The interplay between epigenomic and transcriptomic variation during ecotype divergence in stickleback

Man Luo[1†], Junjie Zhao[1†], Juha Merilä[2,3], Rowan D. H. Barrett[4], Baocheng Guo[5,6*] and Juntao Hu[1*]

Abstract

Background Populations colonizing contrasting environments are likely to undergo adaptive divergence and evolve ecotypes with locally adapted phenotypes. While diverse molecular mechanisms underlying ecotype divergence have been identified, less is known about their interplay and degree of divergence.

Results Here we integrated epigenomic and transcriptomic data to explore the interactions among gene expression, alternative splicing, DNA methylation, and microRNA expression to gauge the extent to which patterns of divergence at the four molecular levels are aligned in a case of postglacial divergence between marine and freshwater ecotypes of nine-spined sticklebacks (*Pungitius pungitius*). Despite significant genome-wide associations between epigenomic and transcriptomic variation, we found largely non-parallel patterns of ecotype divergence across epigenomic and transcriptomic levels, with predominantly nonoverlapping (ranging from 43.40 to 87.98%) sets of differentially expressed, spliced and methylated genes, and candidate genes targeted by differentially expressed miRNA between the ecotypes. Furthermore, we found significant variation in the extent of ecotype divergence across different molecular mechanisms, with differential methylation and differential splicing showing the highest and lowest extent of divergence between ecotypes, respectively. Finally, we found a significant enrichment of genes associated with ecotype divergence in differential methylation.

Conclusions Our results suggest a nuanced relationship between epigenomic and transcriptomic processes, with alignment at the genome-wide level masking relatively independent effects of different molecular mechanisms on ecotype divergence at the gene level.

Keywords DNA methylation, Gene expression, Alternative splicing, MicroRNA, Adaptive divergence, *Pungitius pungitius*

[†]Man Luo and Junjie Zhao contributed equally to this work.

*Correspondence:
Baocheng Guo
guobaocheng@ioz.ac.cn
Juntao Hu
juntao_hu@fudan.edu.cn
[1] Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Institute of Biodiversity Science, Center for Evolutionary Biology, School of Life Sciences, Fudan University, Shanghai, China
[2] Ecological Genetics Research Unit, Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland
[3] Area of Ecology and Biodiversity, The School of Biological Sciences, The University of Hong Kong, Hong Kong SAR, China
[4] Department of Biology, McGill University, Quebec, Canada
[5] Key Laboratory of Animal Biodiversity Conservation and Integrated Pest Management & Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China
[6] University of Chinese Academy of Sciences, Beijing, China

Luo *et al. BMC Biology*   (2025) 23:70

Page 2 of 18

Background

Populations occupying diverse habitats often undergo adaptive evolution and develop into ecotypes that differ in traits shaped by their local environments [1]. Uncovering the molecular underpinnings of adaptive evolution and ecotype divergence is central to evolutionary biology [2]. Many studies have successfully identified adaptive allelic differentiation between ecotypes [3–5], and demonstrated that variation in gene expression can also play key role in adaptation [6–11]. However, non-genetic and post-transcriptional variation underlying ecotype divergence are much less studied than allelic and gene expression variation, especially in non-model species exhibiting local adaptation under natural conditions [12, 13]. For example, epigenetic modifications, especially DNA methylation in vertebrate animals, has been increasingly acknowledged as an important non-genetic regulator for transcription activation and repression as well as structural alteration of transcripts. Changes in DNA methylation have been suggested to contribute to the adaptive phenotypic divergence in marine versus freshwater ecotypes of the three-spined stickleback, sulfidic versus non-sulfidic populations of *Poecilia mexicana*, and urban versus forest populations of a passerine bird [14–18], partially due to the stable transmission of adaptive DNA methylation variation across generations [14, 17, 19].

In addition to DNA methylation, diverse mechanisms have been suggested to be involved in the transcriptional and post-transcriptional regulation of gene expression in a wide range of animal taxa [20]. Most, if not all, of current studies investigating variation in gene expression between ecotypes in natural populations have focused on changes at transcript level, with a number of studies revealing heritable gene expression differences among populations adapted to distinct environments [21–23]. However, emerging evidence suggests that post-transcriptional mechanisms such as alternative splicing can play important role relative to changes at transcript level in contributing to adaptation [24]. Alternative splicing can enable the generation of structurally variable transcripts from a single gene and potentially provides a mechanism for reducing constraints on highly pleiotropic genes [25]. Recent studies have suggested that alternative splicing might be in part modulated by microRNAs (miRNAs), which are short non-coding RNAs (~ 22 nt in length) [26]. Furthermore, miRNAs that downregulate gene expression through mRNA cleavage and/or translational repression could also play role in facilitating adaptation. Indeed, there is evidence of variation in splicing patterns and miRNA expression among locally adapted lineages [27–33]. Despite the importance of gene expression and its regulatory mechanisms in modulating the phenotypic responses of organisms to environmental

change, the interplay between these diverse mechanisms in mediating adaptation to contrasting environments in natural study systems remains largely unexplored. Moreover, regulatory changes have been suggested to play a predominant role during evolution of adaptive phenotypes [34]. Therefore, understanding the ways that different molecular mechanisms influence the ecotype divergence of organisms is crucial for determining their relative contributions to adaptive evolution, and deserves to be investigated systematically.

Theoretical studies have hypothesized that evolution can achieve similar phenotypic outcomes via both parallel and non-parallel processes if changes occur at different levels of molecular organization. Furthermore, the amount of change at each level is assumed to depend on the degree of redundancy between levels [35, 36]. Thus, similarity of evolutionary change at one hierarchical level between populations may not imply similarity at another level. In addition, the complex relationship between levels may also affect the probability of finding similarities at different levels underlying phenotypic convergence. For example, the relationship between genetic and epigenetic variation can vary from completely dependent to largely independent [37, 38]. Consequently, parallel changes may occur if epigenetic variation is under genetic control, while non-parallel changes may be observed when epigenetic variation is autonomous from genetic variation. Yet empirical studies have mainly searched for signatures of parallel molecular change underlying phenotypic convergence at a single molecular level (i.e., allelic, gene expression or protein level) [39–50]. Even in these studies, individual cases have demonstrated large variation in the fraction of genome reused in evolution, including absence of any shared changes, similarity in functional pathways but not genes, reuse of a limited number of genes, and abundant overlap at both gene and functional levels. Despite a few studies analyzing the patterns of phenotypic plasticity across hierarchical levels of molecular organization within a single generation [51–53], we lack empirical information on the extent to which divergence occurs in parallel across levels over longer timescales.

Here, we assess the interplay between epigenomic and transcriptomic variation underlying adaptive evolution and evaluate the degree of divergence across epigenomic (i.e., DNA methylation and miRNA) and transcriptomic (i.e., gene expression and alternative splicing) levels (hereafter referred to as "molecular levels") in the nine-spined stickleback (*Pungitius pungitius*). This species has a circumpolar distribution in the Northern Hemisphere across marine and freshwater habitats. These distinct environmental conditions impose differences in selection pressures among habitats, which has led to repeated

Luo *et al. BMC Biology*   (2025) 23:70

Page 3 of 18

morphological, behavioral, and physiological adaptations to produce distinct marine and freshwater ecotypes [54]. Therefore, the distinct nine-spined stickleback ecotypes provide an opportunity for investigations aimed at disentangling the relative importance of different molecular levels in influencing adaptive phenotypic evolution, and quantifying the degree of parallelism among these molecular levels. We address the aforementioned knowledge gap by measuring gene expression, alternative splicing events, miRNA expression, and DNA methylation in two marine populations and two freshwater populations of *P. pungitius*, and analyzing the differences between marine and freshwater ecotypes across these epigenomic and transcriptomic levels. Our goal was twofold: (1) to provide the first insights into the interplay among molecular mechanisms (viz. gene expression, alternative splicing events, miRNA expression, and DNA methylation) underpinning adaptive evolution, and (2) to assess the degree of parallelism (i.e., the overlap of loci involved in ecotype divergence and the relative strength of divergence) between the four mechanisms underlying adaptive phenotypic divergence.

Results

Extensive changes in gene expression, alternative splicing, miRNA expression, and DNA methylation between ecotypes

For gene expression analysis, we filtered genes with low expression and retained 13,467 genes for downstream analysis. To identify general gene expression patterns, we performed principal component analysis (PCA) on *rlog*-transformed read counts of the 13,467 filtered genes across all samples. PC1 (variance explained: 69.14%) clearly separated brain and liver samples. Within the same tissue, PC2 (variance explained: 10.03%) mainly separated marine and freshwater ecotype samples, with the exception of two marine fish clustering with freshwater fish (Fig. 1a; Additional file 1: Fig. S1a). The quadratic discriminant analysis (QDA) showed that gene expression pattern could accurately separate ecotypes, with accuracy of 0.80 estimated from the cross-validation of the model (Additional file 1: Fig. S2a). We identified 2,088 and 1,667 differentially expressed genes (DEGs) in brain and liver tissues, respectively (Fig. 2a, b), with the number of identified DEGs significantly greater than expected by chance (permutation analysis, $p = 0.011$ in brain; $p = 0.015$ in liver; Additional file 1: Fig. S3a-b). The finding of a greater number of DEGs in brain is consistent with previous studies analyzing ecotype divergence at the gene expression level using these same populations of nine-spined sticklebacks [55]. Our power analysis suggests that the influence of type 2 error on our criteria for calling DEGs based on our sample size is small, with our

sample size and sequencing depth achieving a power of 0.90 for both brain and liver (Additional file 1: Fig. S4a-b). Based on Euclidean distances of relative gene expression levels (i.e., normalized read counts scaled for each gene's expression level, with median read counts equal to 0), individual fish mainly clustered by their ecotypes within each tissue type.

For alternative splicing analysis, we first identified general splicing patterns using PCA based on 29,418 splicing events across all samples. PC1 (variance explained: 38.38%) clearly separated brain and liver samples. Within the same tissue, PC2 (variance explained: 8.85%) separated marine and freshwater ecotype samples, except for one marine sample that clustered with freshwater samples (Fig. 1b; Additional file 1: Fig. S1b). The QDA showed that splicing pattern could accurately separate ecotypes, with accuracy of 1.00 estimated from cross-validation of the model (Additional file 1: Fig. S2b). We then assessed the number of alternative splicing events in each tissue. We detected all five alternative splicing events in both brain and liver tissues. In total, we found 35,744 splicing events in 8,416 genes in brain, and 23,274 splicing events in 5,645 genes in liver. Based on differential splicing events, we identified 1,802 and 1,198 differentially spliced genes (DSGs) in brain and liver, respectively (Fig. 2c, d). In both tissues, intron retention (RI) was the most abundant splicing event, while mutually exclusive exon (MXE) was the least abundant. Between 4.28 and 12.38% of the detected events were differentially spliced, with RI also showing higher number of differential splicing compared to the other four types of splicing events in both tissues (Additional file 1: Fig. S5a-b). The relatively more important role for RI in ecotype divergence was further supported by our PCA-by-splice analyses. Similar to gene expression PCA results, both of our exon- and intron-based analysis showed a clear separation between tissues along PC1 (exon-based: variance explained 15.68%; intron-based 49.23%). However, marine and freshwater ecotypes clustered more clearly along PC2 in intron-based analysis (Additional file 1: Fig. S5c-d). In addition, in contrast to the 438 and 265 differentially spliced intron clusters identified in brain and liver, respectively, we found no differential exon usage in either of the tissues. Thus, the PCA-by-splice analyses supported our event-based results that variation in intron splicing was likely to play larger role in ecotype divergence than variation in exon splicing, with the marine ecotype retaining more introns than the freshwater ecotype in differential splicing (median ΔPSI for all differentially spliced RI in brain: 1.00×10^{-3}; median ΔPSI for all differentially spliced RI in liver: 1.00×10^{-3}; Additional file 1: Fig. S5e-f), although the difference between ecotypes was not significant in brain (brain, Kruskal–Wallis test: $\chi^2 = 0.02$,
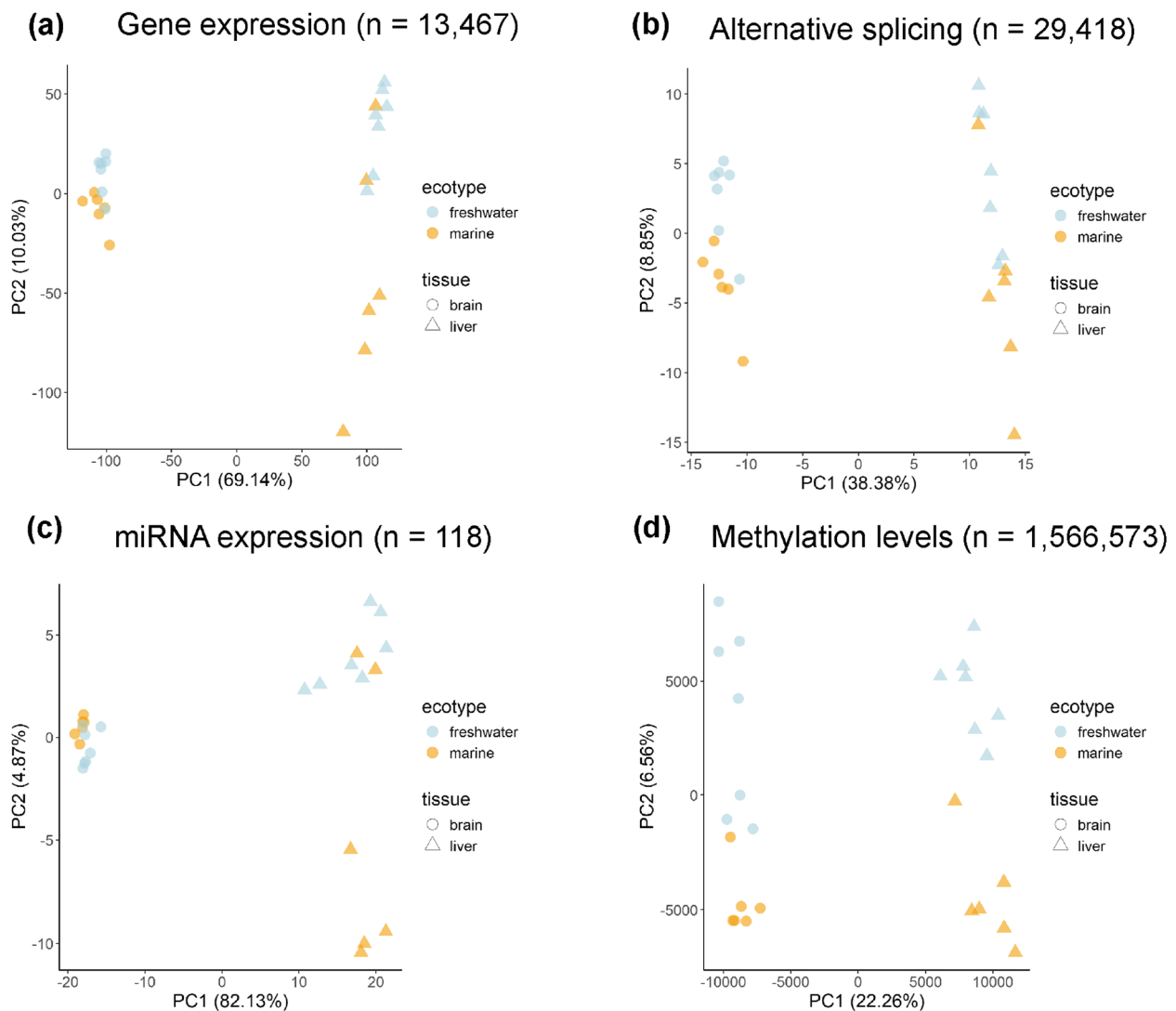
Luo *et al. BMC Biology*   (2025) 23:70

Page 4 of 18

Fig. 1 Epigenomic and transcriptomic divergence patterns between ecotypes. Principal component analysis (PCA) of (**a**) gene expression, (**b**) alternative splicing events, (**c**) miRNA expression, and (**d**) methylation profiles for marine (orange) and freshwater (light blue) in brain (circle) and liver (triangle). The numbers of (**a**) genes, (**b**) splicing events, (**c**) miRNAs, and (**d**) CpG sites used for PCA analyses are indicated in parentheses

df $= 1$, $p = 0.89$; liver, Kruskal–Wallis test: $\chi^2 = 4.68$, df $= 1$, $p = 0.03$). Based on our sample size and sequencing depth, the influence of type 2 error on our criteria for calling splicing events using rMAT should be small (true positive rate of 0.8) [56].

For miRNA expression, we annotated 444 precursor miRNAs, among which 349 (78.60%) were conserved (i.e., with orthologous miRNAs already identified in other species registered in miRBase and our choice of closely related fish species), whereas the remaining 95 (21.40%)

(See figure on next page.)

Fig. 2 Epigenomic and transcriptomic divergence between ecotypes. Heatmaps of (**a**, **b**) differentially expressed genes (DEGs), (**c**, **d**) differentially spliced genes (DSGs), (**e**, **f**) differentially expressed miRNAs (DEmiRs), and (**g**, **h**) differentially methylated cytosines (DMCs) in brain and liver tissues. Each column represents one sample, and each row represents one DEG/DSG/DEmiR/DMC. Columns and rows are both clustered using the "complete" clustering method based on Euclidean distances of samples or DEGs/DSGs/DEmiRs/DMCs. The color bar represents relative expression or methylation level, where darker red indicates higher expression or methylation level, while darker blue indicates lower expression or methylation level in that sample. The numbers of (**a**, **b**) DEGs, (**c**, **d**) DSGs, (**e**, **f**) DEmiRs, and (**g**, **h**) DMCs used for clustering are indicated in parentheses

Luo *et al. BMC Biology*   (2025) 23:70

Page 5 of 18

(a) DEGs (brain, n=2,088)

(b) DEGs (liver, n=1,667)

(c) DSGs (brain, n=1,802)

(d) DSGs (liver, n=1,198)

(e) DEmiRs (brain, n=18)

(f) DEmiRs (liver, n=37)

(g) DMCs (brain, n=1,433)

(h) DMCs (liver, n=1,732)

Fig. 2 (See legend on previous page.)

Luo *et al. BMC Biology*   (2025) 23:70

Page 6 of 18
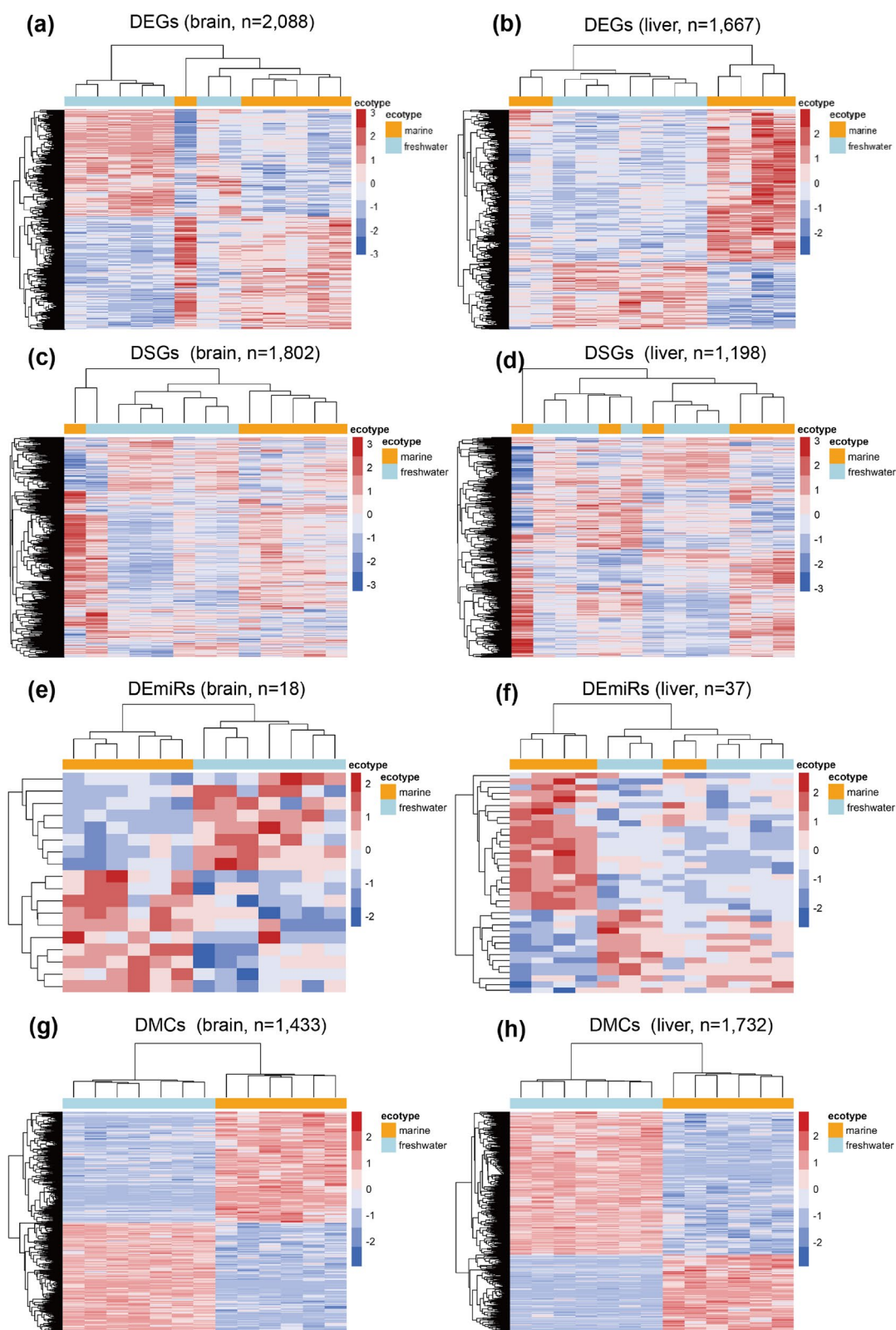
were novel (Additional file 1: Fig. S6a; Additional file 2: Table S1). We found that the total miRNA population was mainly comprised of several highly expressed miRNAs and that their composition varied across tissues (Additional file 1: Fig. S6b). Similar to gene expression analysis, we filtered miRNAs with low expression and retained 118 miRNAs for downstream analysis. As observed for gene expression patterns, when performing PCA on the expression levels of 118 filtered miRNAs, we found that PC1 (variance explained: 82.13%) clearly separated brain and liver tissues. Within the same tissue, PC2 (variance explained: 4.87%) mainly separated marine and freshwater ecotype samples, although there was partial overlap between marine and freshwater ecotype samples (Fig. 1c; Additional file 1: Fig. S1c). The QDA showed that miRNA expression pattern could accurately separate ecotypes, with accuracy of 0.80 estimated from cross-validation of the model (Additional file 1: Fig. S2c). We thus analyzed differential expression of miRNAs between ecotypes separately for brain and liver and identified 18 and 37 differentially expressed miRNAs (DEmiRs) in brain and liver, respectively, with similar numbers of miRNAs up- and downregulated in marine ecotypes relative to freshwater ecotypes in both tissues (brain: 10 up- vs. 8 downregulated; liver: 23 up- vs. 14 downregulated; Additional file 1: Fig. S6c-d). Permutation analyses revealed that in both tissues, the observed numbers of DEmiRs were significantly greater than expected by chance ($p < 0.001$ in brain; $p = 0.011$ in liver; Additional file 1: Fig. S3c-d). Our power analysis suggests that our sample size should provide moderate to high power for calling DEmiRs, with a power of 0.68 and 0.88 achieved for brain and liver samples, respectively (Additional file 1: Fig. S4c-d). Based on Euclidean distances of relative miRNA expression (i.e., normalized read counts scaled for each miRNA's expression level, with the median read equaling 0), individual fish mainly clustered by their ecotypes within each tissue (Fig. 2e, f).

For DNA methylation analysis, PCA on filtered CpGs showed a clear separation between tissues along PC1 (variance explained: 22.26%) as well as a clear separation between ecotypes along PC2 (variance explained: 6.56%) (Fig. 1d; Additional file 1: Fig. S1d). The QDA showed that DNA methylation could accurately separate ecotypes, with accuracy of 1.00 estimated from cross-validation of the model (Additional file 1: Fig. S2d). When analyzing the proximity of differentially methylated cytosines (DMCs) to CpG islands (CpGi), we found 42.36 and 40.13% of DMCs within or proximal to CpGi (i.e., within CpGi shores or CpGi shelves) in brain and liver tissues, respectively, supporting their potential functional roles in gene expression (Additional file 1: Fig. S7a). However, when analyzing the distribution of genomic features

(i.e., promoters, exons, introns, and intergenic regions) built on DMCs, we found no significant enrichment of DMCs in any genomic features in brain and liver tissues compared to the null distribution built on the 1,566,573 filtered CpGs (*G*-test, $p > 0.05$). Based on filtered CpGs, we identified 1,433 and 1,732 DMCs between ecotypes in brain and liver tissues, respectively, with no apparent clustering on specific chromosomes or chromosomal regions (Additional file 1: Fig. S7b-c). Permutation analyses revealed that in both tissues, the observed numbers of DMCs were significantly greater than expected by chance ($p < 0.001$ in both brain and liver; Additional file 1: Fig. S3e-f). Our power analysis suggests that less than 1% of sites are likely to have been falsely identified as non-DMCs (Additional file 1: Fig. S4e) across all effect size groups. This implies that the impact of type 2 error on our criterion for identifying DMCs is limited and would only affect a small number of sites. Individual fish clustered by their ecotype based on Euclidean distance of relative percent DNA methylation (normalized percent DNA methylation scaled for each DMC's percent DNA methylation) in both tissue types (Fig. 2g, h).

Genome-wide and localized interplay between epigenomic and transcriptomic variation

We analyzed the interplay between epigenomic and transcriptomic variation at two scales. First, we quantified the proportion of the total gene expression or splicing variation that is significantly associated with epigenetic variation at a genome-wide scale. For gene expression, when considering the full data set, miRNA and DNA methylation levels jointly explained a significant proportion of total gene expression variation in both brain (adjusted $R^2 = 0.58$, $F_{(6,6)} = 3.74$, $p < 0.01$; Fig. 3a) and liver (adjusted $R^2 = 0.82$, $F_{(6,6)} = 10.07$, $p < 0.01$; Fig. 3b). miRNA and DNA methylation separately explained 50.48% (adjusted $R^2 = 0.50$, $F_{(3,9)} = 5.08$, $p < 0.01$) and 38.66% (adjusted $R^2 = 0.39$, $F_{(3,9)} = 3.52$, $p < 0.01$) of variance in gene expression in brain (Fig. 3a), and 66.06% (adjusted $R^2 = 0.66$, $F_{(3,9)} = 8.78$, $p < 0.01$) and 61.86% (adjusted $R^2 = 0.62$, $F_{(3,9)} = 7.49$, $p < 0.01$) of variance in gene expression in liver (Fig. 3b). When controlling for miRNA effects, DNA methylation explained a significant proportion of variance in gene expression only in liver (adjusted $R^2 = 0.16$, $F_{(3,6)} = 3.64$, $p < 0.05$) but not in brain (adjusted $R^2 = 0.07$, $F_{(3,6)} = 1.52$, $p > 0.05$). Similarly, when controlling for DNA methylation effect, miRNA also explained a significant proportion of variance in gene expression only in liver (adjusted $R^2 = 0.20$, $F_{(3,6)} = 4.33$, $p < 0.05$) but not in brain (adjusted $R^2 = 0.19$, $F_{(3,6)} = 2.36$, $p > 0.05$). For splicing, when considering the full data set, miRNA and DNA methylation levels jointly explained a significant proportion of total splicing variation in both

Luo *et al. BMC Biology*   (2025) 23:70

Page 7 of 18

brain (adjusted $R^2 = 0.62$, $F_{(6,6)} = 4.30$, $p < 0.05$; Fig. 3c) and liver (adjusted $R^2 = 0.79$, $F_{(6,6)} = 8.45$, $p < 0.001$; Fig. 3d). miRNA and DNA methylation separately explained 41.40% (adjusted $R^2 = 0.41$, $F_{(3,9)} = 3.83$, $p < 0.05$) and 33.43% (adjusted $R^2 = 0.33$, $F_{(3,9)} = 3.01$, $p < 0.05$) of variance in splicing in brain (Fig. 3c), and 55.62% (adjusted $R^2 = 0.56$, $F_{(3,9)} = 6.01$, $p < 0.001$) and 79.35% (adjusted $R^2 = 0.79$, $F_{(3,9)} = 16.37$, $p < 0.001$) of variance in splicing in liver (Fig. 3b). However, we only found DNA methylation explained a significant proportion of variance in splicing when controlling for miRNA effect in liver (adjusted $R^2 = 0.23$, $F_{(3,6)} = 4.29$, $p < 0.05$).

Given the significant interplay between epigenomic and transcriptomic variation at the genome-wide level, we next tested for parallelism between the molecular mechanisms underlying ecotype divergence by assessing the overlaps between DEGs, DSGs, target genes of DEmiRs, and genes associated with DMCs. In contrast to the significant correlation between epigenomic variation and gene expression at whole genome scale, the set of genes that were differentially expressed and spliced were largely nonoverlapping with target genes of DEmiRs and DMC-associated genes (Fig. 3e, f). Specifically, we found 20 DEGs regulated by 11 DEmiRs in brain, and 28 DEGs regulated by 14 DEmiRs in liver, suggesting a small proportion (0.96–1.68%) of DEGs in each tissue were putatively regulated by a relatively large proportion (37.84–61.11%) of DEmiRs. In addition, we found three DSGs to be the predicted targets of three DEmiRs in brain, and five DSGs to be the predicted targets of five DEmiRs in liver, suggesting a minor role of miRNAs in regulating differential splicing, although the expression levels of several key genes associated with ecotype divergence [55] were negatively correlated with the expression levels of miRNAs. Finally, we annotated a large number of genes to DMCs in brain ($n = 714$) and liver ($n = 815$) tissues, but only a very small proportion (4.91–9.80%) of these genes overlapped with DEGs and DSGs in respective tissues. Thus, we found small overlaps among differentially expressed, spliced, and methylated genes and target genes with DEmiRs between ecotypes, ranging from 17.37 to 56.60% in brain and 12.02 to 26.67% in liver.

The relative strength of divergence at different molecular levels

When estimating the extent of divergence on DEGs, DSGs, DEmiRs, and DMCs, we found the average effect size of ecotype ranging from 0.22 (DSGs) to 0.79 (DMCs) in brain (Fig. 4a), and from 0.20 (DSGs) to 0.79 (DMCs) in liver (Fig. 4b). A larger ecotype effect size supports a greater relative strength of divergence at the measured molecular level. When comparing across molecular levels, we found significant difference in the extent of divergence between DEGs, DSGs, DEmiRs, and DMCs. Specifically, we found that DMCs and DSGs showed the highest and lowest extent of divergence between ecotypes, respectively, with no significant difference between DEGs and DEmiRs in both tissues. When examining the overlap between loci implicated in ecotype divergence at different molecular levels and genomic loci of high differentiation between ecotypes, we found a significant enrichment of selection signals for genes associated with DMCs in both brain and liver tissues (Fisher's exact tests, $p < 0.001$) and for DEGs in brain (Fisher's exact test, $p = 0.01$), with no enrichment for DSGs and target genes of DEmiRs in both tissues.

Discussion

Many species occupying diverse habitats evolve distinct locally adapted ecotypes. The role of variation at epigenomic and transcriptomic levels in regulating such adaptive divergence has been the subject of increasing attention in recent years [55, 57–59]. Divergence in gene expression has been strongly implicated in the rapid evolution and development of adaptive and divergent phenotypes, particularly as *cis*-regulatory mutations are thought to exhibit fewer deleterious effects than protein-coding changes [11, 60–62]. The contribution of alternative splicing to adaptation is less studied compared to the regulation of gene expression divergence, and the few studies to date have found conflicting results, with either contrasting or redundant roles between alternative splicing and gene expression in phenotypic evolution [33, 63–65]. In addition, the occurrence of phenotypic divergence despite low genetic differentiation suggests a potential role for

(See figure on next page.)

Fig. 3 Interplay between epigenomic and transcriptomic variation. (**a**, **b**) Gene expression variation and (**c**, **d**) alternative splicing variation partitioning, showing the proportions of gene expression or splicing variation explained by CpG methylation and miRNA expression, based on the adjusted R^2 of redundancy analyses (RDA) in (**a**, **c**) brain and (**b**, **d**) liver. Numbers outside the circles refer to total percentage of variation (adjusted R^2) explained by both CpG methylation and miRNA expression (top, full model), or by only CpG methylation or miRNA expression (reduced models). Percentages inside circles indicate pure contributions to gene expression or splicing level resulting from partial RDAs that isolate the effect of one single explanatory variable. Percentages within intersections indicate shared contribution across two variables to gene expression or splicing variation. ANOVA-like permutation tests were calculated on total and pure contributions. NT, non-testable; NS, non-significant; ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$. (**e**, **f**) Venn diagrams of DEGs, DSGs, DEmiR target genes and DMC-associated genes in (**e**) brain and (**f**) liver
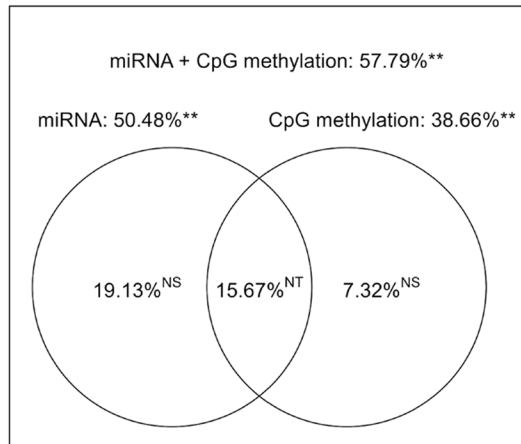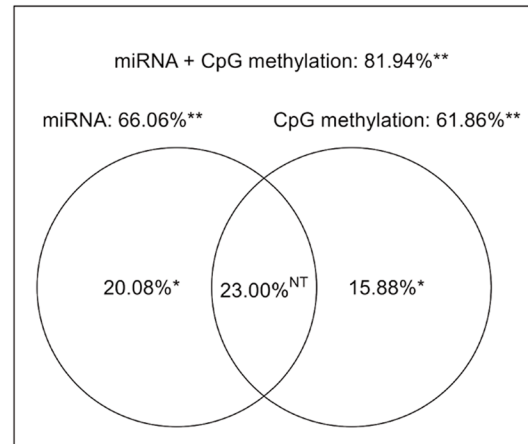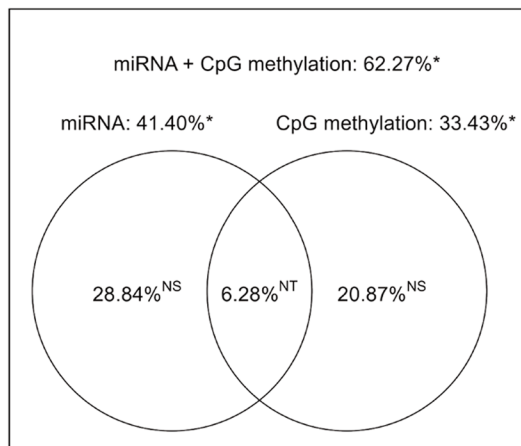
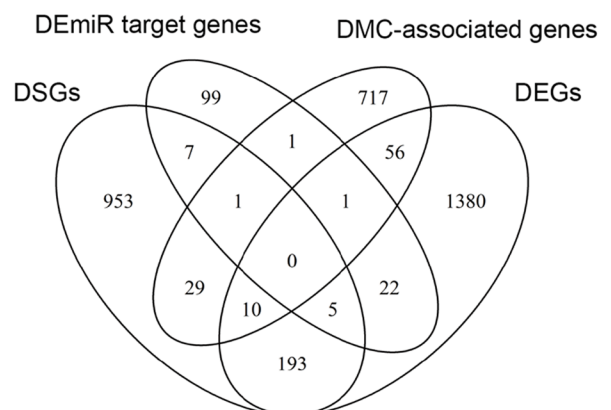Luo *et al. BMC Biology*   (2025) 23:70

Page 8 of 18
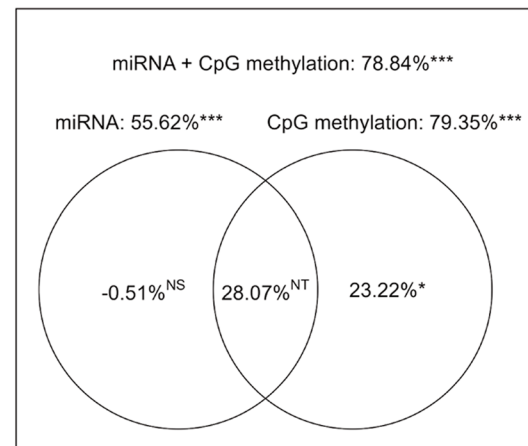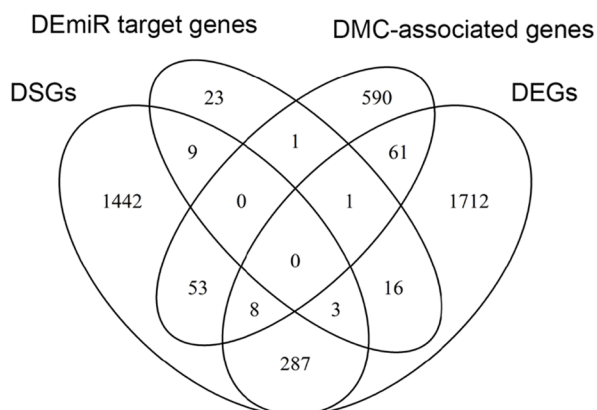
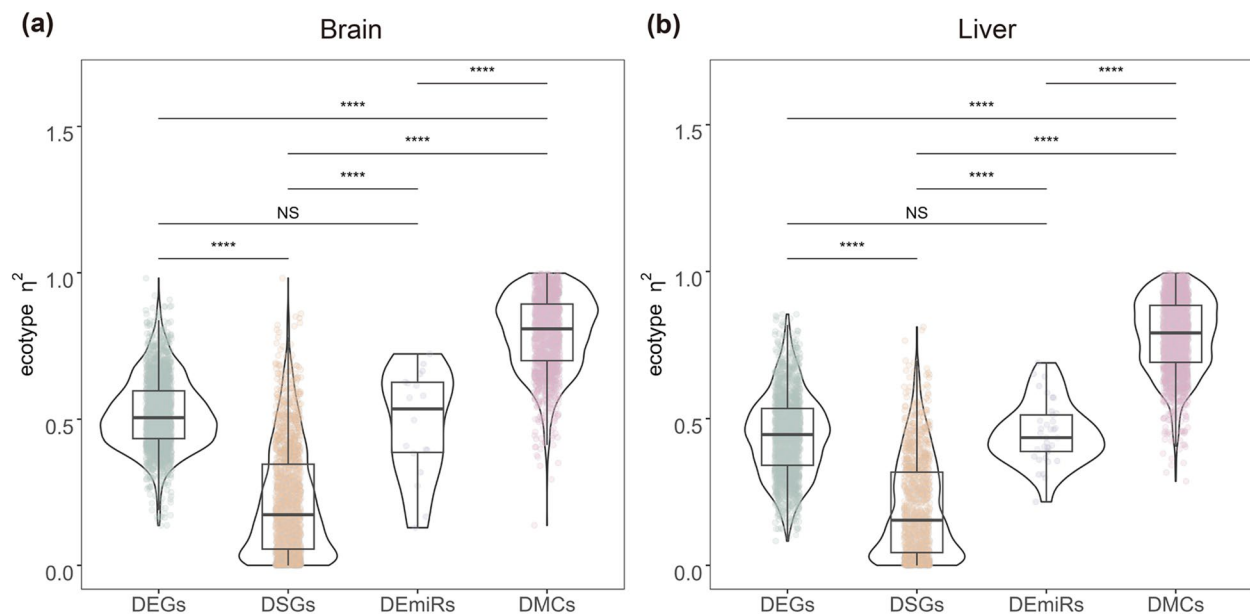**Fig. 3** (See legend on previous page.)

Fig. 4 The extent of divergence across different molecular levels. We examined the relative strength of ecotype divergence in gene expression, alternative splicing, miRNA expression, and DNA methylation in (**a**) brain and (**b**) liver, summarized by the effect size (partial eta, η^2) of ecotype. Width indicates the pooled distribution density of the effect size of ecotype in a given molecular level. Embedded box plots summarize variation in the effect size of ecotype in a given molecular level. The line indicates the median, the box defines the interquartile range, and the whiskers represent the maximum and minimum values of the effect size of ecotype. Significance levels are estimated by the pairwise Wilcoxon rank sum tests with Benjamini–Hochberg correction: ****, $p < 0.0001$; NS, represents non-significance

epigenetic mechanisms such as miRNAs and DNA methylation during adaptation. Noteworthy examples include differentially expressed miRNAs or differentially methylated CpGs between three-spined stickleback ecotypes that contained single-nucleotide polymorphisms (SNPs) with shifted allele frequencies [30, 66] or were associated with genes having potentially adaptive functions [14, 15, 67]. However, the interplay between these molecular levels in adaptive divergence has not been studied in any system and thus remains poorly understood. We quantified gene expression, alternative splicing, miRNA expression, and DNA methylation variation from brain and liver tissues of stickleback ecotypes originating from two contrasting habitat types. While only using female samples restricted our ability to detect sex-specific difference in ecotype divergence, and using limited number of populations along a relatively small environmental gradient hindered our capability to explore marine-freshwater divergence, we found substantial changes at all molecular levels, consistent with previous studies analyzing divergence at a single molecular level. Although we detected a significant correlation between epigenomic variation and gene expression patterns at the genome-wide scale, genes that showed divergence

between ecotypes at different molecular levels were largely nonoverlapping, suggesting non-parallel roles of gene expression, splicing, miRNAs, and DNA methylation in regulating adaptive divergence. Alternatively, it is possible that the small overlap is due to the complex relationship between epigenomic variation and transcriptomic variation, such as involvement of enhancers [68, 69] or the spatiotemporal regulation of gene expression by the epigenome, making direct correlations difficult to find. Pathway analyses such as the Weighted Gene Co-expression Network Analysis might be helpful in connecting epigenomic and transcriptomic changes beyond gene level in future studies [70]. Finally, we found significant differences in the extent of divergence and enrichment of selection signals between these molecular levels, with methylation showing the highest degree of divergence and splicing showing the lowest degree of divergence between ecotypes, and methylation significantly overlapping with loci of high differentiation between ecotypes. Overall, our study provides the first integrated investigation of the diverse epigenomic and transcriptomic mechanisms underlying ecotype divergence in any animal species, opening a window to understand how these different molecular mechanisms can interact and contribute to adaptation.

Luo *et al. BMC Biology*   (2025) 23:70

Page 10 of 18

Genome-wide versus localized interplay between molecular mechanisms

One of the aims of our study was to assess the interplay between epigenomic and transcriptomic variation in ecotype divergence. Consistent with previous studies of phenotypic divergence that showed a significant correlation between DNA methylation, miRNA expression, and gene expression patterns at genome-wide scale [52, 71–74], we also found that epigenomic variation explained a significant proportion of overall gene expression variation. Although the mechanistic underpinnings of the interplay between epigenomic and splicing variation are not well understood in wild populations, there is evidence that epigenetic modifications can regulate genome-wide splicing patterns through the modulation of components in the spliceosome such as RNA splicing factors [26, 75]. Some of the DMC-associated genes and target genes of DEmiRs belong to the spliceosomal complex and splicing factors (e.g., *celf3a*, *nova1*), thus supporting our findings that a significant proportion of splicing variation was explained by epigenomic variation in both tissues. Interestingly, miRNA and DNA methylation together explained a significant proportion of gene expression and splicing variation in both brain and liver. However, when tested alone while controlling for the other epigenetic mechanism, both DNA methylation and miRNAs only explained significant variation of gene expression in liver, suggesting that epigenetic variation may contribute to ecotype divergence to different extents in the two tissues. This finding is consistent with recent studies that have also found that neural tissue responded less strongly to novel environments as compared to other tissues [76, 77]. Therefore, due to the tissue-specific difference in macroscopic traits and molecular variation in fish and other animals [69, 78–82], understanding the interplay between epigenomic and transcriptomic variation based on only one tissue can be challenging. Further studies comparing tissue-specific epigenetic responses within various environmental contexts will be helpful to advance a more comprehensive understanding of the interplay between epigenomic and transcriptomic variation in adaptation.

Due to the significant associations between different molecular levels at a genome-wide scale, we next analyzed whether this interplay between epigenomic and transcriptomic variation was present at the specific loci implicated in ecotype divergence. We found the repeatability of gene usage was low despite high genome-wide correlation among molecular levels, suggesting largely non-parallel roles between epigenomic and transcriptomic variation in ecotype divergence. The non-parallelism between gene expression and alternative splicing in regulating ecotype divergence might be expected because

in vertebrates, including fish, alternative splicing has been suggested to diverge faster than gene expression in intra-specific comparisons and in within-tissue analysis [25, 28]. In addition, it has been suggested that splicing might play a key role at early stages of divergence and ecological speciation by being more evolutionarily flexible than gene expression and thus yielding rapid phenotypic changes [29, 83]. This is consistent with our finding that more genes were differentially spliced than differentially expressed. However, we found little overlap between genes associated with epigenomic variation (i.e., target genes of DEmiRs and genes associated with DMCs) and transcriptomic variation (i.e., DEGs and DSGs). This could be viewed as unexpected given the well-documented role of miRNAs and DNA methylation in regulating gene expression and alternative splicing [26, 27, 84, 85]. Indeed, many evolutionary transitions and key innovations are marked by methylation divergence during early stages of speciation and periods of miRNA gains and losses [58, 59, 86, 87]. Moreover, target genes of DEmiRs and DMCs included a number of genes relevant to responses to environmental changes, for example, osmotic environments, parasite communities, and food composition [54], further suggesting the potential roles of epigenetic variation in ecotype divergence. However, recent studies investigating epigenetic variation in animals exposed to temperature, salinity, and developmental stress have also found a limited role for DNA methylation in differential gene expression [52, 88–90], suggesting that environmentally induced changes in epigenetic variation are not necessarily translated into changes in gene regulation.

Degree of divergence differs between molecular levels

Both theoretical and empirical studies have suggested that the degree of adaptive divergence between molecular levels depends on the extent to which variation at a given molecular level is connected to variation at other levels [91, 92], and thus, we do not necessarily expect to find the same extent of divergence across different molecular levels. Consistent with this expectation, although there was evidence for ecotype divergence at all molecular levels, the degree of this divergence varied between epigenomic and transcriptomic levels, with splicing and DNA methylation showing the highest and lowest extent of divergence between ecotypes, respectively. The inconsistent degree of divergence across levels indicates that the epigenomic and transcriptomic basis of ecotype divergence are likely subject to redundancies, similar to the widespread observation of genetic architecture redundancy underlying repeated phenotypic divergence [93]. Alternatively, this could suggest that epigenomic and transcriptomic variation contribute in tandem to

adaptive divergence, similar to the partially autonomous of epigenetic variation from genetic variation found in populations adapting to heterogeneous environments [14, 19, 94, 95]. Our findings may also suggest that selective pressures may operate independently on variation at each of the molecular levels, which was supported by the differences in the enrichment of selection signals between different molecular levels. This result has also been shown in previous studies comparing the direction of plasticity evolution between gene expression and DNA methylation [51], and the degree of susceptibility to erosion of genetic variation between gene expression and alternative splicing in plasticity [96]. Thus, understanding the relative strength of ecotype divergence at multiple levels can shed light on the constraints on evolvability of each molecular mechanism during ecotype divergence.

Limitations

Finally, the present study has a number of potential caveats that should be noted. First, the number of individuals and populations used in the study was limited. Previous studies have suggested adaptive phenotypic differentiation between stickleback populations from the Baltic Sea due to an environmental gradient [97]. Thus, results from our study should be interpreted with caution because the populations used in our study might not be representative of most marine-freshwater differentiation. Including samples from additional populations along environmental gradients would provide information on inter-population and inter-ecotype level variation, as well as the same loci present in more samples and associated with ecotype divergence likely increasing the power of the analyses. Second, although individuals clustered primarily by ecotype at each molecular level, there were few exceptional individuals. Since individuals were derived from within-population crosses, it is unlikely these individuals were misclassified to a wrong ecotype. It is possible that these outlier individuals are instead due to heterogeneity in ecotype-specific expression and splicing patterns that is not captured by our analyses, or are caused by other molecular mechanisms not studied here. A wider investigation of genome-wide regulatory elements (e.g., Assay for Transposase Accessible Chromatin with high-throughput sequencing (ATAC-seq)) [98] would provide a more comprehensive understanding of the molecular mechanisms underlying ecotype divergence.

Conclusions

In this study, we identified patterns of divergence between ecotypes at gene expression, alternative splicing, miRNA expression, and DNA methylation levels. In addition, we provide the first insights into the interplay between epigenomic and transcriptomic variation in ecotype divergence, and demonstrate contrasting patterns between genome-wide and genetic scales. Finally, we found significant difference in the relative strength of ecotype divergence across molecular levels, with the highest and lowest degree of divergence at DNA methylation and splicing levels, respectively, and a significant enrichment of selection signals in differential methylation. Our findings suggest that different biological levels may respond differently to the same selective pressures and provide distinct routes to adaptive divergence via either complementary or redundant mechanisms.

Methods

Sample collection

We collected adult nine-spined sticklebacks during the breeding season (June–July) of 2013 from four Fennoscandian sites, including two marine [Baltic Sea at Bölesviken in Sweden (SWE-BOL) and Baltic Sea at Helsinki in Finland (FIN-HEL)] and two freshwater populations [Abbortjärn pond in Sweden (SWE-ABB) and Bynästjärnen pond in Sweden (SWE-BYN)] (Additional file 1: Fig. S8). Detailed locality information and sampling methods can be found in Wang et al. [55]. In using a set of populations in which ecotype divergence has been established in previous studies [55, 99, 100], our results build on prior work by explicitly testing the relative roles of multiple underlying molecular mechanisms in a common garden experiment. A randomly selected female and male from each population were artificially crossed and the F_1 offspring from each cross were raised in 5-L tanks in zebrafish rack systems (Aquaneering Inc., San Diego, USA) at 17°C for 180 days after hatching. While we reared marine fish in a freshwater environment, previous studies have found that wild marine nine-spined stickleback populations tend to spawn in or near freshwater habitats [101], and thus our rearing conditions were not expected to significantly bias the epigenomic and transcriptomic patterns of marine fish. Only females identified with the aid of a sex-linked microsatellite marker were selected to avoid sex-specific effects on the transcriptome and epigenome [102]. Fish from the same population were sampled from at least two families. Whole brain and liver from a total of 13 fish (six marine fish, SWE-BOL: $n = 3$, FIN-HEL: $n = 3$; seven freshwater fish, SWE-ABB: $n = 3$, SWE-ABB: $n = 4$) were dissected and immediately frozen in liquid nitrogen and were then transferred to −80°C until DNA and RNA extraction (Additional file 3: Table S2). We selected brain and liver tissues because previous studies have found tissue-specific transcriptomic-wide associations with adaptation to divergent marine and freshwater environments in these

Luo *et al. BMC Biology*   (2025) 23:70

Page 12 of 18

tissues [54, 55]. Therefore, using both brain and liver tissues provided an opportunity to comprehensively analyze the molecular mechanisms underlying ecotype divergence.

RNA and DNA extraction and sequencing

To understand the interplay between diverse molecular mechanisms underlying ecotype divergence, we performed RNA sequencing (RNA-seq), small RNA sequencing (sRNA-seq), and whole genome bisulfite sequencing (WGBS) on each of the 13 fish. Total RNA was extracted using TRIzol reagent (Invitrogen, Calsbad, CA, USA). RNA quality was assessed using Agilent 2100 bioanalyzer (Agilent Technologies). Strand-specific mRNA libraries were constructed and sequenced on an MGISEQ-2000 platform using 150 bp paired-end reads. Each RNA-seq library was sequenced to a mean amount of 23,649,365 ± 1,239,0167 (± SD) reads (Additional file 3: Table S3). sRNA-seq libraries were prepared using BGI UMI small RNA library protocol and sequenced on a BGISEQ-500 platform using 50 bp single-end reads. Each sRNA-seq library was sequenced to a mean amount of 41,415,505 ± 3,409,559 (± SD) reads (Additional file 3: Table S4).

Whole genomic DNA was extracted using the EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, California, USA) following the manufacturer's protocol. DNA quality was assessed using NanoDrop 2000c (Thermo Scientific). WGBS libraries were sequenced on an MGISEQ-2000 platform using 100 bp paired-end reads. Each WGBS library targeted a genome coverage of 15× and was sequenced to a mean amount of 41,428,416 ± 1,822,025 (± SD) reads (Additional file 3: Table S5). The quality control for total DNA and RNA samples, library construction, and sequencing were performed by BGI (Wuhan, China).

Reference genome preparation

We used the chromosome-level nine-spined stickleback reference genome [ENA accession GCA_902500615.3; [103] as the reference in all downstream analyses. Because the chromosome-level nine-spined stickleback reference genome is not annotated, we used Liftoff v1.6.2 [104] with default parameters to lift over genomic features of the scaffold-level nine-spined stickleback reference genome (NCBI accession GCA_902500615.3 NSP_V7) to the chromosome-level genome, followed by phasing the genome using genome-tools v1.6.1 (https://github.com/genometools/genometools). In total, all but seven protein-coding genes were lifted over from the scaffold-level reference genome to the chromosome-level reference genome.

Gene expression analysis

To explore gene expression differences between marine and freshwater ecotypes, we first trimmed raw RNA-seq reads using Trim Galore! v0.6.7 (https://github.com/FelixKrueger/TrimGalore) with default settings after quality checks using FastQC v0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Trimmed reads were then aligned to the nine-spined stickleback reference genome using a two-pass mapping approach implemented in STAR v2.7.10a [105], with only uniquely mapped reads retained for downstream analyses. On average, 95.22% ± 1.01% (± SD) RNA-seq reads across all libraries were uniquely mapped to the reference genome. Retained reads were quantified at gene level using HTseq v0.11.3 [106]. We first filtered genes with no expression across all brain and liver samples and performed principal component analysis (PCA) based on *rlog*-transformed read counts. Based on PCA, we analyzed gene expression difference between ecotypes separately for brain and liver samples (Fig. 1a). In addition to PCAs, we conducted a supervised quadratic discriminant analysis (QDA) on the PCs to classify individuals either as marine or freshwater ecotype, using the R package MASS v7.3–58.2 [107]. Following Coll-Costa et al. [108], we built four models using the training dataset (i.e., randomly selected 80% of samples across tissues), each adding one more principal component (PC): in the first model, we tested the PC1; in the second model, we tested PC1 + PC2; in the third model, we tested PC1 + PC2 + PC3; and in the fourth model, we tested all four PCs. We then predicted ecotypes for individuals in the test dataset (i.e., the remaining 20% of samples). We determined the accuracy of each of the models using leave-one-out cross-validation on the training dataset and retained the models with an accuracy > 0.75. We assigned each individual from the test dataset a score ranging from 0 (marine ecotype) to 1 (freshwater ecotype), which we refer as the probability of being identified as a freshwater ecotype (PFE). Finally, we identified differentially expressed genes using the R package DESeq2 v1.32.0 [109]. We considered genes to be differentially expressed (DEGs) with Benjamini–Hochberg corrected *p*-values below 0.05. To investigate the influence of false negatives (type 2 errors) on our criteria for calling DEGs, we performed a power analysis following Ching et al. [110].

Alternative splicing analysis

To characterize splicing divergence between marine and freshwater ecotypes, we first used the event-based tools

rMATS v4.1.0 [56] to quantify both annotated and novel splicing junctions across all tissues to characterize general splicing patterns, using the same aligned reads from STAR as described above. We then performed differential splicing analysis separately for brain and liver samples. The rMATS programs identified five types of alternative splicing events, including alternative 5' splice site (A5SS), alternative 3' splice site (A3SS), mutually exclusive exon (MXE), intron retention (RI), and skipped exon (SE). Following Tian and Monteiro [111], a sum of 20 counts were required for both inclusion or skipping junctions to support a splice site. We then used rMATS to assess differential splicing patterns between marine and freshwater ecotypes in each tissue. The inclusion level difference (ΔPSI) for each splicing site between ecotypes was reported by rMATS. A spliced site was considered to be differentially spliced (DS) with false discovery rate (FDR) below 0.05. We considered a gene to be differentially spliced (DSG) when it contained at least one DS site between ecotypes. We also conducted a QDA on the PCs of splicing events to classify individuals as either marine or freshwater ecotype.

To further validate splicing patterns revealed by rMATS, we also analyzed alternative splicing patterns between ecotypes using exon- and intron-based methods. To analyze differentially spliced exons between ecotypes, we used the HTSEQ-COUNT script implemented in the R package DEXSEQ v1.38.0 [112] to quantify exon-specific read counts for each sample. We filtered exons with no expression across all brain and liver samples, and estimated differential exon usage separately for brain and liver tissues using the default model structure implemented in DEXSEQ: ~ sample + exon + ecotype:exon, based on normalized exon expression counts. We considered exons to be differentially spliced with Benjamini–Hochberg corrected *p*-values below 0.05. To analyze differentially spliced introns between ecotypes, we used LeafCutter v0.2.9 [113] to perform intron clustering and differential splicing analyses based on differential intron excision ratios (DIE) for each tissue separately (minimum coverage of 10 reads, minimum of six samples supporting an intron, minimum of three samples per ecotype supporting an intron). The difference in usage proportion of intron clusters was measured as the "change of percent spliced in" and intron clusters with Benjamini–Hochberg corrected *p*-values below 0.05 were considered differentially spliced.

Annotation, differential expression and target prediction of miRNAs

To elucidate the miRNA expression patterns between ecotypes, we trimmed raw miRNA reads using Trim Galore! to remove adapters and low-quality reads, and retained reads with length between 18 and 25 nt. To improve the quality of miRNAs, we removed rRNAs by aligning trimmed reads against database implemented in SortMeRNA v2.1b [114]. We further removed tRNAs, snRNAs, and snoRNAs by aligning reads against Rfam database using Bowtie v1.2 [115] (options: –best –strata). The unaligned reads were considered to be clean miR-NAs to be used for downstream analyses.

We used Mirdeep2 v0.1.3 [116] to quantify and annotate clean miRNAs. We first mapped clean miRNA reads to the nine-spined stickleback genome using the mapper.pl module. An average of 86.45% ± 1.78% (± SD) clean miRNA reads across all libraries were mapped to the reference genome. Since miRNA had not yet been annotated in the nine-spined stickleback, we annotated the first set of miRNAs for this species, and then determined how the composition of miRNAs varied across tissues and ecotypes. The hairpin-like miRNA precursors were predicted by the miRDeep2.pl module using all brain and liver samples with default settings. Following Kelley et al. [27], we used all miRNA sequences from chordates in miRBase 22.1 [117], as well as miRNA annotations of the three-spined stickleback [118], the gar [119], the blackfin icefish [120], and the zebrafish [121] as closely related species to guide the annotation. We considered predicted miRNAs to be true positives with miRDeep2 score above 10 and significant randfold *p*-value [122]. To identify conserved and novel miRNAs, the true positive miRNAs were blasted against chordate miRNA annotations from miRBase and the four closely related species, and those with a hit were considered as conserved miR-NAs whereas those with no hit were considered novel miRNAs. All conserved miRNAs were named under the miRNA nomenclature guidelines [123], and a prefix "novel" was used for novel miRNAs. The complete annotation of miRNAs in *P. pungitius* is summarized in Additional file 2: Table S1.

We used the quantifier.pl module implemented in miR-Deep2 to quantify the expression of all conserved and novel miRNAs. We also conducted a QDA on the PCs of miRNA patterns to classify individuals as either marine or freshwater ecotype. Differential expression analysis of miRNAs was performed separately for each tissue, using the R package DESeq2 with the same approach as described above for identifying DEG between ecotypes. We considered miRNAs to be differentially expressed (DEmiRs) with Benjamini–Hochberg corrected *p*-values below 0.05. To predict the putative targets of DEmiRs, we used miRanda v3.3a [124] to select high-confidence miRNA-gene targeting pairs in each tissue. We extracted 3'UTR of all genes from the nine-spined reference genome, and searched for targeting sites of DEmiRs. The resulting miRNA-gene targeting pairs were filtered for an

Luo *et al. BMC Biology*   (2025) 23:70

Page 14 of 18

inverse relationship between expression levels of paired genes and miRNAs in brain or liver tissue. Similar to gene expression analysis, we also performed a power analysis for our miRNA expression analysis following Ching et al. [110].

DNA methylation analysis

To remove adapter contamination and low-quality reads, we performed read quality checks using FastQC, and trimmed low-quality and adapter-contaminated sequences using Trim Galore! with default parameters. We then ran Bismark v0.21.0 [125] with the Bowtie2 v2.3.0 aligner [126] against the nine-spined reference genome with default settings, except for tolerating one non-bisulfite mismatch per read. Duplicated alignments mapping to the same position in the genome were discarded before extracting methylation information. Only CpG context cytosine methylation was analyzed because CpG methylation is the most common functional methylation in vertebrates [127]. We only included reads that mapped uniquely to the reference genome, and cytosines that were covered by at least five reads across all brain and liver samples in downstream analyses. After adapter and quality trimming, an average of $70.25\% \pm 3.34\%$ ($\pm$ SD) WGBS reads across all libraries were uniquely mapped to the reference genome, resulting in an average sequencing depth of $13\times$ per sample, which is comparable to the recommended sequencing depth ($5-15\times$) for providing sufficient statistical power to identify true positives [84]. To improve methylation estimates, we corrected for SNPs, which could have resulted in an incorrect methylation call if C-to-T and G-to-A SNPs were falsely interpreted as unmethylated cytosines [128]. We identified SNPs using the merged BAM format mapping results from the same individual for input to Bis-SNP v0.82.2 [129] with the default parameters. We retained biallelic SNPs with a minor allele frequency above 0.05, a minimum read depth of 5, a maximum read depth of 100, and mapping quality above 20, following steps in Dahms et al. [130]. We produced a list of positions (C-to-T and G-to-A SNPs) for correcting methylation estimates, using custom written Perl scripts from Heckwolf et al. [19]. In total, 1,566,573 CpGs were retained after filtering. We also conducted a QDA on the PCs of methylation levels to classify individuals either as the marine or freshwater ecotype. We identified individual differentially methylated cytosines (DMCs) using the R package methylKit v1.4.1 [131] using CpG sites that passed the filtering step above. We analyzed DMCs separately for brain and liver tissues and considered CpGs to be DMCs with a correction for overdispersion and false discovery rate correction Q-value less than 0.05. We conducted our analyses using CpG sites over regions because the causal relationship

between methylation variation in single CpG sites and expression change of genes has been demonstrated in recent empirical studies [68, 132], and defining methylation regions can be challenging due to the uneven distribution of DNA methylation in vertebrate genomes [84]. To investigate impact of false negatives on our results, we performed a power analysis using a simulated data set in methylKit, following steps in Hu et al. [14]. We conducted simulations of differential methylation for 1% of 1,566,573 filtered CpG sites, with effect sizes of 0, 5, 10, and 15% of differential methylation, respectively. Using the same criteria as for calling DMCs described above, we calculated the proportion of CpG sites falsely identified as non-DMCs (false negatives) across all analyzed sites under each effect size.

To analyze the possible functions of DMCs between ecotypes identified in brains and livers, we first identified the proximity of DMCs to CpG islands (CpGi), which are CpG-rich regions that are usually unmethylated and serve as sites for transcription initiation [133]. We used python scripts (https://github.com/lucasnell/TaJoCGI) that apply the methods described in Takai and Jones [134]. DMCs were considered part of "shores" if they were less than 2 kb away from an island, and part of CpG "shelves" if they were between 2 and 4 kb away from an island. DMCs were considered as "open sea" if they were outside the 4-kb window [135]. We gave the precedence to CpGi > CpGi shores > CpGi shelves > open sea. To analyze the genomic features (promoters/exons/introns/intergenic regions) of DMCs, we compared the distribution of genomic features built on DMCs in brain and liver tissues to the null distribution of genomic features built on all filtered CpG sites. We gave precedence to promoters > exons > introns > intergenic regions when features overlapped, and defined promoter regions as upstream 1 kb and downstream 1 kb from the transcription starting sites [136]. We then identified the positions of DMCs within genomic features and compared the distributions of DMCs to null distributions using *G*-tests.

Interplay between epigenomic and transcriptomic variation at a genome-wide scale

In addition to the detection of differential gene expression, splicing, miRNA expression, and DNA methylation separately, we also assessed the interplay between epigenomic and transcriptomic variation at the whole genome level. We performed a global redundancy analysis (RDA) separately for brain and liver samples, using the *rlog*-transformed gene expression count table or the PSI of all splicing events as the response variable, and the normalized methylation levels of filtered CpG sites and the *rlog*-transformed miRNA expression count table as explanatory variables. Prior to RDA, we performed PCA

Luo *et al. BMC Biology*   (2025) 23:70

Page 15 of 18

on (1) gene expression levels, (2) PSI of all splicing events, (3) DNA methylation levels, and (4) miRNA expression levels, and used the first three PC factors of each of the molecular levels as multivariate summaries. We also performed a partial RDA for the effect of DNA methylation on gene expression or splicing while controlling for the effect of miRNA on gene expression, and a partial RDA for the effect of miRNA on gene expression or splicing while controlling for the effect of DNA methylation on gene expression. We quantified the contributions to the total gene expression variation using the adjusted R^2, and tested the significance of each R^2 by performing an analysis of variance (ANOVA) with 999 permutations [137].

Comparing the extent of parallelism and divergence at different molecular levels

To test for parallelism between the molecular mechanisms underlying ecotype divergence, we analyzed the overlaps between DEGs, DSGs, target genes of DEmiRs, and genes associated with DMCs separately for brain and liver samples. To compare the degree of divergence across molecular levels, we first partitioned the variance of expression levels for DEGs, DSGs, DEmiRs, and methylation levels for DMCs in each tissue identified above, following Winchell et al. [3]. We used the *eta_squared* function implemented in the R package rstatix v0.7.2 (https://rpkgs.datanovia.com/rstatix/) to estimate the effect size (partial eta, η^2) of ecotype effect (i.e., marine and freshwater) and used this to evaluate the relative strength of ecotype divergence across molecular levels, where a stronger effect size of ecotype effect suggests a greater relative strength of divergence at the molecular level. Finally, to investigate the enrichment of selection signals at different molecular levels, we compared DEGs, DSGs, target genes of DEmiRs, and genes associated with DMCs to candidate loci that are associated with divergence between marine and freshwater sticklebacks [55]. We built null distributions using genes that passed the filtering step (for DEGs), genes that contained at least one confident splicing site between ecotypes (for DSGs), targets of miRNAs that passed the filtering step (for targets of DEmiRs), and genes associated with CpGs that passed the filtering step (for genes associated with DMCs).

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12915-025-02176-0.

Additional file 1: Fig. S1. Scree plots showing the eigenvalues for individual principal components based on the results of principal component analysis (PCA) for (a) Gene expression, (b) Alternative splicing events, (c) MiRNA expression and (d) DNA methylation. Fig. S2. Quadratic discriminant analysis (QDA) based on principal components (PCs) of (a) gene expression, (b) alternative splicing, (c) miRNA expression and (d) methylation levels of our filtered dataset. Fig. S3. Event histograms of 1,000 randomly-generated

datasets showing the probability of having the observed numbers of (a-b) DEGs, (c-d) DEmiRs and (e–f) DMCs in brain and liver. Fig. S4. Statistical power as a function of sample size based on simulated (a-b) gene expression and (c-d) miRNA expression datasets in brain and liver, respectively. Fig. S5. Alternative splicing patterns. Fig. S6. miRNA expression pattern. Fig. S7. (a) Proportion of DMCs that are within or proximal to CpG islands (CpGi) or open sea in brain and liver. (b-c) Manhattan plots of the chromosomal positions of methylated CpG loci that differed significantly between marine and freshwater fish in (b) brain or (c) liver. Fig. S8 Map showing the locations of the nine-spined stickleback populations used in this study.

Additional file 2: Table S1. The complete annotation of miRNAs in *Pungitius pungitius*.

Additional file 3: Table S2. Samples used in this study. Table S3. Sample information, mRNA read counts and alignments to the *Pungitius pungitius* genome. Table S4. Sample information, small RNA read counts and alignments to the *Pungitius pungitius* genome. Table S5. Sample information, WGBS read counts and alignments to the *Pungitius pungitius* genome.

Authors' contributions
J.H. and B.G. conceived the study; B.G. performed crossings and collected samples; M.L., J.Z. and J.H. analyzed the data with input from R.D.H.B.; M.L., J.Z. and J.H. wrote the manuscript with input from B.G., R.D.H.B and J.M. All authors read and approved the final manuscript.

Data availability
Raw Illumina sequencing reads for all the analyzed individuals can be downloaded from the NCBI database under accession number PRJNA992673. All codes used for data analyses in this study can be accessed on GitHub (https://github.com/hulab-fdu/nine-spined-sticklebacks).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Hager ER, Harringmeyer OS, Wooldridge TB, Theingi S, Gable JT, McFadden S, et al. A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes. Science. 2022;377(6604):399–405.
2. Barrett RD, Hoekstra HE. Molecular spandrels: tests of adaptation at the genetic level. Nat Rev Genet. 2011;12(11):767–80.

Luo *et al. BMC Biology*   (2025) 23:70

Page 16 of 18

3.   Winchell KM, Campbell-Staton SC, Losos JB, Revell LJ, Verrelli BC, Geneva AJ. Genome-wide parallelism underlies contemporary adaptation in urban lizards. Proc Natl Acad Sci U S A. 2023;120(3):e2216789120.

4.   Ferreira MS, Thurman TJ, Jones MR, Farelo L, Kumar AV, Mortimer SME, et al. The evolution of white-tailed jackrabbit camouflage in response to past and future seasonal climates. Science. 2023;379(6638):1238–42.

5.   Barrett RDH, Laurent S, Mallarino R, Pfeifer SP, Xu CCY, Foll M, et al. Linking a mutation to survival in wild mice. Science. 2019;363(6426):499–504.

6.   Romero IG, Ruvinsky I, Gilad Y. Comparative studies of gene expression and the evolution of gene regulation. Nat Rev Genet. 2012;13(7):505–16.

7.   Huang Y, Chain FJ, Panchal M, Eizaguirre C, Kalbe M, Lenz TL, et al. Transcriptome profiling of immune tissues reveals habitat-specific gene expression between lake and river sticklebacks. Mol Ecol. 2016;25(4):943–58.

8.   Gibbons TC, Metzger DC, Healy TM, Schulte PM. Gene expression plasticity in response to salinity acclimation in threespine stickleback ecotypes from different salinity habitats. Mol Ecol. 2017;26(10):2711–25.

9.   Kusakabe M, Ishikawa A, Ravinet M, Yoshida K, Makino T, Toyoda A, et al. Genetic basis for variation in salinity tolerance between stickleback ecotypes. Mol Ecol. 2017;26(1):304–19.

10. Ishikawa A, Kusakabe M, Yoshida K, Ravinet M, Makino T, Toyoda A, et al. Different contributions of local- and distant-regulatory changes to transcriptome divergence between stickleback ecotypes. Evolution. 2017;71(3):565–81.

11. Verta JP, Jones FC. Predominance of cis-regulatory changes in parallel expression divergence of sticklebacks. eLife. 2019;8:e43785.

12. Adrian-Kalchhauser I, Sultan SE, Shama LNS, Spence-Jones H, Tiso S, Keller Valsecchi CI, et al. Understanding 'non-genetic' inheritance: insights from molecular-evolutionary crosstalk. Trends Ecol Evol. 2020;35(12):1078–89.

13. Angers B, Perez M, Menicucci T, Leung C. Sources of epigenetic variation and their applications in natural populations. Evol Appl. 2020;13(6):1262–78.

14. Hu J, Wuitchik SJS, Barry TN, Jamniczky HA, Rogers SM, Barrett RDH. Heritability of DNA methylation in threespine stickleback (*Gasterosteus aculeatus*). Genetics. 2021;217(1):1–15.

15. Smith G, Smith C, Kenny JG, Chaudhuri RR, Ritchie MG. Genome-wide DNA methylation patterns in wild samples of two morphotypes of threespine stickleback (*Gasterosteus aculeatus*). Mol Biol Evol. 2015;32(4):888–95.

16. Caizergues AE, Le Luyer J, Gregoire A, Szulkin M, Senar JC, Charmantier A, et al. Epigenetics and the city: Non-parallel DNA methylation modifications across pairs of urban-forest Great tit populations. Evol Appl. 2022;15(1):149–65.

17. Kelley JL, Tobler M, Beck D, Sadler-Riggleman I, Quackenbush CR, Arias Rodriguez L, et al. Epigenetic inheritance of DNA methylation changes in fish living in hydrogen sulfide-rich springs. Proc Natl Acad Sci U S A. 2021;118(26):e2014929118.

18. Hu J, Barrett RDH. The role of plastic and evolved DNA methylation in parallel adaptation of threespine stickleback (*Gasterosteus aculeatus*). Mol Ecol. 2023;32(7):1581–91.

19. Heckwolf MJ, Meyer BS, Häsler R, Höppner MP, Eizaguirre C, Reusch TBH. Two different epigenetic information channels in wild three-spined sticklebacks are involved in salinity adaptation. Sci Adv. 2020;6(12):eaaz1138.

20. Hill MS, Vande Zande P, Wittkopp PJ. Molecular and evolutionary processes generating variation in gene expression. Nat Rev Genet. 2021;22(4):203–15.

21. Verta JP, Jacobs A. The role of alternative splicing in adaptation and evolution. Trends Ecol Evol. 2021;37(4):299–308.

22. Hanson D, Hu J, Hendry AP, Barrett RDH. Heritable gene expression differences between lake and stream stickleback include both parallel and antiparallel components. Heredity. 2017;119(5):339–48.

23. Christie MR, Marine ML, Fox SE, French RA, Blouin MS. A single generation of domestication heritably alters the expression of hundreds of genes. Nat Commun. 2016;7:10676.

24. Singh P, Ahi EP. The importance of alternative splicing in adaptive evolution. Mol Ecol. 2022;31(7):1928–38.

25. Wright CJ, Smith CWJ, Jiggins CD. Alternative splicing as a source of phenotypic diversity. Nat Rev Genet. 2022;23(11):697–710.

26. Schorr AL, Mangone M. miRNA-based regulation of alternative RNA splicing in metazoans. Int J Mol Sci. 2021;22(21):11618.

27. Kelley JL, Desvignes T, McGowan KL, Perez M, Rodriguez LA, Brown AP, et al. microRNA expression variation as a potential molecular mechanism contributing to adaptation to hydrogen sulphide. J Evol Biol. 2021;34(6):977–88.

28. Jacobs A, Elmer KR. Alternative splicing and gene expression play contrasting roles in the parallel phenotypic evolution of a salmonid fish. Mol Ecol. 2021;30(20):4955–69.

29. Howes TR, Summers BR, Kingsley DM. Dorsal spine evolution in threespine sticklebacks via a splicing change in MSX2A. BMC Biol. 2017;15(1):1–16.

30. Rastorguev SM, Nedoluzhko AV, Gruzdeva NM, Boulygina ES, Sharko FS, Ibragimova AS, et al. Differential miRNA expression in the three-spined stickleback, response to environmental changes. Sci Rep. 2017;7(1):18089.

31. Franchini P, Xiong P, Fruciano C, Meyer A. The role of microRNAs in the repeated parallel diversification of lineages of midas cichlid fish from Nicaragua. Genome Biol Evol. 2016;8(5):1543–55.

32. Franchini P, Xiong P, Fruciano C, Schneider RF, Woltering JM, Hulsey CD, et al. MicroRNA gene regulation in extremely young and parallel adaptive radiations of crater lake cichlid fish. Mol Biol Evol. 2019;36(11):2498–511.

33. Singh P, Borger C, More H, Sturmbauer C. The role of alternative splicing and differential gene expression in cichlid adaptive radiation. Genome Biol Evol. 2017;9(10):2764–81.

34. Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, et al. The genomic basis of adaptive evolution in threespine sticklebacks. Nature. 2012;484(7392):55–61.

35. Barghi N, Hermisson J, Schlotterer C. Polygenic adaptation: a unifying framework to understand positive selection. Nat Rev Genet. 2020;21(12):769–81.

36. Láruson ÁJ, Yeaman S, Lotterhos KE. The importance of genetic redundancy in evolution. Trends Ecol Evol. 2020;35(9):809–22.

37. Hu J, Barrett RDH. Epigenetics in natural animal populations. J Evol Biol. 2017;30(9):1612–32.

38. Verhoeven KJF, vonHoldt BM, Sork VL. Epigenetics in ecology and evolution: what we know and what we need to know. Mol Ecol. 2016;25:1631–8.

39. Morales HE, Faria R, Johannesson K, Larsson T, Panova M, Westram AM, et al. Genomic architecture of parallel ecological divergence: Beyond a single environmental contrast. Sci Adv. 2019;5(12):eaav9963.

40. Zou Z, Zhang J. No genome-wide protein sequence convergence for echolocation. Mol Biol Evol. 2015;32(5):1237–41.

41. Cooper KL, Sears KE, Uygur A, Maier J, Baczkowski K-S, Brosnahan M, et al. Patterning and post-patterning modes of evolutionary digit loss in mammals. Nature. 2014;511(7507):41–5.

42. Foote AD, Liu Y, Thomas GW, Vinař T, Alföldi J, Deng J, et al. Convergent evolution of the genomes of marine mammals. Nat Genet. 2015;47(3):272–5.

43. Takuno S, Ralph P, Swarts K, Elshire RJ, Glaubitz JC, Buckler ES, et al. Independent molecular basis of convergent highland adaptation in maize. Genetics. 2015;200(4):1297–312.

44. Lim MC, Witt CC, Graham CH, Dávalos LM. Parallel molecular evolution in pathways, genes, and sites in high-elevation hummingbirds revealed by comparative transcriptomics. Genome Biol Evol. 2019;11(6):1573–85.

45. Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE. Convergence in pigmentation at multiple levels: mutations, genes and function. Philos Trans R Soc Lond B Biol Sci. 2010;365(1552):2439–50.

46. Hämälä T, Ning W, Kuittinen H, Aryamanesh N, Savolainen O. Environmental response in gene expression and DNA methylation reveals factors influencing the adaptive potential of *Arabidopsis lyrata*. eLife. 2022;11:e83115.

47. Birkeland S, Gustafsson ALS, Brysting AK, Brochmann C, Nowak MD. Multiple genetic trajectories to extreme abiotic stress adaptation in Arctic Brassicaceae. Mol Biol Evol. 2020;37(7):2052–68.

48. Bohutínská M, Vlček J, Yair S, Kolář F. Genomic basis of parallel adaptation varies with divergence in *Arabidopsis* and its relatives. Proc Natl Acad Sci U S A. 2021;118(21):e2022713118.

Luo *et al. BMC Biology*   (2025) 23:70

Page 17 of 18

49. Bohutínská M, Alston M, Monnahan P, Mandáková T, Bray S, Paajanen P, et al. Novelty and Convergence in Adaptation to Whole Genome Duplication. Mol Biol Evol. 2021;38(9):3910–24.

50. Corbett-Detig R, Russell S, Nielsen R, Losos J. Phenotypic convergence is not mirrored at the protein level in a lizard adaptive radiation. Mol Biol Evol. 2020;37(6):1604–14.

51. Leung C, Grulois D, Quadrana L, Chevin LM. Phenotypic plasticity evolves at multiple biological levels in response to environmental predictability in a long-term experiment with a halotolerant microalga. PLoS Biol. 2023;21(3):e3001895.

52. Leung C, Grulois D, Chevin LM. Plasticity across levels: Relating epigenomic, transcriptomic, and phenotypic responses to osmotic stress in a halotolerant microalga. Mol Ecol. 2022;31(18):4672–87.

53. Leitwein M, Wellband K, Cayuela H, Le Luyer J, Mohns K, Withler R, et al. Strong parallel differential gene expression induced by hatchery rearing weakly associated with methylation signals in adult Coho Salmon (*O. kisutch*). Genome Biol Evol. 2022;14(4):evac036.

54. Merilä J. Nine-spined stickleback (Pungitius pungitius): an emerging model for evolutionary biology research. Ann N Y Acad Sci. 2013;1289:18–35.

55. Wang Y, Zhao Y, Wang Y, Li Z, Guo B, Merilä J. Population transcriptomics reveals weak parallel genetic basis in repeated marine and freshwater divergence in nine-spined sticklebacks. Mol Ecol. 2020;29(9):1642–56.

56. Shen S, Park JW, Lu ZX, Lin L, Henry MD, Wu YN, et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc Natl Acad Sci U S A. 2014;111(51):E5593–601.

57. Fang B, Kemppainen P, Momigliano P, Feng X, Merilä J. On the causes of geographically heterogeneous parallel evolution in sticklebacks. Nat Ecol Evol. 2020;4(8):1105–15.

58. Vernaz G, Hudson AG, Santos ME, Fischer B, Carruthers M, Shechonge AH, et al. Epigenetic divergence during early stages of speciation in an African crater lake cichlid fish. Nat Ecol Evol. 2022;6:1940–51.

59. Vernaz G, Malinsky M, Svardal H, Du M, Tyers AM, Santos ME, et al. Mapping epigenetic divergence in the massive radiation of Lake Malawi cichlid fishes. Nat Commun. 2021;12(1):5870.

60. Alvarez M, Schrey AW, Richards CL. Ten years of transcriptomics in wild populations: what have we learned about their ecology and evolution? Mol Ecol. 2015;24(4):710–25.

61. Campbell-Staton SC, Cheviron ZA, Rochette N, Catchen J, Losos JB, Edwards SV. Winter storms drive rapid phenotypic, regulatory, and genomic shifts in the green anole lizard. Science. 2017;357(6350):495.

62. Jacobs A, Carruthers M, Eckmann R, Yohannes E, Adams CE, Behrmann-Godel J, et al. Rapid niche expansion by selection on functional genomic variation after ecosystem recovery. Nat Ecol Evol. 2019;3(1):77–86.

63. Smith CCR, Tittes S, Mendieta JP, Collier-Zans E, Rowe HC, Rieseberg LH, et al. Genetics of alternative splicing evolution during sunflower domestication. Proc Natl Acad Sci U S A. 2018;115(26):6768–73.

64. Grantham ME, Brisson JA. Extensive differential splicing underlies phenotypically plastic aphid morphs. Mol Biol Evol. 2018;35(8):1934–46.

65. Healy TM, Schulte PM. Patterns of alternative splicing in response to cold acclimation in fish. J Evol Biol. 2019;222(Pt 5):jeb193516.

66. Rastorguev SM, Nedoluzhko AV, Sharko FS, Boulygina ES, Sokolov AS, Gruzdeva NM, et al. Identification of novel microRNA genes in freshwater and marine ecotypes of the three-spined stickleback (*Gasterosteus aculeatus*). Mol Ecol Resour. 2016;16(6):1491–8.

67. Artemov AV, Mugue NS, Rastorguev SM, Zhenilo S, Mazur AM, Tsygankova SV, et al. Genome-wide DNA methylation profiling reveals epigenetic adaptation of stickleback to marine and freshwater conditions. Mol Biol Evol. 2017;34(9):2203–13.

68. Anderson JA, Lin D, Lea AJ, Johnston RA, Voyles T, Akinyi MY, et al. DNA methylation signatures of early-life adversity are exposure-dependent in wild baboons. Proc Natl Acad Sci U S A. 2024;121(11):e2309469121.

69. Anastasiadi D, Esteve-Codina A, Piferrer F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. Epigenetics Chromatin. 2018;11(1):37.

70. Stanford BCM, Clake DJ, Morris MRJ, Rogers SM. The power and limitations of gene expression pathway analyses toward predicting population response to environmental stressors. Evol Appl. 2020;13(6):1166–82.

71. de Carvalho CF, Slate J, Villoutreix R, Soria-Carrasco V, Riesch R, Feder JL, et al. DNA methylation differences between stick insect ecotypes. Mol Ecol. 2023;32(24):6809–23.

72. Sun D, Layman TS, Jeong H, Chatterjee P, Grogan K, Merritt JR, et al. Genome-wide variation in DNA methylation linked to developmental stage and chromosomal suppression of recombination in white-throated sparrows. Mol Ecol. 2021;30(14):3453–67.

73. Fu TT, Sun YB, Gao W, Long CB, Yang CH, Yang XW, et al. The highest-elevation frog provides insights into mechanisms and evolution of defenses against high UV radiation. Proc Natl Acad Sci U S A. 2022;119(46):e2212406119.

74. Rougeux C, Laporte M, Gagnaire P-A, Bernatchez L. The role of genomic vs. epigenomic variation in shaping patterns of convergent transcriptomic variation across continents in a young species complex. bioRxiv. 2019.

75. Foret S, Kucharski R, Pellegrini M, Feng S, Jacobsen SE, Robinson GE, et al. DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. Proc Natl Acad Sci U S A. 2012;109(13):4968–73.

76. Lipshutz SE, Howell CR, Buechlein AM, Rusch DB, Rosvall KA, Derryberry EP. How thermal challenges change gene regulation in the songbird brain and gonad: Implications for sexual selection in our changing world. Mol Ecol. 2022;31(13):3613–26.

77. Ho W-C, Li D, Zhu Q, Zhang J. Phenotypic plasticity as a long-term memory easing readaptations to ancestral environments. Sci Adv. 2020;6(21):eaba3388.

78. Semik-Gurgul E, Pawlina-Tyszko K, Gurgul A, Szmatola T, Rybinska J, Zabek T. In search of epigenetic hallmarks of different tissues: an integrative omics study of horse liver, lung, and heart. Mamm Genome. 2024;35(4):600–20.

79. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T. Identification of tissue-specific microRNAs from mouse. Curr Biol. 2002;12(9):735–9.

80. Naftaly AS, Pau S, White MA. Long-read RNA sequencing reveals widespread sex-specific alternative splicing in threespine stickleback fish. Genome Res. 2021;31(8):1486–97.

81. Metzger DCH, Schulte PM. The DNA Methylation Landscape of Stickleback Reveals Patterns of Sex Chromosome Evolution and Effects of Environmental Salinity. Genome Biol Evol. 2018;10(3):775–85.

82. Leder EH, Cano JM, Leinonen T, O'Hara RB, Nikinmaa M, Primmer CR, et al. Female-biased expression on the X chromosome as a key step in sex chromosome evolution in threespine sticklebacks. Mol Biol Evol. 2010;27(7):1495–503.

83. Marchinko KB. Predation's role in repeated phenotypic and genetic divergence of armor in threespine stickleback. Evolution. 2009;63(1):127–38.

84. Laine VN, Sepers B, Lindner M, Gawehns F, Ruuskanen S, van Oers K. An ecologist's guide for studying DNA methylation variation in wild vertebrates. Mol Ecol Resour. 2022;23(7):1488–508.

85. Gehring NH, Roignant J-Y. Anything but ordinary – Emerging splicing mechanisms in eukaryotic gene regulation. Trends Genet. 2021;37(4):355–72.

86. Berezikov E. Evolution of microRNA diversity and regulation in animals. Nat Rev Genet. 2011;12(12):846–60.

87. Tarver JE, Sperling EA, Nailor A, Heimberg AM, Robinson JM, King BL, et al. miRNAs: small genes with big potential in metazoan phylogenetics. Mol Biol Evol. 2013;30(11):2369–82.

88. Sepers B, Mateman AC, Gawehns F, Verhoeven KJF, van Oers K. Developmental stress does not induce genome-wide DNA methylation changes in wild great tit (*Parus major*) nestlings. Mol Ecol. 2023;32(14):3960–74.

89. Strader ME, Kozal LC, Leach TS, Wong JM, Chamorro JD, Housh MJ, et al. Examining the role of DNA methylation in transcriptomic plasticity of early stage sea urchins: developmental and maternal effects in a kelp forest herbivore. Front Mar Sci. 2020;7:205.

90. Johnson KM, Sirovy KA, Kelly MW. Differential DNA methylation across environments has no effect on gene expression in the eastern oyster. J Anim Ecol. 2021;91(6):1135–47.

91. Storz JF. Causes of molecular convergence and parallelism in protein evolution. Nat Rev Genet. 2016;17(4):239–50.

Luo *et al. BMC Biology*   (2025) 23:70

Page 18 of 18

92. Rennison DJ, Peichel CL. Pleiotropy facilitates parallel adaptation in sticklebacks. Mol Ecol. 2022;31(5):1476–86.
93. Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, et al. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. Am Nat. 2016;188(4):379–97.
94. Gao Y, Chen Y, Li S, Huang X, Hu J, Bock DG, et al. Complementary genomic and epigenomic adaptation to environmental heterogeneity. Mol Ecol. 2022;31(13):3598–612.
95. Venney CJ, Cayuela H, Rougeux C, Laporte M, Mérot C, Normandeau E, et al. Genome-wide DNA methylation predicts environmentally driven life history variation in a marine fish. Evolution. 2022;77(1):186–98.
96. Steward RA, de Jong MA, Oostra V, Wheat CW. Alternative splicing in seasonal plasticity and the potential for adaptation to environmental change. Nat Commun. 2022;13(1):755.
97. Defaveri J, Merila J. Evidence for adaptive phenotypic differentiation in Baltic Sea sticklebacks. J Evol Biol. 2013;26(8):1700–15.
98. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol. 2015;109:21.29.21-21.29.29.
99. Fang B, Kemppainen P, Momigliano P, Merila J. Population Structure Limits Parallel Evolution in Sticklebacks. Mol Biol Evol. 2021;38(10):4205–21.
100. Kemppainen P, Li Z, Rastas P, Loytynoja A, Fang B, Yang J, et al. Genetic population structure constrains local adaptation in sticklebacks. Mol Ecol. 2021;30(9):1946–61.
101. Nelson JS. Salinity tolerance of brook sticklebacks, *Culaea inconstans*, freshwater ninespine sticklebacks, *Pungitius pungitius*, and freshwater fourspine sticklebacks, *Apeltes quadracus*. Can J Zool. 1968;46(4):663–7.
102. Shikano T, Herczeg G, Merilä J. Molecular sexing and population genetic inference using a sex-linked microsatellite marker in the nine-spined stickleback (*Pungitius pungitius*). BMC Res Notes. 2011;4(1):1–6.
103. Kivikoski M, Rastas P, Loytynoja A, Merila J. Automated improvement of stickleback reference genome assemblies with Lep-Anchor software. Mol Ecol Resour. 2021;21(6):2166–76.
104. Shumate A, Salzberg SL. Liftoff: accurate mapping of gene annotations. Bioinformatics. 2021;37(12):1639–43.
105. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.
106. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166–9.
107. Venables B, Ripley B. Modern Applied Statistics With S. New York, NY: Springer; 2002.
108. Coll-Costa C, Dahms C, Kemppainen P, Alexandre CM, Ribeiro F, Zanella D, et al. Parallel evolution despite low genetic diversity in three-spined sticklebacks. Proc Biol Sci. 2024;291(2020):20232617.
109. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):1–21.
110. Ching T, Huang S, Garmire LX. Power analysis and sample size estimation for RNA-Seq differential expression. RNA. 2014;20(11):1684–96.
111. Tian S, Monteiro A. A transcriptomic atlas underlying developmental plasticity of seasonal forms of Bicyclus anynana butterflies. Mol Biol Evol. 2022;39(6):msac126.
112. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. Genome Res. 2012;22(10):2008–17.
113. Li YI, Knowles DA, Humphrey J, Barbeira AN, Dickinson SP, Im HK, et al. Annotation-free quantification of RNA splicing using LeafCutter. Nat Genet. 2018;50(1):151–8.
114. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 2012;28(24):3211–7.
115. Langmead B. Aligning short sequencing reads with Bowtie. Curr Protoc Bioinformatics. 2010;32(1):11.17. 11-11.17. 14.
116. Friedländer MR, Mackowiak SD, Li N, Chen W, Rajewsky N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res. 2012;40(1):37–52.
117. Griffiths-Jones S, Grocock RJ, Van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res. 2006;34(suppl_1):D140–4.
118. Desvignes T, Batzel P, Sydes J, Eames BF, Postlethwait JH. miRNA analysis with Prost! reveals evolutionary conservation of organ-enriched

expression and post-transcriptional modifications in three-spined stickleback and zebrafish. Sci Rep. 2019;9(1):3913.
119. Braasch I, Gehrke AR, Smith JJ, Kawasaki K, Manousaki T, Pasquier J, et al. The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. Nat Genet. 2016;48(4):427–37.
120. Kim B-M, Amores A, Kang S, Ahn D-H, Kim J-H, Kim I-C, et al. Antarctic blackfin icefish genome reveals adaptations to extreme environments. Nat Ecol Evol. 2019;3(3):469–78.
121. Desvignes T, Beam MJ, Batzel P, Sydes J, Postlethwait JH. Expanding the annotation of zebrafish microRNAs based on small RNA sequencing. Gene. 2014;546(2):386–9.
122. Graham AM, Barreto FS. Novel microRNAs are associated with population divergence in transcriptional response to thermal stress in an intertidal copepod. Mol Ecol. 2019;28(3):584–99.
123. Desvignes T, Batzel P, Berezikov E, Eilbeck K, Eppig JT, McAndrews MS, et al. miRNA nomenclature: a view incorporating genetic origins, biosynthetic pathways, and sequence variants. Trends Genet. 2015;31(11):613–26.
124. Enright A, John B, Gaul U, Tuschl T, Sander C, Marks D. MicroRNA targets in Drosophila. Genome Biol. 2003;4(1):1–27.
125. Krueger F, Andrews S. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011;27(11):1571–2.
126. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Meth. 2012;9(4):357–9.
127. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet. 2008;9(6):465–76.
128. Le Luyer J, Laporte M, Beacham TD, Kaukinen KH, Withler RE, Leong JS, et al. Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. Proc Natl Acad Sci U S A. 2017;114(49):12964–9.
129. Liu Y, Siegmund KD, Laird PW, Berman BP. Bis-SNP: combined DNA methylation and SNP calling for Bisulfite-seq data. Genome Biol. 2012;13(7):R61.
130. Dahms C, Kemppainen P, Zanella LN, Zanella D, Carosi A, Merila J, et al. Cast away in the Adriatic: Low degree of parallel genetic differentiation in three-spined sticklebacks. Mol Ecol. 2022;31(4):1234–53.
131. Akalin A, Kormaksson M, Li S, Garrett-Bakelman F, Figueroa M, Melnick A, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol. 2012;13(10):R87.
132. Lea AJ, Vockley CM, Johnston RA, Del Carpio CA, Barreiro LB, Reddy TE, et al. Genome-wide quantification of the effects of DNA methylation on human gene regulation. eLife. 2018;7:e37513.
133. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13(7):484–92.
134. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci U S A. 2002;99(6):3740–5.
135. Baerwald MR, Meek MH, Stephens MR, Nagarajan RP, Goodbla AM, Tomalty KM, et al. Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. Mol Ecol. 2015;25:1785–800.
136. Akalin A, Franke V, Vlahoviček K, Mason CE, Schübeler D. genomation: a toolkit to summarize, annotate and visualize genomic intervals. Bioinformatics. 2015;31(7):1127–9.
137. Legendre P, Legendre L. Numerical ecology. Amsterdam: Elsevier; 2012.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.