Genome-Wide DNA Methylation Patterns in Wild Samples of Two Morphotypes of Threespine Stickleback (*Gasterosteus aculeatus*)

Gilbert Smith,*[1] Carl Smith,[2] John G. Kenny,[3] Roy R. Chaudhuri,[3] and Michael G. Ritchie[2]

[1]Department of Ecology and Evolutionary Biology, University of California, Irvine

[2]School of Biology, University of St Andrews, St. Andrews, Fife, United Kingdom

[3]Centre for Genomic Research, Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom

***Corresponding author:** E-mail: gsmith1@uci.edu.

Associate editor: Naoko Takezaki

Abstract

Epigenetic marks such as DNA methylation play important biological roles in gene expression regulation and cellular differentiation during development. To examine whether DNA methylation patterns are potentially associated with naturally occurring phenotypic differences, we examined genome-wide DNA methylation within *Gasterosteus aculeatus*, using reduced representation bisulfite sequencing. First, we identified highly methylated regions of the stickleback genome, finding such regions to be located predominantly within genes, and associated with genes functioning in metabolism and biosynthetic processes, cell adhesion, signaling pathways, and blood vessel development. Next, we identified putative differentially methylated regions (DMRs) of the genome between complete and low lateral plate morphs of *G. aculeatus*. We detected 77 DMRs that were mainly located in intergenic regions. Annotations of genes associated with these DMRs revealed potential functions in a number of known divergent adaptive phenotypes between *G. aculeatus* ecotypes, including cardiovascular development, growth, and neuromuscular development.

Key words: **adaptation, lateral plate morph, ecotype, epigenetic mechanism, phenotypic plasticity.**

Letter

DNA methylation is a chemical modification to DNA that occurs at cytosine residues, and in vertebrates most often occurs at CpG motifs. DNA methylation plays a number of important biological roles, including gene regulation (expression, silencing, and alternative splicing), the regulation of transposable elements, cell type differentiation, genomic imprinting, and sex chromosome inactivation (Bestor 2000; Edwards and Ferguson-Smith 2007; Lisch 2009; Jones 2012). The effects of DNA methylation on gene expression regulation depend on the genomic context in which they occur (Song et al. 2005; Jones 2012). For example, promoter methylation is most often associated with gene repression, whereas intragenic methylation is correlated with gene expression and is likely to control expression from alternative promoter regions (Maunakea et al. 2010). Contiguous regions of methylated CpGs often regulate gene expression, although even a single CpG within a transcription factor binding site could potentially influence gene regulation (Ziller et al. 2013). DNA methylation patterns can be environmentally responsive suggesting that epigenetic marks could underlie phenotypic plasticity, contributing to rapid adaptive evolution by bridging short-term plastic responses, and more slowly accumulating adaptive genotypic changes (Pal and Miklos 1999; Jaenisch and Bird 2003; Bossdorf et al. 2008; Johnson and Tricker 2010; Richards et al. 2010; Smith et al. 2013).

The repeated parallel evolution of threespine stickleback (*Gasterosteus aculeatus*) ecotypes when invading freshwater environments provides a powerful system for the study of adaptive evolution (Boughman 2007; Jones et al. 2012). Freshwater ecotypes have repeatedly evolved following the invasion of freshwater habitats by marine fish, facilitated by glacial retreat at the end of the last ice age that resulted in continental uplift, trapping coastal populations in freshwater lakes across the north temperate zone (McKinnon and Rundle 2002). Stickleback ecotypes differ in body shape and size, number of lateral plates, osmoregulatory processes, life history, and mating behavior; traits that have evolved convergently across populations numerous times. Phenotypic plasticity has also played a role in the formation of *G. aculeatus* ecotypes and likely contributes to the high level of phenotypic variation seen globally (Boughman 2007). Freshwater ecotypes can evolve rapidly, in as few as ten generations (Klepaker 1993; Bell and Foster 1994), and plasticity has been demonstrated in a number of traits, including trophic morphology, allometry, behavior, and aggression (Day et al. 1994; Day and McPhail 1996; Scotti and Foster 2007; Wund et al. 2008; Garduno-Paz et al. 2010; Svanbäck and Schluter 2012; Wund et al. 2012).

In order to examine genome-wide patterns of DNA methylation within and between morphotypes of stickleback, which commonly distinguish marine and freshwater ecotypes, we quantified DNA methylation patterns at the base-pair level, through reduced representation bisulfite sequencing (RRBS) of two lateral plate morphotypes from a single Scottish stickleback population. We sampled eight female *G. aculeatus* individuals; four exhibiting a completely

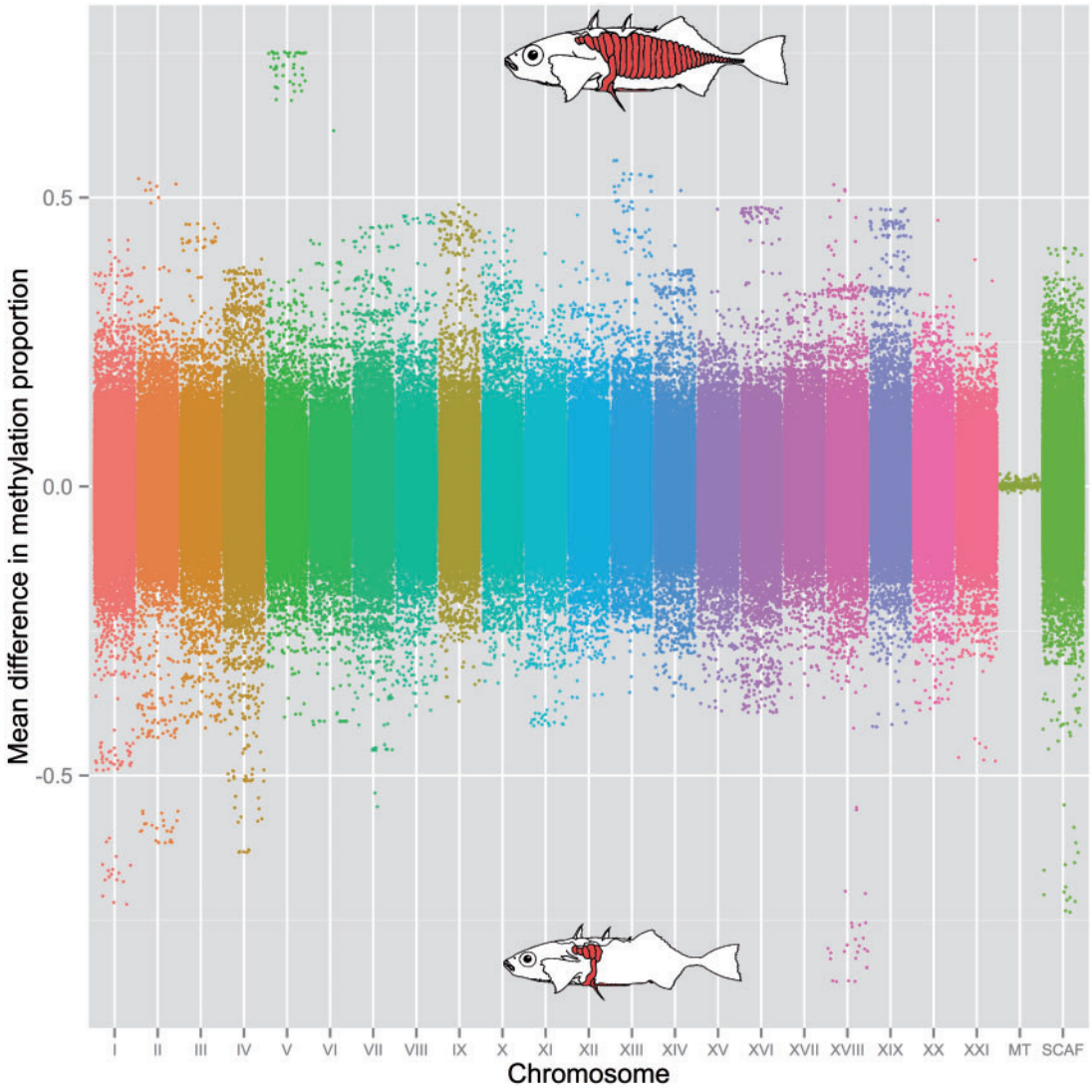

Fig. 1. Methylation differences between complete and low plate morphotypes. Methylation differences are the mean smoothed proportion of methylated Cs at CpG positions across three replicates of each morph. Positive change indicates increased methylation within the complete plate morph (top, lateral plates shown in red), and negative change increased methylation within the low plate morph (bottom, few lateral plates). All chromosomes are presented, including the mitochondrial genome (MT) and genes positioned on scaffolds with unknown chromosomal locations (SCAF).

plated phenotype and four low-plated individuals (see fig. 1 for phenotypes), from a mixed ecotype population in St Andrews, Scotland. In order to minimize differences in cell composition due to sex, age, and tissue sample size, tissues were sampled in a standardized manner from 1+ year class female fish of equivalent size. Further, tissue was sampled from the same location on each individual by removing a sectioned fillet from posterior to the origin of the pectoral fin to the base of the caudal fin, from a single flank. Each sample was thus a mix of several different cell types, encompassing the epidermal, dermal, and subcutaneous tissue, including muscle layers, allowing for the detection of differential methylation across multiple tissue types.

Each of the eight sequencing libraries was mapped to the stickleback genome (assembly BROADS1). On average, 92.8 million cytosines were methylated (sequenced as Cs), representing methylation of 58% of all sequenced cytosines. Only CpG context cytosine methylation was considered further because CpG methylation is the most common functional methylation in vertebrates. To confirm this, we quantified methylation at non-CpG motifs and found on average only 1% of the total non-CpG cytosines were methylated, suggesting a high C to T conversion efficiency. Patterns of DNA methylation data were examined in two ways: 1) to determine highly methylated CpGs within the stickleback genome, and 2) to look for putative differentially methylated CpGs between complete and low morphotypes of stickleback. Differential methylation was carried out with three replicates of completely plated, and three replicates of the low-plated morphs because two samples exhibited very divergent genome-wide methylaton patterns from all other individuals (supplementary fig. S1, Supplementary Material online), and so were excluded from the analysis. This pattern might be linked to the presence of both ecotypes in a mixed

ecotype population. Lateral plate variation is not the ultimate defining feature of marine and freshwater ecotypes, and low frequencies of marine fish exhibit low lateral plating and vice versa.

Highly methylated DNA regions were examined using smoothed methylation proportions calculated in the R package BiSeq (Hebestreit et al. 2013). CpG positions were first clustered into "CpG regions" containing at least ten CpG motifs no more than 100 bp apart. Methylation proportions were then smoothed in order to control for spatial variation in read coverage. DNA methylation in CpG regions was considered high when at least ten cytosines within a CpG region were greater than 90% methylated across all eight samples. The nearest transcription start site (TSS) to each highly methylated CpG region was obtained using the R package GenomicRanges (Lawrence et al. 2013), representing the nearest gene to a region containing highly methylated DNA. The position of methylated regions within genomic features (promoter/exon/intron/intergenic) was determined, giving precedence to promoters > exons > introns > intergenic regions when features overlapped. Enrichment tests of the nearest genes to each highly methylated CpG region were carried out using the R package GOstats (Falcon and Gentleman 2007), using genome annotations from the *G. aculeatus* genome (Jones et al. 2012). GOstats uses a gene set enrichment test to examine significantly overrepresented gene functions based on biological process (BP), molecular function (MF), and cellular component (CC) gene ontology (GO) terms from the annotated *G. aculeatus* genome. *P*-values were corrected using a false discovery rate (FDR) calculated by two methods; *q*-value as per Storey and Tibshirani (2003) and a FDR calculated using SGoF (Carvajal-Rodriguez et al. 2009).

In total, 16,950 CpG positions were found to be highly methylated (> 90%) located within 912 highly methylated CpG regions (containing ten or more methylated CpGs; supplementary table S1, Supplementary Material online). The 912 highly methylated regions were predominantly located in within genes. Significantly more methylated regions were located in exons than expected by chance (*G*-test; $P < 0.001$). Fewer were located in introns (*G*-test; $P < 0.001$) and intergenic regions (*G*-test; $P = 0.002$; fig. 2*A*); however, the former pattern may be due in part to the precedence given to exons when regions overlap an intron/exon boundary. CpG

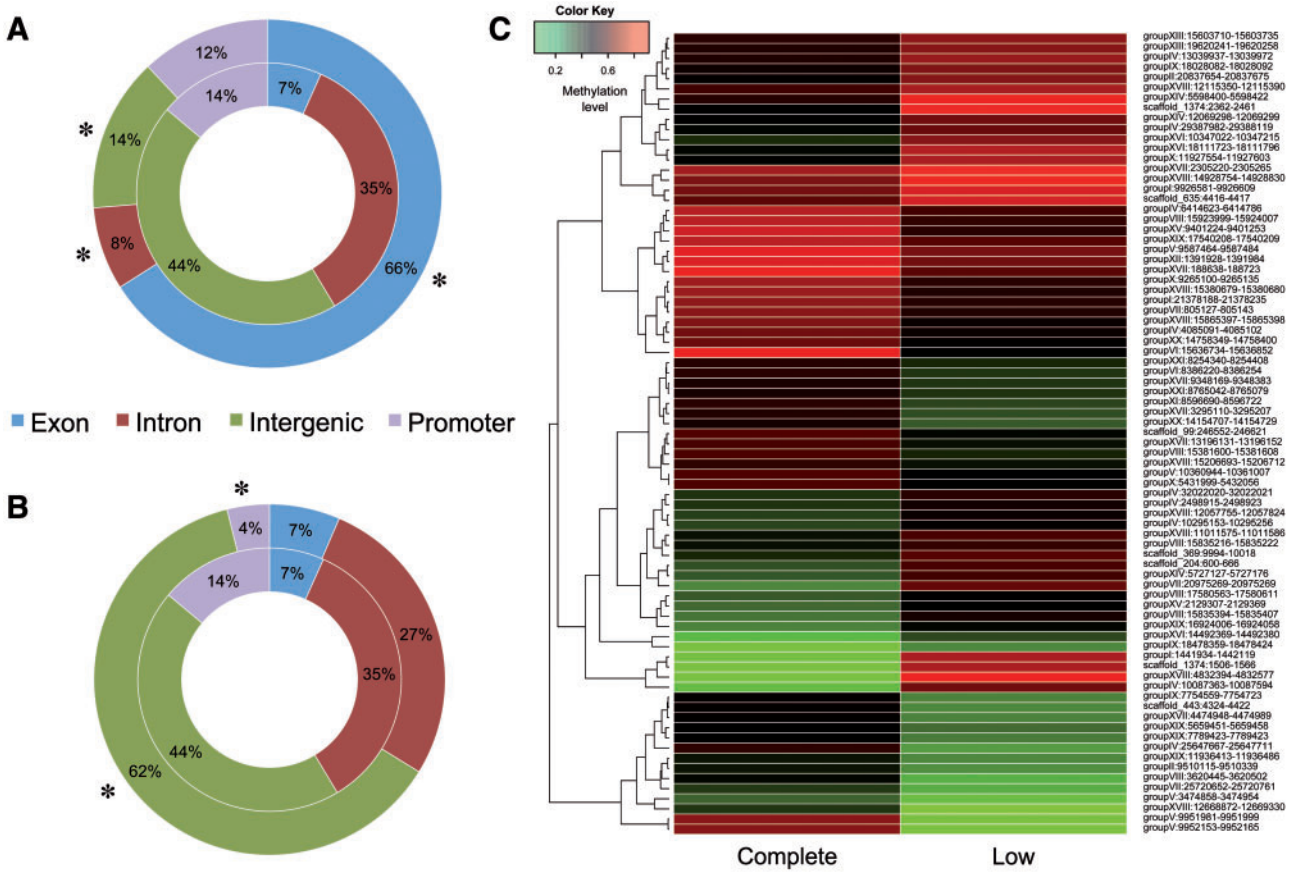

FIG. 2. Comparison of positions of highly methylated regions (*A*), and DMRs (*B*) within genomic features (promoter/exon/intron/intergenic) compared with genome-wide feature content. Outer rings describe the locations of highly methylated regions/DMRs, and inner rings describe the genome-wide feature content. Asterisks denote significant differences between highly methylated regions/DMRs and genome-wide features using a *G*-test at $P < 0.05$. Overlapping genomic features were given the precedence promoters > exons > introns > intergenic. (*C*) Heat map of methylation proportions of the 77 DMRs between complete and low morphotypes. Red indicates higher methylation, green lower methylation, and black 50% methylation. Row labels are the chromosome and genomic coordinates of each DMR.

regions located in promoters were not more likely to be highly methylated than random expectations. A total of 832 unique genes were found to be associated with highly methylated regions, as measured by distance to their TSSs, which were located on average 7,172 bp away from the nearest TSS (supplementary table S2 and fig. S2, Supplementary Material online).

Functional enrichment was performed in GOstats to determine the broad functions of genes near highly methylated regions of DNA. In general, the pattern of functional enrichment was weak with no GO terms being significant after multiple test correction using the *q*-value method of Storey and Tibshirani (2003). However, using corrected *P*-values from SGoF, 20 BP, 17 MF, and 3 CCGO terms were significant at a threshold of FDR < 0.05 (supplementary table S3, Supplementary Material online). BPGO terms included several metabolic and biosynthetic processes, cell adhesion, signaling pathways, blood vessel development, and morphogenesis. MFGO terms included protein binding, metal ion binding, DNA binding, and enzyme regulatory activity. Several genes were involved in epigenetic processes such as acyltransferase and methyltransferase activity involving histones.

Differential methylation analysis was performed using BiSeq (Hebestreit et al. 2013). The raw CpG data were first clustered into contiguous regions and methylation proportions smoothed. Differential methylation was then determined at clustered CpG positions between two sample groups, using a beta regression to model methylation proportions and a Wald test to identify group effects. The correlation between adjacent CpGs was estimated using a variogram, and was used to test for significance under the null hypothesis (with randomly resampled data). BiSeq uses a hierarchical procedure to test for group effects. A weighted FDR controls for multiple tests on the average *P*-value of a region of differentially methylated CpGs (choosing differentially methylated regions [DMRs]), and a second FDR corrects *P*-values of individual differentially methylated positions (DMPs) within a DMR to trim regions and define the DMR boundary. DMRs were defined using an FDR cutoff of 0.15 for DMR choice and 0.15 for DMR trimming, and DMRs were also filtered to retain those with greater than 15% median methylation difference between morphs.

In total, 77 DMRs, comprising 737 differentially methylated positions, were detected between the morphotypes (fig. 1, figs. 2B & C and table 1; supplementary tables S4–S6 and fig. S3, Supplementary Material online). A total of 41 DMRs demonstrated increased methylation in the complete morph and 36 in the low plate morph (table 1). The mitochondrial genome (chromosome MT, fig. 1) is not expected to show a strong signal of DNA methylation in vertebrates, due to the underrepresentation of CpG motifs in mitochondrial genes (Cardon et al. 1994). This lack of methylation was verified here, with low levels of methylation detected across mitochondrial genes and small differences in mitochondrial methylation between morphs.

DMRs were associated with their nearest genes, as measured from the TSS, resulting in 71 unique gene associations.

Table 1. Differential Methylation Results.

	Number of CpGs/DMRs
CpGs tested in each individual	1,445,567
Significant DMRs (composed of DMPs)	77 (737)
Increased methylation in complete plate morph	41
Increased methylation in low plate morph	36
Number of unique nearest genes	71

NOTE.—Table includes the full number of CpG positions (CpGs) considered in the differential methylation analysis, the resulting number of significant DMRs (and positions; DMPs), the number of DMRs with increased methylation in each morphotype, and the number of unique nearest genes to each DMR.

The median distance to the nearest TSS was 9,754 bp (supplementary table S5 and fig. S4, Supplementary Material online). DMRs predominantly lay within intergenic regions (assessed with precedence: promoters > exons > introns > intergenic regions, fig. 2*B*). Significantly more DMRs were discovered within intergenic regions (*G*-test; $P < 0.001$), and fewer in promoter regions (*G*-test; $P < 0.001$), than expected by chance (fig. 2*B*). Smoothed methylation proportions for each DMP within DMRs are available in supplementary table S6, Supplementary Material online.

Functional enrichment analyses were performed in GOstats using the functions of the unique DMR-associated genes. Functional enrichment of DMRs yielded 70 significant BPGO terms using a FDR threshold of <0.05 (table 2; supplementary table S7, Supplementary Material online), and 12 MF terms. CC GO terms were not significant at this level. Significant BP terms involved a range of functions including cardiovascular system development, regulation of cell migration, Notch signaling pathway, insulin-like growth factor pathway, cytoplasmic microtuble organization, amine catabolic process, response to stimulus, and developmental growth. MFGO terms included enzyme regulation and metabolic processes (supplementary table S7, Supplementary Material online).

Functional enrichment of significant DMRs was consistent with phenotypes that differ between marine and freshwater *G. aculeatus* ecotypes. Although our sampling was necessarily based on lateral plate phenotypes, these morphological differences are often consistent between marine and freshwater ecotypes. The population we surveyed is geographically distinct from the best-studied groups; however, recent work has demonstrated a remarkable level of repeated genetic evolution in stickleback freshwater ecotype formation (Jones et al. 2012).

Freshwater stickleback ecotypes have lower maximum oxygen consumption and larger muscle fibers (Dalziel et al. 2012) and heritable differences in hematocrit, ventricle mass, pectoral muscle mass, and pectoral muscle pyruvate kinase activity (Dalziel et al. 2012), which are consistent across independent freshwater populations. Several DMRs between morphs were associated with genes that function in muscle

Table 2. Functional Enrichment Results of Genes Associated with Significant DMRs.

BPGO Term	GOBPID	*P*-Value	FDR	Genes	Total in Genome
Cardiovascular system development	GO:0072358	0.00025315	0.010604432	6	270
Circulatory system development	GO:0072359	0.00025315	0.010604432	6	270
Blood vessel development	GO:0001568	0.000778978	0.021754227	4	125
Vasculature development	GO:0001944	0.001126774	0.023600248	4	138
Angiogenesis	GO:0001525	0.002140618	0.02804114	3	77
Venous endothelial cell migration involved in lymph vessel development	GO:0060855	0.003347	0.02804114	1	1
Negative regulation of blood vessel endothelial cell migration	GO:0043537	0.003347	0.02804114	1	1
Tube development	GO:0035295	0.004001824	0.030479319	3	96
Blood vessel morphogenesis	GO:0048514	0.005706426	0.036907112	3	109
Notch signaling pathway	GO:0007219	0.005726819	0.036907112	2	34
Insulin-like growth factor receptor signaling pathway	GO:0048009	0.006683077	0.038113258	1	2
Regulation of epithelial cell migration	GO:0010632	0.010008266	0.038113258	1	3
Cytoplasmic microtubule organization	GO:0031122	0.010008266	0.038113258	1	3
Negative regulation of cysteine-type endopeptidase activity involved in apoptotic process	GO:0043154	0.010008266	0.038113258	1	3
Regulation of blood vessel endothelial cell migration	GO:0043535	0.010008266	0.038113258	1	3
Regulation of small GTPase mediated signal transduction	GO:0051056	0.014067222	0.042650565	4	281
Venous blood vessel development	GO:0060841	0.016626115	0.042650565	1	5
Regulation of signaling	GO:0023051	0.021833785	0.042650565	5	484
Developmental process	GO:0032502	0.021864641	0.042650565	9	1,268
Negative regulation of hydrolase activity	GO:0051346	0.023200825	0.042650565	1	7
Cellular biogenic amine metabolic process	GO:0006576	0.023200825	0.042650565	1	7
Negative regulation of locomotion	GO:0040013	0.023200825	0.042650565	1	7
Semicircular canal morphogenesis	GO:0048752	0.023200825	0.042650565	1	7
Regulation of vascular endothelial growth factor receptor signaling pathway	GO:0030947	0.023200825	0.042650565	1	7
Regulation of response to stimulus	GO:0048583	0.024351609	0.042650565	5	498
System development	GO:0048731	0.026183545	0.042650565	7	887
Tube morphogenesis	GO:0035239	0.027951258	0.043701764	2	78
Intra-Golgi vesicle-mediated transport	GO:0006891	0.029732669	0.043701764	1	9
Regulation of endopeptidase activity	GO:0052548	0.029732669	0.043701764	1	9
Developmental growth	GO:0048589	0.032746831	0.043861582	2	85
Lymphangiogenesis	GO:0001946	0.032982602	0.043861582	1	10
Spinal cord motor neuron differentiation	GO:0021522	0.03622192	0.046687222	1	11
Lymph vessel development	GO:0001945	0.03622192	0.046687222	1	11
Multicellular organismal development	GO:0007275	0.038855674	0.047901055	8	1,179
Transmembrane receptor protein tyrosine kinase signaling pathway	GO:0007169	0.039349437	0.047901055	2	94

NOTE.—All enriched BP GO terms are significant at FDR < 0.05 and include the GO term ID (GOBPID), the number of genes within the test-set (Genes), and the total number of genes with that function in the stickleback genome (total in genome). Some terms have been removed for brevity (see supplementary table S7, Supplementary Material online, for full results).

development and are expressed in the muscle tissues. For example, *actin binding LIM protein 1b* encodes a cytoskeletal protein that mediates the interactions of actin filaments and cytoplasmic targets (Roof et al. 1997). *Sarcoglycan, delta* is expressed in skeletal and cardiac muscle and is a subunit of sarcoglycan, part of the dystrophin glycoprotein, that forms a link between the F-actin cytoskeleton and the extracellular matrix (Ervasti and Campbell 1993; Hack et al. 2000). DMR-associated genes were also annotated with functions in cardiovascular system development such as *forkhead box C1b*, *delta-like 4*, *T-box 3a*, *kringle containing transmembrane protein 1* and *pleckstrin homology domain containing family A member 7b*.

Changes in osmoregulatory environment are likely to exert a strong selective pressure on the evolution of freshwater *G. aculeatus* ecotypes and can lead to the evolution of an adaptive plastic response in growth rate through genetic accommodation (Robinson 2013). *Gasterosteus aculeatus* ecotypes demonstrate environmental phenotypic plasticity in body size and other morphological characters (Wund et al. 2012). We found several development and growth-related genes associated with DMRs, including *Insulin-like growth*

factor 1b receptor (*Igf1bR*), *teashirt*, *prickle*, *inscrutable*, and *kringle containing transmembrane protein 1* (supplementary table S5, Supplementary Material online). *Igf1bR* plays a role in developmental growth and changes in body mass. In zebrafish (*Danio rerio*), changes in calcium levels can lead to abnormal epithelial cell growth, regulated by interactions between transient receptor potential (TRP) channels and *Igf1R* (Dai et al. 2014). We found a DMR associated with a TRP channel gene (*trpm5*) suggesting that DNA methylation might regulate both *Igf1R* and *trpm5*. *Igf1R* is epigenetically regulated in mice, with DNA methylation in skeletal and cardiac muscle leading to a decrease in gene expression (Nikoshkov et al. 2011). Lastly, one DMR was associated with a calcium regulated signal transduction gene (*RAS guanyl releasing protein 1*).

Several DMR-associated genes involved neuromuscular processes including *katanin*, *NCAM2* and voltage-gated ion channel genes. *Katanin* plays a role in the regulation of synaptic growth at neuromuscular junctions in *Drosophila* (Mao et al. 2014), and *NCAM2* is a neural cell adhesion molecule. Two voltage-gated ion channels were associated with DMRs, which are expressed in neuronal and muscle tissues (*potassium voltage-gated channel, Shaw-related subfamily, member 3b*, and *chloride channel 5*). The *Protocadherin alpha subfamily C, 2* (*PCDHAC2*) gene was associated with a single DMR. PCDH proteins are calcium-dependent neural cell adhesion molecules and are important for segmental plate development, a region of unsegmented paraxial mesoderm from which somites form (embryonic cells giving rise to dermal layers; Murakami et al. 2006). Additionally, two genes (*jagged a1* and *delta-like 4*) were annotated with the Notch signaling pathway, which plays a role in neural differentiation.

We examined genome-wide differential methylation between complete and low plate morphs of *G. aculeatus* across several tissue types and found 77 DMRs associated with genes that play a role in cardiovascular, neuromuscular, and muscular development and growth. It is interesting to note that highly methylated regions were associated with genes that had similar functions to genes associated with DMRs. These common functions included cardiovascular development, cell adhesion, and signaling pathways. It has been suggested that epigenetic mechanisms might underlie phenotypic plasticity (Jaenisch and Bird 2003; Johnson and Tricker 2010; Lira-Medeiros et al. 2010) and that plasticity can enable populations to respond rapidly to novel environments (West-Eberhard 2005; Lande 2009; Scoville and Pfrender 2010). Genetic and environmental contributions to epigenetic variation between morphotypes may have contributed to the patterns we detect. Regardless of the source, our results suggest that the genomes of natural populations of sticklebacks are highly methylated and that variation in methylation patterns may be associated with ecologically important phenotypic differentiation.

Materials and Methods

Gasterosteus aculeatus were collected from Loches Pool, a man-made pond in the University of St Andrews botanic gardens (Latitude 56.3359, Longitude −2.8075; Spence et al. 2012). A census of 369 fish in 2012 showed that 18% of the population expressed the complete plate phenotype. The pond is adjacent (<10 m) to the Kinness Burn, which supports a migratory marine population of *G. aculeatus*, which was the original source of the pond population when it was filled from the river approximately 20 years ago. Despite their close proximity, conditions in the two sites are very different. The burn is a lotic site that supported a small population of brown trout (*Salmo trutta*), and *G. aculeatus* densities are relatively low. The pond is a lentic site from which other fish are absent. The density of *G. aculeatus* is high, and there is a high prevalence of the parasite *Glugea anomala*, with approximately 42% of the population exhibiting at least one externally visible xenoma, some fish hosting greater than 20 xenomas. Avian predators, principally gray herons (*Ardea cinerea*) and kingfishers (*Alcedo atthis*), occur at both sites.

Four complete plate and four low plate phenotype fish were sampled for DNA extraction. All fish were female. DNA was extracted using a DNeasy Blood and Tissue Kit (Qiagen) and assessed for quality and quantity. DNA methylation predominantly occurs on cytosine nucleic acids covalently bonded by a phosphodiester bond to guanine, separated by a phosphorus backbone (CpG motifs). RRBS involves enriching for CpG dense regions of the genome and treating DNA with sodium bisulfite to convert Cs into Ts. If a cytosine is methylated this protects it from conversion and the methylated position is sequenced as a C. DNA was digested using MspI (Fermentas) and treated with sodium bisulfite using the EZ DNA Methylation-Gold kit (Zymo Research, Orange, CA). Library preparation was carried out with the Illumina TruSeq Sample Preparation Kit (Illumina). DNA was end-repaired, A-tailed, and the indexed TruSeq adaptors ligated. Following polymerase chain reaction (PCR), libraries were size selected at a size range of 300–400 bp, quantified by quantitative PCR and pooled for sequencing. Illumina reads for each library are available as fastq files from the European Nucleotide Archive under study accession number PRJEB7912 (http://www.ebi.ac.uk/ena/data/view/PRJEB7912, last accessed December 21, 2014).

Illumina sequencing produced libraries of 100 bp paired-end reads, with an average library size of 59 million reads. Reads were trimmed to remove adaptor sequences and low-quality regions using Trim Galore! (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/, last accessed December 21, 2014) and aligned to the stickleback genome using Bismark (Krueger and Andrews 2011). CpG positions were filtered to remove those with less than 5× coverage in all eight samples. Per base methylation ratios (frequency of Cs and Ts aligned at each cytosine position), for each sample, were quantified, smoothed, and analyzed in the R package BiSeq v1.4.2 (Hebestreit et al. 2013).

Supplementary Material

Supplementary tables S1–S7 and figures S1–S4 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors would like to thank St Andrews Botanic Garden for access to the sample site and Margaret A. Hughes for sequencing assistance, Katja Hebesteit for advice on the analysis, Daniel Barker for general advice, and two anonymous reviewers for their constructive and insightful comments. This work was supported by the Natural Environment Research Council (NERC) through NERC Biomolecular Analysis Facility (NBAF) grant NBAF612.

References

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

Bestor TH. 2000. The DNA methyltransferases of mammals. *Hum Mol Genet.* 9:2395–2402.

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

Boughman JW. 2007. Speciation in sticklebacks. In: Ostlund-Nilsson S, Mayer I, Huntingford FA, editors. Biology of the threespine stickleback. Boca Raton (FL): CRC Press. p. 83–126.

Cardon LR, Burge C, Clayton DA, Karlin S. 1994. Pervasive CpG suppression in animal mitochondrial genomes. *Proc Natl Acad Sci U S A.* 91: 3799–3803.

Carvajal-Rodriguez A, de Una-Alvarez J, Rolan-Alvarez E. 2009. A new multitest correction (SGoF) that increases its statistical power when increasing the number of tests. *BMC Bioinformatics* 10:209–222.

Dai W, Bai Y, Hebda L, Zhong X, Liu J, Kao J, Duan C. 2014. Calcium deficiency-induced and TRP channel-regulated IGF1R-PI3K-Akt signaling regulates abnormal epithelial cell proliferation. *Cell Death Differ.* 21:568–581.

Dalziel AC, Ou M, Schulte PM. 2012. Mechanisms underlying parallel reductions in aerobic capacity in non-migratory threespine stickleback (*Gasterosteus aculeatus*) populations. *J Exp Biol.* 215:746–759.

Day T, McPhail JD. 1996. The effect of behavioural and morphological plasticity on foraging efficiency in the threespine stickleback (*Gasterosteus* sp). *Oecologia* 108:380–388.

Day T, Pritchard J, Schluter D. 1994. A comparison of two sticklebacks. *Evolution* 48:1723–1734.

Edwards CA, Ferguson-Smith AC. 2007. Mechanisms regulating imprinted genes in clusters. *Curr Opin Cell Biol.* 19:281–289.

Ervasti JM, Campbell KP. 1993. A role for the dystrophin-glycoprotein complex as a transmembrane linker between laminin and actin. *J Cell Biol.* 122:809–823.

Falcon S, Gentleman R. 2007. Using GOstats to test gene lists for GO term association. *Bioinformatics* 23:257–258.

Garduno-Paz MV, Couderc S, Adams CE. 2010. Habitat complexity modulates phenotype expression through developmental plasticity in the threespine stickleback. *Biol J Linn Soc.* 100:407–413.

Hack AA, Lam MY, Cordier L, Shoturma DI, Ly CT, Hadhazy MA, Hadhazy MR, Sweeney HL, McNally EM. 2000. Differential requirement for individual sarcoglycans and dystrophin in the assembly and function of the dystrophin-glycoprotein complex. *J Cell Sci.* 113: 2535–2544.

Hebestreit K, Dugas M, Klein H-U. 2013. Detection of significantly differentially methylated regions in targeted bisulfite sequencing data. *Bioinformatics* 29:1647–1653.

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

Johnson LJ, Tricker PJ. 2010. Epigenomic plasticity within populations: its evolutionary significance and potential. *Heredity* 105:113–121.

Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, Swofford R, Pirun M, Zody MC, White S, et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484: 55–61.

Jones PA. 2012. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat Rev Genet.* 13:484–492.

Klepaker T. 1993. Morphological changes in a marine population of threespined stickleback, *Gasterosteus aculeatus*, recently isolated in fresh water. *Can J Zool.* 71:1251–1258.

Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27: 1571–1572.

Lande R. 2009. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *J Evol Biol.* 22: 1435–1446.

Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. 2013. Software for computing and annotating genomic ranges. *PLoS Comp Biol.* 9:e1003118.

Lira-Medeiros CF, Parisod C, Fernandes RA, Mata CS, Cardoso MA, Ferreira PCG. 2010. Epigenetic variation in mangrove plants occurring in contrasting natural environment. *PLoS One* 5:e10326.

Lisch D. 2009. Epigenetic regulation of transposable elements in plants. *Annu Rev Plant Biol.* 60:43–66.

Mao CX, Xiong Y, Xiong Z, Wang Q, Zhang YQ, Jin S. 2014. Microtubule-severing protein *Katanin* regulates neuromuscular junction development and dendritic elaboration in *Drosophila*. *Development* 141: 1064–1074.

Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, et al. 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 466:253–257.

McKinnon JS, Rundle HD. 2002. Speciation in nature: the threespine stickleback model systems. *Trends Ecol Evol.* 17:480–488.

Murakami T, Hijikata T, Matsukawa M, Ishikawa H, Yorifuji H. 2006. Zebrafish protocadherin 10 is involved in paraxial mesoderm development and somitogenesis. *Dev Dyn.* 235:506–514.

Nikoshkov A, Sunkari V, Savu O, Forsberg E, Catrina SB, Brismar K. 2011. Epigenetic DNA methylation in the promoters of the Igf1 receptor and insulin receptor genes in db/db mice. *Epigenetics* 6: 405–409.

Pal C, Miklos I. 1999. Epigenetic inheritance, genetic assimilation and speciation. *J Theor Biol.* 200:19–37.

Richards CL, Bossdorf O, Pigliucci M. 2010. What role does heritable epigenetic variation play in phenotypic evolution? *Bioscience* 60: 232–237.

Robinson BW. 2013. Evolution of growth by genetic accommodation in Icelandic freshwater stickleback. *Proc Biol Sci.* 280:20132197.

Roof DJ, Hayes A, Adamian M, Chishti AH, Li T. 1997. Molecular characterization of abLIM, a novel actin-binding and double zinc finger protein. *J Cell Biol.* 138:575–588.

Scotti MAL, Foster SA. 2007. Phenotypic plasticity and the ecotypic differentiation of aggressive behavior in threespine stickleback. *Ethology* 113:190–198.

Scoville AG, Pfrender ME. 2010. Phenotypic plasticity facilitates recurrent rapid adaptation to introduced predators. *Proc Natl Acad Sci U S A.* 107:4260.

Smith G, Fang Y, Liu X, Kenny J, Cossins A, De Oliveira CC, Etges WJ, Ritchie MG. 2013. Transcriptome-wide expression variation associated with environmental plasticity and mating success in cactophilic *Drosophila mojavensis*. *Evolution* 67:1950–1963.

Song F, Smith JF, Kimura MT, Morrow AD, Matsuyama T, Nagase H, Held WA. 2005. Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc Natl Acad Sci U S A.* 102:3336–3341.

Spence R, Wootton RJ, Przybylski M, Zieba G, Macdonald K, Smith C. 2012. Calcium and salinity as selective factors in plate morph evolution of the three-spined stickleback (*Gasterosteus aculeatus*). *J Evol Biol.* 25:1965–1974.

Storey JD, Tibshirani R. 2003. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A.* 100:9440–9445.

Svanbäck R, Schluter D. 2012. Niche specialization influences adaptive phenotypic plasticity in the threespine stickleback. *Am Nat.* 180: 50–59.

West-Eberhard MJ. 2005. Phenotypic accommodation: adaptive innovation due to developmental plasticity. *J Exp Zool B Mol Dev Evol.* 304: 610–618.

Wund MA, Baker JA, Clancy B, Golub JL, Fosterk SA. 2008. A test of the "Flexible stem" model of evolution: ancestral plasticity, genetic accommodation, and morphological divergence in the threespine stickleback radiation. *Am Nat.* 172:449–462.

Wund MA, Valena S, Wood S, Baker JA. 2012. Ancestral plasticity and allometry in threespine stickleback reveal phenotypes associated with derived, freshwater ecotypes. *Biol J Linn Soc.* 105:573–583.

Ziller MJ, Gu H, Muller F, Donaghey J, Tsai LT, Kohlbacher O, De Jager PL, Rosen ED, Bennett DA, Bernstein BE, et al. 2013. Charting a dynamic DNA methylation landscape of the human genome. *Nature* 500: 477–481.