MOLECULAR ECOLOGY

Molecular Ecology (2016) **25**, 1785–1800 doi: 10.1111/mec.13231

EPIGENETIC STUDIES IN ECOLOGY AND EVOLUTION

Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout

MELINDA R. BAERWALD,* MARIAH H. MEEK,* MOLLY R. STEPHENS,† RAMAN P. NAGARAJAN,‡ ALISHA M. GOODBLA,* KATHARINE M. H. TOMALTY,* GARY H. THORGAARD,§ BERNIE MAY* and KRISTA M. NICHOLS¶

*Department of Animal Science, University of California – Davis, Davis, CA 95616, USA, † School of Natural Sciences, University of California – Merced, Merced, CA 95343, USA, ‡GlaxoSmithKline, Cancer Epigenetics Discovery Performance Unit, Collegeville, PA 19426, USA, §School of Biological Sciences and Center for Reproductive Biology, Washington State University, Pullman, WA 99164, USA, ¶Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd E, Seattle, WA 98112, USA

Abstract

Migration is essential for the reproduction and survival of many animals, yet little is understood about its underlying molecular mechanisms. We used the salmonid *Oncorhynchus mykiss* to gain mechanistic insight into smoltification, which is a morphological, physiological and behavioural transition undertaken by juveniles in preparation for seaward migration. *O. mykiss* is experimentally tractable and displays intra- and interpopulation variation in migration propensity. Migratory individuals can produce nonmigratory progeny and vice versa, indicating a high degree of phenotypic plasticity. One potential way that phenotypic plasticity might be linked to variation in migration-related life history tactics is through epigenetic regulation of gene expression. To explore this, we quantitatively measured genome-scale DNA methylation in fin tissue using reduced representation bisulphite sequencing of F_2 siblings produced from a cross between steelhead (migratory) and rainbow trout (nonmigratory) lines. We identified 57 differentially methylated regions (DMRs) between smolt and resident *O. mykiss* juveniles. DMRs were high in magnitude, with up to 62% differential methylation between life history types, and over half of the gene-associated DMRs were in transcriptional regulatory regions. Many of the DMRs encode proteins with activity relevant to migration-related transitions (e.g. circadian rhythm pathway, nervous system development, protein kinase activity). This study provides the first evidence of a relationship between epigenetic variation and life history divergence associated with migration-related traits in any species.

Keywords: epigenetics, life history variation, *Oncorhynchus mykiss*, plasticity, smoltification

Received 5 January 2015; revision received 30 April 2015; accepted 1 May 2015

Introduction

Migration is a complex phenomenon, with billions of animals annually responding to environmental heterogeneity by moving long distances (Bowlin *et al.* 2010). This allows them to exploit seasonally or geographically available resources (e.g. food, habitat, favourable climate or breeding conditions) or avoid unfavourable conditions (e.g. predators, disease; Alerstam *et al.* 2003). Migratory behaviour occurs across all major branches of the animal kingdom, in groups as diverse as fish, insects, reptiles, crustaceans, birds and mammals (Alerstam *et al.* 2003). It has evolved repeatedly within lineages (e.g. Chinook salmon *Oncorhynchus tshawytscha* (Waples *et al.* 2004), bats (Mammalia: Chiroptera) (Bisson *et al.* 2009)), and in some cases, rapidly (e.g. house finches *Carpodacus mexicanus* (Able & Belthoff 1998)).

Correspondence: Melinda R. Baerwald, Fax: +1 (530) 752-0175; E-mail: mrbaerwald@ucdavis.edu

Mariah H. Meek and Molly R. Stephens contributed equally to this work.

Changes in the propensity, timing and patterns of migration may have large-scale implications for patterns of species connectivity and ecosystem health. Yet, despite human fascination with animal migrations (Nebel 2010), the molecular mechanisms that underlie migratory behaviour are not well understood (Liedvogel *et al.* 2011). With rapidly changing environmental conditions and the associated challenges of managing species for future persistence, it is becoming critical to understand the mechanisms driving migration (Liedvogel *et al.* 2011).

Substantial phenotypic variation in migration trajectory, timing and propensity exists across species and populations; this variability is attributed to a combination of environmental and genetic factors (Liedvogel *et al.* 2011; Dingle 2014). For example, many Pacific salmon species display variable migration timing during juvenile outmigration to the ocean, as well as timing in adult return to freshwater for spawning (Quinn & Myers 2005). These salmon species, and populations within them, also have variable migration trajectories, with some staying in coastal waters and others migrating to the open sea from rivers (Quinn & Myers 2005).

Organisms respond to environmental cues (e.g. photoperiod, temperature, food availability), which can trigger internal cues (e.g. circadian rhythm, hormones, and fat deposits) to initiate migration (Folmar & Dickhoff 1980; McCormick *et al.* 1998). Heritability and genetic mapping studies across diverse taxa suggest there is genetic variation for the propensity to migrate (Berthold 1991; Thrower *et al.* 2004; Roff & Fairbairn 2007; Hecht *et al.* 2015). Additionally, phenotypic plasticity is believed to play a role in producing alternative migratory phenotypes (Hayes *et al.* 2012; Dodson *et al.* 2013). One potential way that phenotypic plasticity might be linked to variation in migratory tactics is through epigenetic regulation of gene expression (e.g. DNA methylation, histone tail modifications, noncoding RNAs). Epigenetic modifications can mediate responses to environmental cues (Bastow *et al.* 2004; Petronis 2010; Feil & Fraga 2012), including behavioural (Champagne 2010, 2012; Mifsud *et al.* 2011) and circadian rhythm responses (DiTacchio *et al.* 2011; Fustin *et al.* 2013; Azzi *et al.* 2014). For example, DNA methylation is involved in transcriptional silencing, alternative splicing and activating intragenic promoters (reviewed in Jones 2012) and acts as a reversible mechanism to drive circadian clock behavioural plasticity in mice (*Mus musculus*; Azzi *et al.* 2014).

Due to its experimental tractability and remarkable variability in migration propensity, the salmonid *Oncorhynchus mykiss* is an excellent model for studying mechanisms driving migration (Zimmerman & Reeves 2000; Riva-Rossi *et al.* 2007; Hayes *et al.* 2012; Courter *et al.* 2013). The species has two distinct migration life history strategies: migratory (anadromous steelhead) and nonmigratory (resident rainbow trout). Both migrant and nonmigrant individuals are born and rear in freshwater to the juvenile stage. Some juveniles then go through a process termed smoltification, during which they undergo morphological, physiological and behavioural changes before migrating to the ocean (Hoar 1976); this migratory or anadromous form of *O. mykiss* is called a steelhead trout. Upon reaching reproductive maturity in the ocean, steelhead migrate back into freshwater to spawn, thus beginning the life cycle again. In contrast, the juveniles that do not undergo smoltification remain in freshwater throughout their lives; this resident form of *O. mykiss* is called a rainbow trout.

Migration-related life history traits are highly heritable within *O. mykiss* (Thrower *et al.* 2004; Hecht *et al.* 2015), but there is also substantial phenotypic variability among families and lineages. Genetic mapping studies have identified quantitative trait loci (QTL) associated with anadromy and residency across the genome. A linkage mapping study using the same F_2 family as this current study found QTL associated with smoltification-related traits across many chromosomes, including a major effect locus on chromosome Omy10 (Nichols *et al.* 2008). A genome scan of two closely related steelhead–rainbow trout populations found non-neutral outlier loci associated with smoltification across 19 chromosomes (Martínez *et al.* 2011). Both linkage mapping and genomewide association studies (GWAS) have found many loci of small effect contributing to migration-related phenotypic differences across the genome using geographically diverse populations (Hecht *et al.* 2012a,b; Hale *et al.* 2013). However, there is evidence that some chromosomes (e.g. Omy5) contain strong associations with smoltification across many distinct populations (Pearse *et al.* 2014), and genetic mapping of this complex trait is ongoing. Genes found proximal to identified QTL include those with functions pertaining to circadian rhythm, sexual maturation, development rate and osmoregulation (Martínez *et al.* 2011; Hecht *et al.* 2012a; Hale *et al.* 2013). However, the specific genes triggering the developmental commitment to migration versus residency are not yet known.

Migration is a threshold trait (Pulido *et al.* 1996; Dodson *et al.* 2013), with manifestation dependent on interactions between the environment and an organism's genetic architecture and development (e.g. body size (Jonsson & Jonsson 1993; Theriault & Dodson 2003)). For *O. mykiss*, environmental and landscape variables such as temperature, photoperiod, precipitation, elevation and upstream distance can influence migration propensity (Zaugg & Wagner 1973; Narum *et al.* 2008;

Kendall *et al.* 2015). There is considerable phenotypic plasticity in life history diversity, with some populations containing individuals of a single life history strategy, while other populations harbour individuals of both strategies (Jonsson & Jonsson 1993; Quinn & Myers 2005; Dodson *et al.* 2013). Each life history type can give rise to the other in both wild (Riva-Rossi *et al.* 2007; Courter *et al.* 2013) and hatchery (Viola & Schuck 1995; McMichael *et al.* 1997) settings. We hypothesize that the observed plasticity may be controlled by an epigenetic response threshold, which integrates parentally inherited factors, environmental factors and developmental history (Yan *et al.* 2014). Therefore, divergent life history phenotypes may be a result of differential epigenetic modifications that link environmental factors and the genome to regulate internal cues (e.g. circadian rhythm). In this work, we evaluate whether epigenetic regulation (DNA methylation) is associated with smoltification, a migration-related phenotype, by comparing smolt and resident phenotypes from an *O. mykiss* F_2 intercross reared in a common environment.

Materials and methods

Samples

The *O. mykiss* F_2 intercross used for this study was previously used for a QTL mapping study of smoltification wherein individual fish were quantitatively classified as smolts or residents (Nichols *et al.* 2008). In that study, steelhead (Clearwater River) and resident (Oregon State University) doubled haploid clonal lines (Nichols *et al.* 2007) were crossed to create F_1 hybrids. Gametes from an F_1 male then underwent androgenesis, to create F_2 progeny, each homozygous across loci, but genetically different from their siblings due to meiotic recombination. At the age of 2, when juveniles often begin to undergo smoltification, metrics strongly associated with smoltification (body size, condition, growth, morphology, osmoregulatory ability and skin reflectance) were measured. The measurements from these phenotypes were compiled to identify a binary phenotype, classifying each individual as either a smolt or resident. Smolts were silver and streamlined in body shape, while residents had darker coloration, dark parr marks on the sides of their bodies, and greater body and caudal peduncle depths (Fig. 1; methods detailed in Nichols *et al.* 2008). Ten individuals (five males and five females) of each phenotype were selected for reduced representation bisulphite sequencing (RRBS). DNA was extracted from caudal fin using the Qiagen DNeasy kit. This tissue type is in intimate contact with the environment and undergoes phenotypic changes during smoltification, such as increased silvering and darkening,

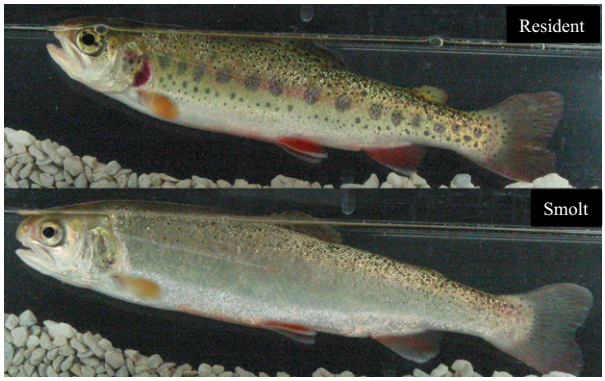

Fig. 1 Phenotypic differences between resident and smolt *Oncorhynchus mykiss* juveniles. Photographs of siblings with alternative life history types were taken on same date and age (June, age 2). Smolts are silvery and streamlined in body shape, while residents are darker in skin and fin coloration and have dark parr marks on the sides of their bodies.

which are associated with metabolic changes that occur during the time frame leading up to migration (McCormick *et al.* 1998). We chose to use this doubled haploid cross in the first analysis of differential methylation between migration-related phenotypes because (i) we could directly compare our differential methylation results with the location of QTL for smoltification using fin tissue previously used for a genetic analysis of smoltification (Nichols *et al.* 2008) and (ii) the doubled haploids come from a homogenous genetic background and were reared in a common controlled environment, reducing any potentially confounding noise from large amounts of genetic diversity or environmental variability.

Reduced representation bisulphite sequencing

Reduced representation bisulphite sequencing libraries were constructed using a previously published protocol (Boyle *et al.* 2012). Briefly, DNA was quantified using the Qubit 2.0 fluorometer (Life Technologies) with the dsDNA BR assay kit. Genomic DNA (100 ng) for each individual was digested using the restriction enzyme *MspI* (New England Biolabs). DNA fragments were end-repaired, A-tailed, and ligated to the methylated Illumina TruSeq adapters. Ten samples were multiplexed per library, for a total of two libraries, with each sample barcoded with unique Illumina adapters to distinguish samples. DNA from fish representing the two life history types and sexes was equalized across libraries. Sodium bisulphite conversion was conducted using the EpiTect Bisulfite kit (Qiagen), following the protocol for formalin-fixed paraffin-embedded samples. After two rounds of bisulphite treatment to ensure complete conversion, the DNA was amplified using

Illumina's TruSeq primers. The optimal number of PCR cycles (18) was determined empirically by comparing band intensities of PCR products on a 5% MetaPhor agarose gel (Lonza) for 10, 13, 16, 19 and 22 total cycles. Gel-free library size selection was performed with Agencourt AMPure XP beads (Beckman Coulter) as described by Boyle *et al.* (2012). Library DNA quantity and quality were assessed using a Qubit fluorometer and 5% MetaPhor agarose gel, respectively. After a quantitative PCR quality control check, the two libraries were sequenced using two lanes (50-bp single-end reads, 10 samples per lane) of an Illumina HiSeq 2000 at the Vincent J. Coates Genomics Sequencing Laboratory (University of California, Berkeley, USA). The RRBS protocol was repeated for nine samples with fewer than 10 million reads after quality filtering. These samples were sequenced as above, but with 100-bp single-end reads. Base calling was performed using Illumina's CASAVA 1.8.0. Initial read quality checks were carried out using FASTQC (available at: http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc).

Trim Galore! (available at http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) removed low-quality and adapter-contaminated sequences, using the 'rrbs' option.

Read alignments to two O. mykiss genome versions

Recently, two versions of the *O. mykiss* genome (the French Genome Project (Berthelot *et al.* 2014) and the U.S. Genome Project (available at http://www.animalgenome.org/repository/aquaculture/) were released to facilitate large-scale genomic studies. Both assemblies used the same DNA source, but were produced using different methods. The French Genome Project (hereafter this project's assembly will be referred to as the French genome) used a whole-genome shotgun sequencing approach, which enables all regions of the genome to be targeted, but has difficulty mapping reads from repetitive elements and paralogous sequences. This is a potential issue since paralogs exist throughout the salmonid genome due to an autotetraploidization event 25–100 million years ago (Allendorf & Thorgaard 1984). Contrastingly, the U.S. Genome Project (hereafter this project's assembly will be referred to as the U.S. genome) used a BAC library sequencing approach, which is better able to uniquely identify and map repetitive and paralogous sequences, but is limited to the regions of the genome contained within the BAC libraries. While these two *O. mykiss* genome assemblies have considerable sequence overlap, they also contain unique sequences. Conducting analyses with both genome assemblies in a comparative fashion is likely the most comprehensive approach until further genome development. Therefore, for the purposes of alignment, both genomes were *in silico* converted to bisulphite genomes using Bismark (Krueger & Andrews 2011) in order to independently align reads to each genome. Although the French genome has longer N50 scaffolds, both genome assemblies contain hundreds to thousands of unassembled contigs. Therefore, to increase mapping efficiency (i.e. ability of reads to uniquely map to a single region in the genome), small scaffolds (<50 kb for the French genome or <10 kb for the U.S. genome) were removed prior to analysis. Bowtie (Langmead *et al.* 2009), implemented in Bismark, was used to align reads for each sample to the genome, tolerating one non-bisulphite mismatch per read, and producing SAM files for further analyses of differential methylation. Only reads that mapped uniquely to at least one of the reference genomes were included in downstream differentially methylated CpG site (DMC) and differentially methylated region (DMR) analyses. The per cent of French genome CpG sites that are encompassed within an *MspI* recognition sequence ('CCGG') was *in silico* predicted using the SimRAD package (Lepais & Weir 2014) within the R computing environment (R Core Team 2014).

Individual differentially methylated CpG (DMC) sites were identified between resident and smolt groups using logistic regression with the R package methylKit (Akalin *et al.* 2012). Read coverage was normalized between samples. A minimum of seven individuals per life history type, each with a minimum of ten reads, were required at a CpG site for that site to be analysed. Default parameters were used (Q-value < 0.01), with a minimum required methylation difference of 25% between smolt and resident phenotypes.

Differentially methylated regions (DMRs) were determined using the R package eDMR (Li *et al.* 2014) with default parameters. To be considered significant, a DMR needed to contain at least three CpG sites within an algorithm-specified genomic distance (see Li *et al.* 2014 for details), with at least one classified as a DMC (Q-value < 0.01), and an absolute mean methylation difference greater than 20% when comparing smolt and resident phenotypes. Visualization of sequence reads within and across individuals, and compared to the reference genomes, was performed with Integrative Genomics Viewer (Robinson *et al.* 2011). Visualization of the differential methylation patterns across individuals was achieved using the heatmap2 function in the R package ggplot2 (Wickham 2009).

CpG island proximity, gene annotation, gene ontology and pathway analysis

For CpG island prediction, EMBL-EBI's Cpgplot (available at http://www.ebi.ac.uk/Tools/seqstats/) was

used with default parameters. DMRs were considered part of CpG shores if they were ≤2 kb away from an island (Irizarry *et al*. 2009) and part of CpG shelves if they were between 2 and 4 kb away from an island, as defined by Illumina.

We used published annotations of the French *O. mykiss* reference genome (available at https://www.genoscope.cns.fr/trout/cgi-bin/gbrowse/truite) for DMRs that aligned to this genome assembly. The U.S. *O. mykiss* genome has yet to be annotated, so for DMRs that only aligned to this genome assembly, we used Bowtie to align the entire scaffold to the French *O. mykiss* genome to identify existing annotations. For the few reads that did not align, we used NCBI's megablast query to determine region-specific annotations for the DMR regions. For all annotations, scaffolds containing DMRs were trimmed to include 2.5 kb upstream and downstream of the DMR. An e-value threshold of <1e-50 was required to assign a gene name.

We used Gene Cards (available at http://genecards.org/) to identify gene ontology terms for the DMRs annotated as part of proximal regulatory regions (5′ or 3′ UTR) or gene bodies. Pathway analysis was conducted using Qiagen's Ingenuity Pathway Analysis (IPA), which identified functional associations among the annotated DMRs, such as particular canonical pathways and common gene networks found across the DMRs.

Results

Reduced representation bisulphite sequencing coverage and mapping

Total reads per library ranged from 135 to 147 million after quality filtering (Table S1, Supporting information). Total number of reads per individual ranged from 4 to 49 million (average: 14 million) after quality filtering. Two of the twenty samples did not have 10 million reads and did not undergo further analysis (samples 19 and 20 in Table S1, Supporting information). The bisulphite conversion was highly efficient, with all samples averaging >99.6% (median: 100%) conversion. One-third (33.2%) of the bisulphite converted reads mapped uniquely to the French version of the *O. mykiss* genome (i.e. at a single location; Table S1, Supporting information). One-quarter (25.9%) of the reads mapped to more than one genomic location. For the U.S. *O. mykiss* genome, 24.4% of the reads mapped uniquely, while 35.7% mapped to multiple locations.

Identification of differentially methylated regions

A total of 117,082 and 158,381 CpG sites met the minimum coverage requirements for the U.S. and French *O. mykiss* genomes, respectively. For context, there were 22.8 million CpG sites in the French *O. mykiss* genome after removal of small contigs (refer to Materials and Methods). Therefore, <1% of all CpG sites were analysed per genome assembly due to (i) reads not uniquely mapping, (ii) reads not being found in or near the *MspI* recognition sequence ('CCGG'), or (iii) low sequence coverage. Single differentially methylated cytosines (DMCs) ranged from 113 (U.S.) to 521 (French). On the basis of these DMCs, we identified 57 DMRs after false discovery rate correction when comparing phenotypically classified resident and smolt life history types (Table 1). Given our lower limit differential methylation cut-off (20%), methylation differences between resident and smolt groups ranged from 20 to 62% per DMR. Individuals clustered primarily by life history type based on the 57 DMRs (Fig. 2). However, one smolt clustered with residents and two residents clustered with smolts. Because adult phenotypes (body morphology, sexual maturation) were recorded after these individuals reached age three, it is unlikely that these individuals were phenotypically misclassified. Specifically, fish that were silvery and streamlined in their appearance were classified as smolts, and those that were dark coloured, more deep-bodied, or producing gametes were classified as resident individuals. The DMRs displayed hyper- and hypomethylation in nearly equal proportions (28 hypermethylated and 29 hypomethylated when comparing residents to smolts). Each DMR contained 3–21 total CpG sites (average = 5), with 1–12 of these CpGs being differentially methylated (average = 3).

Thirteen of the DMRs were identified with both reference genomes, 31 with only the French genome, and 13 with only the U.S. genome, which has a lower N50 scaffold size in comparison with the French genome. Additional details about each of these DMRs, including contig and scaffold positions, can be found in Table S2 (Supporting information).

Genomic features and gene annotation

We mapped 37 DMRs onto 12 chromosomes (Table S2, Supporting information). Chromosomes Omy10 and Omy19 contained the most DMRs, with 7 and 9, respectively (Fig. 4a; Table S2, Supporting information). Five other chromosomes (Omy3, Omy4, Omy7, Omy16 and Omy20) contained two or more DMRs. Five chromosomes each contained a single DMR, and twenty DMRs could not be mapped to a chromosome due to the incomplete nature of the *O. mykiss* reference genome(s).

Twenty-six of the DMRs localized within or proximal to annotated genes, and a subset directly overlapped

Table 1 Differentially methylated regions (DMRs) when comparing resident and smolt juvenile *Oncorhynchus mykiss* phenotypes

Symbol	Entrez gene name	Meth diff[†]	*P*-value	*Q*-value
AATK	Apoptosis-associated tyrosine kinase	−20.7	1.1E-21	1.4E-21
ADHFE1	Alcohol dehydrogenase, iron containing, 1	−38.2	6.2E-29	9.9E-29
ADORA3	Adenosine A3 receptor	−23.5	5.4E-45	2.4E-44
AHR2A	Aryl hydrocarbon receptor 2 alpha	31.9	1.1E-43	3.6E-43
CCDC93	Coiled-coil domain containing 93	20.7	6.8E-18	8.6E-18
CD247	CD247 molecule	22.8	3.2E-34	6.3E-34
CLDN10	Claudin 10	29.5	3.0E-36	6.8E-36
CRYBB3	Crystallin, beta B3	21.3	1.0E-35	2.1E-35
DDX10	DEAD (Asp-Glu-Ala-Asp) box polypeptide 10	29.5	1.2E-31	2.2E-31
GEMIN2	Gem (nuclear organelle) associated protein 2	−34.1	2.5E-30	4.1E-30
GRM1	Glutamate receptor, metabotropic 1	−31.3	1.2E-53	1.0E-52
HES6	Hes family bHLH transcription factor 6	−20.7	2.2E-14	2.6E-14
KIAA1467	KIAA1467	23.2	1.8E-17	2.2E-17
MYOM3	Myomesin 3	−21.7	7.9E-10	8.0E-10
NEUROD4	Neurogenic differentiation 4	−31.4	8.3E-36	1.8E-35
NHLH2	Nescient helix-loop-helix 2	38.0	3.2E-76	1.4E-74
PALLD	Palladin, cytoskeletal-associated protein	24.4	7.2E-62	7.9E-61
PCDHGC5	Protocadherin gamma subfamily C, 5	48.8	3.1E-49	1.7E-48
RASD1	RAS, dexamethasone-induced 1	21.5	1.8E-16	2.2E-16
RGL2	Ral guanine nucleotide dissociation stimulator-like 2	21.2	1.2E-15	1.3E-15
RSAD2	Radical S-adenosyl methionine domain containing 2	−61.8	2.7E-69	5.9E-68
SBK1	Serine/threonine protein kinase SBK1-like	−28.6	7.0E-86	1.9E-84
SETD3	SET domain containing 3	42.5	5.4E-53	4.0E-52
STYK1	Serine/threonine/tyrosine kinase 1	28.3	3.5E-40	1.3E-39
TMEM150A	Transmembrane protein 150A	−23.6	6.4E-24	8.7E-24
TUBA1A	Tubulin, alpha 1a	−26.9	9.1E-34	1.7E-33
—	Uncharacterized protein	42.0	1.2E-30	2.1E-30
—	Uncharacterized protein	−21.1	1.3E-10	1.3E-10
—	Uncharacterized protein	−21.7	2.7E-11	3.0E-11
—	Uncharacterized protein	−31.7	8.3E-36	1.8E-35
—	Uncharacterized protein	−32.3	1.4E-22	1.7E-22
—	Unknown	−21.0	7.9E-29	1.2E-28
—	Unknown	−33.8	1.2E-43	3.6E-43
—	Unknown	20.6	1.1E-09	1.1E-09
—	Unknown	23.7	2.0E-28	3.2E-28
—	Unknown	26.2	2.1E-38	7.1E-38
—	Unknown	−30.0	2.2E-42	8.9E-42
—	Unknown	31.9	8.3E-36	1.8E-35
—	Unknown	29.1	5.3E-46	2.6E-45
—	Unknown	21.3	3.9E-23	5.5E-23
—	Unknown	−20.2	1.4E-31	2.4E-31
—	Unknown	−22.3	6.7E-37	2.0E-36
—	Unknown	−23.7	2.8E-49	1.7E-48
—	Unknown	−24.0	6.8E-38	2.1E-37
—	Unknown	−20.5	1.9E-11	2.1E-11
—	Unknown	30.5	5.3E-23	7.3E-23
—	Unknown	20.6	4.6E-11	4.9E-11
—	Unknown	−21.6	2.1E-23	3.0E-23
—	Unknown	24.4	8.3E-36	1.8E-35
—	Unknown	26.6	6.1E-28	9.3E-28
—	Unknown	−31.3	1.5E-65	2.2E-64
—	Unknown	−31.0	8.3E-36	1.8E-35
—	Unknown	20.1	6.9E-37	1.7E-36
—	Unknown	−23.1	6.1E-33	1.1E-32
—	Unknown	35.4	8.7E-60	1.2E-58
—	Unknown	22.0	8.7E-27	1.2E-26
—	Unknown	−26.3	7.5E-22	1.0E-21

[†]Per cent methylation differences averaged from all CpG sites within the defined region. The comparison is between residents and smolts with positive values representing increased methylation for residents and negative values representing increased methylation for smolts.

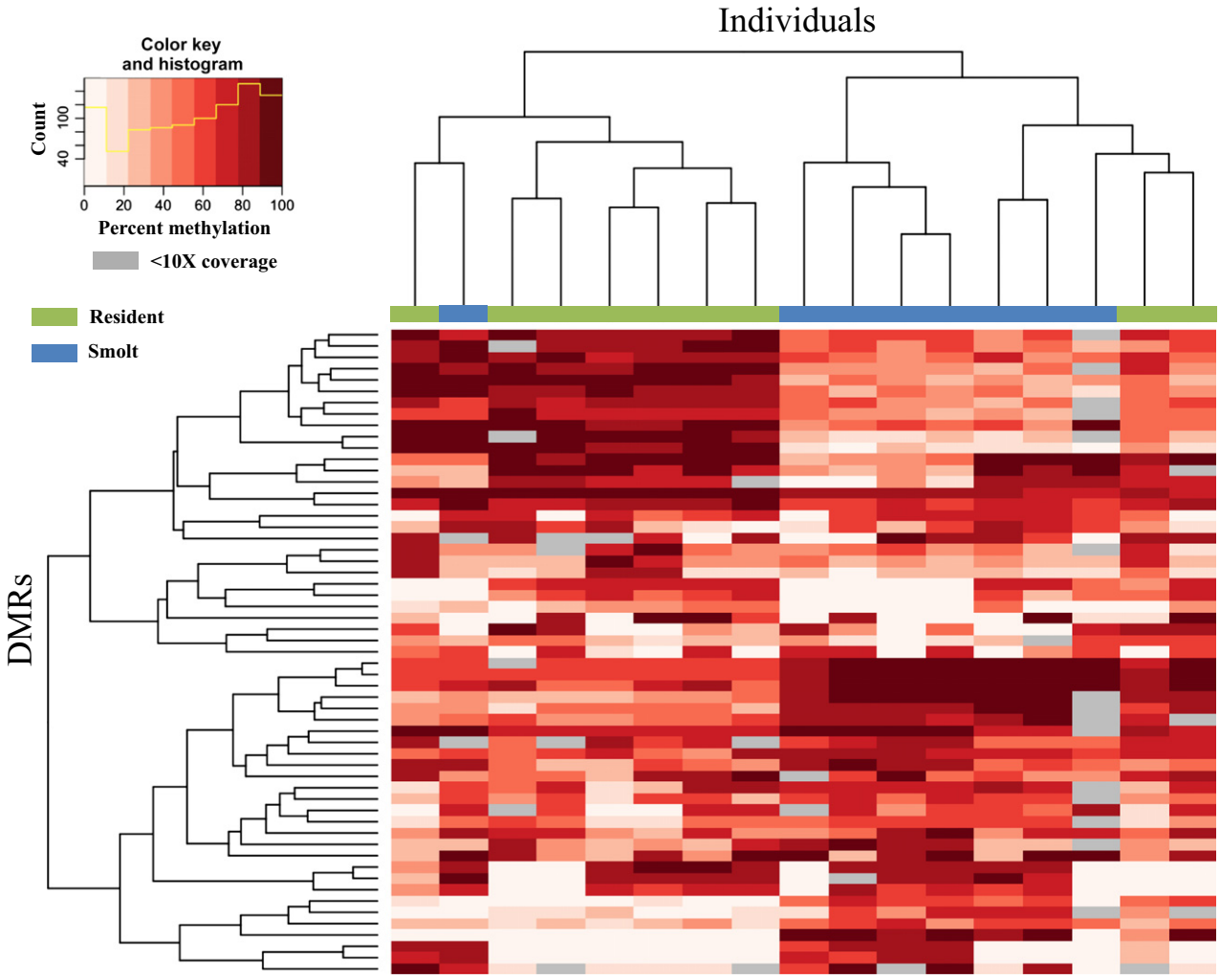

Fig. 2 Hierarchical clustering of hyper- and hypomethylated DMRs for divergent migration-related phenotypes. This heatmap graphically displays the per cent DNA methylation for each individual at the 57 DMRs. Each column represents a colour-coded individual: green for residents and blue for smolts. Each row represents a differentially methylated region when comparing the two life history phenotypes. The darker the red, the more methylated that individual is for that DMR. Individual dendrogram positions are based on their overall methylation patterns across the 57 DMRs. DMRs below a minimum read depth of 10 were excluded and are represented as grey boxes.

genomic regions with potential regulatory function. Over half of the DMRs for this study were found within CpG islands (CGIs), shores or shelves (Fig. 4b). We found five of the annotated DMRs within or proximal to CGIs (i.e. CGI shores, shelves) located within putative promoter regions (Table S2, Supporting information). Six of the annotated DMRs were found in CGIs, shores or shelves within gene bodies (Table S2, Supporting information).

Thirty-one of the 57 DMRs were found within the proximal putative regulatory or coding regions of annotated genes (Fig. 4c; Table S2, Supporting information). Specifically, 15 were found within gene bodies, 8 within 5 kb upstream of the transcription start site (TSS), 4 within 5–10 kb upstream of the TSS, and 4 within 1.5 kb downstream of the transcriptional stop codon. Twenty-six DMRs were located in or proximal to annotated *O. mykiss* genes (Table 1). Of the remaining 31 DMRs, 5 were assigned as uncharacterized proteins (Table 1; Table S2, Supporting information). In addition to proximal regulatory region and gene body methylation differences, 26 of the DMRs were found within intergenic regions.

The most common gene ontology terms found across the 26 annotated DMRs included signal transduction and G-protein-coupled receptor signalling pathway (biological process), protein binding (molecular function), and integral component of membrane, plasma

membrane and nucleus (cellular component; Table S3, Fig. S1, Supporting information).

Pathway analysis

Several molecular pathways associated with smoltification were identified via Ingenuity Pathway Analysis. Six of the top canonical pathways for the gene-associated DMRs were statistically significant (Table S4, Supporting information). We identified two gene networks, each with distinct functional roles as well as a common role relating to nervous system development and function (Table S4, Supporting information). These networks contained 11–15 of the 26 gene-associated DMRs (Fig. S2, Supporting information).

Discussion

Migration plays a central role in the life history of many animals but remains poorly understood at the molecular level, and the involvement of epigenetic mechanisms has not been previously explored for any species. We used a quantitative, high-resolution technique (RRBS) to measure DNA methylation across the *O. mykiss* genome and detected 57 highly significant DMRs relating to smoltification, a migration-related phenotype. Methylation differences of up to 62% per DMR were observed between resident and smolt fin tissue, a level that is on par with differences previously observed for other animal phenotypic changes associated with epigenetic mechanisms (e.g. Herb *et al.* 2012). Not only is the magnitude of methylation changes high, but some of the identified genes have relevance to migratory transitions. Hierarchical clustering based on methylation patterns alone differentiated the life history types, although there were a few individual exceptions (Fig. 2). As it is unlikely these individuals were phenotypically misclassified, it appears that epigenetic heterogeneity is occurring. In these cases, we speculate that gene expression patterns resulting in alternative life history phenotypes may be due to (i) differential methylation at loci not captured in this study or (ii) DNA methylation-independent mechanisms. Thus, distinct epigenetic profiles might underlie similar phenotypic outcomes.

Comparison to MS-AFLP approach and other salmonid epigenetic studies

The vast majority of ecological studies to date have used a methyl-sensitive AFLP (MS-AFLP) approach (Cervera *et al.* 2002), which identifies global methylation changes, but does not provide the nucleotide-level resolution needed to derive functional genomic information and almost certainly would not have detected the small number of DMRs identified in this study. Previous salmonid studies that have used MS-AFLP have identified broad-scale methylation changes associated with sea water adaptation (Morán *et al.* 2013) and early maturation (Morán & Pérez-Figueroa 2011) in Atlantic salmon (*Salmo salar*). Global DNA methylation differences were not found in a study between wild and hatchery *O. mykiss* steelhead (Blouin *et al.* 2010); however, the authors acknowledge the potential for undetected gene-specific methylation changes. The RRBS approach used here has several advantages compared to the MS-AFLP approach (see review by Schrey *et al.* 2013 and citations therein), the most notable being its quantitative nature and its ability to identify particular genomic regions (as opposed to anonymous AFLP loci) that are differentially methylated. Identification of DMRs and their colocalization with genes and regulatory regions allows us to uncover gene annotation and gene ontology patterns as well as to conduct pathway analysis so that results can be interpreted in a more integrative manner with other ecological and genetic studies for the trait of interest. The approach, however, does require a reference genome for alignment purposes. As more species are sequenced, it is our hope that the field of ecological epigenetics becomes more mainstream and is used in tandem with other biological subfields (e.g. genetics, physiology, ecology) to gain more comprehensive and integrative knowledge on the molecular mechanisms underlying complex phenotypic diversity.

Paralogous sequences and tissue type affect interpretation of results

Several of the DMR-associated genes are members of gene families (e.g. CLDN10, MYOM3, PCDHGC5), and the *O. mykiss* genome, as previously mentioned, is highly duplicated. While our analyses exclude reads that aligned to more than one location in each of the genomes, we cannot be certain that all gene family members are contained in the current incompletely assembled versions of the genome. Therefore, it is possible that the specific gene family member with differential methylation may be misidentified in some cases. In the absence of bisulphite cloning to distinguish gene family members, results should be interpreted with caution. Given our requirement of unique mapping, it is likely that differential DNA methylation of many gene family members and other paralogous regions went undetected. Additionally, it should be noted that sampling choices, including tissue or cell type analysed and timing of collection, can greatly affect results for both epigenetic and gene expression studies.

Relating differential DNA methylation to genomic architecture

Over half of the annotated DMRs were found within or proximal to CpG islands (CGIs), strongly suggesting that epigenetic modifications are influencing transcriptional activity of associated genes. CGIs are CpG-rich regions that are predominantly unmethylated and serve as sites for transcription initiation (Deaton & Bird 2011). Approximately 70% of annotated promoters in vertebrates are associated with CGIs (Saxonov *et al.* 2006), and hypermethylation of these regions stabilizes gene silencing in promoter regions (Stein *et al.* 1982; Mohn *et al.* 2008; Payer & Lee 2008). Residents displayed a 32% increase in DNA methylation compared to smolts in a CGI shore of the *aryl hydrocarbon receptor 2 alpha* (*AHR2A*; Fig. 3) distal promoter (<10 kb from TSS), suggesting that *AHR2A* is more stably expressed in smolt than resident fin tissue. AHR is an environmentally responsive transcription factor that mediates toxic responses to environmental pollutants (Hansson *et al.* 2004), circadian rhythm (Anderson *et al.* 2013), cellular differentiation (Abbott & Birnbaum 1989), apoptosis (Gu *et al.* 2000) and immune functioning (Nguyen *et al.* 2013).

In contrast to promoters, CGI methylation in gene bodies is more common and speculated to be associated with regulatory noncoding RNA (Dinger *et al.* 2008; Ørom *et al.* 2010), alternative splicing (Kornblihtt 2006) or intragenic promoter activity (Maunakea *et al.* 2010). An example of gene body differential methylation in this study is *nescient helix-loop-helix 2* (*NHLH2*), with residents having 38% higher DNA methylation levels than smolts within a gene body CGI shore (Fig. 3). NHLH2 is a transcription factor important in modulating body weight and energy availability (Fox *et al.* 2007), which are both believed to be critical factors regulating the occurrence of salmonid smoltification (Thorpe & Metcalfe 1998; Theriault & Dodson 2003) and subsequent migration.

Almost half (46%) of the DMRs were not located within or proximal to known genes. This may be an indication of trans-acting regulatory elements (e.g. enhancers), which are quite commonly differentially methylated in the human genome (Ziller *et al.* 2014), or it may indicate the incomplete annotation of the *O. mykiss* genome precludes identification of additional genes and their proximal regulatory regions.

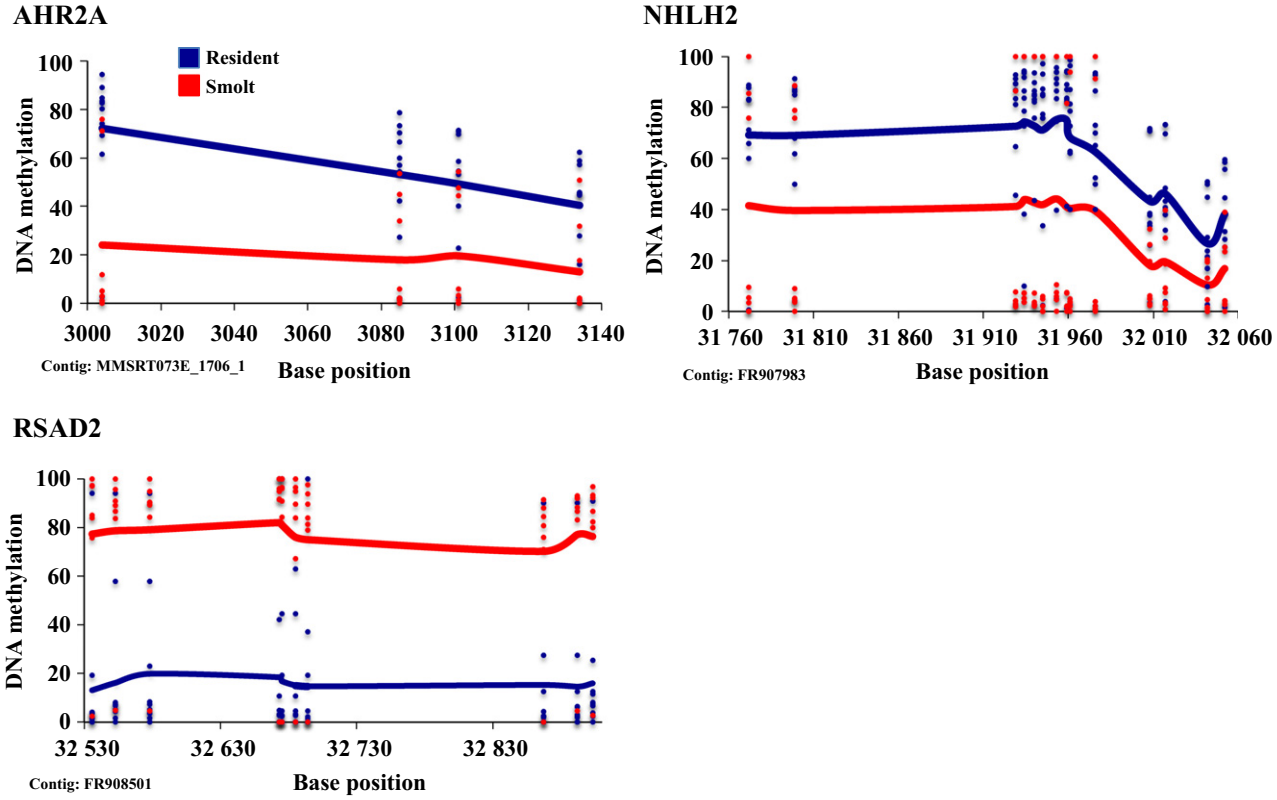

Fig. 3 Differential DNA methylation distinguishes residents and smolts. Three gene-associated DMR examples at the base-pair resolution. Points (blue=resident, red=smolt) represent DNA methylation levels of individual samples at every CpG site within the DMR, and smoothed lines represent average DNA methylation levels for each life history type.

Comparison with previous genetic mapping studies

Our study demonstrates that examining smoltification from an epigenetic perspective uncovers genomic loci not previously associated with migration-related traits and also corroborates loci known to be associated with smoltification or migration propensity. The chromosome identified as having the second highest number of DMRs (Omy10, Fig. 4a) was previously found to have numerous QTL for traits associated with anadromy/residency, including a main-effect QTL for the binary smolt phenotype for this F_2 intercross (Nichols *et al.* 2008). This colocalization of genetic and epigenetic variation onto Omy10 may indicate that genetic and epigenetic factors can act in a coordinated fashion to modify migration-related phenotypes. Further work is warranted to determine whether genetic and epigenetic variation overlap in particular chromosomal regions. This will require considerably higher marker density (and genomic sequence data) than was possible in the original QTL mapping study (Nichols *et al.* 2008), which used microsatellite and amplified fragment length polymorphic markers.

Previous studies have identified QTL for smoltification and migration-related traits on many chromosomes in the *O. mykiss* genome, reflecting the highly complex and polygenic nature of the phenotype. It seems unlikely that all migration-related genetic variation is associated with epigenetic variation; therefore, it is not surprising that we did not identify DMRs on many of the chromosomes containing migration-related QTL. This lack of complete coordination suggests that genetic and epigenetic variation do not always affect the same loci. Additionally, there is considerable variation in the location of QTL regions across populations (Hecht *et al.* 2012a,b; Hale *et al.* 2013), and we expect the same may be true of epigenetic variation, which should be explored in future studies of additional populations. A study in Atlantic salmon further corroborates this idea, providing evidence that resident versus migratory life history types had a complex multigenic basis with little parallelism across populations (Perrier *et al.* 2013).

Comparison with candidate genes from GWAS and gene expression studies

We found similarities between genes associated with some DMRs and those identified in prior GWAS and gene expression studies. For instance, *neurogenic differentiation factor 4* (*NEUROD4*) and *metabotropic glutamate receptor 1* (*GRM1*) were identified as differentially methylated in our current study, both with 31% increased methylation in smolts, and as *F*ST outliers in a previous *O. mykiss* migration-related GWAS (Hale *et al.* 2013). NEUROD4 is a transcriptional activator that mediates neuronal differentiation (Bertrand *et al.* 2002), and GRM1 is a G-protein-coupled receptor for glutamate involved in central nervous system transmission and normal brain function (Frank *et al.* 2011; Yin & Niswender 2014). Glutamate receptors are involved in synaptic plasticity (Kano & Hashimoto 2009; Ohtani *et al.* 2014), which is extremely important for learning and memory.

Multiple DMR-associated genes were members of the same gene families found in a brain gene expression study comparing age-two F_1 progeny of migrant and resident *O. mykiss* crosses from a wild Alaskan population (M. C. Hale, G. J. McKinney, F. P. Thrower & K. M. Nichols, unpublished data). Identification of these gene families (protocadherin, claudin, myomesin) in

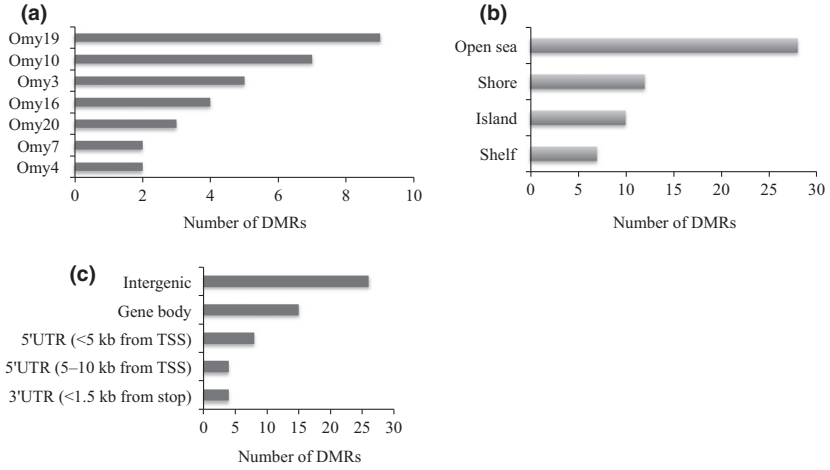

Fig. 4 Genomic features of the 57 migration-related DMRs for *Oncorhynchus mykiss*. (a) Seven chromosomes contain two or more DMRs. (b) Approximately half (29) of the DMRs are found within or proximal to CpG islands. (c) Over half (31) of the DMRs are found within gene bodies or proximal regulatory elements.

our current study suggests that neuronal modifications and tissue development may be epigenetically regulated in association with migration propensity. Specifically, PCDHGC5 is involved in forming and maintaining neuronal connections in the brain (Chen *et al.* 2012) and has 49% increased methylation in residents, CLDN10 can enhance paracellular permeability (Amasheh *et al.* 2010) and has 30% increased methylation in residents, and MYOM3 is involved in muscle tissue formation and contraction (Schoenauer *et al.* 2005) and has 22% increased methylation in smolts. Myomesin was also found to be differentially expressed in Atlantic salmon brain and gill tissue during smoltification (Seear *et al.* 2009). This collection of genes, involved in neuronal and muscle development, may be quite important for migration-related transitions. Extensive neuronal modifications are required for these transitions (Herry *et al.* 2008; Durstewitz *et al.* 2010), including epigenetically regulated modifications (Borrelli *et al.* 2008), while preparation for migration may facilitate increased muscle formation and body mass (Marsh 1984). Only *O. mykiss* that reach a certain size threshold typically transition to a migratory phenotype (Theriault & Dodson 2003; Páez *et al.* 2011). However, the functional link between neuronal and muscle development is less clear when considering the fin tissue used in our current study. It is quite possible in these cases that the genetic variation (which is primarily static across tissue types within an individual) is strongly influencing DNA methylation levels. A study comparing DNA methylation and genetic variation across human populations suggested that differentially methylated CpG sites are 'evolutionarily established mediators between the genetic code and phenotypic variability' (Heyn *et al.* 2013). Therefore, the same epigenetic differences may be detected regardless of the tissue type chosen.

Another category of interest is immune response. The DMR with the greatest differential methylation between divergent life history phenotypes was *radical S-adenosyl methionine domain containing 2* (*RSAD2*), with smolts averaging 62% more methylation at 10 CpG sites in the putative 3′ UTR (Fig. 3). *RSAD2* encodes for a protein with interferon-inducible antiviral and antibacterial properties (Fitzgerald 2011). A previous study found that residents overexpressed antiviral, including interferon-associated, transcripts in gills compared to migrants (Sutherland *et al.* 2014). This may indicate a potential link between immune response and migration propensity, which seems plausible as the immune system is partially controlled by circadian rhythms (Scheiermann *et al.* 2013). Brook charr (*Salvelinus fontinalis*) differentially express immune response genes in gills for alternative migratory tactics, although migrants have increased immune expression (Boulet *et al.* 2012). As skin tissue, including the fin, is a primary site of initial pathogen entry in fish (e.g. Baerwald 2013), it also plays a critical role in immune response. Therefore, epigenetic changes associated with immune response genes in fin tissue are not surprising given these prior links between immune response and migration propensity studied in other tissues.

Potential functional significance of DMRs

Gene ontology, canonical pathway and gene network analyses identified several molecular categories with genes capable of dramatically altering phenotypes, some with clear associations with behavioural phenotypes such as migration. These include protein kinases, G-protein-coupled receptors (GPCRs), and genes linked to the circadian rhythm pathway. Three DMR-associated genes encode for proteins with serine/threonine and tyrosine kinase activity (*AATK*, *SBK1*, *STYK1*; Table S3, Supporting information). Kinase expression and activity is tightly regulated (Huse & Kuriyan 2002), and phosphorylation plays a key role in generating circadian rhythms (Reischl & Kramer 2011). GPCRs sense extracellular molecules to initiate signal transduction cascades (Marinissen & Gutkind 2001), and two DMR-associated genes encode proteins in this functional category (*ADORA3*, *GRM1*; Table S3, Supporting information). GPCR activation can induce tyrosine kinase activity as well as RAS (Luttrell *et al.* 1999), so several of the differentially methylated genes can be connected through common gene networks (see Fig. S2, Supporting information).

We identified multiple differentially methylated genes within the circadian rhythm pathway, which is a primary internal cue for migratory behaviour. Circadian rhythms allow organisms to anticipate and prepare for regular environmental changes. Although circadian rhythms are endogenous, they are adjusted or entrained to the local environment by exogenous environmental cues (Gwinner & Scheuerlein 1998; Lehmann *et al.* 2012). In addition to regular 24-h clock changes, circadian rhythms impact seasonal physiological and behavioural changes, including migratory traits such as migration propensity (Gwinner 1996; Guerra *et al.* 2014). Circadian rhythm gene expression patterns and genotypes affect migration timing in diverse species (e.g. Williams 2001; Belden *et al.* 2007; Mukai *et al.* 2008). Furthermore, circadian rhythms are regulated by DNA methylation, histone modifications and RNA-based mechanisms (DiTacchio *et al.* 2011; Fustin *et al.* 2013; Azzi *et al.* 2014), indicating that epigenetic mechanisms drive circadian clock plasticity. We found two DMR genes associated with circadian rhythm, *AHR2A* and *dexamethasone-induced RAS 1* (*RASD1*). AHR can sense environmental compounds and affect circadian

rhythmicity and responsiveness to environmental light cues (Mukai *et al.* 2008; Anderson *et al.* 2013; Xu *et al.* 2013). *RASD1*, also abbreviated *DEXRAS1*, plays a pivotal role in the circadian clock's responsiveness to external timing cues, including both photic and nonphotic stimuli (Williams 2001; Cheng & Obrietan 2006; Belden *et al.* 2007). Reduced *RASD1* expression can cause nonphotic stimuli to have more influence on clock phasing than photic stimuli (Cheng & Obrietan 2006). In our study, methylation was increased in residents in the putative promoter regions of *AHR2A* and *RASD1* by 32% and 22%, respectively. This may suggest higher expression of these genes in smolts, which then might improve their responsiveness to light cues for clock phasing compared to residents. However, we believe the methylation changes need to display similar patterns between fin tissue and other tissue types that respond to photic stimuli (e.g. brain or optic cells) for these functional associations to be biologically meaningful, which will require further study.

Epigenetically influenced response threshold for triggering migration

Development into differentiated smolts or residents may be at least partially dependent on an epigenetic response threshold. Response thresholds may determine an individual's sensitivity to particular environmental stimuli, with interactions between genetic and epigenetic factors along with environmental cues causing a phenotypic transition when the threshold is achieved (Yan *et al.* 2014). In an ecological context, this may explain the observed plasticity for migration propensity (Kendall *et al.* 2015). Residents may give rise to migratory progeny, and vice versa, due to the complex interactions between the environment, genetics and development leading to alternative phenotypic outcomes that are dependent upon these thresholds. Future research should focus on strengthening the connection between epigenetic and phenotypic diversity by integrating epigenetic results with genetic and gene expression data sets under different environmental conditions to examine the nature and extent of coordination between alternative sources of variation that ultimately drive complex phenotypes. For conservation purposes, the two life history types are often managed separately, which will require careful consideration of environmental and genetic variation, and how they interact to produce a phenotypic response.

Conclusions

The genome-scale methylation data presented here provide the first evidence of a relationship between epigenetic variation and diversity in migration-related traits. It also demonstrates the power of large-scale epigenotyping in ecological studies for identifying loci associated with complex phenotypes. Epigenetic mechanisms link genetic and environmental factors to generate alternative life history strategies. As recently demonstrated by Sloat *et al.* (2014), flexible life history strategies 'can be both heritable and subject to strong environmental influences', and our study indicates that differential DNA methylation at gene regulatory elements may be a critical molecular mechanism allowing interactions between an organism and its environment to balance phenotypic stability and plasticity, and ultimately modify migration-related phenotypes.

Acknowledgements

This work was made possible by funding from the California Department of Fish and Wildlife, contract number P0740017. We thank Michael Miller at University of California, Davis, for assistance with data processing scripts for using the U.S. *O. mykiss* draft genome assembly. We are grateful to Paul Wheeler at Washington State University for his instrumental role in producing original clonal lines and F_2 doubled haploid individuals. We thank Jessica Petersen and three anonymous reviewers for their helpful comments. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303.

References

Abbott BD, Birnbaum LS (1989) TCDD alters medial epithelial cell differentiation during palatogenesis. *Toxicology and Applied Pharmacology*, **99**, 276–286.

Able KP, Belthoff JR (1998) Rapid "evolution" of migratory behaviour in the introduced house finch of eastern North America. *Proceedings of the Royal Society B: Biological Sciences*, **265**, 2063–2071.

Akalin A, Kormaksson M, Li S *et al.* (2012) methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, **13**, R87.

Alerstam T, Hedenstrom A, Akesson S (2003) Long-distance migration: evolution and determinants. *Oikos*, **103**, 247–260.

Allendorf FW, Thorgaard GH (1984) Tetraploidy and the evolution of salmonid fishes. In: *Evolutionary Genetics of Fishes* (ed. Turner BJ), pp. 1–53. Plenum Press, New York.

Amasheh S, Fromm M, Günzel D (2010) Claudins of intestine and nephron - a correlation of molecular tight junction structure and barrier function. *Acta Physiologica*, **201**, 133–140.

Anderson G, Beischlag TV, Vinciguerra M, Mazzoccoli G (2013) The circadian clock circuitry and the AHR signaling pathway in physiology and pathology. *Biochemical Pharmacology*, **85**, 1405–1416.

Azzi A, Dallmann R, Casserly A *et al.* (2014) Circadian behavior is light-reprogrammed by plastic DNA methylation. *Nature Neuroscience*, **17**, 377–382.

Baerwald MR (2013) Temporal expression patterns of rainbow trout immune-related genes in response to Myxobolus cerebralis exposure. *Fish and Shellfish Immunology*, **35**, 965–971.

Bastow R, Mylne JS, Lister C *et al.* (2004) Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature*, **427**, 164–167.

Belden WJ, Larrondo LF, Froehlich AC *et al.* (2007) The band mutation in *Neurospora crassa* is a dominant allele of ras-1 implicating RAS signaling in circadian output. *Genes & Development*, **21**, 1494–1505.

Berthelot C, Brunet F, Chalopin D *et al.* (2014) The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nature Communications*, **5**, 3657.

Berthold P (1991) Genetic control of migratory behaviour in birds. *Trends in Ecology & Evolution*, **6**, 254–257.

Bertrand N, Castro DS, Guillemot F (2002) Proneural genes and the specification of neural cell types. *Nature Reviews Neuroscience*, **3**, 517–530.

Bisson I-A, Safi K, Holland RA (2009) Evidence for repeated independent evolution of migration in the largest family of bats (S Rands, Ed,). *PLoS ONE*, **4**, e7504.

Blouin MS, Thuillier V, Cooper B *et al.* (2010) No evidence for large differences in genomic methylation between wild and hatchery steelhead (*Oncorhynchus mykiss*). *Canadian Journal of Fisheries and Aquatic Sciences*, **67**, 217–224.

Borrelli E, Nestler EJ, Allis CD, Sassone-Corsi P (2008) Decoding the epigenetic language of neuronal plasticity. *Neuron*, **60**, 961–974.

Boulet M, Normandeau E, Bougas B, Audet C, Bernatchez L (2012) Comparative transcriptomics of anadromous and resident brook charr Salvelinus fontinalis before their first salt water transition. *Current Zoology*, **58**, 158–170.

Bowlin MS, Bisson IA, Shamoun-Baranes J *et al.* (2010) Grand challenges in migration biology. *Integrative and Comparative Biology*, **50**, 261–279.

Boyle P, Clement K, Gu H *et al.* (2012) Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biology*, **13**, R92.

Cervera MT, Ruiz-García L, Martínez-Zapater JM (2002) Analysis of DNA methylation in *Arabidopsis thaliana* based on methylation-sensitive AFLP markers. *Molecular Genetics and Genomics: MGG*, **268**, 543–552.

Champagne FA (2010) Epigenetic influence of social experiences across the lifespan. *Developmental Psychobiology*, **52**, 299–311.

Champagne FA (2012) Epigenetics and developmental plasticity across species (D Moore, C Moore, Eds,). *Developmental Psychobiology*, **55**, 33–41.

Chen WV, Alvarez FJ, Lefebvre JL *et al.* (2012) Functional significance of isoform diversification in the protocadherin gamma gene cluster. *Neuron*, **75**, 402–409.

Cheng H-YM, Obrietan K (2006) Dexras1: Shaping the responsiveness of the circadian clock. *Seminars in Cell & Developmental Biology*, **17**, 345–351.

Courter II, Child DB, Hobbs JA *et al.* (2013) Resident rainbow trout produce anadromous offspring in a large interior watershed. *Canadian Journal of Fisheries and Aquatic Sciences*, **70**, 701–710.

Deaton AM, Bird A (2011) CpG islands and the regulation of transcription. *Genes & Development*, **25**, 1010–1022.

Dinger ME, Amaral PP, Mercer TR *et al.* (2008) Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Research*, **18**, 1433–1445.

Dingle H (2014) *Migration: the biology of life on the move*. Oxford University Press, New York, NY.

DiTacchio L, Le HD, Vollmers C *et al.* (2011) Histone lysine demethylase JARID1a activates CLOCK-BMAL1 and influences the circadian clock. *Science*, **333**, 1881–1885.

Dodson JJ, Aubin-Horth N, Thériault V, Páez DJ (2013) The evolutionary ecology of alternative migratory tactics in salmonid fishes. *Biological Reviews*, **88**, 602–625.

Durstewitz D, Vittoz NM, Floresco SB, Seamans JK (2010) Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron*, **66**, 438–448.

Feil R, Fraga MF (2012) Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics*, **13**, 97–109.

Fitzgerald KA (2011) The interferon inducible gene : viperin. *Journal of Interferon & Cytokine Research*, **31**, 131–135.

Folmar LC, Dickhoff WW (1980) The parr-smolt transformation (smoltification) and seawater adaptation in salmonids. *Aquaculture*, **21**, 1–37.

Fox DL, Vella KR, Good DJ (2007) Energy balance pathways converging on the Nhlh2 transcription factor. *Frontiers in Bioscience*, **12**, 3983–3993.

Frank RAW, McRae AF, Pocklington AJ *et al.* (2011) Clustered coding variants in the glutamate receptor complexes of Individuals with schizophrenia and bipolar disorder (T Burne, Ed,). *PLoS ONE*, **6**, e19011.

Fustin J-M, Doi M, Yamaguchi Y *et al.* (2013) RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell*, **155**, 793–806.

Gu Y, Hogenesch JB, Bradfield CA (2000) The PAS superfamily: sensors of environmental and developmental signals. *Annual Review of Pharmacology and Toxicology*, **40**, 519–561.

Guerra PA, Merlin C, Gegear RJ, Reppert SM (2014) Discordant timing between antennae disrupts sun compass orientation in migratory monarch butterflies. *Nature Communications*, **3**, 957–958.

Gwinner E (1996) Circadian and circannual programmes in avian migration. *The Journal of Experimental Biology*, **199**, 39–48.

Gwinner E, Scheuerlein A (1998) Seasonal changes in daylight intensity as a potential zeitgeber of circannual rhythms in equatorial Stonechats. *Journal of Ornithology*, **139**, 407–412.

Hale MC, Thrower FP, Berntson EA, Miller MR, Nichols KM (2013) Evaluating adaptive divergence between migratory and nonmigratory ecotypes of a salmonid fish, *Oncorhynchus mykiss*. *G3 (Bethesda)*, **3**, 1273–1285.

Hansson MC, Wittzell H, Persson K, von Schantz T (2004) Unprecedented genomic diversity of AhR1 and AhR2 genes in Atlantic salmon (*Salmo salar* L.). *Aquatic toxicology*, **68**, 219–232.

Hayes SA, Hanson CV, Pearse DE *et al.* (2012) Should I stay or should I go? The influence of genetic origin on emigration behavior and physiology of resident and anadromous juvenile *Oncorhynchus mykiss*. *North American Journal of Fisheries Management*, **32**, 772–780.

Hecht BC, Campbell NR, Holecek DE, Narum SR (2012a) Genome-wide association reveals genetic basis for the propensity to migrate in wild populations of rainbow and steelhead trout. *Molecular Ecology*, **22**, 3061–3076.

Hecht BC, Hard JJ, Thrower FP, Nichols KM (2015) Quantitative genetics of migration related traits in rainbow and steelhead trout. *G3 (Bethesda)*, **5**, 873–879.

Hecht BC, Thrower FP, Hale MC, Miller MR, Nichols KM (2012b) Genetic architecture of migration-related traits in rainbow and steelhead trout, *Oncorhynchus mykiss*. *G3 (Bethesda)*, **2**, 1113–1127.

Herb BR, Wolschin F, Hansen KD *et al.* (2012) Reversible switching between epigenetic states in honeybee behavioral subcastes. *Nature Neuroscience*, **15**, 1371–1373.

Herry C, Ciocchi S, Senn V *et al.* (2008) Switching on and off fear by distinct neuronal circuits. *Nature*, **454**, 600–606.

Heyn H, Moran S, Hernando-herraez I *et al.* (2013) DNA methylation contributes to natural human variation DNA methylation contributes to natural human variation. *Genome Research*, **23**, 1363–1372.

Hoar WS (1976) Smolt transformation: evolution, behavior, and physiology. *Journal of Fisheries Research Board of Canada*, **33**, 1233–1252.

Huse M, Kuriyan J (2002) The conformational plasticity of protein kinases. *Cell*, **109**, 275–282.

Irizarry RA, Ladd-Acosta C, Wen B *et al.* (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature genetics*, **41**, 178–186.

Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, **13**, 484–492.

Jonsson B, Jonsson N (1993) Partial migration: niche shift versus sexual maturation in fishes. *Reviews in Fish Biology and Fisheries*, **3**, 348–365.

Kano M, Hashimoto K (2009) Synapse elimination in the central nervous system. *Current Opinion in Neurobiology*, **19**, 154–161.

Kendall NW, Mcmillan JR, Sloat MR *et al.* (2015) Anadromy and residency in steelhead and rainbow trout (*Oncorhynchus mykiss*): a review of the processes and patterns. *Canadian Journal of Fisheries and Aquatic Sciences*, **72**, 319–342.

Kornblihtt AR (2006) Chromatin, transcript elongation and alternative splicing. *Nature Structural and Molecular Biology*, **13**, 5–7.

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

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, **10**, R25.

Lehmann M, Spoelstra K, Visser ME, Helm B (2012) Effects of temperature on circadian clock and chronotype: an experimental study on a passerine bird. *Chronobiology International*, **29**, 1062–1071.

Lepais O, Weir JT (2014) SimRAD: an R package for simulation-based prediction of the number of loci expected in RADseq and similar genotyping by sequencing approaches. *Molecular Ecology Resources*, **14**, 1314–1321.

Li S, Garrett-Bakelman FE, Akalin A *et al.* (2014) An optimized algorithm for detecting and annotating regional differential methylation. *BMC Bioinformatics*, **14**, 1–9.

Liedvogel M, Åkesson S, Bensch S (2011) The genetics of migration on the move. *Trends in Ecology & Evolution*, **26**, 561–569.

Luttrell LM, Daaka Y, Lefkowitz RJ (1999) Regulation of tyrosine kinase cascades by G-protein-coupled receptors. *Current Opinion in Cell Biology*, **11**, 177–183.

Marinissen MJ, Gutkind JS (2001) G-protein-coupled receptors and signaling networks: emerging paradigms. *Trends in Pharmacological Sciences*, **22**, 368–376.

Marsh RL (1984) Adaptations of the gray catbird Dumetella carolinensis to long-distance migration: flight muscle hypertrophy associated with elevated body mass. *Physiological Zoology*, **57**, 105–117.

Martínez A, Garza JC, Pearse DE (2011) A microsatellite genome screen identifies chromosomal regions under differential selection in steelhead and rainbow trout. *Transactions of the American Fisheries Society*, **140**, 829–842.

Maunakea AK, Nagarajan RP, Bilenky M *et al.* (2010) Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*, **466**, 253–257.

McCormick SD, Hansen LP, Quinn TP, Saunders RL (1998) Movement, migration, and smolting of Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 77–92.

McMichael GA, Sharpe CS, Pearsons TN (1997) Effects of residual hatchery-reared steelhead on growth of wild rainbow trout and spring chinook salmon. *Transactions of the American Fisheries Society*, **126**, 230–239.

Mifsud KR, Gutièrrez-Mecinas M, Trollope AF *et al.* (2011) Epigenetic mechanisms in stress and adaptation. *Brain Behavior and Immunity*, **25**, 1305–1315.

Mohn F, Weber M, Rebhan M *et al.* (2008) Lineage-specific polycomb targets and *de novo* DNA methylation define restriction and potential of neuronal progenitors. *Molecular Cell*, **30**, 755–766.

Morán P, Marco-Rius F, Megías M, Covelo-Soto L, Pérez-Figueroa A (2013) Environmental induced methylation changes associated with seawater adaptation in brown trout. *Aquaculture*, **392–395**, 77–83.

Morán P, Pérez-Figueroa A (2011) Methylation changes associated with early maturation stages in the Atlantic salmon. *BMC Genetics*, **12**, 86.

Mukai M, Lin TM, Peterson RE, Cooke PS, Tischkau SA (2008) Behavioral rhythmicity of mice lacking AhR and attenuation of light-induced phase shift by 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Journal of Biological Rhythms*, **23**, 200–210.

Narum SR, Zendt JS, Frederiksen C *et al.* (2011) Candidate genetic markers associated with anadromy in *Oncorhynchus mykiss* of the Klickitat River. *Transactions of the American Fisheries Society*, **140**, 843–854.

Narum SR, Zendt JS, Graves D, Sharp WR (2008) Influence of landscape on resident and anadromous life history types of *Oncorhynchus mykiss*. *Canadian Journal of Fisheries and Aquatic Sciences*, **65**, 1013–1023.

Nebel S (2010) Animal migration. *Nature Education Knowledge*, **3**, 77.

Nguyen NT, Hanieh H, Nakahama T, Kishimoto T (2013) The roles of aryl hydrocarbon receptor in immune responses. *International immunology*, **25**, 335–343.

Nichols KM, Broman KW, Sundin K *et al.* (2007) Quantitative trait loci x maternal cytoplasmic environment interaction for development rate in *Oncorhynchus mykiss*. *Genetics*, **175**, 335–347.

Nichols KM, Edo AF, Wheeler PA, Thorgaard GH (2008) The genetic basis of smoltification-related traits in *Oncorhynchus mykiss*. *Genetics*, **179**, 1559–1575.

Ohtani Y, Miyata M, Hashimoto K *et al.* (2014) The synaptic targeting of mGluR1 by its carboxyl-terminal domain is crucial for cerebellar function. *Journal of Neuroscience*, **34**, 2702–2712.

Ørom UA, Derrien T, Beringer M *et al.* (2010) Long noncoding RNAs with enhancer-like function in human cells. *Cell*, **143**, 46–58.

Páez DJ, Brisson-Bonenfant C, Rossignol O *et al.* (2011) Alternative developmental pathways and the propensity to migrate: a case study in the Atlantic salmon. *Journal of Evolutionary Biology*, **24**, 245–255.

Payer B, Lee JT (2008) X chromosome dosage compensation: how mammals keep the balance. *Annual Review of Genetics*, **42**, 733–772.

Pearse DE, Miller MR, Abadia-Cardoso A, Garza JC (2014) Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20140012.

Perrier C, Bourret V, Kent MP, Bernatchez L (2013) Parallel and nonparallel genome-wide divergence among replicate population pairs of freshwater and anadromous Atlantic salmon. *Molecular Ecology*, **22**, 5577–5593.

Petronis A (2010) Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature*, **465**, 721–727.

Pulido F, Berthold P, Van Noordwijk AJ (1996) Frequency of migrants and migratory activity are genetically correlated in a bird population: Evolutionary implications. *Proceedings of the National Academy of Sciences*, **93**, 14642–14647.

Quinn TP, Myers KW (2005) Anadromy and the marine migrations of Pacific salmon and trout: Rounsefell revisited. *Reviews in Fish Biology and Fisheries*, **14**, 421–442.

R Core Team (2014) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Reischl S, Kramer A (2011) Kinases and phosphatases in the mammalian circadian clock. *FEBS Letters*, **585**, 1393–1399.

Riva-Rossi C, Pascual MA, Babaluk JA, García-Asorey M, Halden NM (2007) Intra-population variation in anadromy and reproductive life span in rainbow trout introduced in the Santa Cruz River, Argentina. *Journal of Fish Biology*, **70**, 1780–1797.

Robinson JT, Thorvaldsdóttir H, Winckler W *et al.* (2011) Integrative genomics viewer. *Nature Biotechnology*, **29**, 24–26.

Roff DA, Fairbairn DJ (2007) The evolution and genetics of migration in insects. *BioScience*, **57**, 155–164.

Saxonov S, Berg P, Brutlag DL (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences*, **103**, 1412–1417.

Scheiermann C, Kunisaki Y, Frenette PS (2013) Circadian control of the immune system. *Nature Reviews Immunology*, **13**, 190–198.

Schoenauer R, Bertoncini P, Machaidze G *et al.* (2005) Myomesin is a molecular spring with adaptable elasticity. *Journal of Molecular Biology*, **349**, 367–379.

Schrey AW, Alvarez M, Foust CM *et al.* (2013) Ecological epigenetics: beyond MS-AFLP. *Integrative and Comparative Biology*, **53**, 340–350.

Seear PJ, Carmichael SN, Talbot R *et al.* (2009) Differential gene expression during smoltification of Atlantic salmon (*Salmo salar* L.): a first large-scale microarray study. *Marine Biotechnology*, **12**, 126–140.

Sloat MR, Fraser DJ, Dunham JB *et al.* (2014) Ecological and evolutionary patterns of freshwater maturation in Pacific and Atlantic salmonines. *Reviews in Fish Biology and Fisheries*, **24**, 689–707.

Stein R, Razin A, Cedar H (1982) *In vitro* methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells. *Proceedings of the National Academy of Sciences*, **79**, 3418–3422.

Sutherland BJG, Hanson KC, Jantzen JR, Koop BF, Smith CT (2014) Divergent immunity and energetic programs in the gills of migratory and resident *Oncorhynchus mykiss*. *Molecular Ecology*, **23**, 1952–1964.

Theriault V, Dodson JJ (2003) Body size and the adoption of a migratory tactic in brook charr. *Journal of Fish Biology*, **63**, 1144–1159.

Thorpe JE, Metcalfe NB (1998) Is smolting a positive or a negative developmental decision? *Aquaculture*, **168**, 95–103.

Thrower FP, Hard JJ, Joyce JE (2004) Genetic architecture of growth and early life-history transitions in anadromous and derived freshwater populations of steelhead. *Journal of Fish Biology*, **65**, 286–307.

Viola AE, Schuck ML (1995) A method to reduce the abundance of residual hatchery steelhead in rivers. *North American Journal of Fisheries Management*, **15**, 488–493.

Waples RS, Teel DJ, Myers JM, Marshall AR (2004) Life-history divergence in Chinook salmon: historic contingency and parallel evolution. *Evolution*, **58**, 386–403.

Wickham H (2009) ggplot2: *Elegant Graphics for Data Analysis*. Springer, Berlin.

Williams JA (2001) A circadian output in Drosophila mediated by neurofibromatosis-1 and Ras/MAPK. *Science*, **293**, 2251–2256.

Xu CX, Wang C, Krager SL, Bottum KM, Tischkau SA (2013) Aryl hydrocarbon receptor activation attenuates Per1 gene induction and influences circadian clock resetting. *Toxicological Sciences*, **132**, 368–378.

Yan H, Simola DF, Bonasio R *et al.* (2014) Eusocial insects as emerging models for behavioural epigenetics. *Nature Publishing Group*, **15**, 677–688.

Yin S, Niswender CM (2014) Progress toward advanced understanding of metabotropic glutamate receptors: structure, signaling and therapeutic indications. *Cellular Signalling*, **26**, 2284–2297.

Zaugg WS, Wagner HH (1973) Gill ATPase activity related to parr-smolt transformation and migration in steelhead trout (Salmo gairdneri): influence of photoperiod and temperature. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, **45**, 955–965.

Ziller MJ, Gu H, Müller F *et al.* (2014) Charting a dynamic DNA methylation landscape of the human genome. *Nature*, **500**, 477–481.

Zimmerman CE, Reeves GH (2000) Population structure of sympatric anadromous and nonanadromous *Oncorhynchus mykiss*: evidence from spawning surveys and otolith microchemistry. *Canadian Journal of Fisheries and Aquatic Sciences*, **57**, 2152–2162.

This research was designed by M.R.B., M.H.M., M.R.S., R.P.N., K.M.H.T., B.M. and K.M.N. Laboratory experiments were conducted by A.M.G., and computational and data analyses were performed by M.R.B. G.H.T. provided the clonal line crosses. M.R.B. wrote an initial manuscript draft, which was edited by all authors.

Data accessibility

Demultiplexed Illumina sequencing reads for the 18 analysed individuals are available in the NCBI Gene Expression Omnibus, Accession No. GSE62195. This accession also contains raw DMC and DMR results as well as a readme file containing code for the bioinformatics analyses performed.

Supporting information

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

Fig. S1 Common gene ontology terms for the 26 annotated resident vs. migrant DMRs.

Fig. S2 Gene networks identified using Ingenuity Pathway Analysis.

Table S1 Readcounts and alignments to French and U.S. rainbow trout genomes for resident and smolt individuals.

Table S2 Detalied information for differentially methylates regions (DMRs) when comparing juvenile resident and migratory *O. mykiss* phenotype.

Table S3 Gene ontology full details for each DMR annotated gene when comparing resident and migrant *O. mykiss* juvenile phenotypes

Table S4 Significant canonical pathways, functions, and top networks for the 26 annotated DMRs identified using Ingenuity Pathway Analysis.