ENVIRONMENTAL EPIGENETICS

Environmental Epigenetics, 2023, **9(1)**, 1–8
DOI: https://doi.org/10.1093/eep/dvad001
Advance Access Publication 23 January 2023
Research Article

Context-dependent DNA methylation signatures in animal livestock

Geetha Venkatesh[1], Sina Tönges[1], Katharina Hanna[1], Yi Long Ng[2], Rose Whelan[3], Ranja Andriantsoa[1], Annika Lingenberg[1], Suki Roy[3], Sanjanaa Nagarajan[3], Steven Fong[2], Günter Raddatz[1], Florian Böhl[3] and Frank Lyko[1,*]

[1]Division of Epigenetics, DKFZ-ZMBH Alliance, German Cancer Research Center, Im Neuenheimer Feld 580, Heidelberg 69120, Germany, [2]School of Applied Science, Republic Polytechnic, 9 Woodlands Avenue 9, Singapore 738964, Singapore, [3]Research, Development & Innovation, Evonik Operations GmbH, Rodenbacher Chaussee 4, Hanau-Wolfgang 63457, Germany
*Correspondence address. Division of Epigenetics, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 580, Heidelberg 69120, Germany. Tel: ++49-6221-423800; Fax: ++49-6221-423802; E-mail: f.lyko@dkfz.de

Abstract

DNA methylation is an important epigenetic modification that is widely conserved across animal genomes. It is widely accepted that DNA methylation patterns can change in a context-dependent manner, including in response to changing environmental parameters. However, this phenomenon has not been analyzed in animal livestock yet, where it holds major potential for biomarker development. Building on the previous identification of population-specific DNA methylation in clonal marbled crayfish, we have now generated numerous base-resolution methylomes to analyze location-specific DNA methylation patterns. We also describe the time-dependent conversion of epigenetic signatures upon transfer from one environment to another. We further demonstrate production system-specific methylation signatures in shrimp, river-specific signatures in salmon and farm-specific signatures in chicken. Together, our findings provide a detailed resource for epigenetic variation in animal livestock and suggest the possibility for origin tracing of animal products by epigenetic fingerprinting.

Key words: DNA methylation; epigenetic variation; environment; marbled crayfish; shrimps; chicken

Introduction

DNA methylation is a highly conserved epigenetic modification [1]. In animal genomes, DNA methylation is usually found in the context of CpG dinucleotides, but in complex patterns, that can consist of millions of methylation marks [2]. These methylation patterns are often conserved among closely related species but highly diverse across the animal kingdom [3] and include almost completely methylated genomes, as well as partially, sporadically and unmethylated genomes. Factors that shape animal methylomes include genetic, developmental and environmental cues.

The association of defined DNA methylation signatures with specific environmental factors remains a surprisingly controversial field of research [4, 5]. Despite a large number of published studies, effect sizes are often small and poorly supported by statistical analysis [4]. Additional limitations include the use of artifact-prone methodology, such as antibody-based techniques [6], and confounding effects by genetic polymorphisms. It is therefore of great importance to investigate environmental epigenetic variation using robust methodology and approaches that limit the influences of confounding factors [4, 5].

Marbled crayfish (*Procambarus virginalis*) are a novel freshwater crayfish species [7] and are currently being evaluated as aquaculture livestock [8]. They have a conserved and active DNA methylation system [9] and a clonal genome [10], which eliminate confounding genetic polymorphisms. The analysis of single-base resolution methylomes identified a variable portion of the marbled crayfish methylome that could be used to define population-specific DNA methylation patterns [11]. These epigenetic signatures were validated with independent methodology and with independent samples that were collected 1–2 years later [11]. Together, the findings suggested the existence of location-specific DNA methylation signatures that may reflect the adaptation of the clonal marbled crayfish genome to different environments.

We have now expanded our analyses by generating and analyzing high-quality base-resolution methylomes from numerous additional populations and animal livestock. We also describe the time-dependent conversion of epigenetic signatures upon transfer from one environment to another. We further demonstrate the existence of production system-specific DNA methylation signatures in shrimp, river-specific signatures in salmon and farm-specific signatures in chicken. Together, our findings

Received 26 July 2022; **revised** 6 January 2023; **accepted** 23 January 2023

provide compelling evidence for context-dependent epigenetic variation across animal livestock. Furthermore, our study provides the resources and the conceptual foundation for origin tracing of animal products by epigenetic fingerprinting.

Results

To determine context-dependent DNA methylation patterns in marbled crayfish, Pacific salmon, Pacific whiteleg shrimp and chicken, we acquired published datasets (Table 1) and generated ~5.7 TB of additional sequencing data (Table 2). For chicken and salmon, >80% of the sequencing reads aligned to their respective genomes. Lower mapping rates were observed for marbled crayfish and shrimp, which is likely related to the highly fragmented genome assemblies of these species. An overview of the number of detected CpGs for each of the species is provided in Table S1. After filtering for single nucleotide polymorphisms (SNPs), an average of 1 859 250 (SD = 571 012) and 1 650 571

Table 1: Samples analyzed in this study

Animal	Designation	Description	Number of samples
Marbled crayfish	Madagascar 1	Rice field channel, −18.789167, 48.246067	1
	Madagascar 2	Anjingilo rice field with thermal water [22]	5
	Madagascar 3	Polluted Ihosy river [11]	1
	Madagascar 4	Polluted Ranomaimbo urban lake [22]	1
	Germany 1	Oligotrophic lake [8]	1
	Germany 2	Eutrophic lake [11], Reilingen	5
	Germany 3	Oligotrophic lake [11], Singlis	4
	Germany 4	Eutrophic lake [23]	1
	Ukraine	Flooded quarry [23]	1
	Malta	Pond [23]	4
	Laboratory	DKFZ Heidelberg [11]	1
	Singapore	Aquaculture tank	15
Pacific whiteleg shrimp	Producer 1	Aquaculture tank, climate-controlled facility, Northern Germany	5
	Producer 2	Biological RAS system, climate-controlled facility, Northern Germany	5
	Producer 3	Biofloc RAS system, climate-controlled facility, Central Germany	5
	Producer 4	Aquaculture tank, open facility, Singapore	6
Pacific salmon	Capilano river, Canada	Rapid mountain stream [12]	20
	Quinsam river, Canada	Slow-flowing river [12]	19
Chicken	Australia	Climate-controlled experimental facility, wheat-corn-soybean feed	6
	The USA	Climate-controlled experimental facility, corn-soybean feed	6
	Finland	Climate-controlled experimental barn, wheat-soybean feed	6
	Germany	Climate-controlled experimental barn, corn-wheat-soybean-rapeseed feed	6
	Iran	Commercial barn, corn-soybean feed	6
	Thailand	Negative-pressure research facility, corn-soybean-canola-rice bran feed	6

Table 2: Sequencing results

Animal	Location	Number of samples	Sequencing protocol	Yield (Gbp)	Mapped (%)	Conversion (%)
Marbled crayfish	Madagascar 1	1	WGBS	138	36	99
	Madagascar 2	5	WGBS	758	53	99
	Madagascar 3	1	WGBS	144	51	99
	Madagascar 4	1	WGBS	140	55	99
	Germany 1	1	WGBS	142	52	99
	Germany 2	5	WGBS	762	53	99
	Germany 3	4	WGBS	552	45	99
	Germany 4	1	WGBS	132	57	99
	Ukraine	1	WGBS	136	52	100
	Malta	4	WGBS	594	50	99
	Laboratory	1	WGBS	144	56	99
	Singapore	15	SGBS	66	63	99
Pacific whiteleg shrimp	Producer 1	5	WGBS	364	30	99
	Producer 2	5	WGBS	336	28	99
	Producer 3	5	WGBS	306	28	99
	Producer 4	6	WGBS	868	31	99
Pacific salmon	Capilano river, Canada	20	RRBS	61	83	99
	Quinsam river, Canada	19	RRBS	52	87	99
Chicken	Australia	6	RRBS	10	87	98
	The USA	6	RRBS	8	89	99
	Finland	6	RRBS	17	88	99
	Germany	6	RRBS	14	90	99
	Iran	6	RRBS	9	89	99
	Thailand	6	RRBS	18	87	99

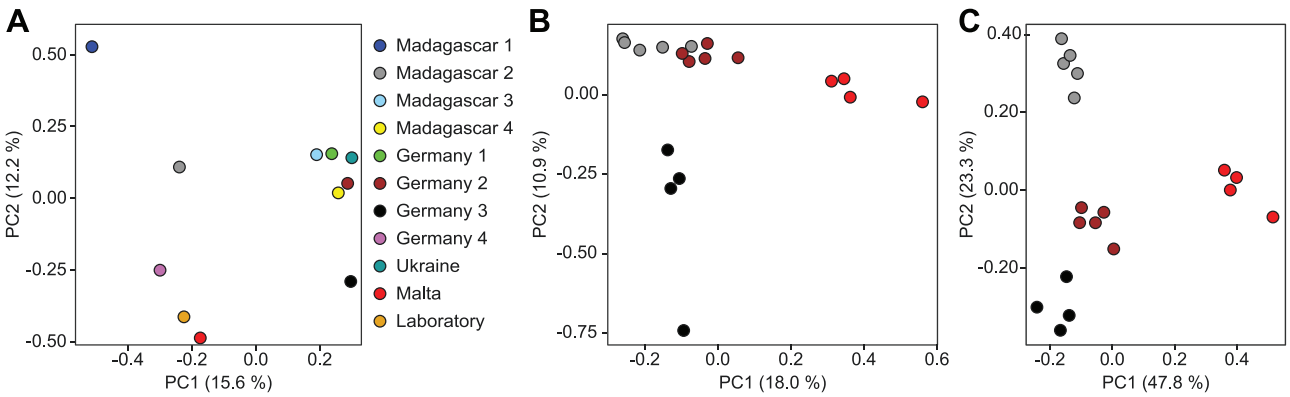

Figure 1: Location-specific epigenetic signatures in marbled crayfish. (A) PCA of abdominal muscle samples from multiple populations (single replicates) based on the methylation levels of the 2000 most variably methylated windows. Madagascar 1: Moramanga, Madagascar 2: Anjingilo, Madagascar 3: Ihosy, Madagascar 4: Ranomaimbo, Germany 1: Murner See, Germany 2: Reilinger See, Germany 3: Singliser See, Germany 4: Moosweiher. (B) PCA of abdominal muscle samples from four populations with multiple replicates based on the methylation levels of the 2000 most variably methylated windows. (C) PCA of four populations with multiple replicates based on the methylation levels of 101 significantly differentially methylated regions.

(SD = 104 022) CpGs per sample were retained for chicken and salmon, respectively. In case of marbled crayfish and shrimp, 1 075 790 and 148 302 CpGs, respectively, were retained after filtering for completely methylated and unmethylated CpGs. The observed variation in the number of retained CpGs is likely related to differences in sequencing depths between reduced representation bisulfite sequencing (RRBS)-based analyses (chicken and salmon) and whole-genome bisulfite sequencing (WGBS)-based analyses (crayfish and shrimp).

Location-specific Methylation Patterns in Marbled Crayfish

Our previous identification of location-specific DNA methylation signatures was based on an analysis of four populations, using sub-genome capture bisulfite sequencing (SGBS) of a variable portion of the marbled crayfish genome [11]. To increase the breadth of this initial analysis, we now collected specimens from 11 different stable wild populations (Table 1). Subsequent WGBS (Table 2) of abdominal muscle samples generated representative single-base resolution maps for every population. Data analysis identified variably methylated 1-kb windows by considering the average methylation variance across CpGs for each window. Windows were then ranked based on their methylation variability, and the top 2000 windows (containing 11 194 CpG sites) were used for a principal component analysis (PCA), revealing a robust separation of individual populations (Fig. 1A). In a follow-up analysis, we used WGBS of abdominal muscle samples to analyze multiple ($N = 4$–5) replicates from four populations (Table 2). Subsequent PCA based on the top 2000 variably methylated windows (11 525 CpG sites) showed robust separation of individual populations (Fig. 1B). Finally, we also performed a differentially methylated region (DMR) analysis using dmrseq for the dataset with multiple replicates. This identified 101 significantly differentially methylated regions ($P < 0.05$; methylation difference >15%) that clearly separated the four locations (Fig. 1C). Together, these findings further confirm and expand the existence of location-specific DNA methylation patterns in marbled crayfish.

Time-dependent Conversion of Environmental Signatures in Marbled Crayfish

To further illustrate the impact of the environment on the methylome of the marbled crayfish, we transferred 50 animals from an indoor aquarium to a closed tropical aquaculture system. Abdominal muscle samples were then collected at different time points (0, 1, 4 and 13 months), and methylation patterns were analyzed using previously established SGBS assay for variably methylated genes [11]. Average methylation was calculated for 1-kb windows, and a Kruskal–Wallis test was applied to identify the differentially methylated windows. This led to the identification of 235 windows with significant ($P < 0.05$, methylation difference >15%) time-dependent methylation changes that allowed a robust separation of the different time points (Fig. 2). These findings demonstrate that changes in environmental parameters affect epigenetic signatures over time.

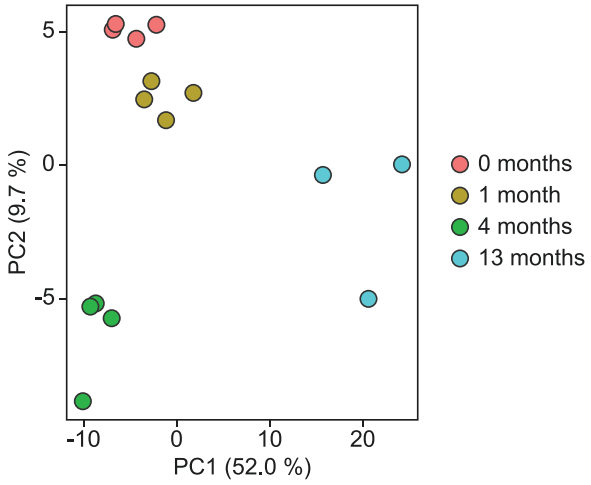

Figure 2: Time-dependent conversion of epigenetic signatures in marbled crayfish. PCA of abdominal muscle samples from four different time points (0, 1, 4 and 13 months) based on the methylation levels of 235 differentially methylated windows

Production System-specific Methylation Signatures in Pacific Whiteleg Shrimp

To confirm the findings from marbled crayfish in a commercially established aquaculture livestock, we analyzed methylation patterns in Pacific whiteleg shrimp (*Litopenaeus vannamei*). We collected multiple (5–6) animals from four independent production

systems (Table 1) and used WGBS (Table 2) to determine their methylation patterns. After filtering of possible genetic polymorphisms, we identified genomic regions that were differentially methylated among the four sources (see Material and Methods for details). A PCA based on 42 significantly differentially methylated regions ($P < 0.05$ and methylation difference >15%) showed robust separation between the different populations (Fig. 3A). Interestingly, animals from two similar, but not identical RAS systems (Table 1) appeared to be poorly separated based on Principal Components 1 and 2 (PC1 and PC2) (Fig. 3A) but could be clearly separated through a second-level analysis based on PC2 and Principal Component 3 (PC3) (Fig. 3B). Our results thus provide further evidence for environment-specific DNA methylation signatures in animals and their potential to discriminate their source.

Context-dependent Methylation Signatures in Pacific Salmon

A previous study [12] suggested that the rearing environment can have an effect on the DNA methylation pattern of white muscle tissue from Pacific salmon (*Oncorhynchus kisutch*). Notably, this study contained RRBS datasets from two distinct rivers, Capilano river and Quinsam river (Table 1). We therefore re-analyzed the published dataset (Table 2), using a logistic regression, with sex and location as covariates. This identified 45 DMRs ($P < 0.05$ and methylation difference >15%) between the two populations. A PCA based on these DMRs allowed a visible separation between the two rearing environments (Fig. 4A), which confirmed the published findings. In subsequent steps, we repeated the analysis using sex and rearing environment as covariates. This identified 193 DMRs ($P < 0.05$ and methylation difference >15%) between the two rivers. A PCA based on these DMRs allowed a very clear separation between the two locations (Fig. 4B), which suggests the presence of location-dependent methylation patterns in an additional aquatic livestock for meat production.

Farm-specific Methylation Signatures in Chicken

Chicken (*Gallus gallus domesticus*) are another important livestock for meat production. In order to analyze whether chicken also show location-dependent methylation signatures, we sequenced samples from six different facilities (Table 1), each with six replicates using RRBS (Table 2). We performed site-specific differential methylation analysis using a logistic regression and found 8215 CpG sites to be significantly differentially methylated between the locations ($P < 0.05$ and methylation difference >25%). A PCA based on these differentially methylated sites allowed a robust separation between the different locations (Fig. 5A). Since the samples from Germany and Finland were clustered closely in the PCA, we performed a separate analysis focusing only on these two locations to further investigate the location-specific methylation

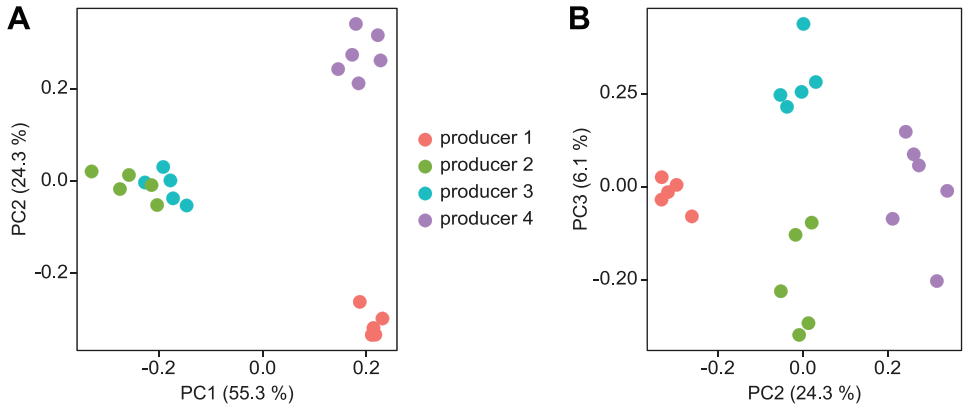

Figure 3: Producer-specific epigenetic signatures in shrimp. (A) PCA of samples from four different producers based on the methylation levels of 42 significantly differentially methylated regions. (B) PC2 and PC3 allow the separation of all four producers

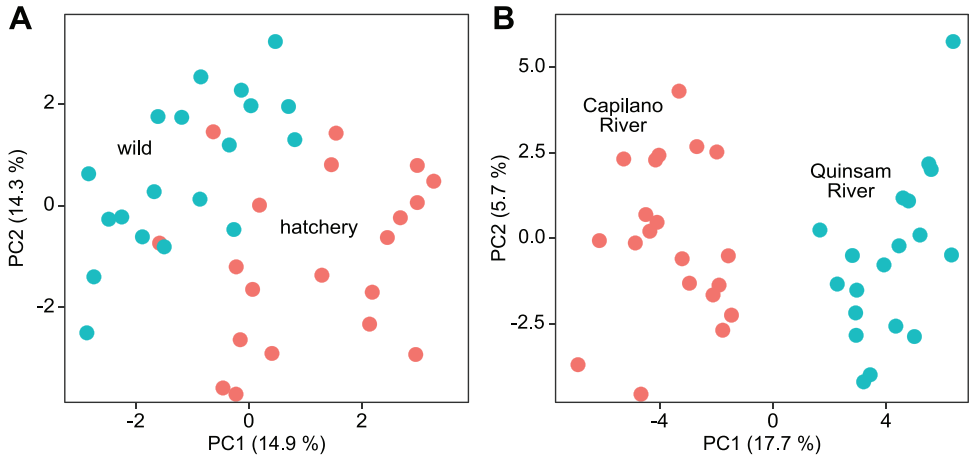

Figure 4: Context-dependent epigenetic signatures in Pacific salmon. (A) PCA of samples from two different rearing environments based on the methylation levels of 45 significantly differentially methylated regions. (B) PCA of samples from two different rivers based on the methylation levels of 193 significantly differentially methylated regions

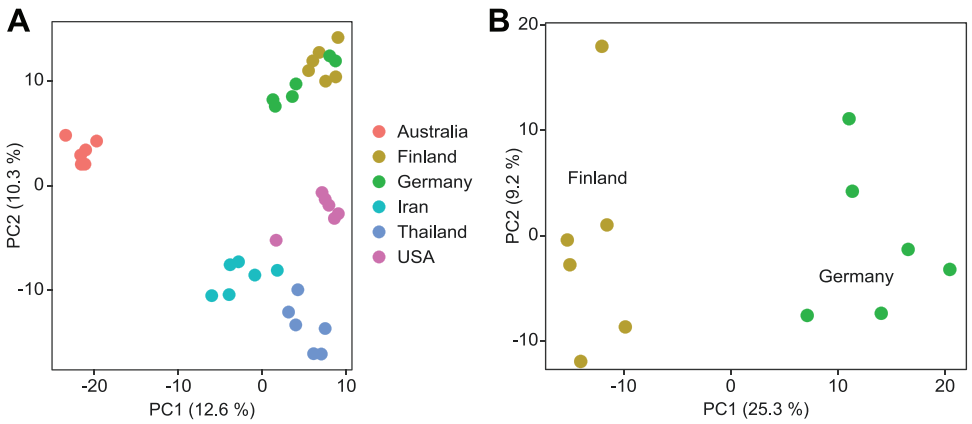

Figure 5: Farm-dependent epigenetic signatures in chicken. (A) PCA of samples from six different chicken farms based on the methylation levels of 8215 CpG sites with farm-specific methylation differences. (B) PCA of samples from Finland and Germany based on the methylation levels of 5987 CpG sites with location-specific methylation differences in a pair-wise analysis

patterns between these two locations. This led to the identification of 5987 significantly differentially methylated CpG sites between the two locations ($P < 0.05$ and methylation difference >15%), and a PCA based on these sites clearly separated the two populations (Fig. 5B). These findings provide an additional example for environment-specific methylation signatures in a land-based livestock for meat production.

Effects and Features of Environment-specific Methylation Signatures

Among the livestock analyzed in this study, chicken are distinguished by a substantially higher level of genome annotation and epigenome characterization. We therefore analyzed whether genes that contain variably methylated CpGs are associated with certain biological functions and pathways. Interestingly, this revealed that the five most highly enriched gene ontologies were related to transcriptional regulation and cellular differentiation (Fig. 6A), consistent with the known roles of DNA methylation in vertebrate genomes. As methylation-dependent modulation of transcription has been shown to be associated with altered promoter methylation, we also investigated the ability of variably methylated promoter-associated CpGs to separate the six different farms. Indeed, a PCA based on the 6939 most variably methylated promoter-associated CpGs demonstrated a rudimentary separation (Fig. 6B), suggesting that promoter methylation dynamics can contribute to environmental methylation signatures. Finally, as methylation variation is often associated with transposable elements (TEs), we also investigated whether variably methylated CpGs are associated with TEs in chicken. The results showed a moderate, but significant ($P = 0.002$, chi-square test) enrichment (Fig. 6C), indicating that methylation variation of TEs can also contribute to environmental methylation signatures. Taken together, these findings provide first insights into the mechanisms that underpin environment-specific methylation signatures in animal livestock.

Discussion

Environmental epigenetic variation has long been postulated [13, 14], but conclusive experimental evidence in support of it remains surprisingly scarce. This is largely attributable to limitations in study design and methodology, resulting in lack of statistical power and/or confounding from a multitude of factors, including genetic variation. Based on our previous findings in marbled crayfish [11], we have now performed a systematic analysis of epigenetic variation in a diverse group of animal livestock, which are characterized by relatively high genetic homogeneity. We also used stringent, sequencing-based methodology and a rigorous analytical pipeline that further removed genetic polymorphisms and other potential confounding factors.

Our results expand previous findings for marbled crayfish [11] by demonstrating location-specific methylation signatures on the genome-wide level and in multiple independent replicates. In addition, we replicate the original findings from marbled crayfish in shrimps from four independent producers, in Pacific salmon from two independent rivers and in chicken flocks from six independent farms. Of particular interest were findings from shrimp and chicken that suggested a close relationship of methylation signatures from similar (artificial) environments that could, however, be resolved by second-level analyses. It was also interesting to note that the river environment had a stronger effect than the rearing environment on the methylation pattern of Pacific salmon.

For the first time, we identified location-specific DMRs that show robust separation of different populations. DMRs are larger genomic regions that are defined by differential methylation between samples from different populations. Interestingly, DMRs were only found in crayfish, shrimps and salmon, while differential methylation was confined to specific individual CpGs in chicken. This suggests that in some species, location-specific DNA methylation patterns are pronounced at the CpG level than at the DMR level. Furthermore, marbled crayfish that were transferred from a climate-controlled aquarium to tropical aquaculture system showed time-dependent changes in their DNA methylation patterns over a period of time. This observation conclusively demonstrates that DNA methylation signatures are affected by environmental parameters and not by other factors.

Making use of the relatively advanced annotation of the chicken (epi)genome, our study also provides first insights into the effects and the features of environment-related methylation signatures. Interestingly, our results show that differentially methylated CpGs are associated with genes involved in transcriptional regulation and cellular differentiation, consistent with known data about DNA methylation function in mammals [15, 16]. We

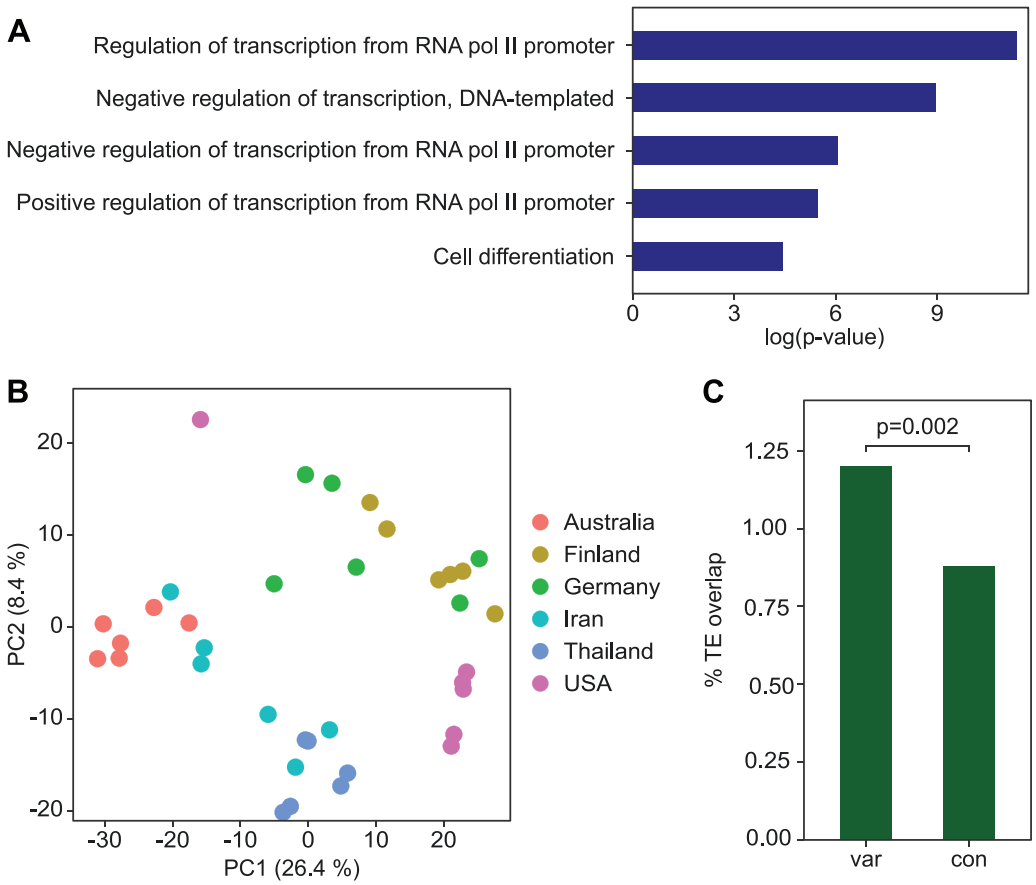

Figure 6: Effects and features of environment-specific methylation signatures. (A) Gene ontology analysis showing the top five significantly enriched biological processes associated with genes containing differentially methylated CpGs. (B) PCA of chicken samples from six farms based on the methylation levels of 6939 differentially methylated CpGs in the promoter regions. (C) Percentage of differentially methylated CpGs (var) and non-differentially methylated CpGs (con) overlapping with TEs in the chicken genome

have also shown previously that the chicken methylome is dynamically methylated at transcription factor binding sites [17], which provides an explanation for the observed promoter methylation variability. Finally, we also observed a moderate, but significant enrichment of TEs in the regions surrounding differentially methylated CpGs. This is similar to previous observations in marbled crayfish [11] and consistent with the known association between TEs and epigenetic variation [18].

Our results also provide a foundation for origin tracing of animal products by epigenetic fingerprinting. The increasing globalization of the food market requires reliable strategies to verify the identity of food and to detect food and feed tampering, which is often used to suggest a higher product quality and safety. However, it is also very difficult to detect analytically, and technological options for retrospective testing are still highly limited. A well-known example is provided by DNA fingerprinting, which allows detection of breed-specific genetic polymorphisms. However, the result simply traces the product to specific breeds and does not provide any information about the environment where the animal was reared. This issue is addressed by trace element fingerprinting [19], which depends on the detection of environment-specific trace element combinations by mass spectrometry, e.g. Ba/Ca, Cd/Ca and Cu/Ca ratios. However, the method has important limitations, including restriction to mineralized samples, such as bivalve shells and fish otoliths, and substantial weather sensitivity, which necessitates frequent (within months) recalibration.

We discovered that animals from different locations showed distinct epigenetic fingerprints. This observation provides the foundation for a novel approach that allows the proof of identity and origin of food. As this information reflects the unique environmental exposures of the animals and is "written" into the epigenome of the animals, it cannot be tampered with. Also, we have previously shown that location-specific epigenetic signatures in marbled crayfish are similar in samples that were collected 1–2 years apart (Tönges *et al.*, 2021b), suggesting considerable stability. For its technical implementation, epigenetic fingerprinting would greatly benefit from the availability of DNA methylation arrays, which provide robust, cost-efficient high-throughput solutions for the analysis of DNA methylation patterns. Such arrays are widely used for the analysis of human DNA and have recently been adapted for the analysis of other mammalian species, including mammalian livestock species [20, 21]. However, they are currently limited to the detection of certain methylation marks that are conserved among mammals. To fully realize the real potential of epigenetic fingerprinting, specific arrays will have to be developed for individual livestock species.

Material and Methods

Sample Collection

Marbled crayfish (*P. virginalis*) were collected from wild populations and the laboratory, as described previously [8, 11, 22, 23].

Samples from Singapore were collected between September 2020 and October 2021 from an aquaculture tank at Republic Polytechnic. All sampled marbled crayfish were adults with comparable sizes. Pacific whiteleg shrimps (*L. vannamei*) were collected from different producers, as specified in Table 1. All sampled shrimps were adults with comparable sizes. Chicken (*G. gallus domesticus*) samples were obtained from different experimental or commercial sources, as specified in Table 1. All chicken samples were obtained from birds of comparable ages (21–42 days), and age distribution did not vary systematically between groups.

DNA Extraction

For marbled crayfish and shrimp, genomic DNA was extracted from abdominal muscle using a TissueRuptor (Qiagen), followed by proteinase K digestion and isopropanol precipitation. Chicken DNA was extracted from breast muscle using the PureLink Genomic DNA Isolation Minikit (Invitrogen). The quality and quantity of isolated genomic DNA were assessed on a NanoDrop (Thermo Scientific) and/or TapeStation (Agilent).

Library Preparation and Sequencing

WGBS libraries were prepared using the ACCEL-NGS METHYL-SEQ DNA Library Kit (Swift Biosciences) and the EpiTect Bisulfite Kit (Qiagen) for bisulfite conversion. Following adapter ligation, six cycles of library amplification were performed. Paired-end (150 bp) sequencing of the resulting libraries was carried out on an Illumina HiSeq platform. Due to the discontinuation of the HiSeq platform, the libraries for Pacific whiteleg shrimp, Producer 4, were sequenced on an Illumina NovaSeq platform. RRBS libraries were prepared for chicken samples using the Zymo-Seq RRBS Library Kit following the manufacturer's instructions. Single-end (50 bp) sequencing of the libraries was carried out on an Illumina HiSeq platform. Sub-genome capture bisulfite sequencing (SGBS) libraries for marbled crayfish samples were prepared as described before [11] using the SureSelectXT Methyl-Seq Target Enrichment System for Illumina Multiplexed Sequencing Protocol, Version D0, July 2015. Paired-end (100 bp) sequencing of the resulting libraries was performed on an Illumina HiSeq platform. In parallel, published datasets for marbled crayfish (http://www.ncbi.nlm.nih.gov/geo, accession number GSE112411) and Pacific salmon (accession number PRJNA389610) were downloaded and integrated into the analysis.

Sequencing Data Processing and Methylation Calling

An overview of the computational workflow is shown in Fig. S1. Briefly, all reads were quality trimmed using Trimmomatic v.0.38 [24] and mapped to the respective genome assemblies using BSMAP v.2.90 [25]. For WGBS data, duplicate reads were filtered out using Picard. Methylation ratio for each CpG site was calculated using the Python script (methratio.py) distributed with the BSMAP package. Only CpGs with a minimum coverage of 10 per strand were considered for further analyses. For RRBS data analysis, the mapping parameter -n 1 was used, as the libraries prepared with Zymo-Seq RRBS library kit are non-directional. Bisulfite conversion rates (Table 2) were calculated as the ratio of unmethylated Cs and the sum of the methylated and unmethylated Cs, using the methylation rates of the mitochondrial genome (WGBS) or non-CpG methylation rates (SGBS and RRBS).

SNP Filtering

To exclude the possibility that C-T polymorphisms would be called as methylation changes, the CpGs that overlapped with C-T SNPs were filtered out. Furthermore, to minimize false methylation calls related to population-level genetic variation in organisms without systematic SNP information (marbled crayfish and shrimp), the CpGs that were completely methylated or completely unmethylated (i.e. average methylation across the samples >0.8 or <0.2) were filtered out. For Pacific salmon, SNPs were identified using whole-genome resequencing data of 20 salmon that were downloaded from the SRA database (accession number PRJNA401427). Variant calling was performed using the Bayesian genetic variant detector, FreeBayes with a minimum mapping quality of 30 and a minimum base quality of 20. Further filtering was based on a minimum read depth of 15, a minimum quality score of 30, a minor allele frequency of 0.5 and a maximum missingness of 0.9 using VCFtools v.0.1.12. The resulting C-T polymorphisms were used as the reference SNPs. For chicken, SNPs in the dbSNP database were used as the reference.

Identification of DMRs

To identify variably methylated regions in crayfish and shrimps, only CpG sites with a coverage of ≥12 were retained. The marbled crayfish scaffolds were divided into 1-kb non-overlapping windows, and average methylation variance across CpGs was calculated for each 1-kb window. Windows were then ranked based on their methylation variability, and the top 2000 windows with at least three CpG sites per window were selected. The PCA was carried out using the prcomp function in R, and the PCA plots were created using the ggfortify R package. Furthermore, the multiple replicates were subjected to DMR analysis using dmrseq [26] with a q-value cutoff of 0.05 and maxPerms = 3 to identify DMRs.

Analysis of Time-dependent Methylation Patterns in Marbled Crayfish

Average methylation was calculated for each 1-kb window, and those windows consisting of at least three CpGs were retained for differential methylation analysis. A Kruskal–Wallis test was used to identify the methylation patterns that were associated with time. The PCA was carried out on significantly differentially methylated windows ($P < 0.05$ and methylation difference >15%).

Identification of DMRs in Pacific Salmon

Methylation levels were computed across the samples using tiling windows approach (window size = 1000, step size = 1000), and DMR analysis was performed using methylKit R package [27]. To identify the rearing environment-specific DMRs, a logistic regression with sex and river of origin as covariates was applied with the "*calculateDiffMeth*" function implemented in the methylKit R package. The DMRs with at least 15% of methylation difference, q-value < 0.05 and containing at least three CpGs were retained. To identify location-specific DMRs, the analysis was repeated with sex and rearing environment as covariates. The PCA was carried out on significant DMRs.

Identification of Differentially Methylated Sites in Chicken

The CpGs that were associated with sex chromosomes were removed from the subsequent analysis. Location-specific differentially methylated CpGs were identified with the "*calculateDiffMeth*" function implemented in the methylKit R package [27]. The CpGs with at least 25% of methylation difference and q-value < 0.05 were retained as differentially methylated CpGs. The PCA was carried out on significantly differentially methylated sites. Furthermore, a pair-wise analysis was conducted for the locations that were

clustered closely in the PCA, using the same function "*calculateDiffMeth*." To identify variably methylated CpGs in promoters, variance in methylation levels was computed for all the CpGs located in promoter regions and the two most variably methylated CpGs were selected per promoter. Promoters were defined as regions 1-kb upstream of transcription start sites.

Gene Ontology Analysis and TE Enrichment

To identify the biological functions associated with genes that contain differentially methylated CpGs, a gene ontology analysis was performed using DAVID [28]. The bar plots were generated using the geom_bar function of ggplot2 in R. Enrichment for TEs among differentially methylated CpGs in chicken was analyzed by identifying TE annotation overlapping with 1-kb window upstream and downstream of differentially methylated CpGs. The *P*-value was calculated using a chi-square test.

Acknowledgements

We thank Walter Pfefferle and Heiko Bender for their support.

Author contributions

Geetha Venkatesh (Formal analysis, Data curation, Writing—original draft), Sina Tönges (Methodology, Data curation, Project administration), Katharina Hanna (Investigation), Yi Long Ng (Investigation), Rose Whelan (Project administration, Resources), Ranja Andriantsoa (Investigation), Annika Lingenberg (Investigation), Suki Roy (Formal analysis), Sanjanaa Nagarajan (Formal analysis), Steven Fong (Supervision, Data curation), Günter Raddatz (Formal analysis, Data curation), Florian Böhl (Project administration, Resources), Frank Lyko (Project administration, Supervision, Writing—review & editing).

Data availability

Sequencing datasets are available in the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE210295.

Supplementary Data

Supplementary data are available at *EnvEpig* online.

Conflict of interest statement. R.W., S.R., S.N., and F.B. are employees of Evonik Operations GmbH. F.L. received consultation fees from Evonik Operations GmbH.

References

1. Schubeler D. Function and information content of DNA methylation. *Nature* 2015;**517**:321–6.
2. Breiling A, Lyko F. Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. *Epigenetics Chromatin* 2015;**8**:24.
3. de Mendoza A, Lister R, Bogdanovic O *et al.* Evolution of DNA methylome diversity in eukaryotes. *J Mol Biol* 2020;**432**:1687–705.
4. Lea AJ, Vilgalys TP, Durst PAP *et al.* Maximizing ecological and evolutionary insight in bisulfite sequencing data sets. *Nat Ecol Evol* 2017;**1**:1074–83.
5. Carneiro VC, Lyko F. Rapid epigenetic adaptation in animals and its role in invasiveness. *Integr Comp Biol* 2020;**60**:267–74.
6. Lentini A, Lagerwall C, Vikingsson S *et al.* A reassessment of DNA-immunoprecipitation-based genomic profiling. *Nat Methods* 2018;**15**:499–504.
7. Lyko F. The marbled crayfish (Decapoda: Cambaridae) represents an independent new species. *Zootaxa* 2017;**4363**:544–52.
8. Tönges S, Masagounder K, Lenich F *et al.* Evaluating invasive marbled crayfish as a potential livestock for sustainable aquaculture. *Front Ecol Evol* 2021;**9**:651981.
9. Gatzmann F, Falckenhayn C, Gutekunst J *et al.* The methylome of the marbled crayfish links gene body methylation to stable expression of poorly accessible genes. *Epigenetics Chromatin* 2018;**11**:57.
10. Gutekunst J, Andriantsoa R, Falckenhayn C *et al.* Clonal genome evolution and rapid invasive spread of the marbled crayfish. *Nat Ecol Evol* 2018;**2**:567–73.
11. Tönges S, Venkatesh G, Andriantsoa R *et al.* Location-dependent DNA methylation signatures in a clonal invasive crayfish. *Front Cell Dev Biol* 2021;**9**:794506.
12. Le Luyer J, Laporte M, Beacham TD *et al.* Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. *Proc Natl Acad Sci USA* 2017;**114**:12964–9.
13. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 2003;**33**:245–54.
14. Feil R, Fraga MF. Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet* 2012;**13**:97–109.
15. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012;**13**:484–92.
16. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet* 2013;**14**:204–20.
17. Raddatz G, Arsenault RJ, Aylward B *et al.* A chicken DNA methylation clock for the prediction of broiler health. *Commun Biol* 2021;**4**:76.
18. Underwood CJ, Henderson IR, Martienssen RA. Genetic and epigenetic variation of transposable elements in *Arabidopsis*. *Curr Opin Plant Biol* 2017;**36**:135–41.
19. Ricardo F, Pimentel T, Genio L *et al.* Spatio-temporal variability of trace elements fingerprints in cockle (*Cerastoderma edule*) shells and its relevance for tracing geographic origin. *Sci Rep* 2017;**7**:3475.
20. Caulton A, Dodds KG, McRae KM *et al.* Development of epigenetic clocks for key ruminant species. *Genes (Basel)* 2021;**13**:96.
21. Arneson A, Haghani A, Thompson MJ *et al.* A mammalian methylation array for profiling methylation levels at conserved sequences. *Nat Commun* 2022;**13**:783.
22. Andriantsoa R, Tonges S, Panteleit J *et al.* Ecological plasticity and commercial impact of invasive marbled crayfish populations in Madagascar. *BMC Ecol* 2019;**19**:8.
23. Maiakovska O, Andriantsoa R, Tonges S *et al.* Genome analysis of the monoclonal marbled crayfish reveals genetic separation over a short evolutionary timescale. *Commun Biol* 2021;**4**:74.
24. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**:2114–20.
25. Xi Y, Li W. BSMAP: whole genome bisulfite sequence mapping program. *BMC Bioinform* 2009;**10**:232.
26. Korthauer K, Chakraborty S, Benjamini Y *et al.* Detection and accurate false discovery rate control of differentially methylated regions from whole genome bisulfite sequencing. *Biostatistics* 2019;**20**:367–83.
27. Akalin A, Kormaksson M, Li S *et al.* methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol* 2012;**13**:R87.
28. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;**4**:44–57.