MOLECULAR ECOLOGY

Molecular Ecology (2013) doi: 10.1111/mec.12322

Recombination in the threespine stickleback genome—patterns and consequences

MARIUS ROESTI, DARIO MOSER and DANIEL BERNER
Zoological Institute, University of Basel, Vesalgasse 1, CH-4051 Basel, Switzerland

Abstract

Heterogeneity in recombination rate may strongly influence genome evolution and entail methodological challenges to genomic investigation. Nevertheless, a solid understanding of these issues awaits detailed information across a broad range of taxa. Based on 282 F_2 individuals and 1872 single nucleotide polymorphisms, we characterize recombination in the threespine stickleback fish genome. We find an average genome-wide recombination rate of 3.11 cM/Mb. Crossover frequencies are dramatically elevated in the chromosome peripheries as compared to the centres, and are consistent with one obligate crossover per chromosome (but not chromosome arm). Along the sex chromosome, we show that recombination is restricted to a small pseudoautosomal domain of *c.* 2 Mb, spanning only 10% of that chromosome. Comparing female to male RAD sequence coverage allows us to identify two discrete levels of degeneration on the Y chromosome, one of these 'evolutionary strata' coinciding with a previously inferred inverted region. Using polymorphism data from two young (<10 000 years old) ecologically diverged lake-stream population pairs, we demonstrate that recombination rate correlates with both the magnitude of allele frequency shifts between populations and levels of genetic diversity within populations. These associations reflect genome-wide heterogeneity in the influence of selection on linked sites. We further find a strong relationship between recombination rate and GC content, possibly driven by GC-biased gene conversion. Overall, we highlight that heterogeneity in recombination rate has profound consequences on genome evolution and deserves wider recognition in marker-based genomic analyses.

Keywords: evolutionary strata, *Gasterosteus aculeatus*, GC content, genetic diversity, pseudoautosomal region, RAD sequencing, sex chromosome

Received 2 March 2013; accepted 17 March 2013

Introduction

Meiotic recombination is a fascinating process because of its pivotal role in multiple biological contexts. For instance, recombination is generally considered instrumental to the proper segregation of homologous chromosomes during meiosis (Mather 1938; Baker *et al.* 1976; Roeder 1997; Smith & Nicolas 1998; Hassold & Hunt 2001). At the same time, recombination breaks the linkage between DNA segments located on the same chromosome. This allows selection to operate more effectively on multiple loci, and hence promotes adaptation (Hill & Robertson 1966; Felsenstein 1974; Otto & Barton 1997; Burt 2000; Otto & Lenormand 2002). Conversely, the suppression of recombination can initiate chromosome degeneration, a process believed to be common during sex chromosome evolution (Bull 1983; Charlesworth & Charlesworth 2000; Charlesworth *et al.* 2005; Wilson & Makova 2009).

Variation in recombination rate may also explain genome-wide heterogeneity in the magnitude of genetic divergence between populations, and genetic diversity within populations. The reason is that linkage between selected loci and their physical neighbourhood is tighter in regions exhibiting relatively low recombination rate. Selectively neutral polymorphisms will therefore be affected by selection more often and more strongly

Correspondence: Daniel Berner, Fax: +41 (0) 61 267 0301;
E-mail: daniel.berner@unibas.ch

when located in low-recombination regions. As a consequence, hitchhiking under positive and background selection is predicted to increase allele frequency shifts between populations, and to reduce genetic diversity within populations, in low-relative to high-recombination genomic regions (Maynard Smith & Haigh 1974; Kaplan *et al.* 1989; Begun & Aquadro 1992; Nordborg *et al.* 1996; Charlesworth *et al.* 1997; Charlesworth 1998; Nachman 2002). Similarly, loci under divergent selection between ecologically distinct habitats should impede neutral gene flow more extensively in low-recombination regions (Barton & Bengtsson 1986; Feder & Nosil 2010). Finally, recombination may have direct effects on the constitution of chromosomes, for instance through biased gene conversion or mutagenesis (Galtier *et al.* 2001; Duret & Galtier 2009; Webster & Hurst 2012).

Despite the recognition of recombination as a major evolutionary factor, our understanding of both the mechanisms governing the process, and its consequences on genome evolution, remains highly incomplete (Nachman 2002; Smukowski & Noor 2011; Webster & Hurst 2012). Moreover, detailed investigations of recombination outside genetic model organisms are needed for the discovery of general patterns. The goal of our study is to provide the first comprehensive analysis of meiotic recombination in threespine stickleback fish (*Gasterosteus aculeatus*).

A thorough understanding of recombination in this powerful model organism for ecological genetics is particularly valuable for two reasons. First, the species has been shown to display a relatively young (<10 Myr old) XY (male-heterogametic) sex determination system (Peichel *et al.* 2004). Information on the extent of XY recombination and associated patterns of Y degeneration, however, remains highly incomplete, but promises exciting insights into sex chromosome evolution (Peichel *et al.* 2004; Ross & Peichel 2008; Shikano *et al.* 2011). Second, performing genome scans in stickleback populations residing contiguously in selectively distinct lake and stream habitats, we have shown recently that population divergence (F_{ST}) is elevated in chromosome centres and argued that this effect is caused by a lower recombination rate within these regions (Roesti *et al.* 2012a). Because robust information on recombination was lacking, however, this hypothesis could not be evaluated definitively. Our study therefore combines single nucleotide polymorphism (SNP) data from a laboratory F_2 cross and natural populations to characterize the stickleback recombination landscape; to explore the role of recombination in sex chromosome evolution; to examine the relationship between recombination rate and the magnitude of divergence among and genetic diversity within populations; and to investigate the association between recombination rate and nucleotide composition.

Materials and methods

Laboratory cross

We generated an F_2 population for linkage map construction by artificially crossing a male and a female from the Central European ROM and CHE populations (described in Berner *et al.* 2010; Moser *et al.* 2012; Roesti *et al.* 2012b) in the spring 2009. The resulting F_1 were raised in two 50-L tanks on a mixed Artemia (live, decapsulated cysts, frozen) and bloodworm diet under 'summer' laboratory conditions (18–20 °C with a 16:8-h day/night photoperiod). After a 'winter' phase (15 °C, 8:16-h photoperiod) of 3 months, summer conditions were re-established in the spring 2010 to initiate reproduction. The F_2 population was generated by performing 20 artificial F_1 full-sib crosses, each involving a unique male–female combination. After 1 year, 282 adult F_2 (140 males, 142 females) were haphazardly chosen, killed with an overdose of MS-222 and stored in absolute ethanol.

Marker generation

DNA was extracted from pectoral fin tissue on a MagNA Pure LC278 extraction robot (Roche) by using the tissue Isolation Kit II. We then prepared restriction site–associated DNA (RAD; Baird *et al.* 2008) libraries, involving *Sbf1* restriction, the fusion of 5-mer individual barcodes and pooling DNA of 62 individuals per library. The final enrichment PCR was performed in duplicate to reduce random amplification variation. Each library was single-end sequenced to 100 base reads in a separate Illumina HiSeq lane. In addition to the F_2 individuals, the two founder individuals of the cross were also sequenced, each twice in different libraries.

The Illumina sequences were sorted according to barcode and aligned to the stickleback reference genome (release Broad S1; Jones et al. 2012) by using Novoalign v2.07.06 (http://novocraft.com), accepting a total of approximately eight high-quality mismatches and/or indels along a read. Alignments were converted to BAM format using SAMtools (Li *et al.* 2009). Each replicate alignment of each grandparent was then screened independently for homozygous RAD loci. A locus qualified as homozygous if it was either invariant or if the binomial probability for the two dominant haplotypes to reflect a heterozygous locus was <0.001. We here ignored loci with <12× coverage (average coverage per locus varied between 31 and 47 among the grandparents

and replicate alignments). RAD loci proving homozygous in *both* replicates of a given grandparent were then screened for SNPs (here subsuming both SNPs and microindels) fixed for different alleles between the grandparents, accepting only one SNP per RAD locus. This conservative SNP detection strategy yielded a total of 2223 markers.

The F_2 population was then genotyped at each SNP detected in the grandparents. We considered a SNP homozygous when only one grandparent allele was present and occurred in at least 20 copies, or heterozygous when both alleles were present in at least 20 copies each (average sequence coverage per RAD locus was 55.6 among the F_2 individuals). Loci not satisfying these criteria received an ambiguous genotype based on the dominant allele or were treated as missing data when the total allele count was below six (<0.5% of all genotypes). Next, we eliminated 58 SNPs displaying clearly skewed allele frequencies across the F_2 and ordered the remaining 2165 markers according to their physical position in the Broad S1 stickleback reference genome.

Genome reassembly

Visual inspection of the genotypes ordered according to the reference genome indicated marker intervals exhibiting extremely high crossover frequency. Without exception, these intervals coincided perfectly with scaffold boundaries, indicating genome assembly errors. This conclusion was also supported by comparing physical and genetic map positions in a low-resolution data set extracted by Roesti *et al.* (2012a) from genetic maps available for North American stickleback (Albert *et al.* 2008; Greenwood *et al.* 2011): markers on scaffolds found to be inverted in the current study showed opposite genetic and physical map order in the latter data set as well (details not presented).

An accurate characterization of recombination thus required genome reassembly. For this, we created *de novo* linkage groups in R/qtl (Broman & Sen 2009) by including markers unanchored to any linkage group in the Broad S1 genome. We used a maximum recombination frequency of 0.3 or less and a LOD score of 8 or greater and further optimized marker order along linkage groups through permutation within a sliding window of seven markers. The resulting genetic map allowed us to invert 13 total scaffolds (size range: 0.7–17.1 Mb; 98.2 Mb in total) within the known linkage groups (hereafter 'chromosomes') and to incorporate 18 previously unanchored scaffolds with a total length of 20.1 Mb into the chromosomes. We ignored unanchored scaffolds smaller than 140 kb, as this was below our average marker resolution. We then recalculated the physical position for every marker. These assembly corrections are described in Appendix S1 (Supporting information). All physical map positions in this study refer to our improved genome assembly, which is available in FASTA format on the Dryad digital repository (doi:10.5061/dryad.846nj).

For final genetic map construction, we first corrected genotyping errors and ambiguous calls by hand, making the common and well-supported assumption that the vast majority of tight double-recombinants reflect genotyping errors. We then clipped the most peripheral marker at each chromosome end because here phase shifts were most difficult to distinguish from genotyping errors. Next, we discarded all markers not assigned to linkage groups, and one (redundant) marker in cases where two SNPs formed a pair derived from sister RAD loci (i.e. loci flanking the same *Sbf1* restriction site). The final data set used for genetic mapping comprised 1872 markers (59–150 per chromosome), with an average spacing of 217 kb. The genotype data are provided as Appendix S2 (autosomes) and S3 (sex chromosome, Supporting information), and linkage map and corresponding physical map positions (the latter before and after reassembly) are listed in Appendix S4 (Supporting information). We note that this data set is expected to slightly underestimate recombination rate along chromosomes. The reason is that with our marker resolution, a few tight double-crossovers may have escaped detection altogether, and a few others may have been captured by one or two markers but taken as genotyping error and eliminated. Moreover, our markers never covered the full physical chromosome span because of the randomness of *Sbf1* restriction sites; because we ignored unanchored scaffolds mapping to one or both ends of many chromosomes when these scaffolds were small and represented by only one to three markers; and because we discarded the peripheral marker on each end of the initially generated linkage groups.

Analysis of recombination

Genetic distances along the 20 autosomes were estimated in R/qtl using the Kosambi map function (assuming crossover interference) and the full F_2 panel. For the sex chromosome (chromosome 19; Peichel *et al.* 2004), final map construction used genotype data from females only ($N = 142$). The reason is that sequence degeneration of the Y relative to the X chromosome precluded reliable genotyping in males (the reference sex chromosome sequence is the X). R/qtl was also used to count the number of crossovers for each individual and chromosome.

We visualized recombination rate along the chromosomes by plotting genetic distance (cM) against physical

distance (Mb). Moreover, we calculated the average recombination rate for every interval between adjacent markers as the ratio of genetic distance to physical distance (cM/Mb) and plotted this rate against the physical midpoint of the marker interval. We also calculated average recombination rate across each chromosome, and across each chromosome arm, using for the latter information on centromere positions extracted from Urton *et al.* (2011). Throughout this paper, effective physical chromosome (and chromosome arm) spans are defined by the position of our most peripheral markers. The only exception is Appendix S1 (Supporting information) where we show the full physical chromosome lengths.

Crossover counts were used to examine the relationship between recombination frequency and chromosome length. We here determined for each chromosome the average crossover number across the 282 F_2 (or the 142 females for chromosome 19) and calculated the correlation coefficient r between this variable and chromosome length. The magnitude of this test statistic was evaluated against its empirical random distribution established by permuting the crossover data 9999 times (Manly 2007; all statistical tests in this study are based on analogous permutation tests). A similar analysis was performed by using chromosome arm length, rather than total chromosome length, as a predictor of crossover number. In this latter analysis, six chromosome arms with low marker coverage were excluded, which had a trivial influence on the results. Also, these analyses were performed with and without the sex chromosome. As this did not materially influence the results, we report the former.

Individual crossover counts across all autosomes were used to test for a difference in overall recombination rate between the sexes, using as test statistic the F-ratio of a linear model with crossover count as response and sex as fixed factor. Analogous tests were also performed to explore sex differences in crossover number for each chromosome separately. Finally, individual crossover counts were used to scan the genome (including chromosome 19) for the presence of quantitative trait loci (QTL) determining recombination rate. We emphasize that our data are not ideal for this purpose; quantifying the recombination phenotype in the F_2 generation would have required crossover data from the F_3 generation or from F_2 gametes. Our scan was thus limited to detecting QTL heterozygous in one or both of the grandparents. The QTL scan was performed in R/qtl using the extended Haley–Knott method (other methods produced very similar results). Significance of LOD peaks was established based on 9999 permutations, following Broman & Sen (2009).

Recombination and divergence within the sex chromosome

Recombination between the X and Y chromosomes was studied by determining which of the 69 SNPs along chromosome 19 occurred homozygous for the grandfather allele in F_2 females. This female genotype necessarily requires XY crossover in the F_1 father.

To explore degeneration of the Y chromosome, we haphazardly selected 100 males and 100 females from the F_2 population. For each sex separately, we determined for every RAD locus along chromosome 19 the total sequence coverage across all individual alignments. For each RAD locus, we then calculated the ratio of female to male coverage. A RAD locus not or little differentiated between the gametologs would display an expected ratio of one because both the X and Y sequences would align to the X reference. At a locus substantially diverged between X and Y, the latter would no longer align to the reference, producing twice the sequence coverage in females relative to males. To reduce noise, we restricted this analysis to RAD loci displaying a minimal total sequence coverage of 4000 in each sex, yielding a total of 1556 informative loci along the X chromosome (average intermarker distance: 13 kb). This analysis of Y degeneration was additionally performed by using a natural population sample from Europe (CHE) and Canada (Boot Lake, see below). These populations are derived independently from Atlantic and Pacific ancestors. Because here sample size was much smaller (N = 13–14 per sex and population), we used a minimal sequence coverage threshold of 50 per sex.

Genetic divergence, genetic diversity and GC content in relation to recombination rate

We tested the prediction of a negative genome-wide correlation between recombination rate and the magnitude of allele frequency shifts by using divergence data from two independent replicate lake–stream population pairs studied in Roesti *et al.* (2012a) (the Boot and Robert's pair; see also Berner *et al.* 2008, 2009). These young (postglacial, <10 000 years old) population pairs are those among the four pairs investigated in Roesti *et al.* (2012a) displaying the strongest divergence in phenotypes and genetic markers between the selectively distinct habitats (genome-wide median F_{ST} is 0.15 and 0.03 for Boot and Robert's; Roesti *et al.* 2012a). Each of the four samples was represented by 27 individuals (balanced sex ratio). Polymorphism data were generated through RAD sequencing, yielding 3930 and 7992 genome-wide SNPs for the Boot and Robert's pair (details on library preparation, sequencing, genotyping,

SNP detection and access to the raw data are given in Roesti *et al.* 2012a,b). The magnitude of divergence between the lake and stream population was quantified by F_{ST} based on haplotype diversity (Nei & Tajima 1981; formula 7), tolerating only informative SNPs with a minor allele frequency of 0.25 or greater (Roesti *et al.* 2012b). F_{ST} was then averaged across the intervals defined by adjacent markers from the mapping cross, resulting in the same resolution as our recombination rate data (see Fig. 1). This allowed us to explore the genome-wide correlation between the magnitude of divergence and recombination rate, using *r* as statistic for significance testing.

Next, we examined the prediction of a positive correlation between recombination rate and levels of genetic diversity within each population. For this, we screened each of the four population samples separately for polymorphisms and calculated genetic diversity (haplotype diversity, Nei & Tajima 1981; singletons were omitted to exclude technical artefacts). RAD loci were allowed to contribute a single SNP only, keeping the one with the highest diversity at loci with multiple SNPs (drawing a SNP at random produced very similar results). The resulting total number of SNPs varied between 4938 and 17 649 among the populations. As a complementary analysis, we also counted the number of polymorphisms (excluding singletons) on each RAD locus, with the number of RAD loci varying between 6440 and 25 186 among the populations. We considered these data, hereafter referred to as SNP density, a valuable alternative genetic diversity metric because selection should not only skew allele frequencies in linked regions, but also reduce the density of polymorphisms in those regions. Both the genetic diversity and SNP density data were averaged to the resolution of the genetic map and tested for an association with recombination rate as described for F_{ST}.

Finally, we investigated a possible association between recombination rate and GC content in an analogous way. However, to maximize precision, we calculated the proportion of GC nucleotides for each marker interval based on the full reference genome sequence rather than our RAD sequences. Moreover, we here detected a clear nonlinear relationship and therefore used as test statistic the ratio of residual to total sum of squares of a nonparametric regression (LOESS—robust locally weighted scatterplot smoothing; Cleveland 1979; a linear fit with *r* as test statistic produced similar results). We note that this analysis assumes that patterns of nucleotide composition in the reference genome, which was built based on a Pacific-derived freshwater stickleback, are also representative of Atlantic-derived European populations. This assumption is justified; repeating the correlation analysis using genome-wide GC content estimated from consensus sequences at 27 396 RAD loci derived from the cross grandmother produced similar results (details not presented).

In the above correlation analyses (F_{ST}, genetic diversity, SNP density, GC content), marker intervals with an extreme recombination rate (below 0.01 and above 40) were excluded, although analyses including *all* intervals produced very similar results. The final data sets thus comprised 1783 genome-wide marker intervals. Also, including or excluding the sex chromosome did not materially influence the analyses; we thus report the former. Apart from sequence alignment and BAM conversion, all analyses and plotting were carried out in the R language (R Development Core Team 2012), benefiting greatly from the Bioconductor packages ShortRead (Morgan *et al.* 2009), Rsamtools and Biostrings. Data smoothing was performed with R's implementation of LOESS.

Results

The 21 stickleback chromosomes accounted for a total genetic map length of 1251 cM, yielding a genome-wide average recombination rate of 3.11 cM/Mb (this number is based on the total physical genome length effectively covered by our markers: 401.8 Mb). However, recombination rate proved highly heterogeneous across the genome: crossovers occurred primarily in the chromosome peripheries, with a greatly reduced rate in the chromosome centres (Fig. 1). Except for two of the smallest chromosomes (5, 21), this pattern was consistent and was particularly pronounced in the larger ones. For instance, the average recombination rate in the first and last 5 Mb of the largest chromosome (4) was 7.8 and 6.8 cM/Mb, whereas the segment ranging from 10 to 25 Mb exhibited an *c.* 20-fold lower rate (0.4 cM/Mb). The general pattern of periphery-biased recombination proved essentially insensitive to centromere position (e.g. compare chromosomes 7 and 8 in Fig. 1). Our data also suggested a tendency for the recombination rate to drop again right at the chromosome ends (e.g. chromosomes 1, 2, 4, 17 in, Fig. 1). Formally testing this observation, however, would have required higher-resolution data, sampling the terminal domains more densely.

Comparing mean crossover number per meiosis among chromosomes revealed a lower limit of approximately one crossover for the chromosomes at the lower end of the size range (around 15 Mb) (Fig. 2A). With increasing chromosome length, the crossover number also increased ($r = 0.92$, permutation $P = 0.0001$), with the largest chromosomes (around 30 Mb) displaying *c.* 1.5 crossovers per meiosis. We also found a positive

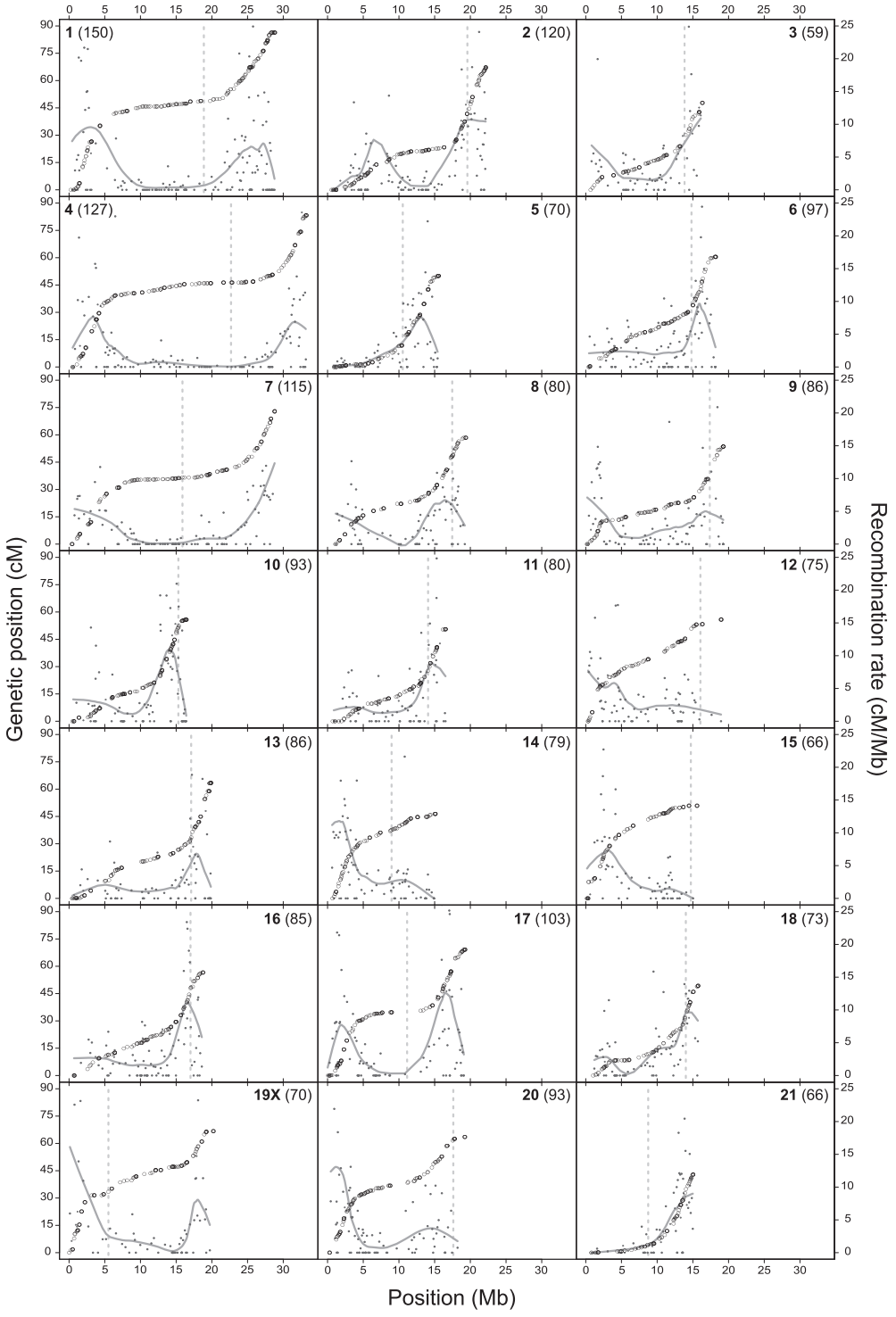

Fig. 1 Recombination along the 20 threespine stickleback autosomes (based on 1872 total markers and 282 F_2 individuals) and along the X chromosome (chromosome 19; based on 142 females). Marker number per chromosome is given in parentheses. The open circles (referring to the left axis) indicate the genetic map position of the markers in Kosambi centimorgan, plotted against their physical position in the genome (in megabases). The smaller grey dots (right axis) represent the average recombination rate in cM/Mb for the intervals defined by pairs of adjacent markers, plotted against the intervals' physical midpoint. The grey curves show the latter data smoothed by LOESS, with a polynomial degree of one and the smoothing span decreasing from 0.33 to 0.149 from the smallest to the largest chromosome to ensure a constant smoothing resolution across the panels. Dashed vertical lines specify centromere positions. Note the striking trend towards elevated recombination rate in the peripheral chromosome regions.

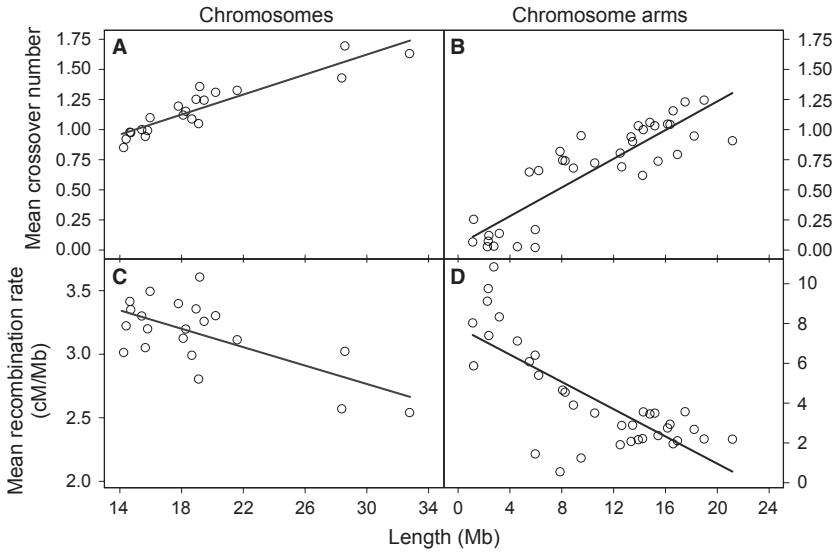

Fig. 2 Stickleback chromosomes display an approximate minimal crossover number of one per meiosis, and crossover number is related positively to chromosome length (A). A similar positive relationship exists between crossover number and chromosome *arm* length (B). Because the increase in crossover number with increasing chromosome length is not proportional, longer chromosomes display a lower average recombination rate (C). The same holds for chromosome arms (D). Note that (C) and (D) have different scales on the *Y*-axis.

association between chromosome *arm* length and crossover number ($r = 0.87$, $P = 0.0001$) (Fig. 2B). Along the short arms of telocentric and acrocentric chromosomes, crossovers occurred rarely. These relationships caused the average recombination rate to be higher on short chromosomes and chromosome arms than on longer ones (Fig. 2C, 2D; chromosomes: $r = -0.66$, $P = 0.0026$; arms: $r = -0.76$, $P = 0.0001$).

We found no indication of an overall difference in recombination rate between the sexes ($P = 0.982$); total autosomal map length was almost identical (1190 cM) for males and females. Analysing each chromosome separately also revealed only trivial sex-related differences in recombination rate (none of them remained significant when correcting for multiple testing). We further detected no significant QTL driving overall recombination rate on any of the 21 chromosomes (maximum LOD = 2.98; $P = 0.261$), keeping in mind the methodological limitations mentioned above.

Recombination and degeneration along the sex chromosome

In our F_2 population, XY recombination never occurred beyond the marker located at 1.75 Mb (Fig. 3). We thus demonstrate the presence of a small pseudoautosomal region (PAR), spanning *c.* 10% of the entire X chromosome only. The comparison of female with male RAD sequence coverage along the sex chromosome clearly revealed Y degeneration outside the PAR (Fig. 3). Moreover, the extent of degeneration was not uniform outside the PAR: within the segment from *c.* 12 Mb to the chromosome end opposed to the PAR, Y sequences generally did not align to the X reference. By contrast, the segment ranging from *c.* 2 to 12 Mb showed weaker degeneration. Despite small sample size and hence more random noise, analogous analyses in the two natural populations produced very similar results supporting identical conclusions (Appendix S5, Supporting information).

Genetic divergence, genetic diversity and GC content

The prediction of a negative genome-wide association between recombination rate and F_{ST} was clearly confirmed (Table 1; Fig. 4A). Shifts in allele frequencies between populations were thus greater in low-recombination regions. As expected, this effect was stronger in the Boot lake–stream pair showing greater overall divergence (and hence higher variance in F_{ST}) than the Robert's pair (see Roesti *et al.* 2012a). The two complementary analyses of genetic diversity within populations also agreed with the prediction: all correlations were positive and generally highly significant (Table 1; Fig. 4B). Genetic diversity was thus reduced in marker intervals exhibiting a relatively low recombination rate. Finally, we found a striking positive broad-scale association between recombination rate and GC content across the genome (Fig. 5; $P = 0.0001$). Marker intervals showing relatively high recombination rates (around 10 cM/Mb or greater) displayed an *c.* 10% higher average GC content than intervals at the lower end of recombination rates (around 1 cM/Mb or lower).

Discussion

A major finding of our analysis of recombination in the threespine stickleback genome is the strong bias of crossover to occur primarily in the chromosome peripheries. This pattern confirms preliminary evidence from

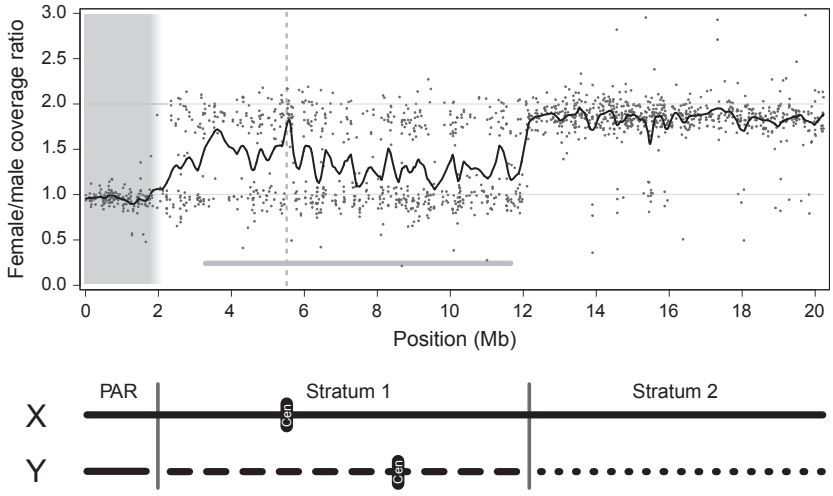

Fig. 3 Patterns of recombination and divergence between the X and Y chromosome. In the top panel, the abscissa gives the physical position along the reference X chromosome. The centromere is indicated by the dashed grey vertical line. The domain on the left shaded in grey indicates the extent of the pseudoautosomal region (PAR) where the gametologs still recombined in our cross (the PAR boundary lies between 1.75 and 2.22 Mb). The dots show the ratio of female to male sequence coverage across 100 individuals per sex for 1556 RAD loci (the black curve shows these data smoothed; degree = 1, span = 0.025). Within the PAR, the coverage ratio approximates unity (lower grey horizontal line), as expected for a DNA segment homochromatic between X and Y. Outside the PAR, many RAD loci display twofold higher sequence coverage in females than males (upper grey horizontal line), consistent with strong degeneration or loss of the X sequence on the Y. Note that two levels of Y degeneration ('evolutionary strata') are indicated (abutting at 12 Mb), the left one coinciding with the minimal size of a pericentric inversion on the Y inferred by Ross & Peichel (2008; visualized as heavy grey horizontal bar). On the bottom, we present the patterns of XY divergence inferred from our data in schematic form. Highly consistent patterns were also found when analysing natural population samples from Europe and Canada (see Appendix S5, Supporting information).

Table 1 Genome-wide associations between recombination rate and genetic population divergence, and between recombination rate and within-population genetic diversity. Divergence was quantified as F_{ST} between the lake and stream sample within the Boot and Robert's population pair. Genetic diversity within each of the four populations was expressed as both haplotype diversity (capturing allele frequency shifts) and the density of single nucleotide polymorphisms (SNPs) per RAD locus. All these metrics were averaged within the physical intervals defined by adjacent markers in the SNP panel used for genetic mapping (N = 1783 intervals)

	Population(s)	r	P
Genetic divergence (F_{ST})	Boot lake—stream	−0.2699	0.0001
	Robert's lake—stream	−0.1127	0.0001
Haplotype diversity	Boot lake	0.1184	0.0001
	Boot stream	0.0925	0.0022
	Robert's lake	0.0400	0.0929
	Robert's stream	0.0665	0.0113
SNP density	Boot lake	0.1873	0.0001
	Boot stream	0.1810	0.0001
	Robert's lake	0.2593	0.0001
	Robert's stream	0.1882	0.0001

low-resolution analyses in the species (Hohenlohe *et al.* 2012; Roesti *et al.* 2012a) and is consistent with results from genome-wide recombination studies in other vertebrates (Borodin *et al.* 2008; Chowdhury *et al.* 2009; Backström *et al.* 2010; Wong *et al.* 2010; Bradley *et al.* 2011; Auton *et al.* 2012; Sandor *et al.* 2012; Tortereau *et al.* 2012), invertebrates (Rockman & Kruglyak 2009; Niehuis *et al.* 2010), plants (Akhunov *et al.* 2003; Anderson *et al.* 2003; but see Salomé *et al.* 2012) and yeast (Barton *et al.* 2008). This striking consistency across taxa implies a common mechanistic basis: crossovers seem to be initiated from the peripheries. Indeed, peripheral clustering of chromosomes during the meiotic prophase I is believed to play a key role in proper homolog pairing and probably also in crossover initiation (Scherthan *et al.* 1996; Roeder 1997; Harper *et al.* 2004; Brown *et al.* 2005; Naranjo & Corredor 2008). Peripheral crossover might also favour proper homolog dissociation (Colombo & Jones 1997; Hassold & Hunt 2001). Whatever the exact cause, the observed periphery bias in the distribution of crossovers in the stickleback genome (and many other genomes) implies strong mechanistic constraints on the distribution of recombination. Therefore, genetic information is reshuffled much more effectively in some genomic regions than in others.

Moreover, taking into account a slight underestimation of recombination (see Materials and methods), our data indicate that stickleback chromosomes display at least one crossover per meiosis. This is consistent with

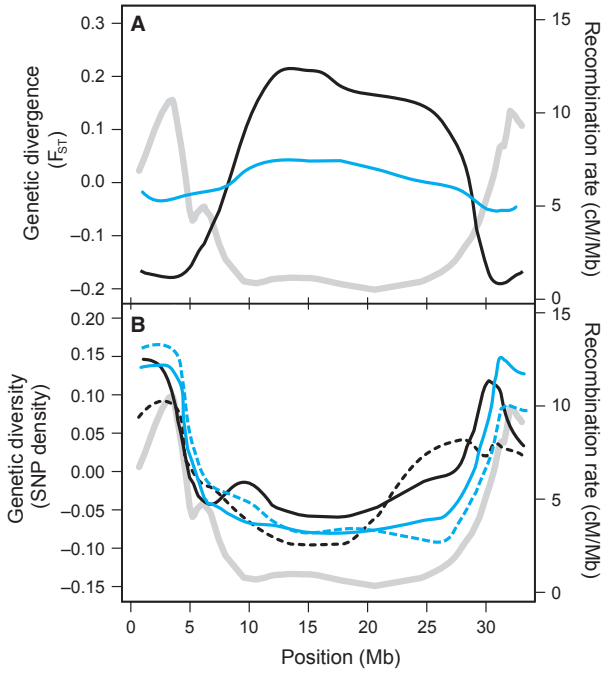

Fig. 4 Genetic divergence and genetic diversity in relation to recombination rate (shown as heavy grey line referring to the right axis in both panels) in natural lake and stream stickleback populations, exemplified for the largest chromosome (4). In (A), we show the smoothed (degree = 0, span = 0.35) magnitude of lake–stream divergence (F_{ST}) for the Boot (black) and Robert's (blue) population pair (for the sake of clarity, the underlying raw data points are not shown). To facilitate comparison, the data were centred to a mean of zero before smoothing. Note that divergence is greatest in the chromosome centre where recombination rate is lowest, an effect more pronounced in the Boot population pair showing much stronger overall divergence. In (B), we display smoothed genetic diversity, quantified as single nucleotide polymorphism density, for the lake (solid line) and stream (dashed line) population in the Boot (black) and Robert's (blue) population pair (data also centred). Note the strong and consistent positive association between genetic diversity and recombination rate.

the notion that one crossover per chromosome and meiosis is generally required for proper homolog segregation (Mather 1938; Baker *et al.* 1976; Roeder 1997; Smith & Nicolas 1998; Hassold & Hunt 2001), and reflects another mechanistic constraint on recombination. The widely accepted idea of one obligate crossover per chromosome *arm*, however, is not supported by our data (see also Borodin *et al.* 2008; Fledel-Alon *et al.* 2009): on acrocentric and telocentric stickleback chromosomes, the shorter arm rarely crosses over. We further find that the number of crossovers beyond one is a function of chromosome length. Standardized by their length, however, large chromosomes still exhibit a lower recombination rate than small chromosomes, the same also being true for chromosome arms.

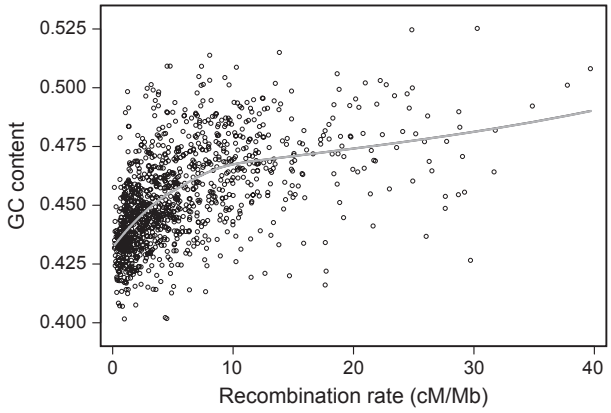

Fig. 5 Genome-wide relationship between GC content and recombination rate. The data points are the intervals defined by adjacent markers in the single nucleotide polymorphism panel used for genetic mapping ($N = 1783$ intervals). GC content was calculated for each marker interval by using information from the stickleback reference genome. The grey line shows the smoothed data (degree = 2, span = 0.85; a standard correlation produces $r = 0.49$).

Sex chromosome evolution

Sex chromosomes are generally thought to evolve from an ordinary pair of homologous autosomes that partly stop crossing over to prevent alleles at loci with sexually antagonistic effects from recombining (Bull 1983; Charlesworth & Charlesworth 2000; Charlesworth *et al.* 2005; Wilson & Makova 2009). This cessation of recombination should initiate the differentiation of the gametologs. While early karyotypic investigations in threespine stickleback found no evidence of heteromorphic sex chromosomes (Chen & Reisman 1970; Cuñado *et al.* 2002), recent investigations have indicated reduced recombination, chromosomal rearrangements and sequence divergence between the X and Y (Peichel *et al.* 2004; Ross & Peichel 2008; Shikano *et al.* 2011; Natri *et al.* 2013). These observations, based on a small number of markers, are greatly refined and extended by our sex chromosome analysis. We confirm that XY recombination is restricted to a small PAR, as suggested by Ross & Peichel (2008). The requirement of at least one crossover per meiosis thus implies a very high average recombination rate (*c.* 25 cM/Mb) across the PAR in males. This agrees with the estimation by Peichel *et al.* (2004) of a much greater distance between markers lying within the PAR in male than in female genetic maps (e.g. the genetic distance in the Paxton cross between the microsatellites Stn303 and Stn186, located at 0.4 and 1.9 Mb, is 27.3 cM in females and 47.7 cM in males).

A consequence of the cessation of recombination along most of the sex chromosome is that the region on the Y outside the PAR occurs in permanently

heterozygous state and at lower population size than the X. Both conditions are predicted to make selection on the Y less effective and hence to promote its degeneration (Felsenstein 1974; Charlesworth & Charlesworth 2000; Otto & Lenormand 2002; Charlesworth *et al.* 2005; Wilson & Makova 2009). Our results strongly support this view: outside the PAR, RAD loci often display only half the sequence coverage in males relative to females, consistent with substantial sequence degeneration (or loss) on the Y. Interestingly, our analysis further indicates two discrete levels of Y degeneration, with much stronger degeneration along the *c.* 8 Mb towards the chromosome end opposed to the PAR than along the *c.* 10 Mb adjacent to the PAR. Such 'evolutionary strata' (Lahn & Page 1999) have been found in mammals (Lahn & Page 1999; Sandstedt & Tucker 2004; Pearks Wilkerson *et al.* 2008), birds (Lawson Handley *et al.* 2004; Nam & Ellegren 2008) and plants (Bergero *et al.* 2007; Wang *et al.* 2012). To our knowledge, we here provide the first evidence for evolutionary strata in fish.

Evolutionary strata are generally taken as evidence that XY recombination ceased simultaneously across large domains of the evolving sex chromosome. An obvious way how this may happen is through chromosomal inversion. Indeed, a recent study using *in situ* fluorescent hybridization argued for a large pericentric inversion on the Y relative to the X, with breakpoints at *c.* 3 and 12 Mb (Ross & Peichel 2008). The evolutionary stratum adjacent to the PAR identified in our work matches this inversion almost perfectly and allows us to refine its physical boundaries. Threespine stickleback thus reinforce the view that recombination suppression along evolving sex chromosomes will primarily occur through inversion rather than crossover rate modifier genes (Charlesworth *et al.* 2005).

It would now be interesting to date the two bouts of recombination suppression underlying the evolutionary strata in the species based on sequence divergence between homologous loci on the X and Y (Lahn & Page 1999; Lawson Handley *et al.* 2004; Nam & Ellegren 2008; Pearks Wilkerson *et al.* 2008; Wang *et al.* 2012). We note that this might be difficult for stratum 2 if the remarkably strong degeneration detected in our analysis actually reflects deletion (Peichel *et al.* 2004). Clearly, however, patterns of XY divergence were already established prior to the split into Pacific and Atlantic stickleback clades (Appendix S5, Supporting information).

Consequences of heterogeneous recombination rate on genome evolution

The rate of recombination within a genomic region determines to which extent selection on a locus influences allele frequencies at neutral loci, and interferes with selection on other loci, in its physical neighbourhood (Hill & Robertson 1966; Maynard Smith & Haigh 1974; Barton & Bengtsson 1986; Kaplan *et al.* 1989; Begun & Aquadro 1992; Nordborg *et al.* 1996; Charlesworth *et al.* 1997; Charlesworth 1998; Nachman 2002; Feder & Nosil 2010). Several types of selection (divergent, positive and background) should therefore increase divergence among populations and reduce genetic diversity within populations in low-recombination genomic regions relative to regions where recombination rate is higher.

Consistent with these predictions, we have recently shown that the magnitude of divergence between neighbouring lake and stream stickleback populations is dramatically biased towards chromosome centres (Roesti *et al.* 2012a). (Note that divergence in these young populations essentially reflects differential sorting of standing variation rather than novel mutations.) Using robust recombination rate data, we here demonstrate that elevated divergence in these population pairs is related to reduced recombination. Because lake and stream stickleback occupy selectively distinct environments (Berner *et al.* 2008, 2009), the divergence–recombination association almost certainly arises from within-chromosome variation in hitchhiking and/or introgression.

The present study further demonstrates reduced within-population genetic diversity in the chromosome centres relative to the peripheries, resulting in a genome-wide positive correlation between diversity and recombination rate. A similar correlation has previously been reported in a broad range of organisms (Begun & Aquadro 1992; Kraft *et al.* 1998; Nachman 2001; Tenaillon *et al.* 2001; Takahashi *et al.* 2004; Roselius *et al.* 2005; McGaugh *et al.* 2012). Given that a positive correlation between recombination rate and genetic diversity may also arise if recombination is directly mutagenic (Spencer *et al.* 2006; Webster & Hurst 2012; but see McGaugh *et al.* 2012), caution is generally warranted when inferring from the above correlation that recombination rate modulates the influence of selection on linked sites across the genome. In our lake–stream stickleback systems, however, the colocalization of elevated population divergence and reduced genetic diversity within young populations residing in selectively distinct environments provides clear support for such an indirect influence of recombination on genome evolution (see also Stoelting *et al.* 2013). The precise selective processes driving these patterns, however, remain to be elucidated.

In addition to these indirect (selective) effects, our study perhaps also points to a direct effect of recombination on stickleback genome evolution: large-scale bias in nucleotide composition. Across the genome, GC content is higher in regions displaying relatively elevated

recombination rate—that is, in the chromosome peripheries. Interestingly, the positive association between GC content and recombination rate seems as widespread as periphery bias in recombination rate; it has been reported in mammals (Jensen-Seaman *et al.* 2004; Spencer *et al.* 2006; Duret & Arndt 2008; Auton *et al.* 2012; Tortereau *et al.* 2012), birds (ICGSC 2004; Backström *et al.* 2010), insects (Niehuis *et al.* 2010; Stevison & Noor 2010; but see Comeron *et al.* 2012), plants (Muyle *et al.* 2011) and yeast (Gerton *et al.* 2000; Birdsell 2002). As hypothesized in other organisms, elevated GC content in the stickleback genome might represent a direct causal consequence of elevated recombination rate, given evidence of GC bias in the machinery correcting nucleotide mismatch in heteroduplex DNA formed during crossover initiation (GC-biased gene conversion; Brown & Jiricny 1987; Bill *et al.* 1998; Galtier *et al.* 2001; Birdsell 2002; Meunier & Duret 2004; Mancera *et al.* 2008; Duret & Galtier 2009; Muyle *et al.* 2011). Our correlational data, however, cannot address this causal hypothesis conclusively; direct experimental evidence is needed.

Methodological implications

In addition to the above influences on genome evolution, heterogeneous recombination rate within the genome has important methodological implications. Marker-based genome scans searching for signatures of divergent selection in the form of locally elevated divergence between ecologically distinct populations (Lewontin & Krakauer 1973; Beaumont & Nichols 1996; Luikart *et al.* 2003; Beaumont 2005; Nielsen 2005; Storz 2005) are becoming commonplace. What is generally ignored is that the distortion between physical and genetic maps will dilute the link between the selection coefficient on a locus and the magnitude of hitchhiking produced in its neutral neighbourhood (Roesti *et al.* 2012a; this study). In other words, a locus under selection is more likely to be detected when located in a low-recombination region where hitchhiking is more extensive. This bias should increase with decreasing marker resolution and with increasing sliding window size. The generality of chromosome periphery-biased recombination rate across taxa therefore raises a potential caveat to the interpretation of differentiation outliers in genome scans when combined physical and genetic map information is missing (i.e. 'anonymous' approaches; for one strategy to alleviate this difficulty when a physical map is available, see Roesti *et al.* 2012a). An analogous issue arises when interpreting the number and effect size of mapped QTL: within low-recombination regions, multiple loci of small effect are more likely to emerge as a single large-effect locus (Noor *et al.* 2001).

Finally, our study highlights the need for a reliably assembled genome for investigations of recombination and linkage. Assembly errors will inflate the genome-wide average crossover frequency, distort the recombination landscape and bias analyses of linkage disequilibrium along chromosomes. For instance, we find that a high-recombination island on chromosome 4 inferred in Hohenlohe *et al.* (2012; Fig. 2b in that study) coincides with the boundary of a scaffold anchored in the wrong sense within that chromosome (see our Appendix S1, Supporting information) and hence represents an artefact. The same assembly error also mimics long-distance linkage disequilibrium along this chromosome (Fig. 4a in Hohenlohe *et al.* 2012).

To summarize, our analysis of recombination in threespine stickleback indicates strong constraints on the frequency and location of crossovers imposed by the mechanistic requirements of meiosis. At the same time, we demonstrate that recombination influences the genome profoundly, both by modulating the consequences of selection across the genome and perhaps by directly influencing nucleotide composition. We anticipate that our characterization of the recombination landscape will facilitate interpretations of genome scans and QTL mapping in the species, promote further investigations on sex chromosome evolution and pave the way for more detailed investigations of the determinants and consequences of recombination.

Acknowledgements

The following contributions made this study possible and are most gratefully acknowledged: F. Hofmann (SFFN—Inspection de la pêche VD) and R. Kistler (fisheries authorities of the canton Thurgau) provided sampling permits for the cross populations, and A.-C. Grandchamp aided sampling. A. Hendry supported sampling of the field populations financially. B. Cresko and C. Peichel provided input on the cross and marker generation design at an early stage. B. Egger, H. Gante, A. Indermaur, W. Salzburger, A. Theys and P. Vonlanthen helped rear the mapping cross. W. Salzburger generously shared laboratory resources and infrastructure. B. Aeschbach and N. Boileau facilitated wet laboratory work. Illumina sequencing was performed by C. Beisel and I. Nissen at the Quantitative Genomics Facility, D-BSSE, ETH Zürich. J. Urton shared information on centromere positions, and M. Noor and two other reviewers provided valuable suggestions to improve the manuscript. MR was supported by a Swiss National Science Foundation (SNF) Sinergia grant (CRSII3_136293) to M. Sanchez, H. Furrer, and W. Salzburger. DB was supported by a SNF Ambizione fellowship (PZ00P3_126391/1) and by the Research Fund of the University of Basel.

References

Akhunov ED, Goodyear AW, Geng S *et al.* (2003) The organization and rate of evolution of wheat genomes are correlated

with recombination rates along chromosome arms. *Genome Research*, **13**, 753–763.
Albert AYK, Sawaya S, Vines TH *et al.* (2008) The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution*, **62**, 76–85.
Anderson LK, Doyle GG, Brigham B *et al.* (2003) High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. *Genetics*, **165**, 849–865.
Auton A, Fledel-Alon A, Pfeifer S *et al.* (2012) A fine-scale chimpanzee genetic map from population sequencing. *Science*, **336**, 193–198.
Backström N, Forstmeier W, Schielzeth H *et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research*, **20**, 485–495.
Baird NA, Etter PD, Atwood TS, *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, **3**, e3376.
Baker BS, Carpenter ATC, Esposito MS, Esposito RE, Sandler L (1976) The genetic control of meiosis. *Annual Review of Genetics*, **10**, 53–134.
Barton N, Bengtsson BO (1986) The barrier to genetic exchange between hybridizing populations. *Heredity*, **57**, 357–376.
Barton AB, Pekosz MR, Kurvathi RS, Kaback DB (2008) Meiotic recombination at the ends of chromosomes in *Saccharomyces cerevisiae*. *Genetics*, **179**, 1221–1235.
Beaumont MA (2005) Adaptation and speciation: what can F-st tell us? *Trends in Ecology & Evolution*, **20**, 435–440.
Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **263**, 1619–1626.
Begun DJ, Aquadro CF (1992) Levels of naturally occurring DNA polymorphism correlate with recombination rates in *Drosophila melanogaster*. *Nature*, **356**, 519–520.
Bergero R, Forrest A, Kamau E, Charlesworth D (2007) Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. *Genetics*, **175**, 1945–1954.
Berner D, Adams DC, Grandchamp A-C, Hendry AP (2008) Natural selection drives patterns of lake-stream divergence in stickleback foraging morphology. *Journal of Evolutionary Biology*, **21**, 1653–1665.
Berner D, Grandchamp A-C, Hendry AP (2009) Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. *Evolution*, **63**, 1740–1753.
Berner D, Roesti M, Hendry AP, Salzburger W (2010) Constraints on speciation suggested by comparing lake-stream stickleback divergence across two continents. *Molecular Ecology*, **19**, 4963–4978.
Bill CA, Duran WA, Miselis NR, Nickoloff JA (1998) Efficient repair of all types of single-base mismatches in recombination intermediates in Chinese hamster ovary cells: competition between long-patch and G-T glycosylase-mediated repair of G-T mismatches. *Genetics*, **149**, 1935–1943.
Birdsell JA (2002) Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. *Molecular Biology and Evolution*, **19**, 1181–1197.
Borodin PM, Karamysheva TV, Belonogova NM *et al.* (2008) Recombination map of the common shrew, *Sorex araneus* (eulipotyphla, mammalia). *Genetics*, **178**, 621–632.
Bradley KM, Breyer JP, Melville DB, Broman KW, Knapik EW, Smith JR (2011) An SNP-based linkage map for zebrafish reveals sex determination loci. *G3: Genes, Genomes, Genetics*, **1**, 3–9.
Broman KW, Sen S (2009) *A Guide to QTL Mapping With R/qtl*. Springer, New York City, New York.
Brown TC, Jiricny J (1987) A specific mismatch repair event protects mammalian cells from loss of 5-methylcytosine. *Cell*, **50**, 945–950.
Brown PW, Judis LA, Chan ER *et al.* (2005) Meiotic synapsis proceeds from a limited number of subtelomeric sites in the human male. *American Journal of Human Genetics*, **77**, 556–566.
Bull JJ (1983) *Evolution of Sex Determining Mechanisms*. Benjamin Cummings, Menlo Park, CA.
Burt A (2000) Sex, recombination, and the efficacy of selection – was Weismann right? *Evolution*, **54**, 337–351.
Charlesworth B (1998) Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution*, **15**, 538–543.
Charlesworth B, Charlesworth D (2000) The degeneration of Y chromosomes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **355**, 1563–1572.
Charlesworth B, Nordborg M, Charlesworth D (1997) The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genetical Research*, **70**, 155–174.
Charlesworth D, Charlesworth B, Marais G (2005) Steps in the evolution of heteromorphic sex chromosomes. *Heredity*, **95**, 118–128.
Chen TR, Reisman HM (1970) Comparative chromosome study of North American species of sticklebacks (Teleostei – Gasterosteidae). *Cytogenetics*, **9**, 321–332.
Chowdhury R, Bois PRJ, Feingold E, Sherman SL, Cheung VG (2009) Genetic analysis of variation in human meiotic recombination. *PLoS Genetics*, **5**, e1000648.
Cleveland WS (1979) Robust locally weighted regression and smoothing scatterplots. *Journal of American Statistical Association*, **74**, 829–836.
Colombo PC, Jones GH (1997) Chiasma interference is blind to centromeres. *Heredity*, **79**, 214–227.
Comeron JM, Ratnappan R, Bailin S (2012) The many landscapes of recombination in *Drosophila melanogaster*. *PLoS Genetics*, **8**, e1002905.
Cuñado N, Barrios J, San Miguel E *et al.* (2002) Synaptonemal complex analysis in oocytes and spermatocytes of threespine stickleback *Gasterosteus aculeatus* (Teleostei, Gasterosteidae). *Genetica*, **114**, 53–56.
Duret L, Arndt PF (2008) The impact of recombination on nucleotide substitutions in the human genome. *PLoS Genetics*, **4**, e1000071.
Duret L, Galtier N (2009) Biased gene conversion and the evolution of mammalian genomic landscapes. *Annual Review of Genomics and Human Genetics*, **10**, 285–311.
Feder JL, Nosil P (2010) The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution*, **64**, 1729–1747.
Felsenstein J (1974) The evolutionary advantage of recombination. *Genetics*, **78**, 737–756.
Fledel-Alon A, Wilson DJ, Broman KW *et al.* (2009) Broad-scale recombination patterns underlying proper disjunction in humans. *PLoS Genetics*, **5**, e1000658.

Galtier N, Piganeau G, Mouchiroud D, Duret L (2001) GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. *Genetics*, **159**, 907–911.

Gerton JL, DeRisi J, Shroff R *et al.* (2000) Global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 11383–11390.

Greenwood AK, Jones FC, Chan YF *et al.* (2011) The genetic basis of divergent pigment patterns in juvenile threespine sticklebacks. *Heredity*, **107**, 155–166.

Harper L, Golubovskaya I, Cande WZ (2004) A bouquet of chromosomes. *Journal of Cell Science*, **117**, 4025–4032.

Hassold T, Hunt P (2001) To err (meiotically) is human: the genesis of human aneuploidy. *Nature Reviews Genetics*, **2**, 280–291.

Hill WG, Robertson A (1966) Effect of linkage on limits to artificial selection. *Genetical Research* **8**, 269–294.

Hohenlohe PA, Bassham S, Currey M, Cresko WA (2012) Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **367**, 395–408.

ICGSC (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, **432**, 695–716.

Jensen-Seaman MI, Furey TS, Payseur BA *et al.* (2004) Comparative recombination rates in the rat, mouse, and human genomes. *Genome Research*, **14**, 528–538.

Jones FC, Grabherr MG, Chan YF *et al.* (2012) The genomic basis of adaptive evolution in three spine sticklebacks. *Nature*, **484**, 55–61.

Kaplan NL, Hudson RR, Langley CH (1989) The hitchhiking effect revisited. *Genetics*, **123**, 887–899.

Kraft T, Sall T, Magnusson-Rading I, Nilsson NO, Hallden C (1998) Positive correlation between recombination rates and levels of genetic variation in natural populations of sea beet (*Beta vulgaris* subsp. *maritima*). *Genetics*, **150**, 1239–1244.

Lahn BT, Page DC (1999) Four evolutionary strata on the human X chromosome. *Science*, **286**, 964–967.

Lawson Handley L-J, Ceplitis H, Ellegren H (2004) Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. *Genetics*, **167**, 367–376.

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of theory of selective neutrality of polymorphisms. *Genetics*, **74**, 175–195.

Li H, Handsaker B, Wysoker A *et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.

Mancera E, Bourgon R, Brozzi A, Huber W, Steinmetz LM (2008) High-resolution mapping of meiotic crossovers and non-crossovers in yeast. *Nature*, **454**, 479–485.

Manly BFJ (2007) *Randomization, Bootstrap and Monte Carlo Methods in Biology*, 3rd edn. Chapman & Hall, Boca Raton.

Mather K (1938) Crossing-over. *Biological Reviews*, **13**, 252–292.

Maynard Smith J, Haigh J (1974) Hitch-hiking effect of a favorable gene. *Genetical Research*, **23**, 23–35.

McGaugh SE, Heil CSS, Manzano-Winkler B *et al.* (2012) Recombination modulates how selection affects linked sites in *Drosophila*. *PLoS Biology*, **10**, e1001422.

Meunier J, Duret L (2004) Recombination drives the evolution of GC-content in the human genome. *Molecular Biology and Evolution*, **21**, 984–990.

Morgan M, Anders S, Lawrence M *et al.* (2009) ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics*, **25**, 2607–2608.

Moser D, Roesti M, Berner D (2012) Repeated lake-stream divergence in stickleback life history within a Central European lake basin. *PLoS ONE*, **7**, e50620.

Muyle A, Serres-Giardi L, Ressayre A, Escobar J, Glemin S (2011) GC-biased gene conversion and selection affect GC content in the *Oryza* genus (rice). *Molecular Biology and Evolution*, **28**, 2695–2706.

Nachman MW (2001) Single nucleotide polymorphisms and recombination rate in humans. *Trends in Genetics*, **17**, 481–485.

Nachman MW (2002) Variation in recombination rate across the genome: evidence and implications. *Current Opinion in Genetics & Development*, **12**, 657–663.

Nam K, Ellegren H (2008) The chicken (*Gallus gallus*) Z chromosome contains at least three nonlinear evolutionary strata. *Genetics*, **180**, 1131–1136.

Naranjo T, Corredor E (2008) Nuclear architecture and chromosome dynamics in the search of the pairing partner in meiosis in plants. *Cytogenetic and Genome Research*, **120**, 320–330.

Natri HM, Shikano T, Merilä J (2013) Progressive recombination suppression and differentiation in recently evolved neo-sex chromosomes. *Molecular Biology and Evolution*, in press.

Nei M, Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. *Genetics*, **97**, 145–163.

Niehuis O, Gibson JD, Rosenberg MS *et al.* (2010) Recombination and its impact on the genome of the haplodiploid parasitoid wasp *Nasonia*. *PLoS ONE*, **5**, e8597.

Nielsen R (2005) Molecular signatures of natural selection. *Annual Review of Genetics*, **39**, 197–218.

Noor MAF, Cunningham AL, Larkin JC (2001) Consequences of recombination rate variation on quantitative trait locus mapping studies: simulations based on the *Drosophila melanogaster* genome. *Genetics*, **159**, 581–588.

Nordborg M, Charlesworth B, Charlesworth D (1996) The effect of recombination on background selection. *Genetical Research*, **67**, 159–174.

Otto SP, Barton NH (1997) The evolution of recombination: removing the limits to natural selection. *Genetics*, **147**, 879–906.

Otto SP, Lenormand T (2002) Resolving the paradox of sex and recombination. *Nature Reviews Genetics*, **3**, 252–261.

Pearks Wilkerson AJ, Raudsepp T, Graves T *et al.* (2008) Gene discovery and comparative analysis of X-degenerate genes from the domestic cat Y chromosome. *Genomics*, **92**, 329–338.

Peichel CL, Ross JA, Matson CK *et al.* (2004) The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Current Biology*, **14**, 1416–1424.

R Development Core Team (2012) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Austria, Vienna.

Rockman MV, Kruglyak L (2009) Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS Genetics*. **5**, e1000419.

Roeder GS (1997) Meiotic chromosomes: it takes two to tango. *Genes & Development*, **11**, 2600–2621.

Roesti M, Hendry AP, Salzburger W, Berner D (2012a) Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. *Molecular Ecology*, **21**, 2852–2862.

Roesti M, Salzburger W, Berner D (2012b) Uninformative polymorphisms bias genome scans for signatures of selection. *BMC Evolutionary Biology*, **12**, 94.

Roselius K, Stephan W, Stadler T (2005) The relationship of nucleotide polymorphism, recombination rate and selection in wild tomato species. *Genetics*, **171**, 753–763.

Ross JA, Peichel CL (2008) Molecular cytogenetic evidence of rearrangements on the Y chromosome of the threespine stickleback fish. *Genetics*, **179**, 2173–2182.

Salomé PA, Bomblies K, Fitz J *et al.* (2012) The recombination landscape in *Arabidopsis thaliana* F2 populations. *Heredity*, **108**, 447–455.

Sandor C, Li W, Coppieters W *et al.* (2012) Genetic variants in REC8, RNF212, and PRDM9 influence male recombination in cattle. *PLoS Genetics*, **8**, e1002854.

Sandstedt SA, Tucker PK (2004) Evolutionary strata on the mouse X chromosome correspond to strata on the human X chromosome. *Genome Research*, **14**, 267–272.

Scherthan H, Weich S, Schwegler H *et al.* (1996) Centromere and telomere movements during early meiotic prophase of mouse and man are associated with the onset of chromosome pairing. *Journal of Cell Biology*, **134**, 1109–1125.

Shikano T, Natri HM, Shimada Y, Merilä J (2011) High degree of sex chromosome differentiation in stickleback fishes. *BMC Genomics*, **12**, 474.

Smith KN, Nicolas A (1998) Recombination at work for meiosis. *Current Opinion in Genetics & Development*, **8**, 200–211.

Smukowski CS, Noor MAF (2011) Recombination rate variation in closely related species. *Heredity*, **107**, 496–508.

Spencer CCA, Deloukas P, Hunt S *et al.* (2006) The influence of recombination on human genetic diversity. *PLoS Genetics*, **2**, 1375–1385.

Stevison LS, Noor MAF (2010) Genetic and evolutionary correlates of fine-scale recombination rate variation in *Drosophila persimilis*. *Journal of Molecular Evolution*, **71**, 332–345.

Stoelting KN, Nipper R, Lindtke D *et al.* (2013) Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. *Molecular Ecology*, **22**, 842–855.

Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671–688.

Takahashi A, Liu YH, Saitou N (2004) Genetic variation versus recombination rate in a structured population of mice. *Molecular Biology and Evolution*, **21**, 404–409.

Tenaillon MI, Sawkins MC, Long AD *et al.* (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 9161–9166.

Tortereau F, Servin B, Frantz L *et al.* (2012) A high density recombination map of the pig reveals a correlation between sex-specific recombination and GC content. *BMC Genomics*, **13**, 586.

Urton JR, McCann SM, Peichel CL (2011) Karyotype differentiation between two stickleback species (Gasterosteidae). *Cytogenetic and Genome Research*, **135**, 150–159.

Wang J, Naa J-K, Yu Q *et al.* (2012) Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 13710–13715.

Webster MT, Hurst LD (2012) Direct and indirect consequences of meiotic recombination: implications for genome evolution. *Trends in Genetics*, **28**, 101–109.

Wilson MA, Makova KD (2009) Genomic analyses of sex chromosome evolution. *Annual Review of Genomics and Human Genetics*, **10**, 333–354.

Wong AK, Ruhe AL, Dumont BL *et al.* (2010) A comprehensive linkage map of the dog genome. *Genetics*, **184**, 595–605.

D.B. and M.R. designed the study and performed the analyses; D.B. produced the cross; M.R. and D.M. generated the R.A.D. libraries; D.B., M.R. and D.M. produced the marker data; D.B. wrote the manuscript, with input from M.R.

Data accessibility

Reassembled threespine stickleback reference genome: Dryad digital repository (doi:10.5061/dryad.846nj).

Genotype data used for final linkage map construction and the corresponding linkage map: online supporting information.

Supporting information

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

Appendix S1 Schematic overview of the stickleback genome reassembly.

Appendix S2 Genotype data used for estimating genetic distances on the 20 autosomes.

Appendix S3 Genotype data used for estimating genetic distances on the sex chromosome (19X).

Appendix S4 Physical and genetic map positions for all 1872 markers.

Appendix S5 Patterns of Y chromosome degeneration revealed by sex-specific RAD locus coverage along the X chromosome in an Atlantic-derived and a Pacific-derived natural population.