roach *Rutilus rutilus*. *Environmental Science & Technology*, *49*(21), 12994–13002. https://doi.org/10.1021/acs.est.5b03777

Hamilton, P. B., Nicol, E., De-Bastos, E. S. R., Williams, R. J., Sumpter, J. P., Jobling, S., … Tyler, C. R. (2014). Populations of a cyprinid fish are self-sustaining despite widespread feminization of males. *BMC Biology*, *12*(1), 1. https://doi.org/10.1186/1741-7007-12-1

Hamilton, P. B., & Tyler, C. R. (2008). Identification of microsatellite loci for parentage analysis in roach *Rutilus rutilus* and eight other cyprinid fish by cross-species amplification, and a novel test for detecting hybrids between roach and other cyprinids. *Molecular Ecology Resources*, *8*(2), 462–465. https://doi.org/10.1111/j.1471-8286.2007.01994.x

Harris, C. A., Hamilton, P. B., Runnalls, T. J., Vinciotti, V., Henshaw, A., Hodgson, D., … Sumpter, J. P. (2011). The consequences of feminization in breeding groups of wild fish. *Environmental Health Perspectives*, *119*(3), 306–311. https://doi.org/10.1289/ehp.1002555

Hashimoto, S., Bessho, H., Hara, A., Nakamura, M., Iguchi, T., & Fujita, K. (2000). Elevated serum vitellogenin levels and gonadal abnormalities in wild male flounder (*Pleuronectes yokohamae*) from Tokyo Bay, Japan. *Marine Environmental Research*, *49*(1), 37–53. https://doi.org/10.1016/S0141-1136(99)00047-1

Heule, C., Salzburger, W., & Böhne, A. (2014). Genetics of sexual development: An evolutionary playground for fish. *Genetics*, *196*(3), 579–591. https://doi.org/10.1534/genetics.114.161158

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLOS Genetics*, *6*(2), e1000862. https://doi.org/10.1371/journal.pgen.1000862

Houston, R. D., Taggart, J. B., Cézard, T., Bekaert, M., Lowe, N. R., Downing, A., … Hamilton, A. (2014). Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*). *BMC Genomics*, *15*(1), 90. https://doi.org/10.1186/1471-2164-15-90

Jalabert, B., Baroiller, J. F., Breton, B., Fostier, A., Le Gac, F., Guiguen, Y., & Monod, G. (2000). Main neuro-endocrine, endocrine and paracrine regulations of fish reproduction, and vulnerability to xenobiotics. *Ecotoxicology*, *9*(1–2), 25–40. https://doi.org/10.1023/A:1008960010754

Jobling, S., Beresford, N., Nolan, M., Rodgers-Gray, T., Brighty, G. C., Sumpter, J. P., & Tyler, C. R. (2002). Altered sexual maturation and gamete production in wild roach (*Rutilus rutilus*) living in rivers that receive treated sewage effluents. *Biology of Reproduction*, *66*(2), 272–281. https://doi.org/10.1095/biolreprod66.2.272

Jobling, S., Coey, S., Whitmore, J. G., Kime, D. E., Van Look, K. J. W., McAllister, B. G., … Sumpter, J. P. (2002). Wild intersex roach (*Rutilus rutilus*) have reduced fertility. *Biology of Reproduction*, *67*(2), 515–524. https://doi.org/10.1095/biolreprod67.2.515

Jobling, S., Nolan, M., Tyler, C. R., Brighty, G., & Sumpter, J. P. (1998). Widespread sexual disruption in wild fish. *Environmental Science & Technology*, *32*(17), 2498–2506. https://doi.org/10.1021/es9710870

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kidd, K. A., Blanchfield, P. J., Mills, K. H., Palace, V. P., Evans, R. E., Lazorchak, J. M., & Flick, R. W. (2007). Collapse of a fish population after exposure to a synthetic estrogen. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(21), 8897–8901. https://doi.org/10.1073/pnas.0609568104

Kovács, B., Egedi, S., Bartfai, R., & Orban, L. (2000). Male-specific DNA markers from African catfish (*Clarias gariepinus*). *Genetica*, *110*(3), 267–276. https://doi.org/10.1023/a:1012739318941

Kurland, S., Wheat, C. W., Paz Celorio Mancera, M., Kutschera, V. E., Hill, J., Andersson, A., … Laikre, L. (2019). Exploring a Pool-seq-only approach for gaining population genomic insights in nonmodel species. *Ecology and Evolution*, *9*(19), 11448–11463. https://doi.org/10.1002/ece3.5646

Lange, A., Katsu, Y., Ichikawa, R., Paull, G. C., Chidgey, L. L., Coe, T. S., … Tyler, C. R. (2008). Altered sexual development in roach (*Rutilus rutilus)* exposed to environmental concentrations of the pharmaceutical 17α-ethinylestradiol and associated expression dynamics of aromatases and estrogen receptors. *Toxicological Sciences*, *106*(1), 113–123. https://doi.org/10.1093/toxsci/kfn151

Lange, A., Paull, G. C., Coe, T. S., Katsu, Y., Urushitani, H., Iguchi, T., & Tyler, C. R. (2009). Sexual reprogramming and estrogenic sensitization in wild fish exposed to ethinylestradiol. *Environmental Science & Technology*, *43*(4), 1219–1225. https://doi.org/10.1021/es802661p

Lange, A., Paull, G. C., Hamilton, P. B., Iguchi, T., & Tyler, C. R. (2011). Implications of persistent exposure to treated wastewater effluent for breeding in wild roach (*Rutilus rutilus*) populations. *Environmental Science & Technology*, *45*(4), 1673–1679. https://doi.org/10.1021/es103232q

Lange, A., Sebire, M., Rostkowski, P., Mizutani, T., Miyagawa, S., Iguchi, T., … Tyler, C. R. (2015). Environmental chemicals active as human antiandrogens do not activate a stickleback androgen receptor but enhance a feminising effect of oestrogen in roach. *Aquatic Toxicology*, *168*, 48–59. https://doi.org/10.1016/j.aquatox.2015.09.014

Leitwein, M., Guinand, B., Pouzadoux, J., Desmarais, E., Berrebi, P., & Gagnaire, P.-A. (2017). A dense brown trout (*Salmo trutta*) linkage map reveals recent chromosomal rearrangements in the Salmo genus and the impact of selection on linked neutral diversity. *G3: Genes, Genomes, Genetics*, *7*(4), 1365–1376. https://doi.org/10.1534/g3.116.038497

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*, 1303.3997v1 [q-bio.GN].

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liu, Q., Goudie, C. A., Simco, B. A., & Davis, K. B. (1996). Sex-linkage of glucosephosphate isomerase-B and mapping of the sex-determining gene in channel catfish. *Cytogenetics and Cell Genetics*, *73*(4), 282–285. https://doi.org/10.1159/000134356

Manousaki, T., Tsakogiannis, A., Taggart, J. B., Palaiokostas, C., Tsaparis, D., Lagnel, J., … Tsigenopoulos, C. S. (2016). Exploring a nonmodel teleost genome through RAD sequencing - Linkage mapping in Common Pandora, *Pagellus erythrinus* and comparative genomic analysis. *G3-Genes Genomes Genetics*, *6*(3), 509–519. https://doi.org/10.1534/g3.115.023432

Marçais, G., & Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*, *27*(6), 764–770. https://doi.org/10.1093/bioinformatics/btr011

Mathers, T. C., Hammond, R. L., Jenner, R. A., Hanfling, B., Atkins, J., & Gomez, A. (2015). Transition in sexual system and sex chromosome evolution in the tadpole shrimp *Triops cancriformis*. *Heredity*, *115*(1), 37–46. https://doi.org/10.1038/hdy.2015.10

Matsuda, M., Nagahama, Y., Shinomiya, A. I., Sato, T., Matsuda, C., Kobayashi, T., … Sakaizumi, M. (2002). DMY is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature*, *417*(6888), 559–563. https://doi.org/10.1038/nature751

McGee, C., Brougham, C., Roche, J., & Fogarty, A. (2012). First report of intersex roach residing in Irish rivers downstream of several wastewater treatment plants. *Biology and Environment-Proceedings of the Royal Irish Academy*, *112B*(1), 69–77. https://doi.org/10.3318/bioe.2011.23

Micheletti, S. J., & Narum, S. R. (2018). Utility of pooled sequencing for association mapping in nonmodel organisms. *Molecular Ecology Resources*, *18*(4), 825–837. https://doi.org/10.1111/1755-0998.12784

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., & Johnson, E. A. (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, *17*(2), 240–248. https://doi.org/10.1101/gr.5681207

Minier, C., Caltot, G., Leboulanger, F., & Hill, E. M. (2000). An investigation of the incidence of intersex fish in Seine-Maritime and Sussex region. *Analusis*, *28*(9), 801–806. https://doi.org/10.1051/analusis:2000280801

Nakamura, M., Kobayashi, T., Chang, X. T., & Nagahama, Y. (1998). Gonadal sex differentiation in teleost fish. *Journal of Experimental Zoology*, *281*(5), 362–372. https://doi.org/10.1002/(SICI)1097-010X(19980801)281:5<362:AID-JEZ3>3.0.CO;2-M

Nanda, I., Kondo, M., Hornung, U., Asakawa, S., Winkler, C., Shimizu, A., … Schartl, M. (2002). A duplicated copy of DMRT1 in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*. *Proceedings of the National Academy of Sciences*, *99*(18), 11778–11783. https://doi.org/10.1073/pnas.182314699

Neethiraj, R., Hornett, E. A., Hill, J. A., & Wheat, C. W. (2017). Investigating the genomic basis of discrete phenotypes using a Pool-Seq-only approach: New insights into the genetics underlying colour variation in diverse taxa. *Molecular Ecology*, *26*(19), 4990–5002. https://doi.org/10.1111/mec.14205

Noaksson, E., Tjärnlund, U., Bosveld, A. T. C., & Balk, L. (2001). Evidence for endocrine disruption in perch (*Perca fluviatilis*) and roach (*Rutilus rutilus*) in a remote Swedish lake in the vicinity of a public refuse dump. *Toxicology and Applied Pharmacology*, *174*(2), 160–176. https://doi.org/10.1006/taap.2001.9194

Nolan, M., Jobling, S., Brighty, G., Sumpter, J. P., & Tyler, C. R. (2001). A histological description of intersexuality in the roach. *Journal of Fish Biology*, *58*(1), 160–176. https://doi.org/10.1111/j.1095-8649.2001.tb00505.x

Olmstead, A. W., Villeneuve, D. L., Ankley, G. T., Cavallin, J. E., Lindberg-Livingston, A., Wehmas, L. C., & Degitz, S. J. (2011). A method for the determination of genetic sex in the fathead minnow, *Pimephales promelas*, to support testing of endocrine-active chemicals. *Environmental Science & Technology*, *45*(7), 3090–3095. https://doi.org/10.1021/es103327r

Palaiokostas, C., Bekaert, M., Davie, A., Cowan, M. E., Oral, M., Taggart, J. B., … Migaud, H. (2013). Mapping the sex determination locus in the Atlantic halibut (*Hippoglossus hippoglossus*) using RAD sequencing. *BMC Genomics*, *14*(1), 566. https://doi.org/10.1186/1471-2164-14-566

Palaiokostas, C., Bekaert, M., Khan, M. G. Q., Taggart, J. B., Gharbi, K., McAndrew, B. J., & Penman, D. J. (2013). Mapping and validation of the major sex-determining region in Nile tilapia (*Oreochromis niloticus* L.) using RAD sequencing. *PLOS One*, *8*, https://doi.org/10.1371/journal.pone.0068389

Palaiokostas, C., Bekaert, M., Khan, M. G., Taggart, J. B., Gharbi, K., McAndrew, B. J., & Penman, D. J. (2015). A novel sex-determining QTL in Nile tilapia (*Oreochromis niloticus*). *BMC Genomics*, *16*(1), 171. https://doi.org/10.1186/s12864-015-1383-x

Pan, Q., Anderson, J., Bertho, S., Herpin, A., Wilson, C., Postlethwait, J. H., … Guiguen, Y. (2016). Vertebrate sex-determining genes play musical chairs. *Comptes Rendus Biologies*, *339*(7), 258–262. https://doi.org/10.1016/j.crvi.2016.05.010

Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for stacks. *Methods in Ecology and Evolution*, *8*(10), 1360–1373. https://doi.org/10.1111/2041-210x.12775

Patino, R. (1997). Manipulations of the reproductive system of fishes by means of exogenous chemicals. *Progressive Fish-Culturist*, *59*(2), 118–128. https://doi.org/10.1577/1548-8640(1997)059<0118:MOTRSO>2.3.CO;2

Paull, G. C., Filby, A. L., & Tyler, C. R. (2009). Growth rate during early life affects sexual differentiation in roach (*Rutilus rutilus*). *Environmental Biology of Fishes*, *85*(4), 277–284. https://doi.org/10.1007/s10641-009-9492-5

Paull, G. C., Lange, A., Henshaw, A. C., & Tyler, C. R. (2008). Ontogeny of sexual development in the roach (*Rutilus rutilus*) and its interrelationships with growth and age. *Journal of Morphology*, *269*(7), 884–895. https://doi.org/10.1002/jmor.10641

Peichel, C. L., Ross, J. A., Matson, C. K., Dickson, M., Grimwood, J., Schmutz, J., … Kingsley, D. M. (2004). The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Current Biology*, *14*(16), 1416–1424. https://doi.org/10.1016/j.cub.2004.08.030

Peng, Y., Leung, H. C. M., Yiu, S. M., & Chin, F. Y. L. (2012). IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, *28*(11), 1420–1428. https://doi.org/10.1093/bioinformatics/bts174

Pitts, C. S. (1994). *Inter-specific hybridization in the fish family Cyprinidae*. PhD thesis, The University Of Hull, UK. Retrieved from https://hydra.hull.ac.uk/resources/hull:5372

Reinecke, M., Björnsson, B. T., Dickhoff, W. W., McCormick, S. D., Navarro, I., Power, D. M., & Gutiérrez, J. (2005). Growth hormone and insulin-like growth factors in fish: Where we are and where to go. *General and Comparative Endocrinology*, *142*(1), 20–24. https://doi.org/10.1016/j.ygcen.2005.01.016

Rellstab, C., Zoller, S., Tedder, A., Gugerli, F., & Fischer, M. C. (2013). Validation of SNP allele frequencies determined by pooled next-generation sequencing in natural populations of a non-model plant species. *PLOS One*, *8*(11), e80422. https://doi.org/10.1371/journal.pone.0080422

Rodgers-Gray, T. P., Jobling, S., Kelly, C., Morris, S., Brighty, G., Waldock, M. J., … Tyler, C. R. (2001). Exposure of juvenile roach (*Rutilus rutilus*) to treated sewage effluent induces dose-dependent and persistent disruption in gonadal duct development. *Environmental Science & Technology*, *35*(3), 462–470. https://doi.org/10.1021/es001225c

Scholz, S., & Klüver, N. (2009). Effects of endocrine disrupters on sexual, gonadal development in fish. *Sexual Development*, *3*(2–3), 136–151. https://doi.org/10.1159/000223078

Strüssmann, C. A., & Nakamura, M. (2002). Morphology, endocrinology, and environmental modulation of gonadal sex differentiation in teleost fishes. *Fish Physiology and Biochemistry*, *26*(1), 13–29. https://doi.org/10.1023/A:1023343023556

Tong, S.-K., Hsu, H.-J., & Chung, B.-C. (2010). Zebrafish monosex population reveals female dominance in sex determination and earliest events of gonad differentiation. *Developmental Biology*, *344*(2), 849–856. https://doi.org/10.1016/j.ydbio.2010.05.515

Trubiroha, A., Kroupova, H., Wuertz, S., Frank, S. N., Sures, B., & Kloas, W. (2010). Naturally-induced endocrine disruption by the parasite *Ligula intestinalis* (Cestoda) in roach (*Rutilus rutilus*). *General and Comparative Endocrinology*, *166*(2), 234–240. https://doi.org/10.1016/j.ygcen.2009.08.010

Truett, G. E., Heeger, P., Mynatt, R. L., Truett, A. A., Walker, J. A., & Warman, M. L. (2000). Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *BioTechniques*, *29*(1), 52–54. https://doi.org/10.2144/00291bm09

Tyler, C. R., Filby, A. L., Bickley, L. K., Cumming, R. I., Gibson, R., Labadie, P., … Hill, E. M. (2009). Environmental health impacts of equine estrogens derived from hormone replacement therapy. *Environmental Science & Technology*, *43*(10), 3897–3904. https://doi.org/10.1021/es803135q

Tyler, C. R., Lange, A., Paull, G. C., Katsu, Y., & Iguchi, T. (2007). The roach (*Rutilus rutilus*) as a sentinel for endocrine disruption in UK rivers. *Environmental Sciences*, *14*(5), 235–253.

Tyler, C. R., van der Eerden, B., Jobling, S., Panter, G., & Sumpter, J. P. (1996). Measurement of vitellogenin, a biomarker for exposure to oestrogenic chemicals, in a wide variety of cyprinid fish. *Journal of Comparative Physiology B-Biochemical Systemic and Environmental Physiology*, *166*(7), 418–426. https://doi.org/10.1007/BF02337886

Vajda, A. M., Barber, L. B., Gray, J. L., Lopez, E. M., Woodling, J. D., & Norris, D. O. (2008). Reproductive disruption in fish downstream from an estrogenic wastewater effluent. *Environmental Science & Technology*, *42*(9), 3407–3414. https://doi.org/10.1021/es0720661

van Aerle, R., Nolan, M., Jobling, S., Christiansen, L. B., Sumpter, J. P., & Tyler, C. R. (2001). Sexual disruption in a second species of wild cyprinid fish (the Gudgeon, *gobio gobio*) in United Kingdom freshwaters. *Environmental Toxicology and Chemistry*, *20*(12), 2841–2847.

Volff, J.-N., Selz, Y., Hoffmann, C., Froschauer, A., Schultheis, C., Schmidt, C., … Schartl, M. (2013). Gene amplification and functional diversification of melanocortin 4 receptor at an extremely polymorphic locus controlling sexual maturation in the Platyfish. *Genetics*, *195*(4), 1337–1352. https://doi.org/10.1534/genetics.113.155952

von Schalburg, K. R., Yazawa, R., de Boer, J., Lubieniecki, K. P., Goh, B., Straub, C. A., … Koop, B. F. (2008). Isolation, characterization and comparison of Atlantic and Chinook salmon growth hormone 1 and 2. *BMC Genomics*, *9*(1), 522. https://doi.org/10.1186/1471-2164-9-522

Wiklund, T., Lounasheimo, L., Lom, J., & Bylund, G. (1996). Gonadal impairment in roach *Rutilus rutilus* from Finnish coastal areas of the northern Baltic Sea. *Diseases of Aquatic Organisms*, *26*(3), 163–171. https://doi.org/10.3354/dao026163

Williams, R. J., Churchley, J. H., Kanda, R., & Johnson, A. C. (2012). Comparing predicted against measured steroid estrogen concentrations and the associated risk in two United Kingdom river catchments. *Environmental Toxicology and Chemistry*, *31*(4), 892–898. https://doi.org/10.1002/etc.1756

Williams, R. J., Johnson, A. C., Smith, J. J. L., & Kanda, R. (2003). Steroid estrogens profiles along river stretches arising from sewage treatment works discharges. *Environmental Science & Technology*, *37*(9), 1744–1750. https://doi.org/10.1021/es0202107

Willing, E. M., Hoffmann, M., Klein, J. D., Weigel, D., & Dreyer, C. (2011). Paired-end RAD-seq for de novo assembly and marker design without available reference. *Bioinformatics*, *27*(16), 2187–2193. https://doi.org/10.1093/bioinformatics/btr346

Wilson, C. A., High, S. K., McCluskey, B. M., Amores, A., Yan, Y.-L., Titus, T. A., … Postlethwait, J. H. (2014). Wild sex in zebrafish: Loss of the natural sex determinant in domesticated strains. *Genetics*, *198*(3), 1291–1308. https://doi.org/10.1534/genetics.114.169284

Withler, R. E., McPhail, J. D., & Devlin, R. H. (1986). Electrophoretic polymorphism and sexual dimorphism in the fresh water and anadromous threespine sticklebacks (Gasterosteus aculeatus) of the Little Campbell River, British Columbia. *Biochemical Genetics*, *24*(9–10), 701–713. https://doi.org/10.1007/bf00499003

Woram, R. A., Gharbi, K., Sakamoto, T., Hoyheim, B., Holm, L.-E., Naish, K., … Danzmann, R. G. (2003). Comparative genome analysis of the primary sex-determining locus in salmonid fishes. *Genome Research*, *13*(2), 272–280. https://doi.org/10.1101/gr.578503

Wyatt, P. M. W., Pitts, C. S., & Butlin, R. K. (2006). A molecular approach to detect hybridization between bream *Abramis brama*, roach *Rutlius rutilus* and rudd *Scardinius erythrophthalmus*. *Journal of Fish Biology*, *69*, 52–71. https://doi.org/10.1111/j.1095-8649.2006.01104.x

Yano, A., Nicol, B., Jouanno, E., Quillet, E., Fostier, A., Guyomard, R., & Guiguen, Y. (2013). The sexually dimorphic on the Y-chromosome gene (sdY) is a conserved male-specific Y-chromosome sequence in many salmonids. *Evolutionary Applications*, *6*(3), 486–496. https://doi.org/10.1111/eva.12032

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, *13*, 134. https://doi.org/10.1186/1471-2105-13-134

Zhang, Q., Nakayama, I., Fujiwara, A., Kobayashi, T., Oohara, I., Masaoka, T., … Devlin, R. H. (2001). Sex identification by male-specific growth hormone pseudogene (GH-Ψ) in Oncorhynchus masou complex and a related hybrid. *Genetica*, *111*(1), 111–118. https://doi.org/10.1023/a:1013799229012

Zheng, B., Liu, R., Liu, Y., Jin, F., & An, L. (2015). Phenolic endocrine-disrupting chemicals and intersex in wild crucian carp from Hun River, China. *Chemosphere*, *120*, 743–749. https://doi.org/10.1016/j.chemosphere.2014.10.049

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Lange A, Paris JR, Gharbi K, et al. A newly developed genetic sex marker and its application to understanding chemically induced feminisation in roach (*Rutilus rutilus*). *Mol Ecol Resour.* 2020;00:1–16. https://doi.org/10.1111/1755-0998.13166