The DNA Methylation Landscape of Stickleback Reveals Patterns of Sex Chromosome Evolution and Effects of Environmental Salinity

David C.H. Metzger* and Patricia M. Schulte

Department of Zoology, University of British Columbia, Vancouver, BC, Canada

*Corresponding author: E-mail: dmetzger@zoology.ubc.ca.

Accepted: February 5, 2018

Data deposition: This project has been deposited at the NCBI sequence read archive (SRA) under the accession SRP127356.

Abstract

Epigenetic mechanisms such as DNA methylation are a key component of dosage compensation on sex chromosomes and have been proposed as an important source of phenotypic variation influencing plasticity and adaptive evolutionary processes, yet little is known about the role of DNA methylation in an ecological or evolutionary context in vertebrates. The threespine stickleback (*Gasterosteus aculeatus*) is an ecological and evolutionary model system that has been used to study mechanisms involved in the evolution of adaptive phenotypes in novel environments as well as the evolution heteromorphic sex chromosomes and dosage compensation in vertebrates. Using whole genome bisulfite sequencing, we compared genome-wide DNA methylation patterns between threespine stickleback males and females and between stickleback reared at different environmental salinities. Apparent hypermethylation of the younger evolutionary stratum of the stickleback X chromosome in females relative to males suggests a potential role of DNA methylation in the evolution of heteromorphic sex chromosomes. We also demonstrate that rearing salinity has genome-wide effects on DNA methylation levels, which has the potential to lead to the accumulation of epigenetic variation between natural populations in different environments.

Key words: epigenetics, *Gasterosteus aculeatus*, bisulfite sequencing, WGBS, BS-Seq.

Introduction

Epigenetic variation has the potential to impact ecological and evolutionary processes, and thus affect species distributions and evolutionary trajectories (Bossdorf et al. 2008; Flores et al. 2013; Jablonka and Raz 2009; Varriale 2014; Franks and Hoffmann 2012). Currently, one of the best-studied mechanisms underlying epigenetic variation is DNA methylation, a heritable epigenetic modification in which a methyl group is added to position 5 of the pyrimidine ring on a cytosine (5mC), most commonly found on cytosine-phosphate-guanine (CpG) dinucleotides in vertebrates (Heard and Martienssen 2014). Changes in DNA methylation can have profound effects on chromatin structure, which can in turn alter gene expression (Klose and Bird 2006; Jaenisch and Bird 2003). The addition or removal of these methyl groups can be dynamically regulated in response to changes in the environment (Boyko et al. 2010; Kucharski et al. 2008; Cooney et al. 2002). Variation in DNA methylation levels have therefore been hypothesized to play a key role in mediating phenotypic responses to environmental change (Bossdorf et al. 2008; Hofmann 2017; Flores et al. 2013), and may represent a dynamic source of heritable variation that can respond to changes in the environment and influence phenotypic variation over multiple time-scales (Richards 2006).

In addition to its potential role in regulating gene expression in response to environmental change, DNA methylation is also critical in regulating gene expression in dosage compensation systems that have evolved to minimize the unequal expression of genes on heteromorphic sex chromosomes (Graves 2016). In older XY sex chromosome systems, such as those found in most mammalian species, DNA methylation is involved in the global silencing of one of the two X

chromosome in females (Graves 2016). In ZW sex chromosome systems (which have a female-specific W chromosome) such as those in birds and some reptiles and fishes, DNA methylation is involved in gene-specific dosage compensation via the activation or suppression of particular dosage sensitive genes (Graves 2016). While epigenetic silencing has been well-established as a mechanism involved in dosage compensation of older heteromorphic sex chromosome systems, patterns of DNA methylation in young sex chromosome systems are less well-understood, but have been hypothesized to play a key role sex chromosome evolution (Gorelick 2003).

The threespine stickleback (*Gasterosteus aculeatus*) has been extensively used to investigate the genetic basis of adaptive evolution to novel environments (Jones et al. 2012a,b). Following the last glaciation, ancestral marine populations of stickleback colonized and adapted to newly available freshwater habitats in the north-temperate zone (Bell and Foster 1994). Adaptation to these novel environments drove the rapid parallel evolution of divergent phenotypes including changes in body shape, armor plate number, gene expression levels, and gene expression plasticity (Jones et al. 2012a; Gibbons et al. 2017; Morris et al. 2014; McCairns and Bernatchez 2009; Ishikawa et al. 2017). Several studies have used reduced representation approaches to characterize variation in DNA methylation patterns between stickleback that vary in their lateral plate morphology, and have suggested that variation in DNA methylation patterns may contribute to the phenotypic divergence observed between marine and freshwater populations (Smith et al. 2015; Trucchi et al. 2016; Artemov et al. 2017). Threespine stickleback also have a relatively young XY sex chromosome pair that has evolved since the species first arose ~13–16 Ma (Bell et al. 2009; Kawahara et al. 2009; Ross et al. 2009), and this species has become a powerful model system to explore the evolution of heteromorphic sex chromosomes and dosage compensation mechanisms (Schultheiß et al. 2015; White et al. 2015). Thus, the threespine stickleback is an ideal model in which to investigate the complementary roles of DNA methylation in both environmental adaptation and the evolution of sex chromosome systems.

In this study, we present the first high-resolution analysis of DNA methylation patterns in the stickleback genome using whole genome bisulfite sequencing (WGBS). This approach allowed us to characterize prominent features in the DNA methylation landscape of stickleback, including variation in DNA methylation patterns between males and females along the entire stickleback sex chromosome, which provides insight into the relationship between epigenetic mechanisms and sex chromosome evolution. By rearing putatively ancestral marine stickleback at both low and high salinity, we also describe the effects of environmental salinity on genomic DNA methylation patterns, and highlight potential salinity responsive genes that may be differentially regulated by DNA methylation.

Materials and Methods

Fish Collection

All animal experimentation was conducted according to University of British Columbia approved animal care protocols (A10-0285 and A11-0372). Adult threespine stickleback (*G. aculeatus*) of the fully plated "marine" ecotype were collected at the beginning of their natural spawning season in May 2013 from Oyster Lagoon, British Columbia in Canada (GPS: 49.6121, -124.0314). Fish were separated into six 110-l glass tanks (20 fish per tank) and acclimated to 21 ppt salt water (dechlorinated Vancouver municipal tap water supplemented with Instant Ocean Sea Salt), 18 °C and 14:10 h light:dark photoperiod. These conditions mimic the natural environmental conditions at the collection location at the time of collection. Fish were fed daily to satiation with Hakari Bio-Pure frozen Mysis Shrimp and were acclimated to laboratory conditions for four weeks.

Fertilization and Rearing Procedure

To determine the impact of salinity on fertilization and hatching, eggs were collected from gravid female stickleback and immediately divided into six different petri dishes containing 5 ml of 2, 7, 14, 21, 28, or 35 ppt saltwater. Testes were collected from males displaying sexually mature characteristics and individually macerated in a 1.75 ml microcentrifuge tube containing 300 µl Ginzberg's fish Ringer's solution. Eggs from a single clutch were fertilized with sperm solution from a single male across all salinities (50 µl of sperm solution for each petri dish at each different salinity). Following fertilization, an additional 10 ml of water at the appropriate salinity was added to each petri dish. This process was repeated ten times creating a total of ten different families, each fertilized at all salinities. Petri dishes were partially covered to prevent water loss via evaporation and to allow for surface gas exchange. Eggs were monitored twice daily during which time any unfertilized eggs were removed and 10 ml of water was changed with sterilized water of the appropriate salinity to prevent mold growth. Percent fertilization and percent hatch were recorded. Percent hatch is recorded as the proportion of fertilized embryos that hatched. The effect of salinity on fertilization and hatching success was analyzed using a logistic regression (Warton and Hui 2011) in the R v3.3.2 base stats package. Tukey post hoc analysis was performed using the glht() function in the *multicomp* v1.4-6 R package.

After all fish in a petri dish had hatched and the yolks had been absorbed (~15 days post fertilization), larvae were transferred to hanging net boxes (Aquaclear) in 110 L glass aquaria containing water at the fertilization salinity. Sponge filters were used for filtration and aeration. Each family was kept separate throughout the experiment. At one-month post hatch whole animals were snap frozen in liquid nitrogen.

WGBS

Genomic DNA was isolated from one-month old whole fish samples from the 2 ppt and the 21 ppt salinity treatments using a Qiagen DNeasy Blood and Tissue Kit following the manufacturer's recommended protocol for RNA-free genomic DNA using RNAase A. The sex of each sample was identified by PCR analysis using primers designed to *idh, gasm6*, and *stn190* following previously described methods (Toli et al. 2016). Genomic DNA samples from three males and three females from 2 ppt and 21 ppt (twelve samples total) were sent to the McGill University and Genome Quebec Innovation Center for DNA quality assessment, library preparation, bisulfite treatment, and 150 base pair paired-end sequencing using an Illumina HiSeqX. The 12 samples were split evenly across three sequencing lanes (4 samples/lane) such that one male and female sample from each of the salinity treatments were represented on each sequencing lane. Average sequencing library size was 102,011,555 reads (± 13,147,873 SD).

WGBS Data Analysis

Reads were mapped to a revised assembly of the stickleback genome (Glazer et al. 2015) obtained from the Dryad Digital Repository (http://datadryad.org/resource/doi: 10.5061/dryad.q018v) and DNA methylation levels were calculated using the bisulfite sequencing plugin v1.2 in CLC Genomics Workbench v10.0. Average mapping efficiency was 89.5% (± 1% SD). DNA methylation data were exported from CLC and analyzed using in R v3.3.2 using *methylKit* package v3.5 (Akalin et al. 2012) following previously recommended guidelines for bisulfite sequence analysis (Ziller et al. 2015; Wreczycka et al. 2017). Sequenced CpG loci were filtered so that only sites with at least 10 reads in each of the 12 samples were retained. Sites that were in the 99.9th percentile of coverage were also removed from the analysis to account for potential PCR bias.

Hierarchical cluster analysis was conducted using Ward's method and was implemented using the clusterSamples() function. Pairwise comparisons between groups were performed using a logistic regression model with a correction for overdispersion using the calculateDiffMeth() function followed by a Chi-square test to identify significantly differentially methylated cytosines (DMCs) between groups. The *P*-values for DMCs were false discovery rate (FDR) corrected using the sliding linear model method (SLIM) with a maximum *q*-value threshold of 0.05 and a minimum change in percent methylation of 10%. For the comparison between males and females, salinity and family were included as covariates. For the comparison between salinity rearing treatments, sex and family were included as covariates. For comparisons between families, sex and salinity were included as covariates. To calculate mean methylation levels across 10 kilobase (kb) genomic regions, the tileMethylCounts() function in methylKit was used to calculate DNA methylation values across sequential 10 kb windows of the genome. All figures were generated in R.

To obtain nearest neighboring gene annotations, the gene coordinates in the annotation file (.gtf) provided by Glazer et al. (2015) in the dryad digital repository (which contains gene coordinates that correspond to the stickleback genome available in Ensembl) were converted to the coordinates in the updated assembly using the R script convertCoordinate.R that is provided by the authors of the revised assembly. Distances to nearest neighboring genes were calculated using the annotatePeakInBatch() function in the R package *ChIPpeakAnno* v3.6.5.

CpG islands for the revised stickleback genome assembly were calculated using python scripts (https://github.com/lucasnell/TaJoCGI) that apply an algorithm based on the methods described by Takai and Jones (2002). The observed distribution of DMCs was compared with the distribution of CpGs across the genome using a Chi-square test.

Candidate Gene Analysis

Previous RNA-seq studies have identified many genes that respond to changes in salinity in stickleback (2,771 in gill tissue, Gibbons et al. 2017 and 1,844 in kidney tissue, Wang et al. 2014). To determine whether DNA methylation could be involved in the differential regulation of these candidate genes we compiled a list of salinity responsive genes from previous studies (4,615 candidate genes) and compared them to genes within 2 kb of DMCs in stickleback reared at different salinities.

Similarly, several studies have also characterized sex-biased gene expression patterns in stickleback. We therefore compiled a list of 2,282 genes that display sex-biased gene expression patterns in brain (1,255 genes, Metzger and Schulte 2016) and liver tissue (1,268 genes, Leder et al. 2010), and compared these to genes within 2 kb of DMCs that were identified between male and female stickleback.

We also compared genes within 2 kb of DMCs between individuals reared a different salinities to genes associated with single nucleotide polymorphisms (SNPs) under positive selection in threespine stickleback from freshwater environments compared with marine environments (Jones et al. 2012a).

Results and Discussion

Characterization of the Stickleback Methylome

We performed WGBS on fish from a marine population of threespine stickleback reared from fertilization to the age of 1 month at a salinity of either 2 ppt or 21 ppt. These salinities represent the widest range that still allows good fertilization and hatching success in this population (supplementary fig. S1, Supplementary Material online). In this study, we utilized a balanced design with WGBS performed on one male and one

female from each of three families and each of the salinity rearing treatments. This design was chosen because genetic variation has been shown to have substantial effects on the divergence of DNA methylation patterns among individuals in both plants and animals (Gertz et al. 2011; McRae et al. 2014; Vidalis et al. 2016). Consistent with this observation, we detected strong effects of family on DNA methylation (supplementary figs. S2 and S3, Supplementary Material online). However, family-level variation in DNA methylation levels could also be indicative of transgenerational environmental or maternal effects (Jablonka and Raz 2009).

CpG loci in the stickleback genome had an average methylation level of 70.3%, which is consistent with whole genome assessments of methylation in other fish species (Feng et al. 2010; Shao et al. 2014; Zemach et al. 2010). However, there were several hypomethylated regions (<40% methylation) that are indicative of DNA methylation canyons or valleys (Jeong et al. 2014; Xie et al. 2013) with the most prominent of these located on chromosomes 4, 10, 11, 12, and 16 (fig. 1, supplementary fig. S4, Supplementary Material online). While the factors that determine the size of DNA methylation canyons remains unknown, larger hypomethylated canyons such as those described here in stickleback have been shown to be under strong transcriptional suppression due to increased abundance of repressive histone H3 lysine 27 methylation that interacts with hypomethylated DNA (Nakamura et al. 2014). This mechanisms of transcriptional repression is thought to maintain these regions in a "poised" transcriptional state to allow rapid activation of gene transcription in these regions at specific times during embryonic development (Nakamura et al. 2014), but whether these hypomethylated canyons play a functional role in adults is unknown. However, genes that are essential for proper development typically dominate these regions (Jeong et al. 2014; Nakamura et al. 2014; Xie et al. 2013). Consistent with this pattern, genes located in the hypomethylated canyons in the stickleback genome include protocadherins on chromosome 4 and homeobox genes on chromosomes 10, 11, 12, and 16 (supplementary table S1, Supplementary Material online), suggesting a conserved role of hypomethylated canyons across vertebrates.

Sex-Biased DNA Methylation Patterns

We identified a total of 18,564 DMCs between males and females (fig. 2, supplementary data set S1, Supplementary Material online). Although relatively few studies have examined differential methylation patterns between males and females at the whole genome level in fishes, a study in tilapia detected a similar number of DMCs between males and females in muscle tissue (17,112 DMCs; Wan et al. 2016), whereas a study of sex-specific differential methylation in zebrafish brain detected only 914 DMCs (Chatterjee et al. 2016). These data suggest that the extent of sexually dimorphic methylation may be highly variable among teleosts, consistent with the wide range of sex-determining mechanisms in this group (Devlin and Nagahama 2002).

The distribution of DMCs across genomic features (e.g., promoters, exons, CpG islands) did not differ from the relative proportions of these features in the genome (supplementary table S2, Supplementary Material online). No DMCs were identified between males and females in the mitochondrial genome. The majority of DMCs (90%; 16,626 DMCs) between males and females showed a bias towards higher methylation in females relative to males suggesting that female stickleback genome is hypermethylated relative to male stickleback genome.

The most striking pattern in these data is the apparent hypermethylation of chr19 (the threespine stickleback sex chromosome) in females relative to males, which is where 65% of the putative DMCs identified between males and females are located (12,112 DMCs). Chr19 also had the highest proportion of DMCs relative to the number of CpG sites on the chromosome (5%) compared with the rest of the autosomes where ~0.07% of CpG loci were differentially methylated between the sexes (fig. 2).

Three distinct regions (strata) have been characterized on chr19 based on the extent of divergence in these regions between the X and Y chromosome (Ross and Peichel 2008; White et al. 2015), and two of these strata no longer recombine between the X and Y: Stratum two (the younger evolutionary stratum located between ~2.5 Mb and 12 Mb), and stratum one (the older evolutionary stratum located from ~12 Mb to the end of the chromosome). There is also a pseudoautosomal region (PAR) located in the first ~2.5 Mb of chr19 that is thought to still recombine between the X and Y chromosomes (White et al. 2015). To assess whether these evolutionary strata are also associated with sex-specific DNA methylation, we divided chr19 into 10 kb consecutive nonoverlapping bins and calculated the frequency of CpG loci that were putatively identified as hypermethylated or hypomethylated in female stickleback relative to male stickleback (fig. 3). This analysis revealed clear patterns that correspond to the evolutionary strata of chr19 for loci that were hypermethylated in females relative to males, with the greatest apparent hypermethylation in stratum two and the least in the PAR (fig. 3*B*).

Identifying Differential Methylation on the Sex Chromosome

One of the challenges for unambiguously determining levels of sex-specific DNA methylation on stickleback sex chromosomes is that there is currently no publically available sequence for the Y chromosome and the published sequence for chr19 is predominately derived from X chromosome sequence. Because DNA methylation is detected as sequence differences between bisulfite-treated DNA and the reference sequence at CpG sites, both divergence between the X and Y

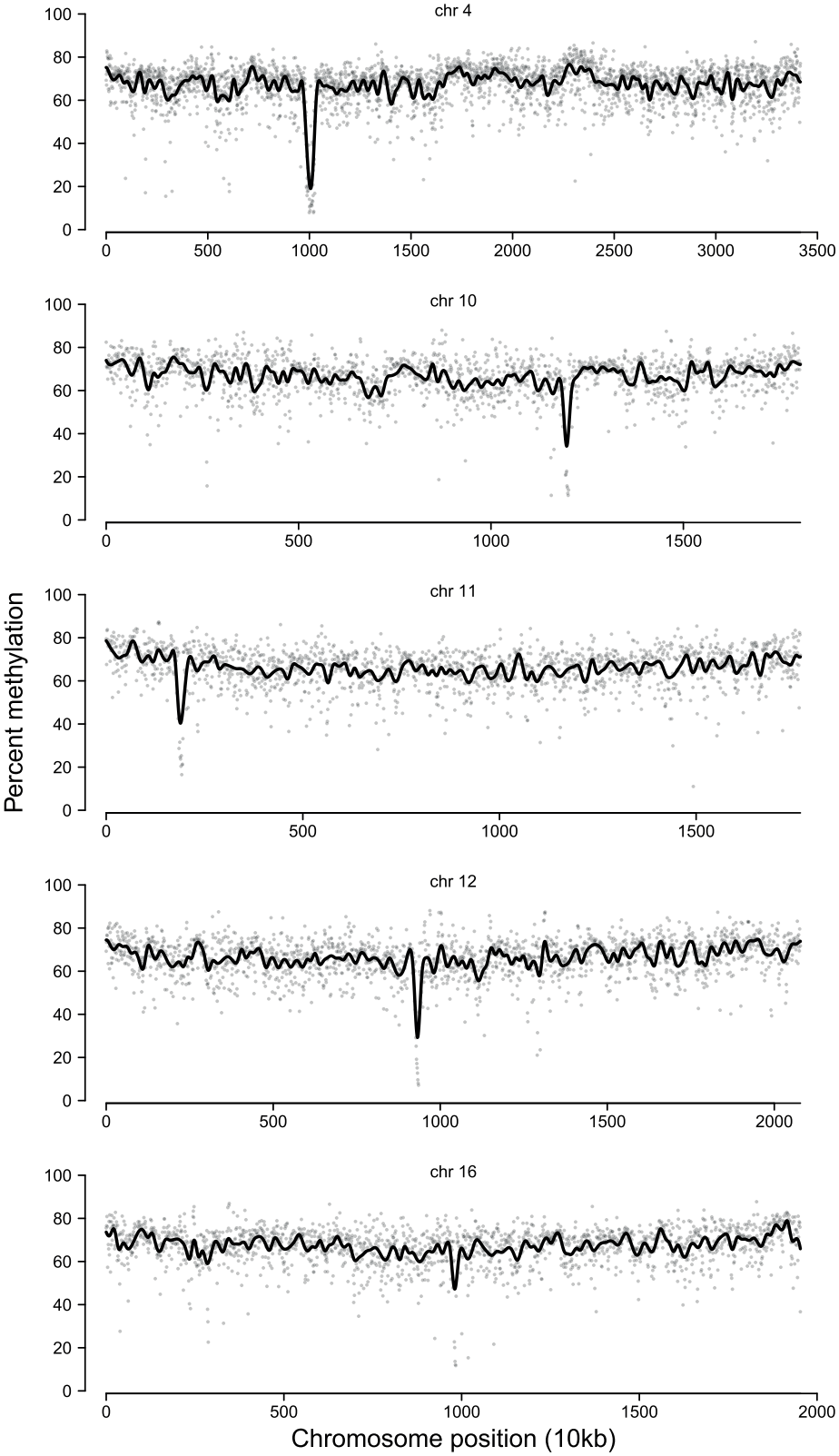

Fig. 1.—Mean CpG methylation level across chromosomes 4, 10, 11, 12, and 16. Each point represents the mean methylation level across all twelve individuals for a single 10 kb window. The solid line represents the smoothed spline fit to these data. Position along the *x*-axis represents the base position along the chromosome. The *y*-axis is the average DNA methylation level.

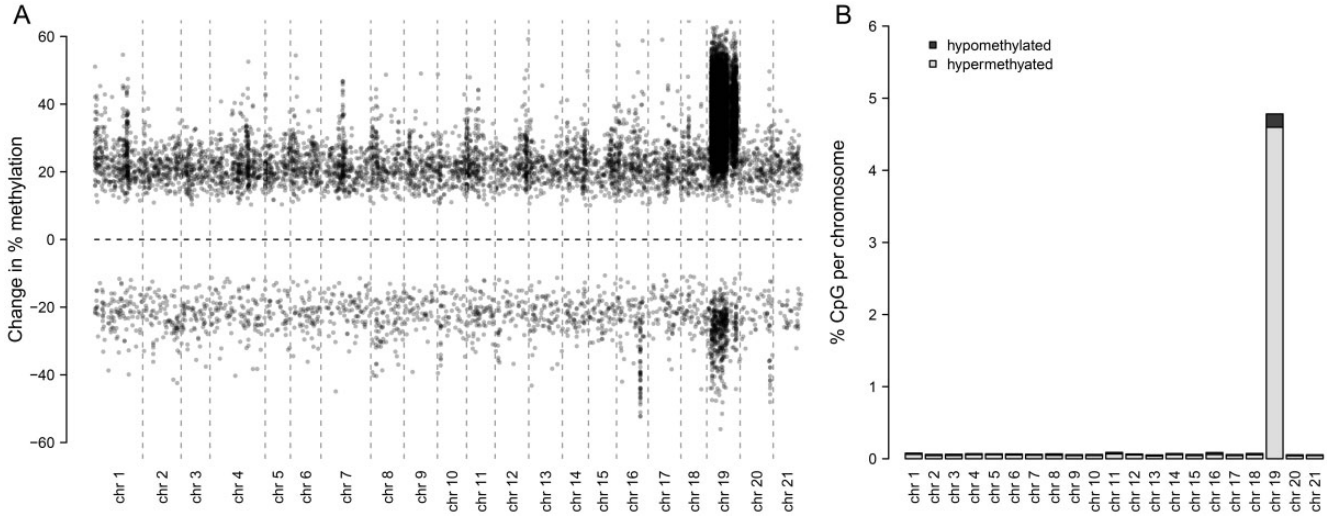

Fig. 2.—Differentially methylated CpG loci between male and female stickleback. (*A*) Differentially methylated CpG (DMC) loci between male and female threespine stickleback. Each point represents an individual DMC. The *y*-axis indicates the percent difference in methylation between males and females. A positive value on the *y*-axis indicates a DMC that is hypermethylated in females relative to males. A negative value on the *y*-axis indicates a DMC that is hypomethylated in females relative to males. The *x*-axis indicates the position of the DMC in the stickleback genome. Chromosome boundaries are represented by vertical dashed lines. Only DMCs for which a change in methylation of >10% are presented. (*B*) The percentage of CpG loci on a given chromosome that were differentially methylated between male and female threespine stickleback. The light shading represents DMCs that are hypermethylated and the dark shading represents DMCs that are hypomethylated in female stickleback compared with males.

chromosome at CpG sites and differential methylation can result in the same patterns in the sequence data. Thus, the signal of differential methylation that we observe could be attributed to one of three possible mechanisms: 1) identical sequences but differential methylation between the X and Y chromosome, 2) differential methylation between X chromosomes in males and females, or 3) sequence divergence between the X and Y chromosomes resulting in alteration of CpG sites.

To address this issue, we again divided chr19 in to 10 kb consecutive nonoverlapping segments and compared the number of reads that mapped to chr19 in males and females for each 10 kb segment. If chr19 reads map uniquely to the published X chromosome then we would expect half the number of reads to map to chr19 in males compared with females. Given that the PAR is known to recombine between the X and Y chromosomes, suggesting low levels of divergence in this region, we predicted that sequencing reads derived from both the X and Y chromosomes would map to the reference sequence, resulting in a ratio of one for the number of reads mapped in males and females in the PAR. The results from the read coverage analysis matched this prediction (fig. 3*D*).

For the younger, less diverged stratum (stratum two), we predicted a read count ratio between 0.5 and 1 because sequence similarity between the X and Y chromosome would result in reads from both chromosomes mapping to the X chromosome reference sequence. The results from the read coverage analysis matched this prediction (fig. 3*D*). Thus, we cannot unambiguously determine whether differential methylation or X Y polymorphism is the cause of the apparent hypermethylation in females in stratum two.

Stratum one is thought to be the most divergent region of the sex chromosome, and thus we predicted that few reads from the Y chromosome would map to the chr19 reference, resulting in a ratio of 0.5 for the number of reads mapped in males and females. Again, the results from the read coverage analysis mostly matched this prediction (fig. 3*D*), although there were specific regions where the read count ratio was close to one. These regions in stratum one may correspond to regions that are thought to be under purifying selection to maintain dosage sensitive genes from being lost on the Y chromosome (White et al. 2015). We also detected apparent hypermethylation in females in these regions (fig. 3*B* and *D*). Thus, we cannot conclusively determine whether this apparent hypermethylation of chr19 in females is due to differential methylation or from the accumulation of TG polymorphisms on the Y that are being interpreted as unmethylated loci.

While we are unable to unambiguously determine the ultimate cause of the apparent DNA methylation differences between males and females on chr19, whether the patterns we observe are the result of sequence polymorphism between the X and Y chromosome that alters CpG sites, or are due to differential methylation of conserved sequences between males and females in chr19, the ultimate effect would be differences in methylation between the sex chromosomes. Thus, taken together, the patterns of putative differential methylation that we observe suggest that divergence

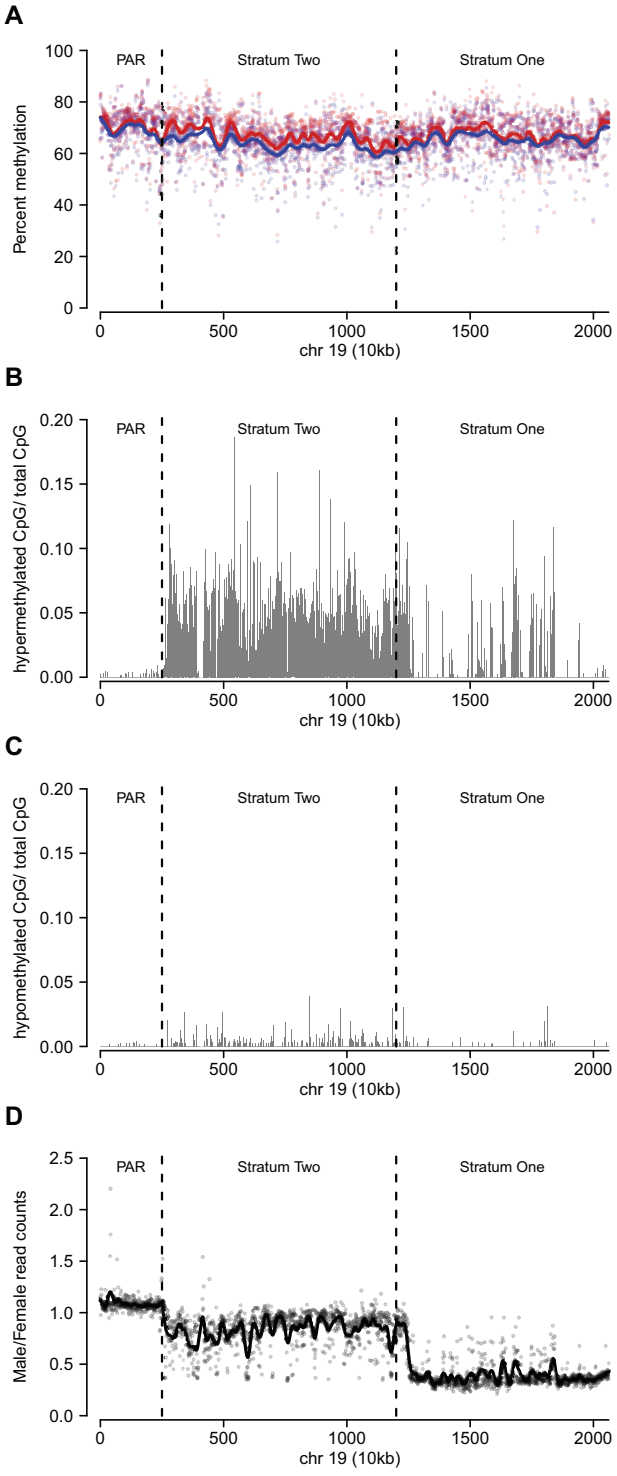

Fig. 3.—Differential methylation between sexes on chromosome 19 (chr19). (*A*) Mean DNA methylation levels for CpG loci along chromosome chr19. Each point represents the mean DNA methylation level in a 10 kb window for six individual stickleback that were either male (blue) or female (red). Solid lines represent the smooth spline fit for the DNA methylation levels in males (blue) and females (red). (*B*, *C*) Proportion of DMCs along chr 19 that are hypermethylated (*B*) or hypomethylated (*C*) in female stickleback compared with male stickleback. Values on the *y*-axis represent the total number of DMCs in a 10 kb window relative to the number of CpG loci in that same 10 kb window. (*D*) Ratio of mapped read counts for males relative to females along chr19. Each point represents the ratio of mean counts for a 10 kb window in males compared with females. The solid black line is the smooth spline fit. Vertical dashed lines represent the boundaries between the three evolutionary strata on chr19: The pseudoautosomal region (PAR), stratum two, and stratum one.

in DNA methylation patterns between males and females on the stickleback sex chromosome are closely associated with the known evolutionary history of this chromosome.

DNA Methylation and Sex Chromosome Evolution

Sex chromosome evolution from autosomes is thought to involve recombination suppression in sex determining regions, followed by the accumulation of deleterious mutations and the degeneration of the sex-specific (e.g., Y) chromosome (Graves 2016). Degradation of the sex-specific heteromorphic sex chromosome following recombination suppression has the potential to cause imbalances in gene expression. Many taxa with heteromorphic sex chromosome pairs have evolved dosage compensation mechanisms to resolve this effect, but the nature and extent of these dosage compensation mechanisms varies greatly among taxa (Graves 2016). DNA methylation has been proposed as a key mechanism responsible for regulating every step of the evolution of sex chromosomes from recombination suppression in the early stages of sex chromosome evolution to dosage compensation in more derived sex chromosome systems (Gorelick 2003); however, there has been little empirical evidence to test this hypothesis. Taxa with relatively "young" heteromorphic sex chromosomes, such as the threespine stickleback, provide an opportunity to investigate the potential role of DNA methylation in regulating sex chromosome recombination and the evolution of dosage compensation mechanisms. In the following section, we discuss the apparent differential methylation between males and females on chr19 in the context of the different stages of sex chromosome evolution in stickleback.

DNA methylation promotes the formation of heterochromatin (Mirouze et al. 2012; Yelina et al. 2015; Melamed-Bessudo and Levy 2012), and it is thought to play a role in suppressing recombination of sex chromosomes in plants (Zhang et al. 2008). In stickleback we observed apparent hypermethylation of the younger evolutionary stratum in females (stratum two), and relatively less differential methylation between males and females along the older evolutionary stratum (stratum one) and the PAR. The apparent hypermethylation of stratum two on the X chromosome in females (hypomethylated in males) corresponds to the region hypothesized to have undergone the first chromosomal inversion during the evolution of the Y chromosome (Ross and Peichel 2008). The apparent differential methylation in

stratum two between males and females (hypomethylated in males and hypermethylated in females) could have played a role in suppressing recombination between male and female sex chromosomes and in establishing the boundaries in which this inversion first occurred.

The next stage in sex chromosome evolution, following recombination suppression, is thought to be the accumulation of genetic variation and degradation in the nonrecombining region(s). If methylated cytosines on the female X chromosome correspond to thymines on the male Y chromosome as previously discussed, then it is possible that the accelerated mutation rate of methylated cytosines, which can be deaminated to become thymines (Coulondre et al. 1978; Shen et al. 1994), could play an important role in the divergence between X and Y chromosomes. Alternatively, instead of being a direct result of C to T polymorphisms, the observed increase in hypermethylated loci in females could be closely linked to the accumulation of genetic polymorphisms. The frequency at which SNPs occur in genomes has been shown to be higher near methylated CpG loci (Qu et al. 2012). The CGCG motif has been identified as a candidate cis-element associated with this observation and is enriched in hypomethylated regions (Qu et al. 2012). Therefore, individuals with higher DNA methylation levels at particular loci would be predicted to have a higher degree of sequence divergence near those loci relative to individuals with lower methylation levels. The PAR had the highest frequency of the CGCG motif (1.98/1 kb) whereas stratum one and stratum two had a lower frequency of the CGCG motif (1.22/1 kb and 1.18/1 kb respectively). These data suggest that DNA methylation may also be associated with the accumulation of genetic variation between the nonrecombining regions of the male and female sex chromosomes.

We next explored whether the apparent differential methylation patterns between males and females on chr19 are consistent with the regulation of dosage sensitive genes. In stickleback, there are two conflicting hypotheses regarding the existence of a dosage compensation system. One hypothesis is that there is locally confined partial dosage compensation in stratum one in males that is also associated with a hypertranscription of genes in stratum one in females (Schultheiß et al. 2015). The second hypothesis suggests that dosage compensation has not evolved in the stickleback but that there is purifying selection to maintain dosage sensitive genes in stratum one of the Y chromosome (White et al. 2015). The differential methylation patterns between sexes along the X chromosome that we observe are not entirely consistent with either of these prevailing hypotheses. We observed apparent hypermethylation in stratum two in females. This might be expected to result in reduced transcription or partial silencing of genes in this region, which has not been observed in stickleback (Schultheiß et al. 2015; White et al. 2015). The less extensive and highly localized pattern of hypermethylation in stratum one that we observe is suggestive of gene-specific regulation, which is not consistent with a generalized hypertranscription of genes in stratum one in females (Schultheiß et al. 2015). Because the male to female coverage ratio is similar in these localized regions in stratum one, this localized pattern may be more consistent with the potential preservation of dosage sensitive genes in these regions (White et al. 2015); however, it is also possible that the differential methylation in these regions is caused by the accumulation of C to T polymorphisms which is less consistent with purifying selection acting in these regions (White et al. 2015).

Taken together, the apparent differential methylation between male and female stickleback described in this study is consistent with the proposed role of DNA methylation in the evolution of sex chromosomes (Gorelick 2003). Thus differential DNA methylation could be playing a role in the evolution of this "young" heteromorphic sex chromosome system, either through influencing patterns of recombination or potentially through mediating the early stages of the development of dosage compensation.

DNA Methylation and Sex-Biased Gene Expression

To determine whether the apparent variation in DNA methylation patterns between males and females could be influencing previously described sex-biased gene expression patterns in stickleback, we compared the list of genes near DMCs between males and females (supplementary data set S2, Supplementary Material online) to previously identified genes that exhibit sex-biased gene expression patterns (Metzger and Schulte 2016; Leder et al. 2010). Of the 2,282 genes that have been shown to exhibit sex biased gene expression patterns from these studies, 490 overlapped with genes near DMCs in our study of which 269 are on chr19 (supplementary data set S3, Supplementary Material online) including genes located in the region considered to be tightly linked to sex determination in stickleback such as *sema4ba* and *idh2* (Peichel et al. 2004). This pattern is consistent with differential methylation between males and females playing a role in regulating sex-biased patterns of gene expression.

Effects of Environmental Salinity on DNA Methylation

Variation in DNA methylation has also been suggested to be an important component of an organism's response to environmental change (Bossdorf et al. 2008; Hofmann 2017; Flores et al. 2013). Changes in environmental salinity are known to cause substantial changes in gene expression in many species of fish, including stickleback (Gibbons et al. 2017; Zhang et al. 2017; Wang et al. 2014). In order to explore whether changes in DNA methylation may be involved in environmental regulation of gene expression, we identified differentially methylated loci in stickleback reared at two salinities (2 and 21 ppt), and we compared genes near DMCs identified in fish reared at different salinities to genes that

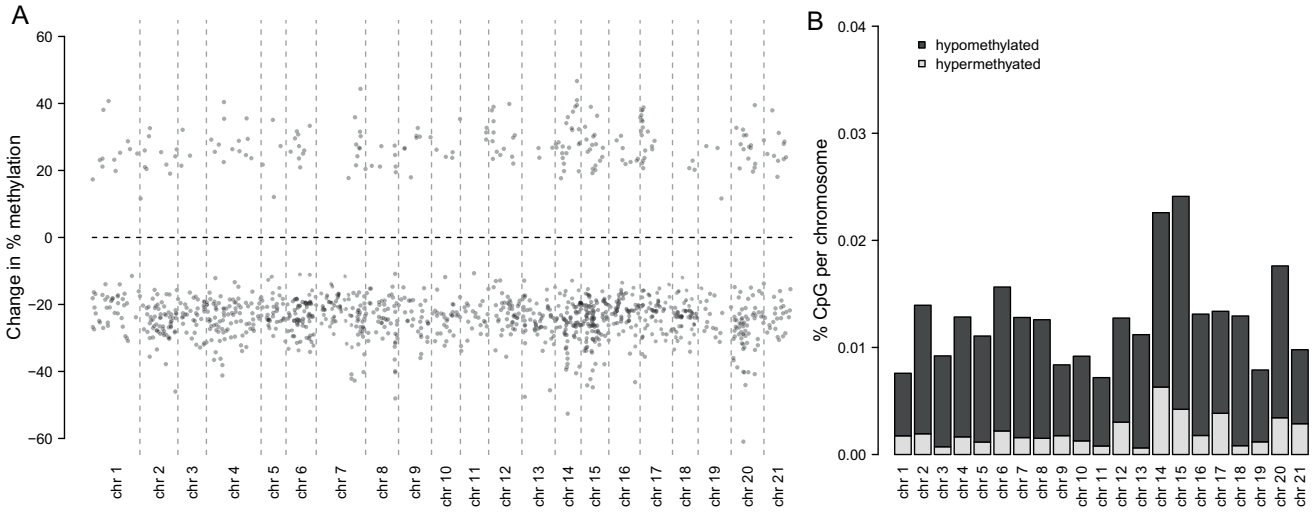

Fig. 4.—Differentially methylated CpG loci between stickleback reared at low and high salinity. (*A*) Differentially methylated CpG (DMC) loci between stickleback reared at a salinity of 2 ppt compared with 21 ppt. Each point represents an individual DMC. The *y*-axis indicates the percent difference in methylation between salinity rearing treatments. A positive value on the *y*-axis indicates a DMC this is hypermethylated in stickleback reared at 21 ppt relative 2 ppt. A negative value on the *y*-axis indicates a DMC that is hypomethylated in stickleback reared at 21 ppt relative to 2 ppt. The *x*-axis indicates the position of the DMC in the stickleback genome. Chromosome boundaries are represented by vertical dashed lines. Only DMCs for which a change in methylation of >10% are presented. (*B*) The percentage of CpG loci on a given chromosome that were detected as being differentially methylated between stickleback reared at 21 ppt and 2 ppt. The light shading represents DMCs that are hypermethylated and dark shading represents DMCs that are hypomethylated in stickleback reared at 21 ppt compared with 2 ppt respectively.

have been previously identified as salinity-responsive using RNA-seq (Gibbons et al. 2017; Wang et al. 2014).

1,259 CpG loci were differentially methylated between salinity treatments (supplementary data set S4, Supplementary Material online), and these DMCs were distributed across all chromosomes, with an average of 0.01% of the CpG loci on each chromosome being differentially methylated (fig. 4). No DMCs were identified between individuals from low and high salinities in the mitochondrial genome. The distribution of DMCs across genomic features (e.g., promoters, exons, CpG islands) did not differ from the relative proportions of these features in the genome (supplementary table S2, Supplementary Material online). The majority of DMCs (1,051) was hypomethylated in individuals from high salinity relative to low salinity. Analysis of the genes located close to these DMCs revealed several genes known to be involved in the response to salinity in fish (supplementary data set S5, Supplementary Material online). However, GO enrichment analysis did not detect significant enrichment for any GO categories following FDR correction. The ten GO terms with the lowest *P*-values are listed in supplementary tables S3–S5, Supplementary Material online.

Comparison of genes located near DMCs in fish reared at different salinities to previously identified as salinity-responsive using RNA-seq (Gibbons et al. 2017; Wang et al. 2014) identified 126 candidate genes with changes in both expression and methylation in response to salinity (supplementary data set S6, Supplementary Material online). Among the candidate genes that we identified are ion channels that are important for regulating cellular ion concentrations in hyper and hypoosmotic conditions such as the calcium pump *atp2b4*, the sodium/chloride cotransporter *slc12a3*, and the sodium/potassium/2 chloride cotransporter *slc12a1*. Taken together, these data suggest that changes in DNA methylation could play a role in facilitating the transition between marine and freshwater environments.

In stickleback, a variety of genomic regions have been identified as having been subject to positive selection following colonization of freshwater habitats by ancestral marine fish (Jones et al. 2012b). Because epigenetic variation has been suggested to be a driver of adaptive evolution (Flores et al. 2013), we screened our data set of salinity responsive DMCs in marine fish to identify those associated with genes found in regions under positive selection in freshwater populations (Jones et al. 2012b). Very few of the DMCs identified in our study were near these genes (supplementary data set S7, Supplementary Material online), suggesting that salinity-responsive changes in DNA methylation are unlikely to have played a role in driving genetic divergence in these regions between marine and freshwater populations of stickleback.

Conclusions

In this study we used whole-genome bisulfite sequencing to identify novel DNA methylation features in the stickleback epigenome. Apparent hypermethylation of stratum two on the female X chromosome compared with levels in males suggests that DNA methylation could play

an important role in the suppressing recombination between the X and Y chromosome, and potentially in regulating sex-biased gene expression patterns. We also detected significant changes in DNA methylation in response to rearing salinity, some of which were associated with genes known to be differentially regulated in response to changes in environmental salinity. This epigenetic change reflects a response to environmental salinity that could facilitate the accumulation of epigenetic variation between natural populations, and thus be implicated in long-term responses to environmental change.

Data Accessibility

The sequencing FASTQ files from the whole genome bisulfite sequencing can be downloaded from the NCBI sequence read archive (SRA accession: SRP127356).

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We thank Timothy M. Healy for insightful discussions and comments during the preparation of this manuscript and for assistance with the analysis in R. Support for this research was provided by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant to P.M.S., a UBC Faculty of Graduate Studies Four Year Fellowship to D.C.H.M, and a Zoology Graduate Fellowship to D.C.H.M.

Literature Cited

Akalin A, et al. 2012. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol. 13(10):R87.

Artemov AV, et al. 2017. Genome-wide DNA methylation profiling reveals epigenetic adaptation of stickleback to marine and freshwater conditions. Mol Biol Evol. 34(9):2203–2213.

Bell MA, Foster SA. 1994. The evolutionary biology of the threespine stickleback. New York: Oxford University Press.

Bell MA, Stewart JD, Park PJ. 2009. The world's oldest fossil threespine stickleback fish. Copeia 2009(2):256–265.

Bossdorf O, Richards CL, Pigliucci M. 2008. Epigenetics for ecologists. Ecol Lett. 11(2):106–115.

Boyko A, et al. 2010. Transgenerational adaptation of *Arabidopsis* to stress requires DNA methylation and the function of dicer-like proteins. PLoS One 5(3):e9514.

Chatterjee A, et al. 2016. Sex differences in DNA methylation and expression in zebrafish brain: a test of an extended 'male sex drive' hypothesis. Gene. 590:307–316.

Cooney CA, Dave AA, Wolff GL. 2002. Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. J Nutr. 132(8 Suppl):2393S–2400S.

Coulondre C, Miller JH, Farabaugh PJ, Gilbert W. 1978. Molecular basis of base substitution hotspots in *Escherichia coli*. Nature 274(5673):775–780.

Devlin RH, Nagahama Y. 2002. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. Aquaculture 208(3–4):191–364.

Feng S, et al. 2010. Conservation and divergence of methylation patterning in plants and animals. Proc Natl Acad Sci U S A. 107(19):8689–8694.

Flores KB, Wolschin F, Amdam GV. 2013. The role of methylation of DNA in environmental adaptation. Integr Comp Biol. 53(2):359–372.

Franks SJ, Hoffmann AA. 2012. Genetics of climate change adaptation. Annu Rev Genet. 46:185–208.

Gertz J, et al. 2011. Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. PLoS Genet. 7(8):e1002228.

Gibbons TC, Metzger DCH, Healy TM, Schulte PM. 2017. Gene expression plasticity in response to salinity acclimation in threespine stickleback ecotypes from different salinity habitats. Mol Ecol. 26(10):2711–2725.

Glazer AM, Killingbeck EE, Mitros T, Rokhsar DS, Miller CT. 2015. Genome assembly improvement and mapping convergently evolved skeletal traits in sticklebacks with genotyping-by-sequencing. G3 5(7):1463–1472.

Gorelick R. 2003. Evolution of dioecy and sex chromosomes via methylation driving Muller's ratchet. Biol J Linn Soc. 80(2):353–368.

Graves JAM. 2016. Evolution of vertebrate sex chromosomes and dosage compensation. Nat Rev Genet. 17(1):33–46.

Heard E, Martienssen RA. 2014. Transgenerational epigenetic inheritance: myths and mechanisms. Cell 157(1):95–109.

Hofmann GE. 2017. Ecological epigenetics in marine metazoans. Front Mar Sci. 4:doi:10.3389/fmars.2017.00004.

Ishikawa A, et al. 2017. Different contributions of local- and distant-regulatory changes to transcriptome divergence between stickleback ecotypes. Evolution (N Y) 71(3):565–581.

Jablonka E, Raz G. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. Q Rev Biol. 84(2):131–176.

Jaenisch R, Bird A. 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 33(3s):245–254.

Jeong M, et al. 2014. Large conserved domains of low DNA methylation maintained by Dnmt3a. Nat Genet. 46(1):17–23.

Jones FC, et al. 2012a. The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484(7392):55–61.

Jones FC, et al. 2012b. A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. Curr Biol. 22(1):83–90.

Kawahara R, Miya M, Mabuchi K, Near TJ, Nishida M. 2009. Stickleback phylogenies resolved: evidence from mitochondrial genomes and 11 nuclear genes. Mol Phylogenet Evol. 50(2):401–404.

Klose RJ, Bird AP. 2006. Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci. 31(2):89–97.

Kucharski R, Maleszka J, Foret S, Maleszka R. 2008. Nutritional control of reproductive status in honeybees via DNA methylation. Science 319(5871):1827–1830.

Leder EH, et al. 2010. Female-biased expression on the X chromosome as a key step in sex chromosome evolution in threespine sticklebacks. Mol Biol Evol. 27(7):1495–1503.

McCairns RJS, Bernatchez L. 2009. Adaptive divergence between freshwater and marine sticklebacks: insights into the role of phenotypic plasticity from an integrated analysis of candidate gene expression. Evolution 64(4):1029–1047.

McRae AF, et al. 2014. Contribution of genetic variation to transgenerational inheritance of DNA methylation. Genome Biol. 15(5):R73.

Melamed-Bessudo C, Levy AA. 2012. Deficiency in DNA methylation increases meiotic crossover rates in euchromatic but not in heterochromatic regions in *Arabidopsis*. Proc Natl Acad Sci U S A. 109(16):E981–E988.
Metzger DCH, Schulte PM. 2016. Maternal stress has divergent effects on gene expression patterns in the brains of male and female threespine stickleback. Proc R Soc B Biol Sci. 283(1839):20161734.
Mirouze M, et al. 2012. Loss of DNA methylation affects the recombination landscape in *Arabidopsis*. Proc Natl Acad Sci. 109(15):5880–5885.
Morris MRJ, et al. 2014. Gene expression plasticity evolves in response to colonization of freshwater lakes in threespine stickleback. Mol Ecol. 23(13):3226–3240.
Nakamura R, et al. 2014. Large hypomethylated domains serve as strong repressive machinery for key developmental genes in vertebrates. Development 141(13):2568–2580.
Peichel CL, et al. 2004. The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. Curr Biol. 14(16):1416–1424.
Qu W, et al. 2012. Genome-wide genetic variations are highly correlated with proximal DNA methylation patterns. Genome Res. 22(8):1419–1425.
Richards EJ. 2006. Inherited epigenetic variation—revisiting soft inheritance. Nat Rev Genet. 7(5):395–401.
Ross JA, Peichel CL. 2008. Molecular cytogenetic evidence of rearrangements on the Y chromosome of the threespine stickleback fish. Genetics 179(4):2173–2182.
Ross JA, Urton JR, Boland J, Shapiro MD, Peichel CL. 2009. Turnover of sex chromosomes in the stickleback fishes (Gasterosteidae). PLoS Genet. 5(2):e1000391.
Schultheiß R, Viitaniemi HM, Leder EH. 2015. Spatial dynamics of evolving dosage compensation in a young sex chromosome system. Genome Biol Evol. 7(2):581–590.
Shao C, et al. 2014. Epigenetic modification and inheritance in sexual reversal of fish. Genome Res. 24(4):604–615.
Shen JC, Rideout WM, Jones PA. 1994. The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. Nucleic Acids Res. 22(6):972–976.
Smith G, Smith C, Kenny JG, Chaudhuri RR, Ritchie MG. 2015. Genome-wide DNA methylation patterns in wild samples of two morphotypes of threespine stickleback (*Gasterosteus aculeatus*). Mol Biol Evol. 32(4):888–895.
Takai D, Jones PA. 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci U S A. 99(6):3740–3745.
Toli E-A, Calboli FCF, Shikano T, Merilä J. 2016. A universal and reliable assay for molecular sex identification of three-spined sticklebacks (*Gasterosteus aculeatus*). Mol Ecol Resour. 16(6):1389–1400.
Trucchi E, et al. 2016. BsRADseq: screening DNA methylation in natural populations of non-model species. Mol Ecol. 25(8):1697–1713.
Varriale A. 2014. DNA methylation, epigenetics, and evolution in vertebrates: facts and challenges. Int J Evol Biol. 2014:1–7.
Vidalis A, et al. 2016. Methylome evolution in plants. Genome Biol. 17(1):264.
Wan ZY, et al. 2016. Genome-wide methylation analysis identified sexually dimorphic methylated regions in hybrid tilapia. Sci Rep. 6:35903.
Wang G, et al. 2014. Gene expression responses of threespine stickleback to salinity: implications for salt-sensitive hypertension. Front Genet. 5:312.
Warton DI, Hui FKC. 2011. The arcsine is asinine: the analysis of proportions in ecology. Ecology 92(1):3–10.
White MA, Kitano J, Peichel CL. 2015. Purifying selection maintains dosage-sensitive genes during degeneration of the threespine stickleback Y chromosome. Mol Biol Evol. 32(8):1981–1995.
Wreczycka K, et al. 2017. Strategies for analyzing bisulfite sequencing data. bioRxiv doi:10.1101/109512.
Xie W, et al. 2013. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. Cell 153(5):1134–1148.
Yelina NE, et al. 2015. DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in *Arabidopsis*. Genes Dev. 29(20):2183–2202.
Zemach A, McDaniel IE, Silva P, Zilberman D. 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science 328(5980):916–919.
Zhang W, Wang X, Yu Q, Ming R, Jiang J. 2008. DNA methylation and heterochromatinization in the male-specific region of the primitive Y chromosome of papaya. Genome Res. 18(12):1938–1943.
Zhang X, et al. 2017. RNA-Seq analysis of salinity stress–responsive transcriptome in the liver of spotted sea bass (*Lateolabrax maculatus*). PLoS One 12(3):e0173238.
Ziller MJ, Hansen KD, Meissner A, Aryee MJ. 2015. Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. Nat Methods. 12(3):230–232.

Associate editor: Kateryna Makova