DTU Library

Genome-wide methylation in the panmictic European eel (Anguilla anguilla)

Liu, Shenglin; Tengstedt, Aja Noersgaard Buur; jacobsen, Lars Magnus W.; Pujolar, Jose Martin; Jónsson, Bjarni; Lobón-Cervià, Javier; Bernatchez, Louis; Hansen, Michael Møller

[Link back to DTU Orbit](https://orbit.dtu.dk)

Genome-wide methylation in the panmictic European eel (Anguilla anguilla)

Genome-wide methylation in the panmictic European eel (*Anguilla anguilla*)

Shenglin Liu[1][#], Aja Noersgaard Buur Tengstedt[1], Magnus W. Jacobsen[2], Jose Martin Pujolar[3], Bjarni Jónsson[4][§], Javier Lobón-Cervià[5], Louis Bernatchez[6], Michael M. Hansen[1][#]

[1] Department of Biology, Aarhus University, Aarhus, Denmark

[2] Section for Marine Living Resources, National Institute of Aquatic Resources, Technical University of Denmark, Silkeborg, Denmark

[3] Centre for Gelatinous Plankton Ecology and Evolution, National Institute of Aquatic Resources, Technical University of Denmark, Kongens Lyngby, Denmark

[4] North West Iceland Nature Center, Saudárkrókur, Iceland

[5] National Museum of Natural Sciences (CSIC), Madrid, Spain

[6] IBIS (Institut de Biologie Intégrative et des Systèmes), Université Laval, Québec, Canada

[#] Corresponding authors.

Shenglin Liu, Department of Biology, Aarhus University, Ny Munkegade 114-116, DK-8000 Aarhus C, Denmark. E-mail liushenglin1222@gmail.com

Michael M. Hansen, Department of Biology, Aarhus University, Ny Munkegade 114-116, DK-8000 Aarhus C, Denmark. E-mail mmh@bio.au.dk

[§] Current address: The Icelandic Parliament, Reykjavík, Iceland

31 **Abstract**
32 The role of methylation in adaptive, developmental and speciation processes has attracted
33 considerable interest, but interpretation of results is complicated by diffuse boundaries
34 between genetic and non-genetic variation. We studied whole genome genetic and
35 methylation variation in the European eel, distributed from subarctic to subtropical
36 environments, but with panmixia precluding genetically based local adaptation beyond
37 single-generation responses. Overall methylation was 70.9%, with hypomethylation
38 predominantly found in promoters and first exons. Redundancy analyses involving juvenile
39 glass eels showed 0.06% and 0.03% of the variance at SNPs to be explained by localities and
40 environmental variables, respectively, with GO terms of genes associated with outliers
41 primarily involving neural system functioning. For CpGs 2.98% and 1.36% of variance was
42 explained by localities and environmental variables. Differentially methylated regions
43 particularly included genes involved in developmental processes, with hox clusters featuring
44 prominently. Life stage (adult versus glass eels) was the most important source of inter-
45 individual variation in methylation, likely reflecting both ageing and developmental
46 processes. Demethylation of transposable elements relative to pure European eel was
47 observed in European X American eel hybrids, possibly representing postzygotic barriers in
48 this system characterized by prolonged speciation and ongoing gene flow. Whereas the
49 genetic data are consistent with a role of single-generation selective responses, the
50 methylation results underpin the importance of epigenetics in the life cycle of eels and
51 suggests interactions between local environments, development and phenotypic variation
52 mediated by methylation variation. Eels are remarkable by having retained eight hox clusters,
53 and the results suggest important roles of methylation at hox genes for adaptive processes.
54
55 **Key words:**
56
57 Adaptive processes, *Anguilla anguilla*, epigenetics, hox clusters, hybridization, methylation.
58
59

Introduction

60
61   It is increasingly appreciated that epigenetics, defined as modifications of DNA that affects
62   expression of genes but without changing the DNA sequence (Dupont *et al.* 2009), is highly
63   important in developmental, adaptive and evolutionary processes (Adrian-Kalchhauser *et al.*
64   2020; Anastasiadi *et al.* 2021; Gore *et al.* 2018; Greenberg & Bourc'his 2019; Jablonka 2017;
65   Jablonka & Raz 2009; Lind & Spagopoulou 2018; Stajic *et al.* 2019; Verhoeven *et al.* 2016).
66   Epigenetics mechanisms include DNA methylation, histone modifications and small RNAs
67   (Deans & Maggert 2015; Feil & Fraga 2012; Law & Jacobsen 2010), where DNA
68   methylation has so far attracted the most interest. This involves addition of a methyl group to
69   a nucleotide, in most cases Cytosin (C), which in animals primarily occurs at CpG sites (C
70   followed by G in the genome sequence). Gain or loss of methylation at CpG sites in
71   regulatory regions can lead to silencing or reactivation of genes, with hypomethylation of
72   promoter regions generally leading to increased transcription in vertebrates (Christensen *et*
73   *al.* 2021; Greenberg & Bourc'his 2019; Jones 2012; Law & Jacobsen 2010; Moore *et al.*
74   2013).

75

76   Methylation plays key roles in developmental processes and cell differentiation (Greenberg &
77   Bourc'his 2019). In mammals this involves two events of reprogramming during
78   embryogenesis, but general knowledge about major changes in methylation is scarce in other
79   animals, including those that undergo extensive metamorphosis. Importantly, methylation can
80   be environmentally induced and transferable across cell divisions (Feil & Fraga 2012). There
81   is furthermore some evidence, particularly in plants but less clear-cut in animals that
82   epigenetic marks can be transferable across generations. This can ultimately lead to traits
83   being inherited despite not being coded by the DNA sequence (Anastasiadi *et al.* 2021; Gapp
84   *et al.* 2014; Richards 2006; Schmitz *et al.* 2013; Skvortsova *et al.* 2018).

85

86   The environmental inducibility of methylation states raises the possibility that this could
87   represent rapid adaptive mechanisms in response to spatial and temporal environmental
88   variation (Angers *et al.* 2020; Feil & Fraga 2012), and it is assumed to be a major component
89   in phenotypic plasticity and may also be involved in knock-on effects, that is early perceived
90   environmental cues leading to phenotypic change later in life (Jonsson *et al.* 2022). Indeed,
91   some studies report significant methylation differences of functional importance associated
92   with environmental variation (Artemov *et al.* 2017; Gugger *et al.* 2016; Heckwolf *et al.* 2020;
93   Le Luyer *et al.* 2017; Merondun *et al.* 2019; Metzger & Schulte 2018; Schmitz *et al.* 2013;
94   Wogan *et al.* 2020) A distinction has usually been made between genetic adaptation (encoded
95   by DNA) resulting from evolution across generations; and phenotypic plasticity, within-
96   generation responses of individuals to environmental conditions, e.g. by adjustments in
97   physiology (Gienapp *et al.* 2008; Kawecki & Ebert 2004). Whereas methylation from this
98   perspective could be viewed as a source of non-genetic adaptation, it is increasingly realized
99   that the boundaries between genetic and non-genetic factors are unclear and involve complex
100 interactions, also in the case of epigenetics (Adrian-Kalchhauser *et al.* 2020; Taudt *et al.*
101 2016; Verhoeven *et al.* 2016). In particular, a proportion of methylation may be under genetic
102 control (Richards 2006), as for instance demonstrated by different non-recombining
103 chromosome inversions showing different methylation (Sun *et al.* 2021). Differences in
104 methylation patterns between populations could therefore ultimately reflect individual- and
105 population-level genetic differences in genes controlling methylation (Anastasiadi *et al.*
106 2021; Dubin *et al.* 2015; Richards 2006; Taudt *et al.* 2016). It would therefore be of
107 significant interest to study epigenetic patterns and its association with geographical and
108 environmental variation in species where genetically based adaptation can be ruled out.

109

110   Methylation also has the important role to repress transposable elements (TEs), thus
111   preventing deleterious proliferation of TEs in the genome (Jones 2012; Slotkin &
112   Martienssen 2007). Derepression of TEs by demethylation has been found to occur in some
113   cases of hybridization, leading to harmful reactivation and proliferation of transposons
114   (Laporte *et al.* 2019; Michalak 2009; O'Neill *et al.* 1998; Ungerer *et al.* 2006). This could
115   potentially represent postzygotic barriers, but the importance of TE derepression relative to
116   other postzygotic barriers remains unclear. It would therefore be of interest to study this in
117   cases where environmental conditions experienced by species are similar and gene flow
118   between species is still ongoing.
119
120   The European eel (*Anguilla anguilla*) represents an excellent model for further increasing our
121   knowledge about the role of methylation in adaptive and developmental processes. It is
122   distributed across environmental conditions spanning from subarctic climates in Iceland to
123   Sub-tropical environments in North Africa (Tesch 2003). It spawns in the Southern Sargasso
124   Sea in partial sympatry with its sister species American eel (*A. rostrata*) (Kleckner *et al.*
125   1983; Munk *et al.* 2010), which shows a very similar life history and is distributed along
126   similar environmental gradients along the American Atlantic coast (Tesch 2003). European
127   eel larvae are transported by ocean currents towards the European and North African coastal
128   regions. The recently arrived juveniles metamorphose into so-called glass eels, settle in
129   freshwater and coastal marine habitats and go through an additional stage of metamorphosis
130   until they mature as silver eels and undertake their > 5,000 km spawning migration back to
131   the Sargasso Sea (Schmidt 1923; Tesch 2003). Previous results based on anonymous
132   methylation markers suggest major differences between life stages (Trautner *et al.* 2017), but
133   it is unknown if these differences represent functionally important methylation in relation to
134   developmental stages or merely ageing effects (Anastasiadi & Piferrer 2020; Horvath & Raj
135   2018).
136
137   Both European and American eel are remarkable by being panmictic species, that is despite
138   being distributed across a wide range of environmental conditions they mate randomly in the
139   Sargasso Sea (Als *et al.* 2011; Côté *et al.* 2013; Enbody *et al.* 2021; Palm *et al.* 2009; Pujolar
140   *et al.* 2014b); though see contrasting views by e.g. Baltazar-Soares *et al.* (2014).Signals of
141   spatially varying selection have been detected in both European and American eel (Babin *et
142   al.* 2017; Gagnaire *et al.* 2012; Pavey *et al.* 2015; Pujolar *et al.* 2014b; Williams *et al.* 1973),
143   but this is expected to be a single generation effect as individuals may end up in
144   environments that differ considerably from those of their parents (Gagnaire *et al.* 2012;
145   Pujolar *et al.* 2014b). The resulting absence of genetically based local adaptation suggests
146   that phenotypic plasticity could play an important role in the species' ability to persist in
147   different environments, although there is also evidence for a role of weak, polygenic selection
148   to occur (Côté *et al.* 2014; Pavey *et al.* 2015). For instance, American glass eels sampled at
149   geographically and environmentally different localities showed different growth rates and
150   transcriptomic reaction norms in common garden settings when exposed to different salinities
151   (Côté *et al.* 2014; Côté *et al.* 2009). This suggests that phenotypic plasticity interacts with
152   processes that have occurred at local scales at a very early stage coinciding with the arrival of
153   glass eels at the sites; either genetic variation shaped by within-generation selection (as the
154   species is panmictic) and/or epigenetic imprints. However, knowledge about differences in
155   methylation across geography and environments is currently lacking.
156
157   Finally, European and American eel can hybridize (Albert *et al.* 2006; Avise *et al.* 1990;
158   Jacobsen *et al.* 2017; Pujolar *et al.* 2014a), as is also the case for other Anguillid species
159   (Barth *et al.* 2020). Genomic analyses suggest a protracted speciation process by involving

160   episodes of isolation and secondary contact and with ongoing gene flow (Nikolic *et al.* 2020).
161   Given the incomplete speciation process it is of interest to assess if TE derepression occurs in
162   hybrids between European and American eel.
163
164   Here, we applied both whole genome sequencing and bisulphite sequencing of eels from
165   geographical locations ranging from Iceland to Morocco. As the functional significance of
166   methylation depends on the specific genomic categories (e.g. promotors and exons) being
167   methylated, we first made use of whole genome information to characterize the general
168   methylation landscape of European eel. Subsequently, we 1) tested the hypothesis that
169   differences in methylation is present among glass eels from different localities and are
170   furthermore associated with differences in environmental parameters. Given the panmictic
171   nature of the species we further predicted that such methylation differences should vastly
172   exceed genetic differences, even if within-generation selection occurs. 2) We assessed if
173   methylation differences are present between glass and adult eels, and if so if this can be
174   ascribed to the pronounced stages of metamorphosis or to mere ageing effects (Horvath &
175   Raj 2018). 3) By analysing methylation in European x American eel hybrids we tested the
176   hypothesis that transposon methylation does not differ from pure European eel, reflecting the
177   prolonged and incomplete speciation process.
178
179   **Materials and Methods**
180
181   Samples
182   A total of 50 European eels were analyzed, representing seven locations in Europe and
183   Northwestern Africa, spanning 30 degrees of latitude (Fig. 1, Table 1, Supporting
184   Information, Table S1). The samples were collected between 2001 and 2016 (Table 1).
185   Icelandic samples were collected for the present study, whereas the remaining samples have
186   previously been analyzed using RAD sequencing (Pujolar *et al.* 2014b; Pujolar *et al.* 2015).
187   Thirty-nine individuals were glass eels (juvenile eels recently arrived at the coasts), whereas
188   11 adult individuals (silver eels, i.e. about to undertake their spawning migration) were
189   included from two locations (Burrishoole, Ireland; Valencia, Spain). Three individuals (two
190   from Iceland and one from Ireland) were detected as hybrids with American eel (one F1
191   hybrid and two backcrosses in the direction of European eel), determined using species-
192   diagnostic SNPs (Pujolar *et al.* 2014a). For silver eels, tissues consisted of muscle, whereas
193   for glass eels DNA was extracted from the tail end, composed primarily of muscle. DNA was
194   extracted using the E.Z.N.A.® Tissue DNA Kit (OMEGA, Bio-tek, CA, USA) following the
195   manufacturer's recommendations. Whole-genome sequencing (WGS) and whole-genome-
196   bisulfite sequencing (WGBS) was outsourced to Novogene Europe (Cambridge, UK).
197   Sequencing was conducted 150 bp paired-end on the Illumina HiSeq platform and aimed for
198   a minimum coverage of 10X.
199
200   Mapping WGS reads and calling SNPs
201   The WGS reads were filtered using Trim Galore v0.4.1
202   (https://github.com/FelixKrueger/TrimGalore) and mapped to a recent chromosome level
203   European eel genome assembly (Rhie *et al.* 2021) (GenBank accession: GCA_013347855.1)
204   using the BWA-MEM algorithm of BWA v0.7.17 (Li & Durbin 2009). The resulting SAM
205   files were sorted by coordinate and were converted to BAM format using samtools v1.9 (Li *et*
206   *al.* 2009). A VCF file of SNPs encompassing all 50 individuals was generated from the BAM
207   files using bcftools v1.9 (Li *et al.* 2009), constraining the minimum mapping quality to 20.
208   Only biallelic SNPs with minimum variant quality of 20 and with combined coverage falling

209   between 500 and 750 were kept. The coverage thresholds were decided upon inspecting the
210   coverage distribution of the SNPs (Supporting Information, Fig. S1).
211
212   Genome-wide H_O (observed heterozygosity) of the individuals was calculated from the VCF
213   file by dividing the number of heterozygous sites with adjusted genome lengths. We used this
214   measure to corroborate the hybrid status of individuals and as a quality check of the data; in a
215   panmictic species with expectedly almost no inbreeding, genome-wide H_O should be very
216   similar across individuals. The genome length was adjusted per individual by correcting for
217   the missing sites generated by the SNP calling process. A PCA aimed at analyzing genetic
218   relationships among sampled individuals was conducted using the R function "prcomp" (R
219   Core Team 2018) on the genotype table of the individuals, where the genotypes were denoted
220   as the number of alternative alleles.
221
222   <u>Mapping WGBS reads and calling methylation</u>
223   A total of 36 individuals succeeded in WGBS (Supporting Information, Table S1) with
224   degradation and insufficient yields of DNA causing failure in sequencing of the remaining
225   individuals. The WGBS reads were filtered using Trim Galore by allowing "--trim1" and
226   were mapped to genomes using Bismark v0.22.3 (Krueger & Andrews 2011). Reads were
227   mapped to the individual genomes obtained from the WGS data instead of the general
228   reference genome. This was considered necessary due to the exceptionally high genetic
229   diversity of the European eel (Pujolar *et al.* 2013), leading to lower mapping success when
230   using the reference genome. Default parameters were used except for a relaxed gap penalty
231   ("--rdg 2,1 --rfg 2,1").
232
233   We subsequently ran "bismark_methylation_extractor" and "bismark2bedGraph" (Krueger &
234   Andrews 2011) to extract all the sequenced CpG sites together with their methylation status.
235   The information was stored in the COVERAGE files in the output. During the extraction
236   process, the first two base pairs of all the Read 2 files were removed based on the M-bias
237   plots. CpG sites containing mutations were excluded. As CpG is palindromic and
238   complementary CpGs are synchronized in methylation due to dnmt1 activity during DNA
239   replication, complementary CpGs were merged. The COVERAGE files of all the individuals
240   were merged using a custom script. This generated a file where the CpGs of all the
241   individuals were aligned by coordinate. Within each individual, the CpGs with coverage
242   lower than five were marked as missing. CpGs missing in more than half of the individuals
243   were filtered out. CpGs whose combined coverage (across individuals) fell outside the range
244   between 115 and 539 were removed. These coverage thresholds were decided from the
245   coverage distribution (Supporting Information, Fig. S2).
246
247   To assess differences in global methylation among individuals, PCA was conducted using the
248   R function "prcomp" on the methylation matrix. The data points were the individuals and the
249   variables were the methylation level of the CpG sites.
250
251   <u>Methylation in genomic categories</u>
252   We used the gene annotation file provided with the reference genome sequence, and for each
253   transcript we defined two potential promoter regions, one is from 1 bp to 500 bp upstream
254   from the TSS (transcription start site), referred to as "promoter_1", and the other from 501 bp
255   to 1000 bp upstream, referred to as "promoter_2. We also identified CpG islands using
256   cpgplot in the EMBOSS v6.6.0.0 package (Madeira *et al.* 2019). We further predicted and
257   annotated the repetitive sequences for the reference genome using RepeatModeler v1.0.11
258   and RepeatMasker v4.0.9-p2 (http://www.repeatmasker.org). Based on the annotation, we

259   divided the reference genome of European eel into 17 non-exclusive categories (Supporting
260   Information, Table S2) and examined methylation patterns within each of these categories.
261   The aligned CpGs were assigned into the genomic categories using the "intersect" command
262   of bedtools v2.29.0 (Quinlan & Hall 2010).
263
264   Correlation between methylation and gene expression
265   We anticipated that distinct methylation patterns of the first exons and the promoters of the
266   transcripts would indicate a functional role in regulating gene expression. We furthermore
267   also considered first introns, as Anastasiadi *et al.* (2018) reported inverse relationships
268   between methylation of this genomic category and gene expression. Gene expression profiles
269   were not generated in the present study, but we used a published transcriptome dataset of the
270   European eel (Bracamonte *et al.* (2019b) NCBI BioProject: PRJNA419718 and
271   PRJNA547691) as an approximate measure. This means that we could assess patterns of
272   general association between gene expression and methylation in genomic categories, but not
273   specific changes in gene expression as a result of differential methylation. The dataset is
274   comprised of Illumina paired-end reads from 30 experiments (20 individuals). The reads were
275   filtered using Trim Galore v0.4.1 and were mapped to the reference genome with the
276   guidance of the gene annotation using HISAT2 v2.1.0 (Kim *et al.* 2019). The expression
277   profile of all the transcripts was called for each experiment using StringTie v2.0 (Kovaka *et
278   al.* 2019). Upon inspecting the expression profile and visually examining the transcripts in
279   IGV v2.7.2 (Robinson *et al.* 2011), five experiments were removed due to low numbers of
280   expressed transcripts.
281
282   The expression levels (measured using FPKM, Fragments Per Kilobase of transcript per
283   Million mapped reads) of the transcripts were averaged across the experiments, and the
284   average values were used to correlate with the methylation levels of the first exons, first
285   introns and the promoters. The methylation level for each first exon, first intron or promoter
286   was represented with the average methylation level across all the individuals.
287
288   Identification of lowly-methylated sites (LMSs)
289   Due to the importance of lowly-methylated sites (LMSs) in activating genes in an otherwise
290   globally methylated genome (Nakamura *et al.* 2014), we inspected the genome-wide
291   distribution pattern of LMSs. In order to include all meaningful LMSs across all individuals,
292   a CpG with a methylation level lower than 0.05 in at least two individuals was defined as an
293   LMS. This threshold was chosen because the CpGs of this methylation level showed the
294   highest tendency of clustering together (Supporting Information Fig. S3). The criterion of a
295   minimum of two individuals aimed to decrease the possibility of false positives caused by
296   modest sequencing coverage.
297
298   Association of SNPs and methylation with localities and environmental parameters
299   We used redundancy analysis (RDA, (Forester *et al.* 2018; Legendre & Legendre 2012)) to
300   study association of SNPs and methylation, respectively, with local environments. The
301   analysis was conducted in the "vegan" package (v2.5-6 (Oksanen *et al.* 2008)) in R. Only
302   non-hybrid glass eels were included in the analyses, encompassing 32 individuals for the
303   SNP and 25 for the methylation data. Two rounds of RDA were implemented. The first round
304   had sampling locations (dummy variables) as explanatory variables. This analysis aimed to
305   examine the genetic or methylation response to the sampling locations regardless of their
306   environmental composition. The second round had sea surface temperature (SST30),
307   chlorophyll concentration and mean day length (MDL) as explanatory variables (Table 1), all
308   encompassing means of 30 days prior to the date of sampling. This analysis can be regarded

309 as representing a targeted subset of the (unknown) environmental composition represented by
310 the sampling locations. Remotely sensed sea surface temperatures encompassing a resolution
311 of 0.25 degree latitude × 0.25 degree longitude on a global grid and measured for each day,
312 were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA,
313 (http://www.esrl.noaa.gov/psd/) and retrieved using the function extractOISSTdaily from the
314 R script NOAA_OISST_ncdf4.R (http://lukemiller.org/index.php/2014/11/extracting-noaa-
315 sea-surface-temperatures-with-ncdf4/). Data on chlorophyll concentration were extracted
316 from the CCI-OC Data Portal (https://www.oceancolour.org/) (Sathyendranath *et al.* 2019).
317 Mean day length data were obtained from the Photoperiod Calculator at
318 https://www.ou.edu/research/electron/internet/solarjav.html. SST30 and chlorophyll
319 concentration were included to reflect basic abiotic and biotic properties of the environments,
320 whereas MDL was included to represent diurnal and seasonal variation among localities,
321 possibly associated with e.g. genetic or methylation variation at circadian genes. By including
322 mean values for the 30 days preceding sampling, we aimed to capture as much as possible the
323 environmental conditions the glass eels were exposed to either at the site or close to the site
324 during the last stages of oceanic transport.
325
326 RDA was conducted for the SNP and methylation data separately. For the SNP data, we
327 filtered out loci with overall minor allele count lower than four and removed SNPs containing
328 missing values, thus retaining 18,337,468 SNPs. For the methylation data, we filtered out
329 CpGs with more than five missing values across individuals, retaining 1,934,985 CpGs.
330 Missing values in the methylation data were replaced with cross-individual mean values of
331 the corresponding CpGs. The p-value of each RDA was calculated through 5000
332 permutations and p-values of the RDA axes were calculated using 2000 permutations each.
333 For each significant RDA axis, we extracted the loadings of the SNPs or the CpGs. SNPs or
334 CpGs with extreme loadings were defined as outliers. For SNPs, we used four times the
335 standard deviation away from the mean as the threshold. For CpGs, we used three times the
336 standard deviation away from the mean as the threshold. This difference in thresholds was
337 used to obtain comparable number of outliers between the two datasets.
338
339 <u>Methylation related to developmental stage and hybridity</u>
340 We used P_{ST} combined with methylation difference to search for methylation functionally
341 related to developmental stage and hybridity. P_{ST} is a measure of phenotypic differentiation
342 between groups (Leinonen *et al.* 2013; Pujol *et al.* 2008), here accommodated to evaluate
343 methylation divergence between groups in developmental stage or hybridity and calculated
344 using a custom script in R. We adopted criteria of P_{ST} higher than 0.8 and methylation
345 difference higher than 0.35 to define outliers. For the developmental stage, we compared
346 adults and glass eels. For each group, each CpG had to be scored for at least two individuals.
347
348 The outliers were assigned to the 17 genomic categories defined above to check for
349 enrichment. The outlier enrichment in a category was calculated as:

350
$$outlier\ enrichment = \log_2 \frac{O(frq)}{E(frq)}$$

351 where *O(frq)* was the observed frequency of the outliers in the category, and *E(frq)* was the
352 expected frequency by random chance. *E(frq)* was obtained by assigning all the CpGs to the
353 genomic categories. A positive enrichment implies overrepresentation of the outliers in the
354 category, and a negative enrichment indicates underrepresentation. The significance of the
355 enrichment was tested by comparison to the confidence interval of null hypothesis, i.e, no
356 enrichment. The confidence interval was defined using binomial distribution.
357

358   GO term enrichment analysis for the outliers
359   For each set of outliers (both SNPs and methylation data), we extracted all genes that
360   overlapped with the outliers within a 3000 bp range upstream and downstream. The 3000-bp
361   threshold was decided according to the median length of the intergenic regions (Supporting
362   Information, Table S2, Fig. S4). The resulting gene list was tested for GO (gene ontology)
363   term enrichment using the "weight01" algorithm of the "topGO" package (Alexa *et al.* 2006;
364   Schulz *et al.* 2007) in R. The p-values of the GO terms were adjusted following Benjamini-
365   Hochberg procedure (Benjamini & Hochberg 1995). The GO IDs of the genes were retrieved
366   by blasting the genes against the Swiss-Prot database (Release 2021_03, The UniProt
367   Consortium 2019).
368
369   Defining DMRs from the outliers
370   We developed a method for identifying DMRs (differentially methylated regions) from a set
371   of methylation outliers based on the distribution of the neighboring distances among the
372   outliers. Such distributions show two peaks. The peak with a higher mean is a geometric
373   distribution, representing neighbouring distances of randomly distributed outliers with
374   uncertain functional value and not easily separable from noise. In contrast, the peak with
375   lower mean represents regions with multiple closely located outliers, hence strongly
376   indicative of functional roles. We identified these regions as DMRs and obtained them by
377   grouping the outliers in the small-mean peak according to distance. A threshold, *K*, was set
378   for the minimum number of outliers required in a DMR, in order to filter out the noise
379   generated by the large-mean peak. *K* was obtained from the following inequality.

380
$$\frac{\sum_{X=K}^{\infty} E(a_X)}{\sum_{X=K}^{\infty} O(a_X)} < FDR$$

381
382   Here, a_X represents the number of DMRs containing X outliers. FDR is the false discovery
383   rate, which we set as 0.01 in this study. $O(a_X)$ is the observed series of a_X, and $E(a_X)$ is the
384   expected series of a_X. $E(a_X)$ was calculated as:

$$E(a_X) = N \cdot p^{X-1} \cdot (1-p)^2$$

386   where N is the number of neighboring distances in the large-mean peak assuming a geometric
387   distribution, and p is the expected ratio of neighboring distances in the small-mean peak
388   under this distribution.
389
390   **Results**
391
392   Genetic variation
393   An overview of whole genome sequencing and mapping statistics for each individual is
394   provided in Supporting Information, Table S3. A total of 74,040,803 SNPs were obtained
395   from the WGS data. The majority of individuals showed highly similar levels of
396   heterozygosity (Supporting Information Fig. S5), ranging from 0.00981 to 0.01084 (mean:
397   0.01028), whereas the three admixed individuals exhibited higher levels (0.01156 to
398   0.01232). Four individuals from Morocco showed higher heterozygosity (0.01161 to
399   0.01229), almost similar to the hybrids. However, species-diagnostic SNPs (Pujolar *et al.*
400   2014a) confirmed them not to be hybrids. Such variation in heterozygosity would not be
401   expected in a panmictic species with high effective population size, and we found that the
402   elevated heterozygosity was most likely due to cross-sample contamination (see Supporting
403   Information, Note S1), with contamination rate ranging from 3.4% to 6.0%. We therefore
404   excluded these individuals from all subsequent analyses, although we note that we observed
405   no noticeable impacts on the methylation-based analyses. A Principal Components Analysis

406 (PCA) based on SNPs (Fig. 2.a) showed virtually no divergence between the majority of
407 individuals, consistent with the assumption of panmixia of the species (Als *et al.* 2011;
408 Enbody *et al.* 2021; Palm *et al.* 2009; Pujolar *et al.* 2014b), whereas the three hybrids showed
409 separation along PC1 and PC2.
410
411 <u>Global methylation</u>
412 An overview of whole genome bisulphite sequencing, mapping statistics and methylation for
413 each individual is provided in Supporting Information, Table S4. Methylation was analyzed
414 in 33 individuals (after removal of contaminated individuals). A total of 7,484,974 CpG sites
415 were identified (out of 24,369,391 sites in the reference genome). The overall methylation of
416 70.9% is within the range observed in vertebrates (Head 2014) and of similar magnitude as
417 observed in fishes like three-spined stickleback (*Gasterosteus aculeatus*) (70.3%) (Metzger &
418 Schulte 2018) and zebrafish (*Danio rerio*) (80%) (Feng *et al.* 2010). The methylation level
419 distribution exhibited a U-shape with the two peaks at the ends of the range of methylation
420 (corresponding to hyper and hypomethylation; Fig. 2.c), and this pattern was highly
421 consistent across individuals. A PCA based on methylation of CpG sites separated adult and
422 glass eels along PC1 and PC2 (Fig. 2.b).
423
424 <u>Methylation in genomic categories</u>
425 We divided the reference genome into 17 non-exclusive functional categories according to
426 genomic annotation (Supporting Information, Table S2) and found methylation in most
427 categories, including first introns to be high (Fig. 3). However, promoters and the first exons
428 were notable exceptions and showed the lowest methylation levels. Visual inspection of the
429 methylation level across the chromosomes also indicated a high frequency of methylation
430 valleys at the beginning of genes (see Supporting Information, Fig. S6 for an example from
431 Chr_01). The methylation level of each CpG was highly consistent across individuals for the
432 promoters and the first exons (Fig. 3).
433
434 Comparison of methylation and gene expression using the transcriptome dataset by
435 Bracamonte *et al.* (2019a) showed that genes with hypomethylated first exons were overall
436 highly expressed, and those with hypermethylated first exons generally showed lower
437 expression levels (Fig. 4.a). This contrasted with first introns (Fig. 4.b), where only weak
438 association was found between hypo- and hypermethylation and gene expression. For
439 promoter 1 and 2 regions there was association between hypo- and hypermethylation and
440 gene expression, although the patterns were weaker than for first exons (Fig. 4.c and d).
441
442 We also found that repetitive regions, especially transposable elements (TEs), showed very
443 low numbers of hypomethylated CpGs (Supporting Information, Fig. S7), likely related to TE
444 silencing. This was further supported by the observation that the CpGs in TEs exhibited the
445 lowest methylation variation (Fig. 3).
446
447 We found that the CpG islands were overall highly methylated (Fig. 3), in accordance with
448 other studies (Deaton & Bird 2011) and in contrast to previous notions that the
449 hypomethylated CpGs are mainly confined to the CpG islands in promoter regions (Saxonov
450 *et al.* 2006). The CpG islands were not enriched in promoters (Supporting Information, Fig.
451 S8), but were enriched in CDS and LTR, both being highly methylated.
452
453 PCA performed for each of the genomic categories generally separated adults and glass eels
454 (Supporting Information, Fig. S9), similar to genome-wide methylation patterns (Fig. 2.b).

455   This suggests that the methylation differences between developmental stages exist in all
456   categories and is the dominant source of variance.
457
458   <u>Characteristics of lowly-methylated sites (LMSs)</u>
459   We identified 1,099,209 lowly-methylated CpG sites (LMSs), which tended to cluster into
460   local groups (Supporting Information, Fig. S3). They were highly enriched in the promoters
461   and the first exons and were underrepresented in the repetitive sequences (Supporting
462   Information, Fig. S10). Some genomic regions exhibited particularly high density of LMSs
463   (Supporting Information, Fig. S11). The gene clusters and genes identified in these regions
464   included: all eight hox clusters of the European eel (see Supporting Information Fig. S11 and
465   Table S5 for genomic coordinates), the two largest protocadherin clusters (12 copies on
466   Chr_03 and 6 copies on Chr_09), the two largest olfactory receptor clusters (110 copies on
467   Chr_09 and 125 copies on Chr_12) and a zscan2 cluster (5 copies on Chr_08). The genes
468   included: zic gene pairs, tbx, tfap2 and homeobox genes other than hox. The hox and
469   protocadherin clusters, homeobox genes, zic pairs and tbx have previously been reported in
470   medaka fish (*Oryzias latipes*) and/or threespine stickleback to reside in large hypomethylated
471   domains (Metzger & Schulte 2018; Nakamura *et al.* 2014), hence consistent with their high
472   density of LMSs found in this study. We provide an extended description of the distribution
473   of LMSs in Supporting Information, Note S2.
474
475
476   <u>Genetic and methylation response to local environments</u>
477   Redundancy analysis [RDA (Forester *et al.* 2018; Legendre & Legendre 2012)] was used to
478   study association of genetic or methylation variation with local environments experienced by
479   glass eels. That left 32 individuals for the genetic data and 25 for the methylation data. Two
480   rounds of RDA were implemented, with the first having the sampling locations (dummy
481   variables) as explanatory variables, thus examining the genetic or methylation response to the
482   sampling locations encompassing a range of (undefined) environmental parameters. For the
483   genetic data (SNPs), 0.06% of the variance was explained by the location variables, and for
484   the methylation data (CpG sites), 2.98% was explained (Fig.5a). The results from both
485   datasets were significant (p-value: 0.0028 and 0.0004, respectively). We further extracted
486   outliers from the loading scores of the first axes, since for both datasets only the first axes
487   were significant (p-value: 0.0075 and 0.0170, respectively). For the SNP data, 11949 outliers
488   were obtained. GO term enrichment analysis for the surrounding genes revealed high
489   abundance of genes involved in nervous system development and functioning (Supporting
490   Information, Table S6). We inspected the genome-wide distribution of the outliers and found
491   regions with high outlier density (Fig. 5b). However, the genes in these regions do not show
492   much functional overlap with the major GO terms.
493
494   For the methylation data, 23912 outliers were found. After grouping the outliers into regions
495   according to their neighbouring distances, this led to the identification of 1523 DMRs
496   (Supporting Information, Table S7). Genes in these regions were enriched with functions
497   related to developmental processes (Supporting Information, Table S8). The genome-wide
498   distribution of the outliers displayed multiple high-density regions (Fig. 5b). Genes in these
499   regions exhibited high correlation with the major GO terms. In particular, seven of the eight
500   hox clusters turned out to be high-density regions.
501
502   The second round of RDA had sea surface temperature (SST30), mean day length (MDL) and
503   chlorophyll concentration as explanatory variables, all representing means over the last 30
504   days prior to the date of sampling. Similar to the first round of RDA, less variance (0.03%)

505 was explained by environmental variables for the genetic data as compared to the methylation
506 data (1.36%; Fig. 6a). The results were significant for both datasets (p-value: 0.0396 and
507 0.0056, respectively), but only the first axes were significant or marginally significant (p-
508 value: 0.0915 and 0.0325, respectively). A total of 12124 outliers were obtained from the
509 genetic data, with surrounding genes enriched with the same major GO terms as in the RDA
510 involving localities (Supporting Information, Table S9). Similar to the first round of RDA,
511 the genome-wide distribution of the outliers showed some high-density regions (Fig. 6b), but
512 with no functional overlap of major GO terms observed for the genes in these regions. For the
513 methylation data, 19311 outliers were found, from which 803 DMRs were identified
514 (Supporting Information, Table S10). Genes in these DMRs were enriched with functions
515 related to developmental processes (Supporting Information, Table S11), and the outliers
516 showed several high-density regions (Fig. 6b), encompassing among others six of the hox
517 clusters. Inspection of outliers associated with the individual environmental parameters did
518 not reveal obvious differences (Supporting Information, Table S12); the major GO terms
519 were primarily associated with development regardless of the environmental parameter. A
520 total of 9% percent of the SNP outliers (1092 SNPs) and 39.5% (7632 CpGs) of the
521 methylation outliers overlapped with those identified in the first round of RDA, reflecting the
522 fact that the three environmental parameters represent a subset of the total environmental
523 variation among sites.
524 A total of 30 GO terms overlapped between outlier SNPs and DMRs for the RDA involving
525 locality as explanatory variable (6.9% of DMR GO terms), and 7 GO terms overlapped for
526 SNPs and DMRs identified using environmental variables as explanatory variables (2.3% of
527 DMR GO terms). The overlapping GO terms are highlighted in Supporting Information,
528 Table S6 and S9.
529
530 <u>Methylation related to developmental stage and hybridization</u>
531 For developmental stage, the P_{ST} distribution between glass eels and adults indicated that the
532 two groups were divergent in methylation throughout most of the genome, and the
533 methylation difference distribution showed adults to be overall hypomethylated compared to
534 glass eels (Fig. 7.a). Using criteria of $P_{ST} > 0.8$ and methylation difference > 0.35, we
535 identified 10767 hypomethylated and 3411 hypermethylated outliers in adults relative to
536 glass eels. According to the neighbouring distance distribution of the outliers,
537 hypomethylated outliers seemed to be more randomly distributed across the genome whereas
538 hypermethylated outliers were highly targeted (Fig. 7.b). We examined the enrichment of the
539 outliers in the 17 genomic categories defined above. The hypomethylated outliers were
540 significantly enriched in intergenic regions, and hypermethylated outliers were significantly
541 enriched in the intergenic regions, the promoters and the first introns (Fig. 7.c). From the
542 hypomethylated outliers, 389 DMRs were identified (Supporting Information, Table S13). No
543 significantly enriched GO terms were found from the genes residing in these DMRs
544 (Supporting Information, Table S14). From the hypermethylated outliers, 577 DMRs were
545 found (Supporting Information, Table S13), enriched with transcription regulators and
546 development-related genes (Supporting Information, Table S14).
547
548 The genome-wide distribution of the outliers corroborated the neighbouring distance
549 distribution (Fig. 7.d). The outliers hypomethylated in adults were primarily randomly
550 distributed. Hence, despite their abundance being three times higher than hypermethylated
551 outliers, only six high-density regions were identified, encompassing four genes that were all
552 related to development and cell differentiation. In comparison, the hypermethylated outliers
553 encompassed 25 high-density regions. The three highest peaks correspond to two copies of
554 nfix and one pura. They are involved in the initiation of DNA replication and transcription

555 (Gronostajski 2000). Two copies of nfic and one nfia were also among the high-density
556 regions. They have similar functions as nfix, and together they cover five of the six nuclear
557 factor I genes in the European eel genome. Interestingly, the two copies of nfix and nfic were
558 found in both hypomethylated and hypermethylated DMRs (Supporting Information, Table
559 S13). This may be related to alternative splicing of these genes, as the isoforms of these
560 genes tend to have different or even opposite effects (Gronostajski 2000). The genome
561 contains two copies of zic1/zic4 gene pair, and both copies were found within the high-
562 density regions. These genes are crucial for nervous system development in embryos. Three
563 copies of igf2bp genes were found, and there are in total five copies in the genome. However,
564 the hypermethylated DMRs represented all five copies (Supporting Information, Table S13).
565 Igf2bp genes are RNA-binding factors. They play direct roles in transport and translation of
566 mRNAs and protect them from endonuclease and miRNA attacks. Igf2bp plays important
567 role in nervous system development. Two copies of bmi1 were found in the high-density
568 regions, out of the three copies in the genome. They are components of PRC1 complex,
569 which induce gene repression through histone modification. PRC1 regulate many genes
570 during development, including the hox clusters. For the rest of the genes in the high-density
571 regions, most are related to developmental processes. We also note that in contrast to outliers
572 associated with local environments, none of the hox clusters were found in high-density
573 regions, and among DMRs only hoxDa on Chr_03 was visibly hypermethylated (Supporting
574 Information, Table S13).

576 All adult samples were from Bur and Val, with each locality represented by three adults and
577 three glass eels. We repeated the analysis above for each locality in order to assess
578 parallelism in methylation differences between adults and glass eels. For Bur, 37727 outliers
579 were found, and 6007 of them overlapped with the outliers found above (42.37%; the
580 percentages are relative to the number from the analysis encompassing all adult versus all
581 glass eels). For Val, 36739 outliers were found of which 5411 (38.16%) overlapped. A total
582 of 1860 (13.12%) outliers were shared by all three sets. The percentages remained similar
583 when only hypomethylated or hypermethylated outliers were considered. Under a null
584 hypothesis that the three sets of outliers are uncorrelated, the expected percentages would be
585 1.95%, 1.90%, 0.037%.

587 For hybridity, we defined hybrids and the non-hybrids as two groups. The methylation
588 profiles of the two groups based on P_{ST} were only mildly divergent (Fig. 8.a). We obtained
589 8577 hypomethylated and 667 hypermethylated outliers in hybrids relative to non-hybrids.
590 Both sets of outliers were largely randomly distributed in the genome (Fig. 8.b). The
591 hypomethylated outliers were significantly enriched in intergenic regions, repetitive regions
592 and DNA TEs (Fig. 8.c), with patterns in the last category suggesting demethylation of
593 transposons in hybrids. The hypermethylated outliers were significantly enriched in
594 intergenic regions and promoters (Fig. 8.c). A total of 129 hypomethylated and 90
595 hypermethylated DMRs were found (Fig. 8.d, Supporting Information, Table S15), with no
596 significant enrichment of any GO terms found for either set of DMRs (Supporting
597 Information, Table S16).

599 **Discussion**

601 The results of this study, along with other recent papers (Christensen *et al.* 2021; Leitwein *et*
602 *al.* 2021; Wellband *et al.* 2021), represent some of the first data on methylation at the whole
603 genome level in fishes, and we observed a complex methylation landscape that is associated
604 with the general functional roles of methylation. In relation to our specific research objectives

605 we found 1) outlier SNPs in glass eels with respect to geographical location and
606 environmental parameters, possibly reflecting within-generation selection. 2) In parallel with
607 outlier SNPs we found differentially methylated regions in glass eels associated with
608 geographical locations and local environments, indicating a role in local adaptive responses.
609 Neighbouring genes particularly represented functions related to development, and especially
610 Hox genes were prominent. 3) Life stage (glass and adult eels) was the overall strongest
611 determinant of methylation differences among individuals, and a considerable portion of
612 methylation differences was associated with genes of importance to developmental processes.
613 4) TEs were highly represented among methylation outliers between hybrids and non-hybrid
614 European eel, and were hypomethylated in hybrids, indicating that TE derepression also
615 occurs in this system of incomplete speciation. We discuss these findings in more detail in
616 the following.
617
618 Methylation landscape in European eel
619 The general methylation landscape was in accordance with findings in other vertebrates
620 (Brenet *et al.* 2011; Feng *et al.* 2010; Head 2014; Metzger & Schulte 2018), including overall
621 high levels of genome-wide methylation but with promotors and first exons showing
622 distinctly lower methylation and also higher variance in methylation. Moreover, the
623 association between gene expression and methylation at first exons and promoters suggested
624 a functional role of methylation in these genomic categories (Brenet *et al.* 2011; Jones 2012).
625 We found only weak association between methylation at first introns and gene expression,
626 hence somewhat contrasting with the results by Anastasiadi *et al.* (2018). On the other hand,
627 however, first introns showed enrichment among hypermethylated outliers in adult as
628 compared to glass eels, suggesting a functional role of this genomic category. Some notable
629 patterns were found in the analysis of lowly methylated sites (LMS), where specific gene
630 clusters and genes, particularly hox, protocadherin and olfactory receptor clusters coincided
631 with large lowly methylated domains, in parallel to findings in medaka and three-spine
632 stickleback (Metzger & Schulte 2018; Nakamura *et al.* 2014). It has been previously
633 suggested that such large hypomethylated domains act to suppress transcription of genes
634 while at the same time retaining flexibility for transcription during development (Nakamura
635 *et al.* 2014). This illustrates the complexity of patterns of methylation, as hypomethylation of
636 promoter regions *per se* is otherwise positively associated with increased transcription
637 (Moore *et al.* 2013).
638
639 Genetic and methylation response to local environments
640 The unique life history of European eel causes it at the same time to be panmictic and yet to
641 be distributed across geographically and environmentally highly divergent localities (Als *et
642 al.* 2011; Enbody *et al.* 2021; Pujolar *et al.* 2014b), hence providing opportunities to assess
643 epigenomic responses to environmental variables independent of population-specific genetic
644 variation. However, despite panmixia, redundancy analyses nevertheless showed a low but
645 significant proportion of genetic variance at the level of SNPs that was explained by locality
646 and environmental variables (0.06 and 0.03%, respectively), and > 10,000 outlier SNPs were
647 identified. The finding of genomic regions with high densities of outlier SNPs and
648 enrichment of GO-terms associated with nervous system development and functioning
649 further lends credibility to these findings representing genuine biological signals rather than
650 false positives. In that sense, the findings are in accordance with previous findings in both
651 European and American eel (Babin *et al.* 2017; Gagnaire *et al.* 2012; Pavey *et al.* 2015;
652 Pujolar *et al.* 2014b; Williams *et al.* 1973), ascribed to within-generation selection and
653 involving polygenic selection. The finding of specific genomic outlier regions as opposed to
654 more even distribution of outliers is puzzling, as this would not be likely to occur as a result

655   of selection within a single generation; panmixia would imply that offspring may end up in
656   very different geographic localities as compared to their parents. Pavey *et al.* (2015) found
657   evidence for polygenic divergence between freshwater and brackish/saltwater ecotypes of
658   American eel and suggested genotype-dependent habitat choice as one possible explanation.
659   It is possible that genotype-dependent habitat choice could also explain our results, but
660   addressing this issue would require other studies and sampling designs.
661
662   A significant proportion of the variance in methylation was explained by locality and
663   environmental parameters (2.96 and 1.36%, respectively). The fact that the specific
664   environmental variables mean day length, sea surface temperature and chlorophyll
665   concentration accounted for considerably less of the variance, as compared to locality,
666   suggests that other environmental factors that varies across localities (and/or across years)
667   may have important effects on methylation. One such factor could be local population
668   density; for instance, sex determination in Anguillid eels is principally or exclusively
669   environmentally determined (Geffroy & Bardonnet 2016). It has been suggested that high
670   population density of glass eels leads to predominance of initially fast-growing males and
671   low density predominance of initially slow-growing females (Davey & Jellyman 2005),
672   although the mechanisms and factors involved may be considerably more complex (Côté *et*
673   *al.* 2015; Geffroy & Bardonnet 2016). Although specific information about density of glass
674   eels at the sampling localities was not available, recruitment is known to vary considerably
675   across the distributional range of the species (Bornarel *et al.* 2018; Dekker 2003; Tesch
676   2003), with the Bay of Biscay and the Iberian Coast (here represented by the Gironde
677   sample) accounting for a major part of total recruitment (Bornarel *et al.* 2018). Hence, we
678   find it plausible that methylation differences across samples could at least partly reflect
679   differences in population density, also considering the high representation of developmental
680   processes among significant GO-terms for DMRs. Indeed, it has previously been found in a
681   bird species that changes in a social environment leading to increased competition also led to
682   altered patterns of methylation (Rodriguez-Martinez & Galvan 2019).
683
684   The finding of DMRs associated with Chlorophyll a concentration and sea surface
685   temperature is not surprising, given the biological importance of these factors. The former of
686   these can be considered a proxy of productivity, thereby affecting feeding and growth of
687   glass eels. It is well established in humans and other vertebrates that diet can affect
688   methylation (Lea *et al.* 2016; Zhang & Kutateladze 2018). It is therefore biologically
689   meaningful that most DMRs were associated with GO-terms related to growth and
690   developmental processes.
691
692   The life cycle and behaviour of European eel, including feeding and locomotory activity, is
693   strongly affected by light regimes (Lopez-Olmeda *et al.* 2012; Tesch 2003). Given the
694   variation in mean day length across sampling sites and dates, we anticipated a certain
695   representation of circadian-related genes in DMRs associated with this environmental
696   parameter. In fact, a previous population genomics study found a significant correlation
697   between latitude and the circadian clock gene period (*Per*) indicative of within-generation
698   selection (Pujolar *et al.* 2014b). At the methylation level in the present study, however, no
699   DMRs were associated with circadian genes, although we note that this could be a result of
700   analyzing muscle tissue, whereas circadian genes would be expected to be functionally most
701   important in brain tissue (Baras *et al.* 1998). Similar to the other environmental parameters,
702   significant GO-terms were instead dominated by developmental processes. It is possible that
703   different activity schemes associated with different light regimes could influence

704 development, and mean day length could also be correlated with other environmental factors
705 affecting development.
706
707 Remarkably, hox clusters featured prominently among high-density regions of outliers, both
708 in relation to localities and specific environmental parameters. Hence, hox clusters are not
709 only generally situated in large hypomethylated domains, but also represent some of the
710 genomic regions showing most pronounced differential methylation among individuals from
711 different localities and environments. Hox genes are of fundamental importance in
712 developmental processes, notably with respect to determining body plans (Carroll 2008;
713 Duboule 2007). It has previously been found that eels have retained a surprisingly large
714 repertoire of duplicated hox clusters and this has been suggested to underlie the two different
715 body plans of the leaf-shaped larval stage (leptocephalus) and the glass and adult eel stages
716 (Henkel *et al.* 2012). Our results suggest that variation in methylation of hox genes (and by
717 inference their regulation) could also be associated with phenotypic variation that develops in
718 response to local environmental conditions. Our data do not allow for specifically associating
719 methylation with phenotypic traits. However, examples of morphological variation exist in
720 adult eels such as distinct narrow-headed and broad-headed types that exhibit different
721 feeding preferences; these morphs are associated with different transcriptomic profiles
722 already at the glass eel stage (De Meyer *et al.* 2017).
723
724 Other explanations should, however, also be considered, as glass eels arriving at different
725 localities could show different ages and development, for instance involving different
726 methylation of hox clusters. Hence, if recruitment of glass eels exclusively occurs via the
727 Gulf Stream, then glass eels would be expected to be younger in northern as opposed to
728 southern localities. However, other ocean currents than the Gulf Stream are assumed to be
729 involved in transport of larvae (Munk *et al.* 2010), age determination of glass eels is
730 generally considered controversial (Bonhommeau *et al.* 2010), and it has been suggested that
731 distance from inshore regions to the Continental Shelf could be the primary factor affecting
732 age of newly recruited glass eel (Lecomte-Finiger 1992). Hence, this scenario merits
733 consideration, but is not possible to assess with the data and knowledge of recruitment
734 patterns at hand.
735
736 Our results show that already at the early life stage of glass eel, where individuals settle in
737 their future nursery and foraging areas, differential methylation is present that is associated
738 with geographical locations and/or environmental parameters. These differences have the
739 potential to affect gene expression and phenotypes also later in life, and the results raise the
740 possibility that epigenetics could in fact underlie differences in growth rates and
741 transcriptomic reaction norms as observed in American eels from different localities (Côté *et*
742 *al.* 2014; Côté *et al.* 2009). Nevertheless, it is a complex question if the methylation
743 differences lead to phenotypic plasticity of adaptive value. This would require that
744 environmental factors affecting methylation should also be predictive of the environmental
745 conditions encountered later in life (Bateson *et al.* 2014). Analysis of methylation in older
746 (yellow) eels from different localities and environments could shed further light on the role of
747 methylation in adaptive processes in eels, and if the same cohorts could be followed from the
748 glass eel stage this could allow for assessing the temporal stability and adaptive significance
749 of methylation differences induced in early life. Finally, although our focus on a panmictic
750 species should minimize interactions between genetic variation and methylation, the results
751 also show some genetic variation associated with environmental factors. A genetic influence
752 on methylation patterns can therefore not be ruled out entirely, although we note that the
753 functional overlap between outlier SNPs and DMRs was limited. Richards (2006)

754   distinguished between different categories of epigenetic variation, where genetic variation
755   controls (*obligatory*) or influences (*facilitated*) epigenetic variation, whereas in the *pure*
756   category epigenetic variation is independent of genetic variation. Indeed, empirical evidence
757   now exists from a range of organisms showing that at least a part of methylation variation
758   interacts with or is controlled by genetic variation (i.e. *obligatory* or *pure* epigenetic
759   variation) that may furthermore interact with environmental conditions (Berbel-Filho *et al.*
760   2019; Dubin *et al.* 2015; Teh *et al.* 2014). It would be an interesting future research question
761   if a higher proportion of epigenetic variation associated with environmental factors is *pure* in
762   panmictic eels as compared to other species showing genetic differentiation and local
763   adaptation across populations. However, this would obviously require a deeper understanding
764   of interactions at the genetic and epigenetic levels, along with comparable data from relevant
765   species.
766
767   <u>Differences in methylation between life stages</u>
768   Patterns of global methylation clearly separated juveniles (glass eels) from adults (silver
769   eels). Samples from juvenile eels could encompass other tissues than muscle, but muscle
770   would nevertheless constitute the bulk of tissue analyzed. We therefore find it less likely that
771   the patterns of methylation should reflect different tissues as opposed to different life stages.
772   Hence, with this caveat in mind life stage was the most important source of inter-individual
773   variation in methylation, and it is noteworthy that this pattern showed high parallelism across
774   the two environments from which adult eels were sampled. The European eel life cycle is
775   characterized by several metamorphoses; from larvae to glass eel, from glass eel to yellow
776   (adult) eel, and from yellow eel to mature silver eel, all involving distinct morphological and
777   physiological changes (Tesch 2003). The extensive methylation differences observed could
778   reflect extensive change of methylation associated with metamorphosis, as previously found
779   in both vertebrate and invertebrate species (Covelo-Soto *et al.* 2015; Gegner *et al.* 2021;
780   Kyono *et al.* 2020), but could also represent more gradual age-related changes in methylation
781   (Horvath 2013; Horvath & Raj 2018; Issa 2014). In that sense it was interesting that outliers
782   that were hypomethylated in adult eels showed a relatively random genomic distribution,
783   whereas hypermethylated regions showed a more targeted genomic distribution with
784   enrichment of promoter regions and first introns. Genome-wide hypomethylation and
785   hypermethylation of promoters is in fact a general pattern of methylation associated with
786   ageing (Johnson *et al.* 2012).
787
788   On the other hand, the strong representation of developmental processes among GO terms for
789   hyper-methylated outliers supports links to metamorphic processes. Moreover, whereas our
790   study analyzed methylation in muscle tissue, a previous study of European eel using
791   methylation-sensitive amplified polymorphisms (MSAP) and comparing life stages found
792   little divergence in liver tissue but larger differences in gill and brain tissues (Trautner *et al.*
793   2017). These tissue-specific differences argue against merely age-related effects and support
794   methylation differences being due to specific traits and environmental conditions encountered
795   by the life stages, e.g. fresh or brackish water in yellow eels and oceanic salinities to be
796   encountered during the spawning migration of silver eels. In the case of muscle tissue,
797   important differences in metabolic capacity and power output also develop between the
798   yellow and silver eel stages (Egginton 1986; Ellerby *et al.* 2001), ascribed to their long
799   spawning migration. However, since our sampling included glass and silver eels, but not
800   yellow eels, it remains uncertain exactly at which life stages the observed methylation
801   differences have occurred. In total, it is possible that the distinct differences in methylation
802   between glass and silver eels could reflect both ageing and metamorphosis, and it would
803   require more extensive analysis of individuals at different age stages to fully resolve this.

804
805 Interestingly, whereas hox clusters represented some of the most distinctive methylation
806 outlier regions between glass eels from different localities and environments, they were not
807 represented among methylation outlier regions between juvenile and adult eels, despite their
808 importance in developmental processes. We do not rule out that differential methylation
809 could exist between earlier life stages, notably leptocephali (larvae) and glass eels (as
810 implicitly suggested by Henkel *et al.* (2012)). However, for the life stages covered in this
811 study, differential methylation of hox genes appears almost exclusively associated with
812 environments. This decreases the possibility that the results obtained from glass eels could
813 represent artefacts such as subtle differences in ages and developmental stages among
814 individuals from different localities, as discussed previously.
815
816 Methylation in European x American eel hybrids
817 Transposable elements (TEs) can be considered genomic parasites, and free proliferation of
818 TEs in the genome is harmful. Hence, TEs are inactivated in particular by methylation
819 mediated by small piRNA interacting with PIWI proteins (Goodier 2016). Evolutionary
820 "arms races" between TEs and genes in the PIWI-piRNA pathway has led to rapid evolution
821 and divergence between species, that again results in incompatibilities in hybrids leading to
822 derepression of TEs (Aravin *et al.* 2007; Simkin *et al.* 2013). The resulting remobilization of
823 TEs has been suggested as an important postzygotic reproductive isolation mechanism, even
824 in cases of recent speciation (Laporte *et al.* 2019; Michalak 2009; O'Neill *et al.* 1998;
825 Ungerer *et al.* 2006).
826
827 The speciation history of European and American eel is complex and prolonged involving an
828 initial period of reproductive isolation, presumably due to vicariance, followed by secondary
829 contact and ongoing gene flow (Nikolic *et al.* 2020). Genomic outlier regions separating the
830 species primarily represent genes related to energy and development, consistent with
831 differences in length of spawning migration and larval phase duration of the two species
832 (Jacobsen *et al.* 2014). Our results suggest, however, that postzygotic isolation does not only
833 involve selection at ecologically important genes, but could also encompass intrinsic
834 incompatibilities leading to demethylation of TEs, even despite ongoing gene flow. Even
835 though we reject the hypothesis that transposon methylation does not differ between pure
836 European eel and hybrids, we stress that the results presented here are preliminary and do not
837 involve a comparison with the epigenome of pure American eel. Analysis of higher numbers
838 of F1 hybrids and backcrosses could shed further light on TE demethylation and the extent to
839 which it decays with each generation of backcrossing.
840
841 Conclusions
842 Our study of a panmictic species where genetically based local adaptation cannot occur
843 yielded important insights into the much debated issue of the ecological and adaptive role of
844 methylation (Bossdorf *et al.* 2008; Flores *et al.* 2013; Rey *et al.* 2020; Verhoeven *et al.*
845 2016). In the absence of genetic differentiation and at most a limited degree of within-
846 generation selection, variance in methylation between life stages, between hybrids and non-
847 hybrids and between glass eels from different localities and environments was pronounced.
848 Whereas the variance associated with life stages and hybridization concerns innate properties
849 of the species, the association of methylation with localities and environmental variables does
850 suggest that the genomes of eels can respond epigenetically to local conditions. We cannot
851 entirely rule out the possibility that the genetic variation found to be associated with local
852 environments could also interact with methylation. It can therefore not be concluded directly
853 that methylation "substitutes" genetically based local adaptation, and this would also require

854   demonstration of the phenotypic effects of methylation and its adaptive value. However,
855   there is certainly the possibility that at least some environmentally induced methylation at the
856   glass eel stage is of adaptive value later in life. There are as yet few comparable studies of
857   wild species quantifying variation of methylation in response to environmental factors.
858   However, the 2-3% of methylation variation associated with localities and environments in
859   the early life stage of glass eel is considerably higher than the ca. 0.01% of methylation
860   associated with different salinities in experiments with three-spined sticklebacks (Metzger &
861   Schulte 2018), but lower than the ca. 16% associated with river and hatchery environments in
862   Coho salmon (*Oncorhynchus kisutch*) (Le Luyer *et al.* 2017).
863
864   In total, our study of a panmictic species shows that despite no genetic differentiation a
865   portion of epigenetic variation is associated with local conditions and may contribute to
866   adaptation of individuals. Along with other studies focusing on asexual species or species
867   almost devoid of genetic variation (Angers *et al.* 2010; Berbel-Filho *et al.* 2019; Leung *et al.*
868   2016; Liu *et al.* 2019; Verhoeven & Preite 2014), or analyzing methylation-environment
869   association while controlling for genetic structure (Gugger *et al.* 2016; Wogan *et al.* 2020),
870   this provides evidence for the biological significance of epigenetic variation while controlling
871   for aspects of genetic variation.
872
873
874
875   **Acknowledgements**
876   We thank Annie Brandstrup for technical assistance, Russel Poole, Eric Feunteun, Françoise
877   Daverat, Gregory Maes and Håkan Wickström for providing samples and two anonymous
878   reviewers for constructive comments and suggestions. This work was funded by The Danish
879   Council for Independent Research, Natural Science (Grant No. 7014-00167B to MMH) and
880   MarGen II, an Interreg project under the Øresund-Kattegat-Skagerrak program.
881
882   **Author contributions**
883   MMH, SL and LB conceived and designed the study, SL conducted bioinformatics and
884   statistical analyses, ANBT and MWJ identified and validated hybrid individuals, MP, BJ and
885   JLC provided samples and information, SL and MMH wrote the manuscript with input from
886   all other authors. All authors read and approved the final version of the manuscript.
887
888   **Data Accessibility and Benefit Sharing**
889   Raw data files for both WGS and WGBS are available at NCBI (The National Center for
890   Biotechnology Information) with accession number PRJNA812038. Data files with SNPs
891   (VCF) and methylated sites and scripts used for analyzing the data are available through
892   DRYAD (https://doi.org/10.5061/dryad.q2bvq83nm) (Liu *et al.* 2022).
893
894   **References**
895

896   Adrian-Kalchhauser I, Sultan SE, Shama LNS*, et al.* (2020) Understanding 'Non-genetic'
897   Inheritance: Insights from Molecular-Evolutionary Crosstalk. *Trends in Ecology &*
898   *Evolution* **35**, 1078-1089.
899   Albert V, Jónsson B, Bernatchez L (2006) Natural hybrids in Atlantic eels (*Anguilla anguilla*,
900   *A. rostrata*): evidence for successful reproduction and fluctuating abundance in
901   space and time. *Molecular Ecology* **15**, 1903-1916.

902 Alexa A, Rahnenfuhrer J, Lengauer T (2006) Improved scoring of functional groups from
903   gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**, 1600-
904   1607.
905 Als TD, Hansen MM, Maes GE*, et al.* (2011) All roads lead to home: panmixia of European
906   eel in the Sargasso Sea. *Molecular Ecology* **20**, 1333-1346.
907 Anastasiadi D, Esteve-Codina A, Piferrer F (2018) Consistent inverse correlation between
908   DNA methylation of the first intron and gene expression across tissues and species.
909   *Epigenetics & Chromatin* **11**.
910 Anastasiadi D, Piferrer F (2020) A clockwork fish: Age prediction using DNA methylation-
911   based biomarkers in the European seabass. *Molecular Ecology Resources* **20**, 387-
912   397.
913 Anastasiadi D, Venney CJ, Bernatchez L, Wellenreuther M (2021) Epigenetic inheritance and
914   reproductive mode in plants and animals. *Trends Ecol Evol* **36**, 1124-1140.
915 Angers B, Castonguay E, Massicotte R (2010) Environmentally induced phenotypes and DNA
916   methylation: how to deal with unpredictable conditions until the next generation
917   and after. *Molecular Ecology* **19**, 1283-1295.
918 Angers B, Perez M, Menicucci T, Leung C (2020) Sources of epigenetic variation and their
919   applications in natural populations. *Evolutionary Applications* **13**, 1262-1278.
920 Aravin AA, Hannon GJ, Brennecke J (2007) The Piwi-piRNA pathway provides an adaptive
921   defense in the transposon arms race. *Science* **318**, 761-764.
922 Artemov AV, Mugue NS, Rastorguev SM*, et al.* (2017) Genome-wide DNA methylation
923   profiling reveals epigenetic adaptation of stickleback to marine and freshwater
924   conditions. *Molecular Biology and Evolution* **34**, 2203-2213.
925 Avise JC, Nelson WS, Arnold J*, et al.* (1990) The evolutionary genetic status of Icelandic eels.
926   *Evolution* **44**, 1254-1262.
927 Babin C, Gagnaire PA, Pavey SA, Bernatchez L (2017) RAD-seq reveals patterns of additive
928   polygenic variation caused by spatially-varying selection in the American eel
929   (*Anguilla rostrata*). *Genome Biology and Evolution* **9**, 2974-2986.
930 Baltazar-Soares M, Biastoch A, Harrod C*, et al.* (2014) Recruitment Collapse and Population
931   Structure of the European Eel Shaped by Local Ocean Current Dynamics. *Current
932   Biology* **24**, 104-108.
933 Baras E, Jeandrain D, Serouge B, Philippart JC (1998) Seasonal variations in time and space
934   utilization by radio-tagged yellow eels *Anguilla anguilla* (L.) in a small stream.
935   *Hydrobiologia* **372**, 187-198.
936 Barth JMI, Gubili C, Matschiner M*, et al.* (2020) Stable species boundaries despite ten million
937   years of hybridization in tropical eels. *Nature Communications* **11**, 1433.
938 Bateson P, Gluckman P, Hanson M (2014) The biology of developmental plasticity and the
939   Predictive Adaptive Response hypothesis. *Journal of Physiology-London* **592**, 2357-
940   2368.
941 Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful
942   approach to multiple testing. *Journal of the royal statistical society. Series B
943   (Methodological)* **57**, 289-300.
944 Berbel-Filho WM, Rodriguez-Barreto D, Berry N, De Leaniz CG, Consuegra S (2019)
945   Contrasting DNA methylation responses of inbred fish lines to different rearing
946   environments. *Epigenetics* **14**, 939-948.
947 Bonhommeau S, Castonguay M, Rivot E, Sabatie R, Le Pape O (2010) The duration of
948   migration of Atlantic *Anguilla* larvae. *Fish and Fisheries* **11**, 289-306.

949   Bornarel V, Lambert P, Briand C, *et al.* (2018) Modelling the recruitment of European eel
950   (*Anguilla anguilla*) throughout its European range. *Ices Journal of Marine Science* **75**,
951   541-552.
952   Bossdorf O, Richards CL, Pigliucci M (2008) Epigenetics for ecologists. *Ecology Letters* **11**,
953   106-115.
954   Bracamonte SE, Johnston PR, Knopf K, Monaghan MT (2019a) Experimental infection with
955   *Anguillicola crassus* alters immune gene expression in both spleen and head kidney
956   of the European eel (*Anguilla anguilla*). *Marine Genomics* **45**, 28-37.
957   Bracamonte SE, Johnston PR, Monaghan MT, Knopf K (2019b) Gene expression response to
958   a nematode parasite in novel and native eel hosts. *Ecology and Evolution* **9**, 13069-
959   13084.
960   Brenet F, Moh M, Funk P, *et al.* (2011) DNA methylation of the first exon Is tightly linked to
961   transcriptional silencing. *Plos One* **6**, e14524.
962   Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: A genetic theory of
963   morphological evolution. *Cell* **134**, 25-36.
964   Christensen KA, Le Luyer J, Chan MTT, *et al.* (2021) Assessing the effects of genotype-by-
965   environment interaction on epigenetic, transcriptomic, and phenotypic response in a
966   Pacific salmon. *G3-Genes Genomes Genetics* **11**, jkab021.
967   Côté CL, Castonguay M, Kalujnaia MS, Cramb G, Bernatchez L (2014) In absence of local
968   adaptation, plasticity and spatially varying selection rule: a view from genomic
969   reaction norms in a panmictic species (*Anguilla rostrata*). *Bmc Genomics* **15**, 403.
970   Côté CL, Castonguay M, Verreault G, Bernatchez L (2009) Differential effects of origin and
971   salinity rearing conditions on growth of glass eels of the American eel *Anguilla*
972   *rostrata*: implications for stocking programmes. *Journal of Fish Biology* **74**, 1934-
973   1948.
974   Côté CL, Gagnaire PA, Bourret V, *et al.* (2013) Population genetics of the American eel
975   (*Anguilla rostrata*): FST = 0 and North Atlantic Oscillation effects on demographic
976   fluctuations of a panmictic species. *Molecular Ecology* **22**, 1763-1776.
977   Côté CL, Pavey SA, Stacey JA, *et al.* (2015) Growth, female size, and sex ratio variability in
978   American eel of different origins in both controlled conditions and the wild:
979   implications for stocking programs. *Transactions of the American Fisheries Society*
980   **144**, 246-257.
981   Covelo-Soto L, Saura M, Moran P (2015) Does DNA methylation regulate metamorphosis?
982   The case of the sea lamprey (*Petromyzon marinus*) as an example. *Comparative*
983   *Biochemistry and Physiology B-Biochemistry & Molecular Biology* **185**, 42-46.
984   Davey AJH, Jellyman DJ (2005) Sex determination in freshwater eels and management
985   options for manipulation of sex. *Reviews in Fish Biology and Fisheries* **15**, 37-52.
986   De Meyer J, Maes GE, Dirks RP, Adriaens D (2017) Differential gene expression in narrow-
987   and broad-headed European glass eels (*Anguilla anguilla*) points to a transcriptomic
988   link of head shape dimorphism with growth rate and chemotaxis. *Molecular Ecology*
989   **26**, 3943-3953.
990   Deans C, Maggert KA (2015) What do you mean, "epigenetic"? *Genetics* **199**, 887-896.
991   Deaton AM, Bird A (2011) CpG islands and the regulation of transcription. *Genes &*
992   *Development* **25**, 1010-1022.
993   Dekker W (2003) On the distribution of the European eel (*Anguilla anguilla*) and its fisheries.
994   *Canadian Journal of Fisheries and Aquatic Sciences* **60**, 787-799.

995   Dubin MJ, Zhang P, Meng DZ*, et al.* (2015) DNA methylation in Arabidopsis has a genetic
996   basis and shows evidence of local adaptation. *Elife* **4**.
997   Duboule D (2007) The rise and fall of Hox gene clusters. *Development* **134**, 2549-2560.
998   Dupont C, Armant DR, Brenner CA (2009) Epigenetics: definition, mechanisms and clinical
999   perspective. *Seminars in Reproductive Medicine* **27**, 351-357.
1000 Egginton S (1986) Metamorphosis of the American eel, *Anguilla rostrata* Leseur .1. Changes
1001   in metabolism of skeletal muscle. *Journal of Experimental Zoology* **237**, 173-184.
1002 Ellerby DJ, Spierts ILY, Altringham JD (2001) Slow muscle power output of yellow- and silver-
1003   phase European eels (*Anguilla anguilla* L.): Changes in muscle performance prior to
1004   migration. *Journal of Experimental Biology* **204**, 1369-1379.
1005 Enbody ED, Pettersson ME, Sprehn CG*, et al.* (2021) Ecological adaptation in European eels
1006   is based on phenotypic plasticity. *Proc Natl Acad Sci U S A* **118**, e2022620118.
1007 Feil R, Fraga MF (2012) Epigenetics and the environment: emerging patterns and
1008   implications. *Nature Reviews Genetics* **13**, 97-109.
1009 Feng SH, Cokus SJ, Zhang XY*, et al.* (2010) Conservation and divergence of methylation
1010   patterning in plants and animals. *Proceedings of the National Academy of Sciences of*
1011   *the United States of America* **107**, 8689-8694.
1012 Flores KB, Wolschin F, Amdam GV (2013) The Role of methylation of DNA in environmental
1013   adaptation. *Integrative and Comparative Biology* **53**, 359-372.
1014 Forester BR, Lasky JR, Wagner HH, Urban DL (2018) Comparing methods for detecting
1015   multilocus adaptation with multivariate genotype-environment associations.
1016   *Molecular Ecology* **27**, 2215-2233.
1017 Gagnaire PA, Normandeau E, Côté C, Hansen MM, Bernatchez L (2012) The genetic
1018   consequences of spatially varying selection in the panmictic American eel (*Anguilla*
1019   *rostrata*). *Genetics* **190**, 725-733.
1020 Gapp K, von Ziegler L, Tweedie-Cullen RY, Mansuy IM (2014) Early life epigenetic
1021   programming and transmission of stress-induced traits in mammals. *Bioessays* **36**,
1022   491-502.
1023 Geffroy B, Bardonnet A (2016) Sex differentiation and sex determination in eels:
1024   consequences for management. *Fish and Fisheries* **17**, 375-398.
1025 Gegner J, Vogel H, Billion A, Forster F, Vilcinskas A (2021) Complete metamorphosis in
1026   *Manduca sexta* involves specific changes in DNA methylation patterns. *Frontiers in*
1027   *Ecology and Evolution* **9**.
1028 Gienapp P, Teplitsky C, Alho JS, Mills JA, Merilä J (2008) Climate change and evolution:
1029   disentangling environmental and genetic responses. *Molecular Ecology* **17**, 167-178.
1030 Goodier JL (2016) Restricting retrotransposons: a review. *Mobile DNA* **7**, 16.
1031 Gore AV, Tomins KA, Iben J*, et al.* (2018) An epigenetic mechanism for cavefish eye
1032   degeneration. *Nature Ecology & Evolution* **2**, 1155–1160.
1033 Greenberg MVC, Bourc'his D (2019) The diverse roles of DNA methylation in mammalian
1034   development and disease. *Nature Reviews Molecular Cell Biology* **20**, 590-607.
1035 Gronostajski RM (2000) Roles of the NFI/CTF gene family in transcription and development.
1036   *Gene* **249**, 31-45.
1037 Gugger PF, Fitz-Gibbon S, Pellegrini M, Sork VL (2016) Species-wide patterns of DNA
1038   methylation variation in *Quercus lobata* and their association with climate gradients.
1039   *Molecular Ecology* **25**, 1665-1680.
1040 Head JA (2014) Patterns of DNA Methylation in Animals: An Ecotoxicological Perspective.
1041   *Integrative and Comparative Biology* **54**, 77-86.

1042 Heckwolf MJ, Meyer BS, Hasler R*, et al.* (2020) Two different epigenetic information
1043   channels in wild three-spined sticklebacks are involved in salinity adaptation. *Science*
1044   *Advances* **6**, eaaz1138.
1045 Henkel CV, Burgerhout E, de Wijze DL*, et al.* (2012) Primitive duplicate Hox clusters in the
1046   European eel's genome. *Plos One* **7**, e32231.
1047 Horvath S (2013) DNA methylation age of human tissues and cell types. *Genome Biology* **14**,
1048   3156.
1049 Horvath S, Raj K (2018) DNA methylation-based biomarkers and the epigenetic clock theory
1050   of ageing. *Nat Rev Genet* **19**, 371-384.
1051 Issa JP (2014) Aging and epigenetic drift: a vicious cycle. *Journal of Clinical Investigation* **124**,
1052   24-29.
1053 Jablonka E (2017) The evolutionary implications of epigenetic inheritance. *Interface Focus* **7**,
1054   20160135.
1055 Jablonka E, Raz G (2009) Transgenerational epigenetic inheritance: prevalence, mechanisms,
1056   and implications for the study of heredity and evolution. *Quarterly Review of Biology*
1057   **84**, 131-176.
1058 Jacobsen MW, Pujolar JM, Bernatchez L*, et al.* (2014) Genomic footprints of speciation in
1059   Atlantic eels (*Anguilla anguilla* and *A. rostrata*). *Molecular Ecology* **23**, 4785–4798.
1060 Jacobsen MW, Smedegaard L, Sorensen SR*, et al.* (2017) Assessing pre- and post-zygotic
1061   barriers between North Atlantic eels (*Anguilla anguilla* and *A. rostrata*). *Heredity*
1062   **118**, 266-275.
1063 Johnson AA, Akman K, Calimport SRG*, et al.* (2012) The role of DNA methylation in aging,
1064   rejuvenation, and age-related disease. *Rejuvenation Research* **15**, 483-494.
1065 Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond.
1066   *Nature Reviews Genetics* **13**, 484-492.
1067 Jonsson B, Jonsson N, Hansen MM (2022) Knock-on effects of environmental influences
1068   during embryonic development of ectothermic vertebrates. *The Quarterly Review of*
1069   *Biology* **97**, 95-139.
1070 Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters* **7**, 1225-
1071   1241.
1072 Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and
1073   genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* **37**, 907-915.
1074 Kleckner RC, McCleave JD, Wippelhauser GS (1983) Spawning of American eel, *Anguilla*
1075   *rostrata*, relative to thermal fronts in the Sargasso Sea. *Environmental Biology of*
1076   *Fishes* **9**, 289-293.
1077 Kovaka S, Zimin AV, Pertea GM*, et al.* (2019) Transcriptome assembly from long-read RNA-
1078   seq alignments with StringTie2. *Genome Biology* **20**, 278.
1079 Krueger F, Andrews SR (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-
1080   Seq applications. *Bioinformatics* **27**, 1571-1572.
1081 Kyono Y, Raj S, Sifuentes CJ*, et al.* (2020) DNA methylation dynamics underlie metamorphic
1082   gene regulation programs in *Xenopus* tadpole brain. *Developmental Biology* **462**,
1083   180-196.
1084 Laporte M, Le Luyer J, Rougeux C*, et al.* (2019) DNA methylation reprogramming, TE
1085   derepression, and postzygotic isolation of nascent animal species. *Science Advances*
1086   **5**, eaaw1644.
1087 Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation
1088   patterns in plants and animals. *Nature Reviews Genetics* **11**, 204-220.

1089   Le Luyer J, Laporte M, Beacham TD, *et al.* (2017) Parallel epigenetic modifications induced
1090   by hatchery rearing in a Pacific salmon. *Proceedings of the National Academy of*
1091   *Sciences of the United States of America* **114**, 12964-12969.
1092   Lea AJ, Altmann J, Alberts SC, Tung J (2016) Resource base influences genome-wide DNA
1093   methylation levels in wild baboons (*Papio cynocephalus*). *Molecular Ecology* **25**,
1094   1681-1696.
1095   Lecomte-Finiger R (1992) Growth history and age at recruitment of European glass eels
1096   (*Anguilla anguilla*) as revealed by otolith microstructure. *Marine Biology* **114**, 205-
1097   210.
1098   Legendre P, Legendre L (2012) *Numerical ecology*, Third English edition. edn. Elsevier,
1099   Amsterdam.
1100   Leinonen T, McCairns RJS, O'Hara RB, Merila J (2013) Q(ST)-F-ST comparisons: evolutionary
1101   and ecological insights from genomic heterogeneity. *Nature Reviews Genetics* **14**,
1102   179-190.
1103   Leitwein M, Laporte M, Le Luyer J, *et al.* (2021) Epigenomic modifications induced by
1104   hatchery rearing persist in germ line cells of adult salmon after their oceanic
1105   migration. *Evolutionary Applications* **14**, 2402-2413.
1106   Leung C, Breton S, Angers B (2016) Facing environmental predictability with different
1107   sources of epigenetic variation. *Ecology and Evolution* **6**, 5234-5245.
1108   Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler
1109   transform. *Bioinformatics* **25**, 1754-1760.
1110   Li H, Handsaker B, Wysoker A, *et al.* (2009) The Sequence Alignment/Map format and
1111   SAMtools. *Bioinformatics* **25**, 2078-2079.
1112   Lind MI, Spagopoulou F (2018) Evolutionary consequences of epigenetic inheritance.
1113   *Heredity* **121**, 205-209.
1114   Liu S, Tengstedt ANB, Jacobsen MW, *et al.* (2022) Dataset for: Genome-wide methylation in
1115   the panmictic European eel (*Anguilla anguilla*). *DRYAD*,
1116   https://doi.org/10.5061/dryad.q5062bvq5083nm.
1117   Liu SL, Aagaard A, Bechsgaard J, Bilde T (2019) DNA Methylation Patterns in the Social
1118   Spider, *Stegodyphus dumicola*. *Genes* **10**, 137.
1119   Lopez-Olmeda JF, Lopez-Garcia I, Sanchez-Muros MJ, *et al.* (2012) Daily rhythms of digestive
1120   physiology, metabolism and behaviour in the European eel (*Anguilla anguilla*).
1121   *Aquaculture International* **20**, 1085-1096.
1122   Madeira F, Park YM, Lee J, *et al.* (2019) The EMBL-EBI search and sequence analysis tools
1123   APIs in 2019. *Nucleic Acids Research* **47**, W636-W641.
1124   Merondun J, Murray DL, Shafer ABA (2019) Genome-scale sampling suggests cryptic
1125   epigenetic structuring and insular divergence in Canada lynx. *Molecular Ecology* **28**,
1126   3186-3196.
1127   Metzger DCH, Schulte PM (2018) The DNA methylation landscape of stickleback reveals
1128   patterns of sex chromosome evolution and effects of environmental salinity.
1129   *Genome Biology and Evolution* **10**, 775-785.
1130   Michalak P (2009) Epigenetic, transposon and small RNA determinants of hybrid
1131   dysfunctions. *Heredity* **102**, 45-50.
1132   Moore LD, Le T, Fan GP (2013) DNA methylation and its basic function.
1133   *Neuropsychopharmacology* **38**, 23-38.

1134   Munk P, Hansen MM, Maes GE, *et al.* (2010) Oceanic fronts in the Sargasso Sea control the
1135   early life and drift of Atlantic eels. *Proceedings of the Royal Society Series B:*
1136   *Biological Sciences* **277**, 3593-3599.
1137   Nakamura R, Tsukahara T, Qu W, *et al.* (2014) Large hypomethylated domains serve as
1138   strong repressive machinery for key developmental genes in vertebrates.
1139   *Development* **141**, 2568-2580.
1140   Nikolic N, Liu SL, Jacobsen MW, *et al.* (2020) Speciation history of European (*Anguilla*
1141   *anguilla*) and American eel (*A. rostrata*), analysed using genomic data. *Molecular*
1142   *Ecology* **29**, 565-577.
1143   O'Neill RJW, O'Neill MJ, Graves JAM (1998) Undermethylation associated with retroelement
1144   activation and chromosome remodelling in an interspecific mammalian hybrid.
1145   *Nature* **393**, 68-72.
1146   Oksanen J, Kindt R, Legendre P, O'Hara RB (2008) vegan: Community ecology package.
1147   Version 1.7–81, http://cran.r-project.org/.
1148   Palm S, Dannewitz J, Prestegaard T, Wickstrom H (2009) Panmixia in European eel revisited:
1149   no genetic difference between maturing adults from southern and northern Europe.
1150   *Heredity* **103**, 82-89.
1151   Pavey SA, Gaudin J, Normandeau E, *et al.* (2015) RAD sequencing highlights polygenic
1152   discrimination of habitat ecotypes in the panmictic American eel. *Current Biology* **25**,
1153   1666-1671.
1154   Pujol B, Wilson AJ, Ross RIC, Pannell JR (2008) Are Q_{ST}-F_{ST} comparisons for natural
1155   populations meaningful? *Molecular Ecology* **17**, 4782-4785.
1156   Pujolar JM, Jacobsen MW, Als TD, *et al.* (2014a) Assessing patterns of hybridization between
1157   North Atlantic eels using diagnostic single-nucleotide polymorphisms. *Heredity* **112**,
1158   627-637.
1159   Pujolar JM, Jacobsen MW, Als TD, *et al.* (2014b) Genome-wide single-generation signatures
1160   of local selection in the panmictic European eel. *Molecular Ecology* **23**, 2514-2528.
1161   Pujolar JM, Jacobsen MW, Bekkevold D, *et al.* (2015) Signatures of natural selection
1162   between life cycle stages separated by metamorphosis in European eel. *Bmc*
1163   *Genomics* **16**.
1164   Pujolar JM, Jacobsen MW, Frydenberg J, *et al.* (2013) A resource of genome-wide single-
1165   nucleotide polymorphisms generated by RAD tag sequencing in the critically
1166   endangered European eel. *Molecular Ecology Resources* **13**, 706-716.
1167   Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic
1168   features. *Bioinformatics* **26**, 841-842.
1169   R Core Team (2018) R: A language and environment for statistical computing. R Foundation
1170   for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
1171   Rey O, Eizaguirre C, Angers B, *et al.* (2020) Linking epigenetics and biological conservation:
1172   Towards a conservation epigenetics perspective. *Functional Ecology* **34**, 414-427.
1173   Rhie A, McCarthy SA, Fedrigo O, *et al.* (2021) Towards complete and error-free genome
1174   assemblies of all vertebrate species. *Nature* **592**, 737-746.
1175   Richards EJ (2006) Opinion - Inherited epigenetic variation - revisiting soft inheritance.
1176   *Nature Reviews Genetics* **7**, 395-U392.
1177   Robinson JT, Thorvaldsdottir H, Winckler W, *et al.* (2011) Integrative genomics viewer.
1178   *Nature Biotechnology* **29**, 24-26.

1179 Rodriguez-Martinez S, Galvan I (2019) Exposure to a competitive social environment
1180   activates an epigenetic mechanism that limits pheomelanin synthesis in zebra
1181   finches. *Molecular Ecology* **28**, 3698-3708.
1182 Sathyendranath S, Brewin RJW, Brockmann C, *et al.* (2019) An ocean-colour time series for
1183   use in climate studies: the experience of the ocean-colour climate change initiative
1184   (OC-CCI). *Sensors* **19**, 4285.
1185 Saxonov S, Berg P, Brutlag DL (2006) A genome-wide analysis of CpG dinucleotides in the
1186   human genome distinguishes two distinct classes of promoters. *Proceedings of the
1187   National Academy of Sciences of the United States of America* **103**, 1412-1417.
1188 Schmidt J (1923) The breeding places of the eel. *Philosophical Transactions of the Royal
1189   Society B: Biological Sciences* **211**, 179-208.
1190 Schmitz RJ, Schultz MD, Urich MA, *et al.* (2013) Patterns of population epigenomic diversity.
1191   *Nature* **495**, 193-198.
1192 Schulz WA, Alexa A, Jung V, *et al.* (2007) Factor interaction analysis for chromosome 8 and
1193   DNA methylation alterations highlights innate immune response suppression and
1194   cytoskeletal changes in prostate cancer. *Molecular Cancer* **6**, 14.
1195 Simkin A, Wong A, Poh YP, Theurkauf WE, Jensen JD (2013) Recurrent and recent selective
1196   sweeps in the piRNA pathway. *Evolution* **67**, 1081-1090.
1197 Skvortsova K, Iovino N, Bogdanovic O (2018) Functions and mechanisms of epigenetic
1198   inheritance in animals. *Nature Reviews Molecular Cell Biology* **19**, 774-790.
1199 Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of
1200   the genome. *Nature Reviews Genetics* **8**, 272-285.
1201 Stajic D, Perfeito L, Jansen LET (2019) Epigenetic gene silencing alters the mechanisms and
1202   rate of evolutionary adaptation. *Nature Ecology & Evolution* **3**, 491-498.
1203 Sun D, Layman TS, Jeong H, *et al.* (2021) Genome-wide variation in DNA methylation linked
1204   to developmental stage and chromosomal suppression of recombination in white-
1205   throated sparrows. *Molecular Ecology* **30**, 3453-3467.
1206 Taudt A, Colome-Tatche M, Johannes F (2016) Genetic sources of population epigenomic
1207   variation. *Nature Reviews Genetics* **17**, 319-332.
1208 Teh AL, Pan H, Chen L, *et al.* (2014) The effect of genotype and in utero environment on
1209   interindividual variation in neonate DNA methylomes. *Genome Research* **24**, 1064-
1210   1074.
1211 Tesch F (2003) *The Eel* Blackwell Science Ltd, Oxford.
1212 Trautner JH, Reiser S, Blancke T, Unger K, Wysujack K (2017) Metamorphosis and transition
1213   between developmental stages in European eel (*Anguilla anguilla*, L.) involve
1214   epigenetic changes in DNA methylation patterns. *Comparative Biochemistry and
1215   Physiology D-Genomics & Proteomics* **22**, 139-145.
1216 Ungerer MC, Strakosh SC, Zhen Y (2006) Genome expansion in three hybrid sunflower
1217   species is associated with retrotransposon proliferation. *Current Biology* **16**, R872-
1218   R873.
1219 Verhoeven KJF, Preite V (2014) Epigenetic variation in asexually reproducing organisms.
1220   *Evolution* **68**, 644-655.
1221 Verhoeven KJF, Vonholdt BM, Sork VL (2016) Epigenetics in ecology and evolution: what we
1222   know and what we need to know. *Molecular Ecology* **25**, 1631-1638.
1223 Wellband K, Roth D, Linnansaari T, Curry RA, Bernatchez L (2021) Environment-driven
1224   reprogramming of gamete DNA methylation occurs during maturation and is
1225   transmitted intergenerationally in Atlantic Salmon. *G3-Genes Genomes Genetics* **11**.

1226   Williams GC, Koehn RK, Mitton JB (1973) Genetic differentiation without isolation in
1227   American eel, *Anguilla rostrata*. *Evolution* **27**, 192-204.
1228   Wogan GOU, Yuan ML, Mahler DL, Wang IJ (2020) Genome-wide epigenetic isolation by
1229   environment in a widespread Anolis lizard. *Molecular Ecology* **29**, 40-55.
1230   Zhang Y, Kutateladze TG (2018) Diet and the epigenome. *Nature Communications* **9**, 3375.
1231
1232

1233 **Figure legends.**

1234

1235 Fig. 1. Map showing the sample localities of the analyzed European eels. The colour coding
1236 for the localities applies to all the figures in this paper unless specified otherwise.

1237

1238 Fig. 2. Principal Component Analysis (PCA) based on a) SNPs and b) methylation,
1239 respectively, showing the relationships among individuals, with numbers in parentheses
1240 indicating the percentage of variance explained by each PC. The colours indicate the
1241 sampling localities, as shown in Fig. 1. Squares denote adult eels and circles denote glass
1242 (juvenile) eels. c) Distribution of methylation level across CpG sites. The blue lines represent
1243 single individuals. The red line represents the average across all individuals.

1244

1245 Fig. 3. Methylation in genomic categories. Each line represents one individual, and the colour
1246 indicates sampling location as specified in Fig. 1. a) Average methylation level of all the
1247 sequenced CpG sites in each category, and b) b) the standard deviation. c) Inter-individual
1248 variation of methylation in genomic categories are shown in the lower panel. Each violin
1249 represents the distribution of standard deviation of methylation level calculated across
1250 individuals per CpG site. Yellow dots indicate medians.

1251

1252 Fig. 4. Correlation between methylation and gene expression, the latter based on gene
1253 expression data from Bracamonte *et al.* (2019a) and Bracamonte *et al.* (2019b). Methylation
1254 was divided into three categories, i.e., overall, hypomethylated (methylation level <= 0.2) and
1255 hypermethylated (> 0.2). We added a value 1e-7 to the expression level to visualize the
1256 unexpressed genes in the logarithmic scale. a) Correlation between exon_1 methylation and
1257 gene expression. b) Correlation between intron_1 methylation and gene expression. c)
1258 Correlation between promoter_1 methylation and gene expression. d) Correlation between
1259 promoter_2 methylation and gene expression.

1260

1261 Fig. 5. a) Redundancy analyses (RDA) of SNPs and methylation (CpGs), respectively, using
1262 location as explanatory variable. b) Genomic distribution of RDA outliers for SNPs and
1263 methylation (CpGs), respectively. Genes associated with high density regions (peaks) are
1264 indicated.

1265

1266 Fig. 6. a) Redundancy analyses (RDA) of SNPs and methylation (CpGs), respectively, using
1267 environmental parameters (mean day length [MDL], sea surface temperature [SST30] and
1268 Chlorophyll concentration [Chlorophyll]) as explanatory variables. b) Genomic distribution
1269 of RDA outliers for SNPs and methylation (CpGs), respectively. Genes associated with high
1270 density regions (peaks) are indicated.

1271

1272 Fig. 7. Analysis of methylation differences between life stages. a) Joint distribution of P_{ST}
1273 and methylation difference between adults and glass eels. The methylation difference was
1274 calculated as the average methylation level of adults minus the average methylation level of
1275 glass eels. b) Distribution of neighbouring distance (logarithmically scaled) among CpG
1276 outliers related to the developmental stage. The left panel represents outliers hypomethylated
1277 in adults, and the right panel outliers hypermethylated in adults. c) Outlier enrichment in
1278 genomic categories. Positive values indicate overrepresentation of the outliers in the category
1279 and negative values implicate underrepresentation. Asterisks indicate significance (alpha =
1280 0.05). d) Genome-wide distributions of hypo- and hypermethylated outliers, respectively.
1281 Genes associated with high density regions (peaks) are indicated.

1282

1283   Fig. 8. Analysis of methylation differences between hybrids and non-hybrids. a) Joint
1284   distribution of P_{ST} and methylation difference between hybrids and non-hybrids eels. The
1285   methylation difference was calculated as the average methylation level of hybrids minus the
1286   average methylation level of non-hybrids. b) Distribution of neighbouring distance
1287   (logarithmically scaled) among CpG outliers related to hybrid or non-hybrid status. The left
1288   panel represents outliers hypomethylated in hybrids, and the right panel outliers
1289   hypermethylated in hybrids. c) Outlier enrichment in genomic categories. Positive values
1290   indicate overrepresentation of the outliers in the category and negative values implicate
1291   underrepresentation. Asterisks indicate significance (alpha = 0.05). d) Genome-wide
1292   distributions of hypo- and hypermethylated outliers, respectively. Genes associated with high
1293   density regions (peaks) are indicated.

Table 1. Overview of analyzed individuals and environmental parameters. SST30 denotes mean sea surface temperature (degrees C), MDL mean day length (in minutes) and Chlorophyll mean Chlorophyll concentration (mg per m^3) across 30 days prior to sampling. N(WGS) denotes sample size for whole genome sequencing and N(WGBS) sample size for whole genome bisulphite sequencing.

Location	Abbreviation	Time of sampling	Latitude	Longitude	SST30	MDL	Chlorophyll	N(WGS)	N(WGBS)
Glass eels									
Ellidaar, Iceland	Ell	2016-05-16	64.1268	-21.8419	6.26	1015.7	1.371	5	5
Stokkseyri, Iceland	Sto	2016-05-15	63.8324	-21.0603	7.13	1000.2	1.905	6	6
Burrishoole, Ireland	Bur	2008-03-15	53.8989	-9.5742	9.23	640.6	4.48	5	3
Ringhals, Sweden	Rin	2005-03-14	57.2633	12.1025	4.12	633.9	18.175	5	5
Gironde, France	Gir	2008-04-16	45.1193	-0.693	11.33	766.9	11.055	5	3
Valencia, Spain	Val	2010-01-15	39.4724	-0.3107	15.2	562.6	3.17	6	3
Oved Sebou, Morocco	Seb	2001-04-28	34.2698	-6.654	17.5	776.3	2.665	7	5
Adults									
Burrishoole, Ireland	Bur	2010	53.8989	-9.5742	NA	NA	NA	7	3
Valencia, Spain	Val	2012	39.4724	-0.3107	NA	NA	NA	4	3
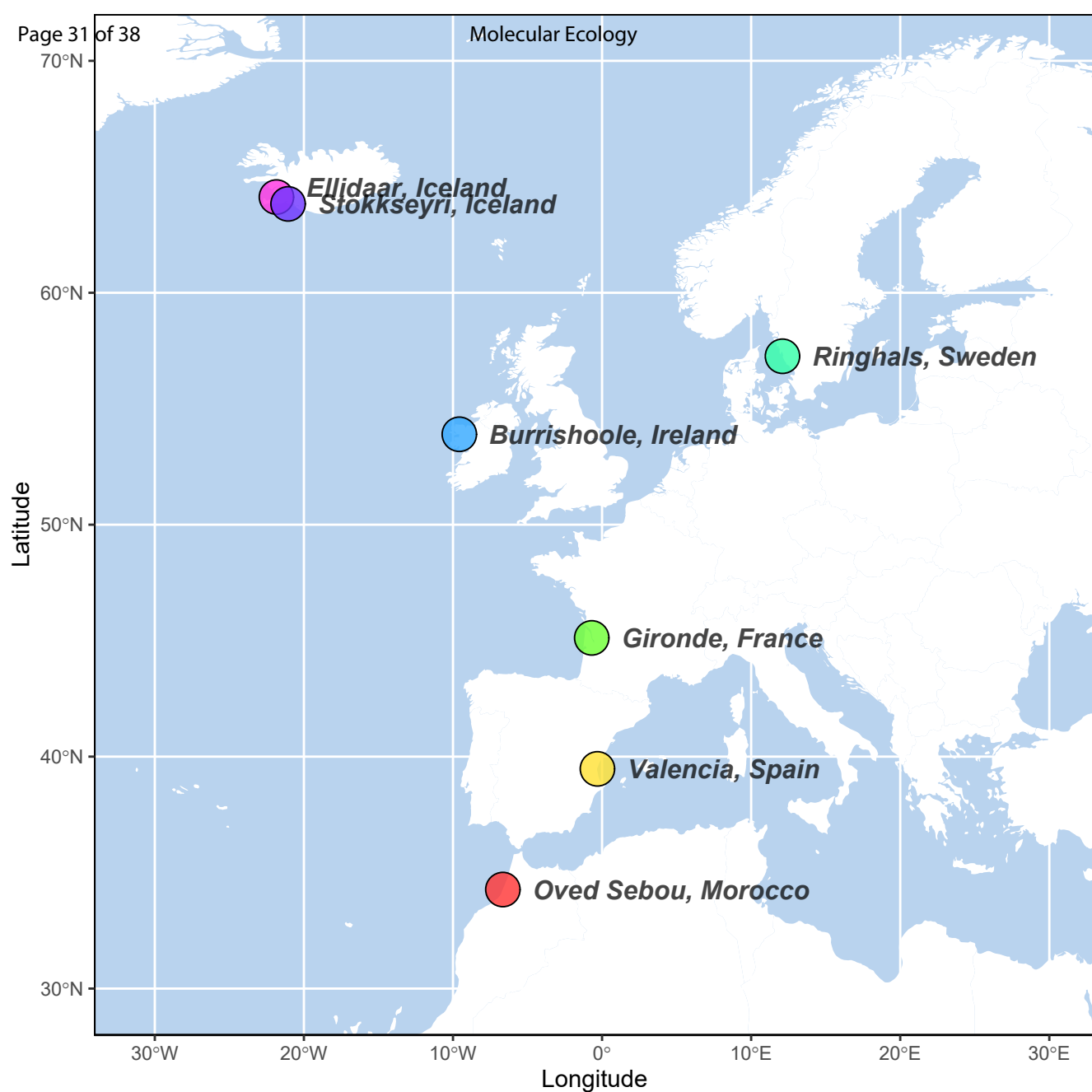
a) **PCA of SNPs**
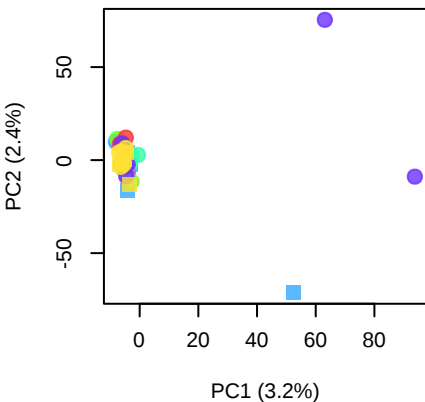
b) **PCA of methylation**
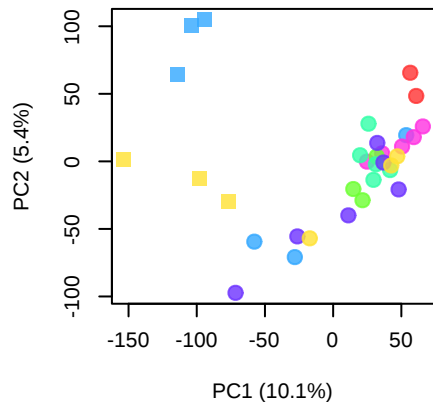
c) **Methylation level distribution**
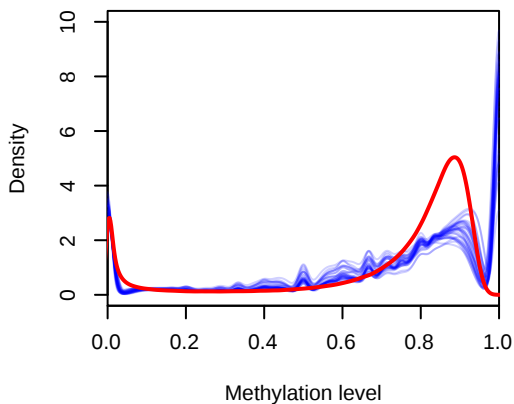
Methylation of genomic categories

Molecular Ecology

a)
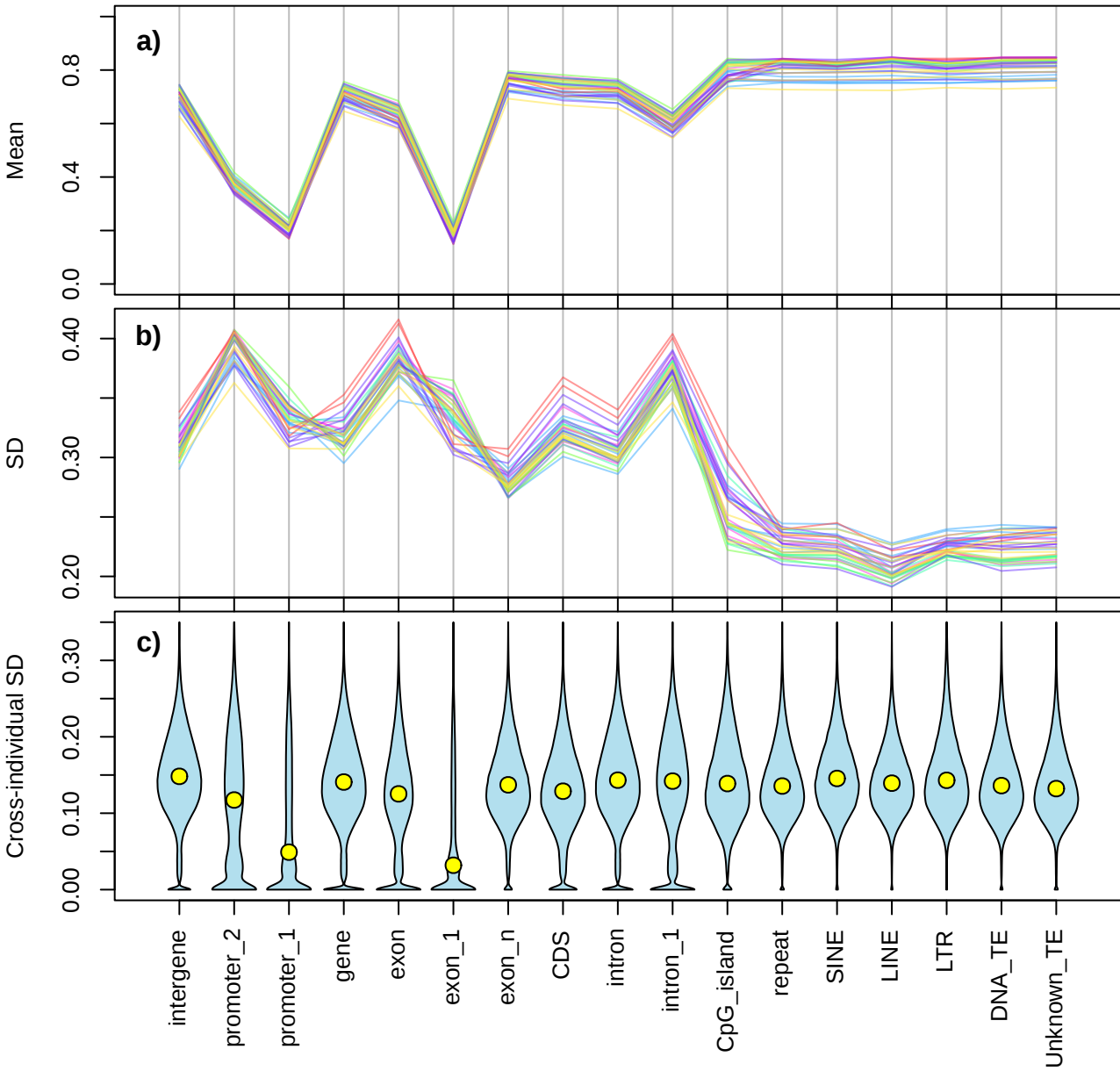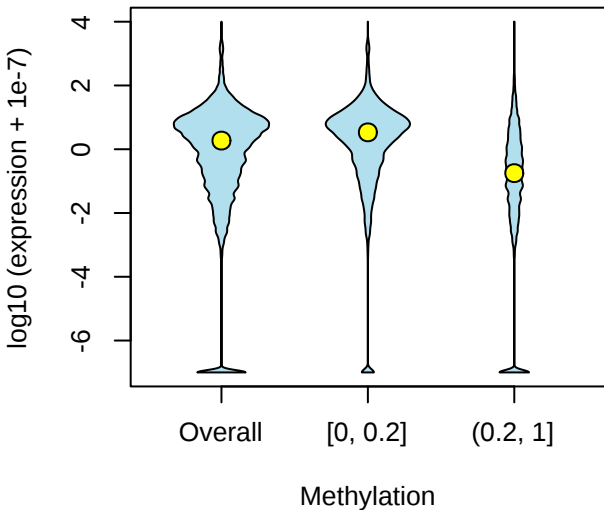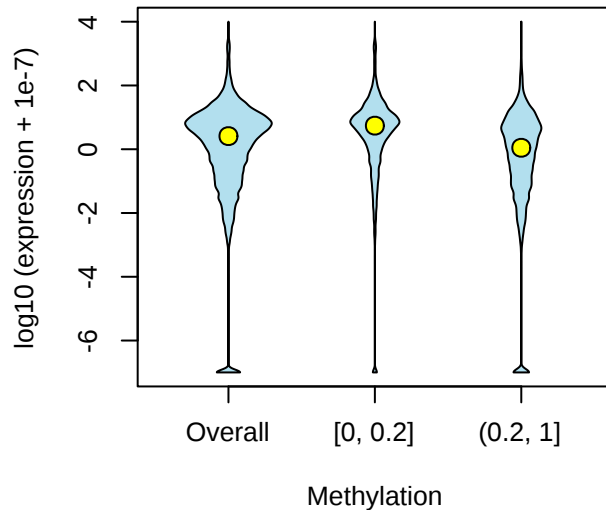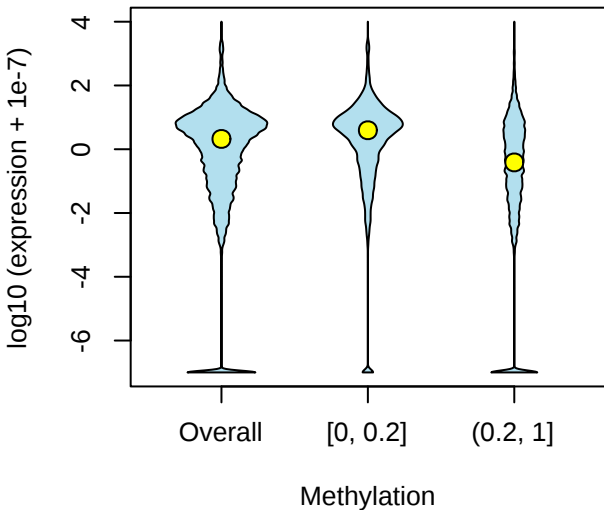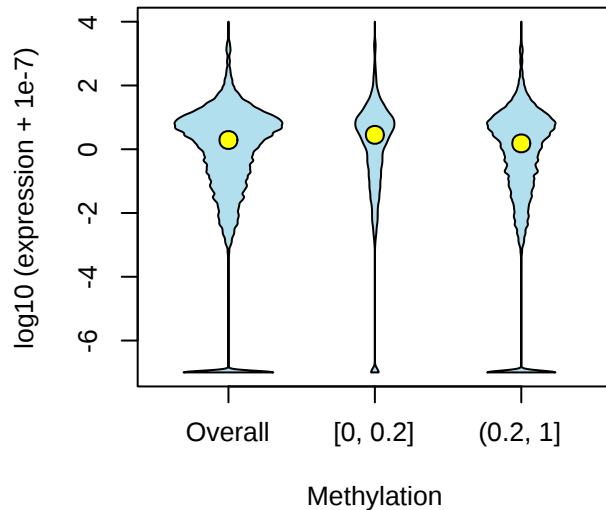
RDA of SNPs
(R^2: 0.0006; p-value: 0.0028)
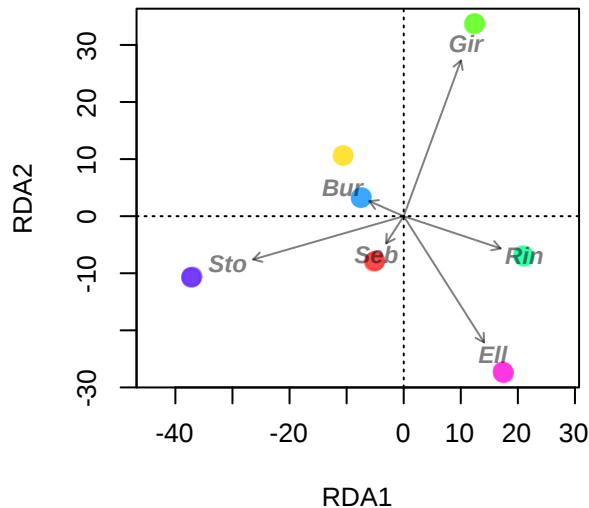
RDA of methylation
(R^2: 0.0298; p-value: 0.0004)

b)

Outlier distribution

a)
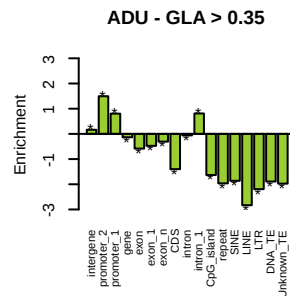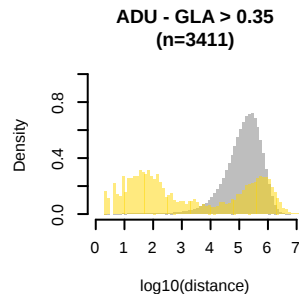
RDA of SNPs
(R^2: 0.0003; p-value: 0.0396)

RDA of methylation
(R^2: 0.0136; p-value: 0.0056)

b)

Outlier distribution

a) **Adults vs. glass eels**

b) **ADU - GLA < -0.35 (n=10767)** / **ADU - GLA > 0.35 (n=3411)**

c) **ADU - GLA < -0.35** / **ADU - GLA > 0.35**
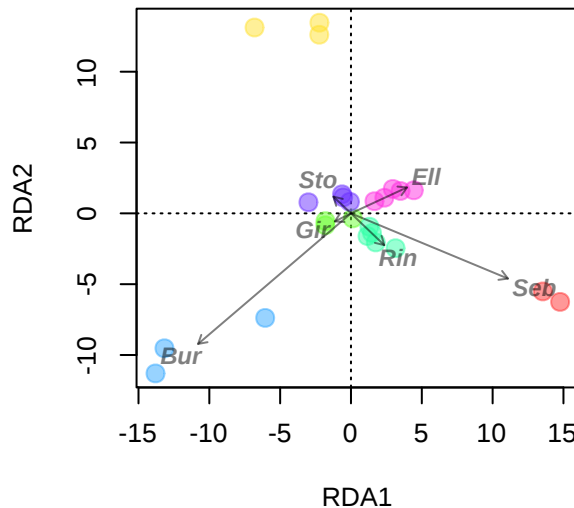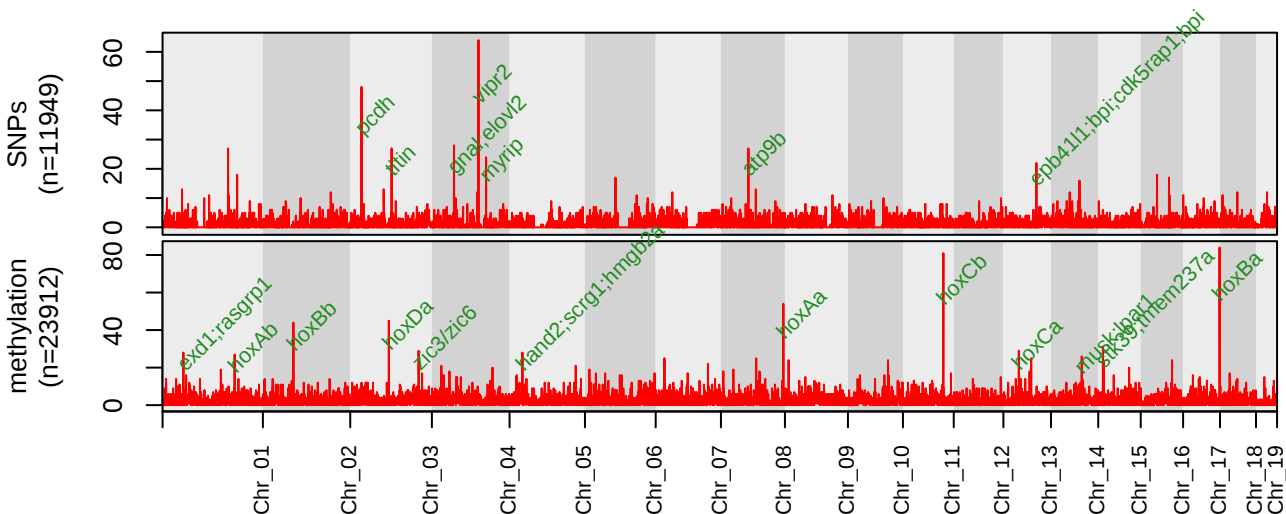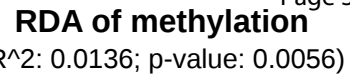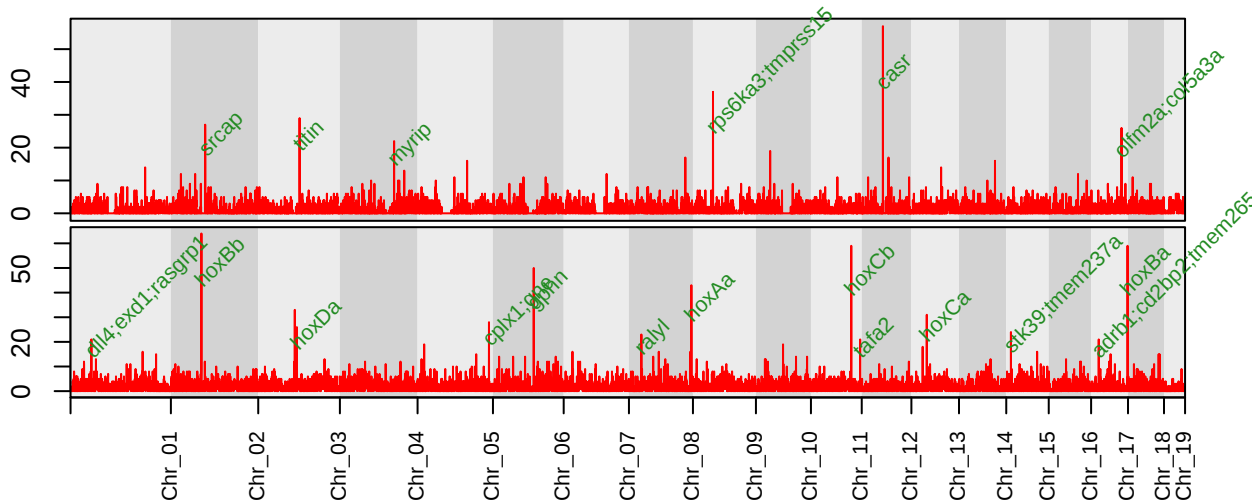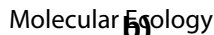
d) **Outlier distribution**
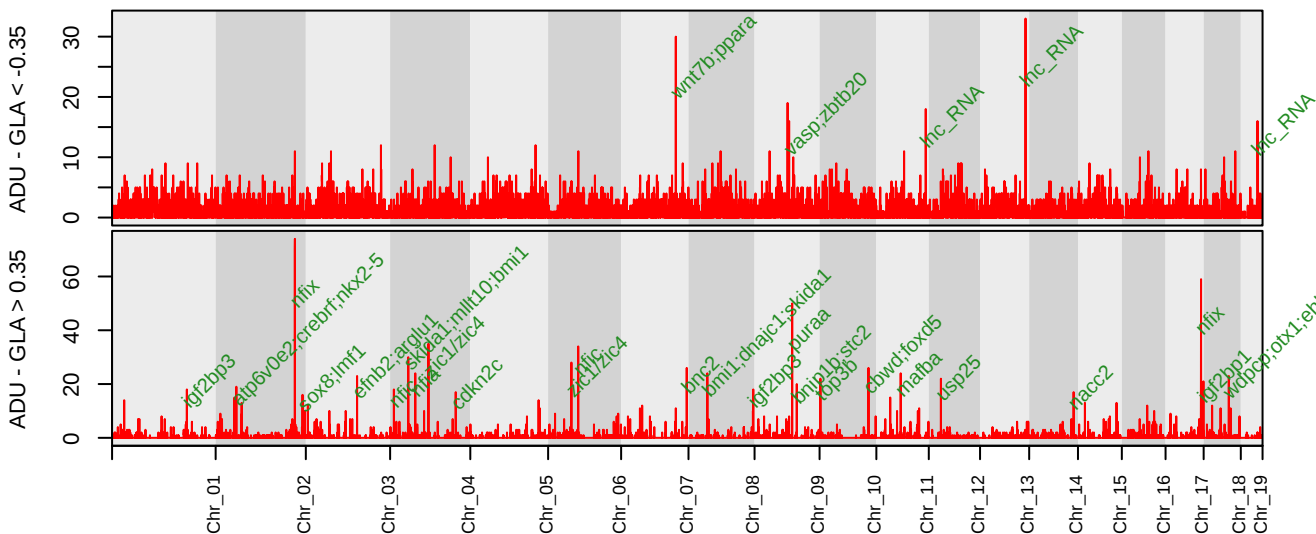
a) **Hybrids vs. non-hybrids**
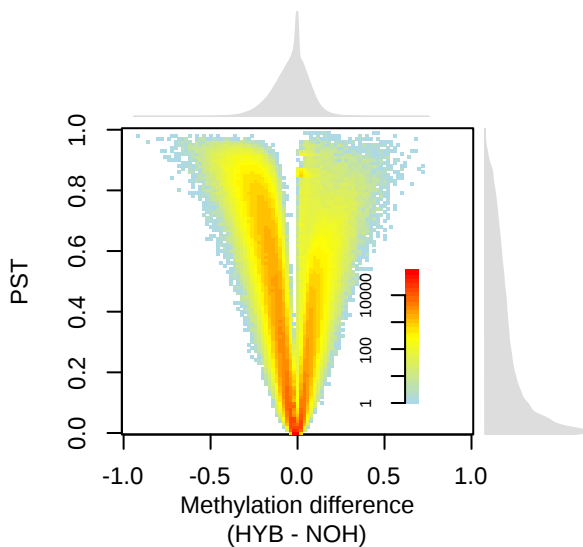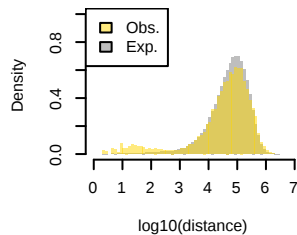
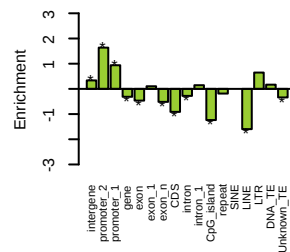b) HYB - NOH < -0.35 (n=8577)   HYB - NOH > 0.35 (n=667)
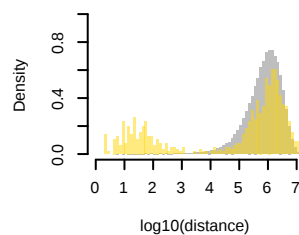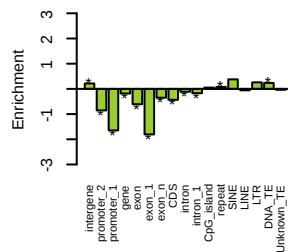
c) HYB - NOH < -0.35   HYB - NOH > 0.35

d) **Outlier distribution**