Title

Epigenetics and island-mainland divergence in an insectivorous small mammal

Authors

Marie-Laurence Cossette[1*], Donald T. Stewart[2], Amin Haghani[3], Joseph A. Zoller[4], Aaron B.A. Shafer[1,5*], Steve Horvath[3,4,6]

Affiliations:

[1]Dept. of Environmental Life Sciences Graduate Program, Trent University, Peterborough, ON, Canada

[2]Dept. of Biology, Acadia University, Wolfville, NS, Canada

[3]Dept. of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

[4]Dept. of Biostatistics, Fielding School of Public Health, University of California, Los Angeles, CA, USA

[5]Dept. of Forensic Science, Trent University, Peterborough, ON, Canada

[6]Altos Labs, San Diego, CA, USA

[*]Corresponding authors: mcossette@trentu.ca; aaronshafer@trentu.ca

Abstract

Geographically isolated populations, specifically island-mainland counterparts, tend to exhibit phenotypic variation in many species. The so-called island syndrome occurs when different environmental pressures lead to insular divergence from mainland populations. This phenomenon can be seen in an island population of Nova Scotia masked shrews (*Sorex cinereus*), which have developed a specialized feeding habit and digestive enzyme compared to their mainland counterparts. Epigenetic modifications, such as DNA methylation (DNAm), can impact phenotypes by altering gene expression without changing the DNA sequence. Here, we used a *de novo* masked shrew genome assembly and a mammalian methylation array profiling 37 thousand conserved CpGs to investigate morphological and DNA methylation patterns between island and mainland populations. Island shrews were morphologically and epigenetically different than their mainland counterparts, exhibiting a smaller body size. A gene ontology enrichment analyses of differentially methylated CpGs implicated developmental and digestive system related pathways. Based on our shrew epigenetic clock, island shrews might also be aging faster than their mainland counterparts. This study provides novel insight on phenotypic and epigenetic divergence in island-mainland mammal populations and suggests an underlying role of methylation in island-mainland divergence.

Keywords

epigenetic clock, DNA methylation, masked shrew, genome assembly, island divergence, island syndrome

Introduction

Insular populations tend to exhibit divergent traits compared to their continental counterparts (Adler & Levins, 1994; Foster, 1964; Goltsman et al., 2005; Meröndun et al., 2019; Novosolov et al., 2013; van Valen, 1973). This phenomenon is known as the island syndrome and refers to ecological and morphological differences between island and mainland populations (Baeckens & van Damme, 2020). Altered predation, competition, environment, and geographic isolation are thought to lead to phenotypic differences between mainland and insular populations (Baeckens & van Damme, 2020; Benítez-López et al., 2021; Lomolino et al., 2013).

Relatively little is known about the specific mechanisms causing phenotypic changes in island-mainland scenarios, whether it be genetic drift, natural selection, or plasticity (Baeckens & van Damme, 2020; Benítez-López et al., 2021). Miller et al. (2021) found heritable variations for deer mice for both mass and length associated with outlier loci between island-mainland populations and suggested that natural selection underlies such phenotypic divergence. Island-mainland populations, however, do not always exhibit clear signal of nuclear genetic divergence with divergent phenotypes (e.g., Meröndun et al., 2019; Vernaz, Malinsky, et al., 2021). In contrast, different epigenetic signatures, which include chemical modifications that regulate gene expression (Al Aboud et al., 2021), have been associated with substantial levels of divergence between species and populations exhibiting different phenotypes (Meröndun et al., 2019; Vernaz, Malinsky, et al., 2021) and unlike DNA, epigenetic modifications can respond quickly to environmental pressures (Anastasiadi et al., 2021). For example, differential methylation patterns associated with phenotypic traits have been noted for body size in sheep (*Ovis ammon*) (Cao et al., 2015) and lynx (*Lynx canadensis*) (Meröndun et al., 2019), for pigmentation levels in rabbits

69 (*Oryctolagus cuniculus*) (Chen et al., 2021) and for breeding success in barn swallows (*Hirundo*
70 *rustica*) (Saino et al., 2017).

71 Epigenetic modifications may take a variety of forms, these include DNA methylation,
72 histone modifications and non-coding RNA activity (Moore et al., 2013), all of which can regulate
73 gene expression without requiring changes in the DNA sequence (Razin & Riggs, 1980).
74 Differential patterns of epigenetic modification have also been associated with early stages of
75 speciation (Quadrana & Colot, 2016; Vernaz, Hudson, et al., 2021; Vernaz, Malinsky, et al., 2021;
76 Vogt, 2017). DNA methylation (DNAm) involves the transfer of a methyl group to a cytosine base
77 (Moore et al., 2013) and can occur throughout the genome but mostly occurs on CpG dinucleotide
78 sites (Moore et al., 2013). Depending on where in the genome these modifications occur, they can
79 either supress or activate gene expression (Dhar et al., 2021; Wan et al., 2015) and in some cases,
80 methylation of only one CpG can inhibit the binding of transcription factors (Iguchi-Ariga &
81 Schaffner, 1989; Watt & Molloy, 1988). CpG islands (CGI) are regions of the genome that contain
82 a high frequency of CpGs relative to CpGs found in other regions of the mammalian genome
83 (Illingworth & Bird, 2009). CGIs are also known to overlap most promoter regions (Gardiner-
84 Garden & Frommer, 1987; Jones, 1999). These sites might be evolutionarily important in
85 mammals since many CGIs are conserved between mice and humans (Deaton & Bird, 2011;
86 Illingworth et al., 2010; Moore et al., 2013). It is theorized that over time differential methylation
87 can lead to permanent heritable (cis) nuclear changes, referred to as genetic assimilation
88 (Anastasiadi et al., 2021). This can happen when environmental conditions persist through
89 generations favouring a specific phenotype, and over time resulting in selection for a permanent
90 genetic basis (e.g., point mutations) for the phenotype previously produced from the plastic
91 epigenetically-influenced phenotype (Anastasiadi et al., 2021).

In addition, DNAm levels, particularly hypermethylation, on specific CpG sites changes throughout an individual's lifetime, making it possible to age individuals based on their methylation levels (Horvath, 2013; Lu et al., 2021; Prado et al., 2021; Wilkinson et al., 2021). These epigenetic *clocks* can predict chronological age in different mammalian species (Jasinska et al., 2021; Lu et al., 2021; Prado et al., 2021; Raj et al., 2021; Schachtschneider et al., 2021), and can be used to validate current aging techniques based on field observations (Robeck et al., 2021). Epigenetic clocks can also be used to predict demographic trends for conservation efforts (Prado et al., 2021) and have been shown to be elevated in inbred individuals, suggesting that inbreeding can lead to aging at a faster rate (Larison et al., 2021). Likewise, metabolic syndrome in humans (i.e., obesity) also results in increased epigenetic aging (Lee & Park, 2020; Nannini et al., 2019). It stands to reason that the differential selection pressures underlying island-mainland patterns might also alter the DNAm aging rate.

Shrews of the family Soricidae appeared about 30 to 40 million years ago (Churchfield, 1990). Hundreds of extant shrew species are now found on most continents (Churchfield, 1990; George, 1986, 1988) and have evolved unique traits such as echolocation (Chai et al., 2020; Forsman & Malmquist, 1988; Tomasi, 1979), venomous saliva (Kita et al., 2004) and the ability to dive underwater (Catania et al., 2008; Mendes-Soares & Rychlik, 2009). Shrews have the fastest mass-specific metabolic rate of all mammals (Ochocińska & Taylor, 2005) and are one of the smallest and shortest-lived mammals having a maximum lifespan of ~17 months (Churchfield, 1990), making them important to study from an aging perspective (Lu et al., 2021). Shrews of the genus *Sorex* show differing trends in body size variation across their range; populations of some species at higher latitudes have smaller body sizes, which is counter to the predictions of Bergmann's rule (Ochocińska & Taylor, 2003; Yom-Tov & Yom-Tov, 2005), whereas populations

of other species are larger at more northerly latitudes (Huggins & Kennedy, 2018). *Sorex* shrews also exhibit morphological plasticity; the Eurasian common shrew (*Sorex araneus*), for example, manifests seasonal and reversible changes in skull morphology, body mass and organ size (Lázaro et al., 2017; Schaeffer et al., 2020; Taylor et al., 2013). Whether plastic or genetic in nature, these phenotypic variations throughout winter and latitude are generally explained as a response to limited resource availability (Lázaro et al., 2017; Ochocińska & Taylor, 2003; Schaeffer et al., 2020; Taylor et al., 2013; Yom-Tov & Yom-Tov, 2005).

Masked or Cinereus shrews (*Sorex cinereus*) are the most abundant shrew in North America (Nagorsen, 1996). The range of this species spans much of the northern United States and most of Canada (Nagorsen, 1996) with a varied diet consisting primarily of invertebrates (Bellocq et al., 1994; Churchfield, 1990; Whitaker John O. Jr. & French, 1984). One population of island dwelling masked shrews off the southern coast of Nova Scotia became isolated from the mainland during the late Wisconsin glaciation around 20,000 years ago (Roland, 1982; Stewart & Baker, 1992) and has developed a specialized feeding habit in response to their littoral environment (Stewart et al., 1989). Masked shrews on Bon Portage Island (BPI) are the only population of *Sorex* known to mainly feed on sand fleas (*Platorchestia platensis*) (MacPherson & Stewart, 2003; Stewart et al., 1989). *Platorchestia platensis* is a type of amphipod typically not consumed by insectivores, although Churchfield (1990) observed that the population of the lesser white-toothed shrew (*Crocidura suaveolens)* on St. Martin's, Isles of Scilly, in the United Kingdom, also feed on amphipods. In the case of *S. cinereus*, electrophoresis assays by Stewart and Baker (1992) found that the BPI masked shrew population expressed a unique (slow) allele for Peptidase A and showed evidence of drift, but genetic differentiation across populations was low (Nei's $D \cong 0.01$). It was later hypothesized that due to the importance of sand fleas on BPI, the shrew's digestive enzyme,

138 Peptidase A, has undergone an adaptive response (MacPherson & Stewart, 2003). Here, using a
139 *de novo* assembled and annotated genome, we built a masked shrew epigenetic clock and
140 quantified differential methylation patterns between populations to examine the association
141 between geographical isolation and epigenetic divergence. We hypothesized that island shrews
142 will exhibit different phenotypes than mainland shrews based on the island syndrome and that
143 divergent methylation patterns could underly these differences.
144

145 **Materials and methods**

146 **Sampling, sexing, and aging shrews**

147 Masked shrew samples were collected (Animal Care Certificate No. 26234) from five locations:
148 Peterborough County (44°53′N, 78°05′W) in Ontario, Canada, and Sandy Cove (44°48′N,
149 66°08′W), North Mountain (45°13′N, 64°22′W), Long Island (44°19′N, 66°16′W) and Bon
150 Portage Island (BPI) (43°28′N, 65°45′W) in Nova Scotia, Canada (Figure S1). All shrews were
151 trapped between mid-July and late August to limit seasonal morphological differences, such as
152 skull and body size (Lázaro et al., 2017, 2019, 2021). DNA was extracted from liver as this is
153 where digestive enzymes are found; tail and fetal tissue were selected to help build the epigenetic
154 clock (see Lu et al., 2021). Liver and tail samples produce high DNA yields and are easy to locate
155 in shrews. We used a DNeasy Blood & Tissue Kit from QIAGEN. Shrews were left in dermestid
156 beetle tanks for a week to remove all remaining tissue from the skulls. A Leica EZ4 microscope
157 was used to assess age class for each shrew based on teeth wear according to the method of Pruitt
158 (Pruitt, 1954; Rudd, 1955). Age in months was estimated based on the trapping date, the age class,
159 and the known reproductive season of masked shrews assuming a maximum lifespan of ~17
160 months (Churchfield, 1990). Shrews were sexed by gel electrophoresis of PCR amplicons of the

161 Y chromosome-linked SRY gene (Cervantes et al., 2013; Matsubara et al., 2001) using custom
162 primers (Table S1); sex assignments were subsequently validated with the methylation assay.
163
164 **Genome assembly and annotation**
165 A single masked shrew from Peterborough County, Ontario had DNA and RNA extracted from
166 the heart and liver immediately after euthanasia. We used the MagAttract HMW DNA Kit from
167 QIAGEN and the Monarch Total RNA Miniprep Kit from New England BioLabs, respectively,
168 and sent the extracts to The Centre for Applied Genomics, Toronto, Ontario, Canada, for
169 sequencing. Short-read (SR) libraries were generated using Truseq PCR-free preparation and
170 sequenced on two lanes on an Illumina HiSeq X platform with 150 bp paired-end reads. Linked
171 reads (LR) were prepared using the 10X genome library after selecting fragments >20Kb using
172 BluePippin. The linked reads were sequenced on one Illumina HiSeq X Lane. Twenty-five million
173 RNA reads were sequenced on a HiSeq2500 with 126bp paired-end reads.
174 We used BBMap v. 35.8 (Bushnell, n.d.) to remove adapters and trim low-quality data
175 from the fastq files. Contaminant screening was performed using Kraken2 v. 2.0.9-beta (Wood &
176 Salzberg, 2014) and all unclassified sequences were kept. Kmergenie v. 1.7048 (Chikhi &
177 Medvedev, 2014) was used to find the kmer size for downstream analysis. The genome was
178 assembled using a tiered approach: 1) using only the SR data and Meraculous v.2.2.4 (Chapman
179 et al., 2011) and 2) with only the LR data using Supernova v.2.1.1 (Weisenfeld et al., 2018) using
180 the --maxreads filter set to "all". The pseudohap2 style was selected for the assembly output.
181 Following the assembly, both SR and LR assembly versions were merged via quickmerge v.0.3
182 (Chakraborty et al., 2016) with the LR as the backbone. Scaff10x v.5.0 (Ning et al., n.d.) was used
183 to further polish the *S. cinereus* genome. Genome completeness was assessed using Benchmarking

184 Universal Single-copy Orthologs (BUSCO) v.3.0.2 (Simão et al., 2015) by comparing the hybrid
185 genome to highly conserved genes in mammals based on the *mammalia_odb9* dataset. The
186 mitochondrial genome was assembled using MitoZ v.2.4-alpha (Meng et al., 2019). Genome
187 annotation was conducted via GenSAS v.6.0 (Humann et al., 2019) using the assembled genome
188 and RNAseq data. Here, repeat regions of the genome were identified using RepeatModeler
189 v.1.0.1.1 (Smit & Hubley, n.d.-b) and RepeatMasker v.4.0.7 (Smit & Hubley, n.d.-a). Masked
190 shrew liver and heart RNA reads were mapped to the genome using HISAT2 v.2.1.0 (Kim et al.,
191 2019). The resulting BAM file was used by BRAKER v.2.1.1 (Hoff et al., 2019) and
192 EVidenceModeler v1.1.1 (Haas et al., 2008) to predict features in the genome. Functional
193 annotation to find common protein sequences between the masked shrew and the common shrew,
194 *Sorex araneus* (GenBank assembly accession: GCA_000181275.2), and other vertebrate mammals
195 was done using BLAST v.2.11.0 (Altschup et al., 1990), SwissProt v.2.11.0 (Bairoch & Apweiler,
196 1998) and InterProScan v.5.44-79.0 (Jones et al., 2014) databases.
197
198 **Epigenetic clocks**
199 DNA methylation profiles for each sample (n = 48, Table S2) were assayed on the
200 HorvathMammalMethylChip40, a novel Infinium array that profiles up to 37,000 highly conserved
201 CpGs across the genome (Arneson et al., 2022). The SeSaMe normalization method was used to
202 define beta values for each probe (Zhou et al., 2018). Unsupervised hierarchical clustering based
203 on inter-array correlation coefficients revealed that samples cluster by tissue type. The methylation
204 array measurements can assess the sex of each sample based on the X chromosomal CpGs. One
205 fetal sample was removed due to inaccurate sexing. Only probes that aligned to the masked shrew
206 genome (bwa-mem) were retained for downstream analysis (n = 29,609). Elastic net regression

207 models (alpha = 0.5, 10-fold internal cross-validation) were applied to the DNAm data (covariates)
208 and the chronological age estimate (dependent variable). The resulting multivariate regression
209 model will be referred to informally as the masked shrew epigenetic clock. The alpha value of the
210 glmnet R software package (Friedman et al., 2010) was set to 0.5 to be midpoint between a ridge
211 and lasso type regression. Elastic net regressions help prevent overfitting the model to the data by
212 selecting variables (CpGs) and shrinking them simultaneously (Zou & Hastie, 2005). Three
213 different models were used as per Lu et al. (2021); 1) no transformation to the chronological age,
214 2) log-transformed chronological age (log(x+1)), and 3) square root transformed chronological age
215 (sqrt(x+1)) with offsets of 1 year to avoid negative values for our fetal samples. To assess the
216 accuracy of our epigenetic clock models we used a leave-one-out (LOO) cross-validation method.
217 This method consists of training the epigenetic clock models with a subset of the data and testing
218 how well it can predict other observed data. We report the LOO estimates of the Pearson
219 correlation between age and its DNAm based estimates as well as the median absolute error (MAE)
220 between DNAm age and observed age (estimated chronological age in years).
221

222 **Statistical analyses and functional gene enrichment**

223 We ran a series of models in R v.1.3.959 (R Core Team, 2021). Using linear models, we separately
224 regressed body size, weight, and skull length against location, age and sex (i.e., phenotype ~
225 location + age + sex) of 33 masked shrews (Table S2). Here, location was a categorical variable
226 of island or mainland (see Table 1). Additionally, we extracted the residuals from the epigenetic
227 clock for each sample and ran a Welch Two Sample T-test using location and sex to test if different
228 groupings epigenetically aged at comparable rates. Positive residuals were indicative of an
229 epigenetically older individual than the linear model would predict.

230 Epigenome-wide association studies (EWAS) were conducted using the limma package
231 v.3.44 (Ritchie et al., 2015) with 44 methylomes (Table S2); note we removed the fetal samples
232 from this analysis to avoid biased results associated with developmental genes and pathways.
233 Using the β values of aligned probes, we ran the following model: DNAm ~ location + sex + age
234 + tissue including animal ID as a random effect. This model was run two times; once with BPI and
235 Long Island grouped together as 'Island' in the model with all other locations assigned as
236 'Mainland' and once with BPI versus all other populations. The latter grouping was meant to
237 isolate possible unique dietary adaptations of BPI shrews. The island versus mainland model was
238 also run with only Long Island samples to test for a potential bias of having more BPI samples in
239 the island grouping.
240 A p value < 2e-6 was set according to the Bonferroni-corrected threshold as the cutoff for
241 significance to be conservative (Johnson et al., 2010). The location of probes in relation to the gene
242 annotations were identified using the ChIPseeker (Yu et al., 2015) and GenomicFeatures
243 (Lawrence et al., 2013) packages. Enriched pathways associated with the significant CpGs for the
244 EWAS models were identified with rGREAT v.1.20 (McLean et al., 2010) using the human Hg19
245 genome annotation to allow for comparisons (Jasinska et al., 2021; Prado et al., 2021;
246 Schachtschneider et al., 2021). We accounted for bias resulting from the design of the mammalian
247 array by using 14,290 CpGs that mapped to our shrew gene annotations as background. Following
248 Schachtschneider et al. (2021), from these probes we filtered and kept the top 500 hypermethylated
249 and top 500 hypomethylated sites based on p value to use for the enrichment analysis.
250

251 **Results**

252 **Genome assembly and annotations**

253 We assessed the quality and completeness of our genome assembly to ensure confidence in our
254 downstream analyses. The SR and LR assembly statistics are provided in Table S3; the final hybrid
255 genome was 2.66 Gb with an N50 of 1.99 Mb and L50 of 291 scaffolds. The GC content was
256 42.88% with over 86% of the genome being in scaffolds over 50 Kb and the largest scaffold being
257 around 22 million base pairs. Close to 95% of BUSCOs were identified in this assembly (Table
258 S3): 3504 of 4101 were complete single-copies, 74 were complete duplicates and 301 were
259 fragmented. Repeats represented 38.60% of the genome (Table S4) with long interspersed nuclear
260 elements (LINEs) and unclassified repeats representing 20.17% and 15.39% of the genome
261 respectively. The functional annotation pipeline identified 20,067 protein coding genes (Table S3).
262 The metrics assessed were comparable to other genome assemblies using similar sequencing
263 approaches (e.g., Etherington et al., 2020; Wolf et al., 2022). The raw sequence data have been
264 deposited in the Short Read Archive (SRA) under accession number PRJNA826195. The assembly
265 has been deposited at DDBJ/ENA/GenBank (accession no. pending).
266
267 **Epigenetic clocks**
268 We built a shrew methylation clock to assess epigenetic aging rates between different shrew
269 populations (n = 24 from islands versus n = 12 from mainland). The methylation data consisted of
270 33 tail, 11 liver, and 3 fetal tissue DNA samples selected from 36 different individuals (Table S2).
271 One fetal sample was dropped due to quality control metrics (wrong sex). The average age of our
272 samples was seven months; the youngest shrews being fetuses and the oldest one being an
273 estimated 16 months (Table 1). Out of the 37,000 probes, 29,609 were aligned to our masked shrew
274 genome. We constructed and cross-validated epigenetic clocks for different masked shrew tissue
275 types, which included liver, tail, fetus, and pan-tissue clock. The final clocks were based on the

methylation profiles of 26 CpGs, similar to clocks from other species such as elephants or cats (e.g., Prado et al., 2021; Raj et al., 2021), our clocks were highly accurate with a LOO cross validation estimate of the Pearson correlation between age and its predicted value r >= 0.91 and a median absolute error of MAE = 0.13 years for the multi-tissue clock, representing approximately a 7-week difference between the estimated DNAm age and the teeth-based chronological age. The liver DNAm clock had the lowest MAE of 0.07 years (~3 weeks), followed by the multi-tissue and tail clocks with MAE around 0.1 (~5 weeks) (Figure 1).

Island-mainland differences

Morphological and epigenetic age differences between 21 island and 12 mainland masked shrew populations (Table S2) were examined using linear regression. Mass, body length and skull length differed between the mainland and island samples (Figure 2). Our models including age, sex, and island-mainland covariates explained a high percentage of variance: skull length (Adj. $R^2 = 0.79$), weight (Adj. $R^2 = 0.53$) and body length (Adj. $R^2 = 0.49$; Table 2). In all models, location was the variable that correlated most to the change in body size and weight (p value < 0.01; Table 2). All size metrics were negatively correlated to being on an island (Table 2, Figure 2). Age was most often positively correlated to body size whereas sex did not have an effect (Table 2).

Epigenetic clock residuals were slightly elevated on islands indicative of individuals epigenetically older than predicted (t = -1.95, df = 39.73, p value = 0.06); residual variation was higher on the island (mean of 0.01) compared to the mainland (mean of -0.02). There was no difference in residuals among the sexes (t = -0.97, df = 40.2, p value = 0.34) or age groups (yearlings versus older) (t = -0.21, df = 11.79, p value = 0.84).

299 **EWAS Island effects**

300 EWAS were conducted using the methylation data consisting of the 33 tail, 11 liver samples (Table
301 S2) to identify differentially methylated genes and pathways between island and mainland masked
302 shrews. A total of 1067 differentially methylated CpG sites (482 hypermethylated and 585
303 hypomethylated relative to mainland) were identified (*p* value < 2e-6) between our island and
304 mainland populations (Figure 3A). The most divergent CpG was on the *EWS* exon (*p* value =
305 3.35e-21) a gene that encodes a multifunctional protein involved in various cellular processes such
306 as gene expression, cell signaling, and RNA processing and transport. Top differentially
307 methylated CpGs were located on or near genes associated to developmental and adipose-related
308 processes: *MEST* exon (*p* value = 3.34e-14), *SIM1* intergene (*p* value = 3.79e-13), *PHIP* intergene
309 (*p* value = 1.03e-12) and metabolism: *ZBTB7C* promotor (*p* value = 8.45e-14) and *ATPA* exon (*p*
310 value = 6.37e-16) (Table 3, Figure 3A, 3B). The top hypermethylated CpGs were associated to
311 three different fibroblast growth factor receptors pathways (*p* value < 5e-16) (Figure 3C). Other
312 significantly enriched pathways included synaptic and adipose-related processes (Figure 3C).

313 Isolating BPI identified 485 significant CpGs (*p* value < 2e-6), 308 were hypomethylated
314 and 177 were hypermethylated compared to other populations (Figure 4A). The top significant
315 hypomethylated CpG in BPI samples was on the *SMO* exon (*p* value = 1.42e-30; Table 3, Figure
316 4A, 4B). Top CpGs were most often hypomethylated in BPI island samples (Table 3, Figure 4A,
317 4B). Pathway analysis highlighted that differentially methylated sites in BPI shrews compared to
318 other populations seem to affect digestive and intestinal phenotypes and functions (Figure 4C).
319 Additionally, other phenotypes related to skull morphology, more specifically the mandible, teeth,
320 and ear canal showed pathway enrichment in BPI shrews (Fig 4C). When BPI was removed from

321 the island-mainland model (i.e. only Long Island), digestive enzymes were not identified as being
322 a significantly enriched pathway (Figure S2).
323
324 **Discussion**
325 The island syndrome describes the distinctive morphological differences between island-mainland
326 populations due to different ecological and environmental pressures (Baeckens & van Damme,
327 2020). We examined whether phenotypic divergence was associated with epigenetic
328 modifications, specifically DNAm, which alter gene expression without changing the DNA
329 sequence (Razin & Riggs, 1980). Various authors have argued that methylated cytosines reduce
330 gene expression in promotors by causing chromatin structures to condense which influences the
331 binding of transcription factors (Moore et al., 2013; Pappalardo & Barra, 2021). However,
332 methylation in other parts of the genome, e.g., intragenic, or distal regions, can both repress and
333 induce gene expression (Anastasiadi et al., 2018). Our analysis showed that differential methylated
334 patterns are associated with distinct phenotypes between island and mainland masked shrew
335 populations. Over 29,000 CpG methylation probes aligned to our genome, consistent with other
336 mammalian studies (Larison et al., 2021; Pinho et al., 2021; Schachtschneider et al., 2021; Sugrue
337 et al., 2021), and a robust epigenetic clock was generated (Figure 1). The differential methylation
338 between island-mainland populations showed a remarkable overlap between enriched pathways
339 and measured phenotypic differences. While some association assessments suffer from false
340 positives and "story-telling" (Pavlidis et al., 2012), there was an unambiguous connection to the
341 measured body phenotypes (Figure 2), dietary differences (MacPherson & Stewart, 2003;
342 McAlpine, 2009; Stewart et al., 1989), and the observed enriched pathways (Figure 3C, 4C). Thus,

343 the clear morphological and dietary island-mainland differences appear to be, in part, under
344 epigenetic influence.

345 The island syndrome describes the distinctive morphological differences between island-
346 mainland populations due to different ecological and environmental pressures (Baeckens & van
347 Damme, 2020). We examined both phenotypic divergence and epigenetic modifications in masked
348 shrews, specifically DNAm which has the potential to increase phenotypic plasticity within an
349 individual (Vogt, 2017). Highly diversified DNAm methylation among populations can be a
350 strategy to take advantage of unpredictable environments (Anastasiadi et al., 2021); moreover,
351 epigenetic modifications have been shown to allow isogenic lines to adjust their phenotype to
352 maximize fitness (van Egeren et al., 2018). Bon Portage Island masked shrew populations have
353 lower genetic diversity than their mainland counterparts (Stewart & Baker, 1992), but likely have
354 a broader ecological niche due to reduced interspecific competition; masked shrews on mainland
355 Nova Scotia are broadly sympatric with six other species of shrews, and diet studies demonstrate
356 that these soricids differentially exploit available invertebrate food resources (Whitaker & French,
357 1984). Without standing genetic variation, differential methylation among offspring could be a
358 bet-hedging strategy to allow for exploration and exploitation of a range of food resources on the
359 island (Anastasiadi et al., 2021). Further, DNAm can theoretically lead to genetic modifications
360 through the process of genetic assimilation where methylated cytosines have a higher frequency
361 of mutagenesis resulting in novel alleles (Anastasiadi et al., 2021). Differential methylation is an
362 effective strategy that individuals in genetically impoverished populations may use to generate
363 phenotypic diversity. Thus, we hypothesize that BPI shrews are in the process of methylation-
364 induced genetic assimilation, with the high frequency of the unique peptidase A allele in this
365 population being one outcome of this process.

366

367 **Phenotypic and epigenetic divergence**

368 DNA methylation regulates gene expression thereby contributing to phenotypic plasticity (al
369 Aboud et al., 2021; Kooke et al., 2015; Razin & Riggs, 1980), which can lead to phenotypic
370 variation among individuals subject to different environmental conditions (Anastasiadi et al., 2021;
371 Kooke et al., 2015; Meröndun et al., 2019). Here, insular masked shrews exhibited different
372 phenotypes than their mainland counterparts; morphologically, this presented as decreased body
373 length, mass, and skull length on islands. Body size is a heritable polygenic trait in mammals
374 (Posbergh & Huson, 2021; Postma, 2014) with many underlying mechanisms, often making it
375 difficult to clearly understand the role of each gene or pathway involved (e.g., Anderson et al.,
376 2022). Cytosine methylation levels measured in highly conserved stretches of DNA have been
377 linked to mammalian traits such as maximum lifespan (Haghani et al., 2021) and has been
378 suggested as a universal mechanism to regulates body mass and morphological development in
379 animals in other association studies (Cao et al., 2015; Haghani et al., 2021). Genetic variation in
380 quantitative trait loci (QTL) almost certainly play a role in phenotypic variations between insular
381 and mainland populations – and might also influence methylation (J. T. Bell et al., 2011; Smith et
382 al., 2014). We suggest that pooled genome sequencing approaches (e.g., island versus mainland
383 samples) are a feasible option to test for nuclear genome associations of this nature.

384 Differential methylated pathways between island and mainland shrew populations for all
385 tissue types included fibroblast growth factors (FGF), which are crucial for animal body
386 development (Oulion et al., 2012), and adipose-related genes and processes. Haghani *et al.* (2021)
387 also found that across mammals, species weight was correlated to development and adipose-related
388 pathways, possibly due to DNAm affecting the expression of adipose tissue and regulation of lipid

389 storage (Ma & Kang, 2019). In masked shrews, CpGs on the homolog to mesoderm-specific
390 transcript protein (*MEST*) exon and single-minded homolog 1 (*SIM1*) intergenic region were hypo
391 and hypermethylated respectively in island shrews compared to mainland shrews. *MEST* is known
392 to be correlated to adipocytes in humans (Karbiener et al., 2015) and mice (Takahashi et al., 2005),
393 while *SIM1* is associated with food intake regulation and obesity (Holder et al., 2004; Yang et al.,
394 2006), both possibly affecting body mass in shrews. Differential methylation of genes such as
395 *SIM1* might affect shrews' metabolism and ultimately body size by regulating the expression of
396 hormones involved in signaling for food intake. Further, island shrews had a hypomethylated CpG
397 in the pleckstrin homology domain-interacting protein (*PHIP*) intergenic region, a probable
398 regulator of insulin-like growth factor signaling pathways (Farhang-Fallah et al., 2002), which
399 reinforces the idea that insulin growth factors are involved in mammalian life-history variation
400 (Swanson & Dantzer, 2014). The overall associations between DNAm and genes related to body
401 composition, possibly indicates that methylation inhibits or activates various genes during
402 developmental stages in shrews, but future studies are needed to better understand such
403 mechanisms.

404 Masked shrew insular populations appear to have decreased body mass compared to their
405 mainland counterparts; BPI masked shrews occur in high densities (Downie, 1986; Stewart et al.,
406 1989; Telfer, 1984), which should lead to small animals having larger body size to better compete
407 for resources (Juette et al., 2020). However, the discrepancy with this expected phenotype for
408 island shrews might be due to their diet. Previous studies have shown that food scarcity contributes
409 to the selection of smaller body size in Palearctic shrews, which might also be the case for Nearctic
410 populations (Ochocińska & Taylor, 2003; Yom-Tov & Yom-Tov, 2005). Churchfield (2002)
411 found that in locations where higher biomass resources are scarce, shrews tend be smaller in size

412 and survive by consuming less nutritional but more abundant sources of food. This might be the
413 case for BPI shrews feeding on sand fleas. Supporting this idea, top EWAS genes involved in
414 energy homeostasis and metabolism were differentially methylated between mainland-island
415 populations (Table 3). For example, ATP Synthase F1 Subunit Alpha (*ATPA*) catalyzes the
416 synthesis of ATP, the main molecule used to store and transfer energy in cells (Schapira, 2006).
417 Similarly, the zinc finger and BTB domain–containing 7c (*ZBTB7C)* gene promotor was also
418 differentially methylated in island shrews. *ZBTB7C* is an important metabolic regulating gene for
419 blood-glucose homeostasis during fasting (Choi et al., 2019).

420 BPI shrews also showed differentially methylated sites associated with various processes
421 related to the digestive system, including peptidase activity. This is of significance as previous
422 studies noted that masked shrews on BPI have developed specialized littoral feeding habits
423 (MacPherson & Stewart, 2003; Stewart et al., 1989). Previous work showed that BPI shrews
424 possess a unique polymorphism for the digestive enzyme Peptidase A (Stewart & Baker, 1992),
425 which potentially allows shrews to easily digest the sand fleas' exoskeleton (McAlpine, 2009).
426 Digestive system related pathways being in the top enrichment category indicate that these dietary
427 habits seem to be one of the most significant molecular differences between this insular population
428 and others and support earlier allozyme work (MacPherson & Stewart, 2003; McAlpine, 2009;
429 Stewart & Baker, 1992). Our work remains a correlative, but future genomic assays can be used
430 to test our genetic assimilation hypothesis, while mouse models or common-garden experiments
431 could be used to assess the localization of these CpGs and phenotypic associations to provide
432 information on possible mechanism of enriched regions on masked shrew methylomes.

433

434 **Epigenetic age**

435 DNA methylation and chronological age can be used to build epigenetic clocks to identify genes
436 associated with aging and predict how methylation patterns change throughout an individual's
437 lifetime (Larison et al., 2021; Lu et al., 2021; Prado et al., 2021; Wilkinson et al., 2021). Epigenetic
438 clocks have been constructed for a wide range of species, showing consistent trends with CpGs
439 associated to developmental genes and pathways (Lu et al., 2021; Pinho et al., 2021; Raj et al.,
440 2021; Schachtschneider et al., 2021). Second generation clocks, such as the DNAm GrimAge,
441 account for variables such as chronological age, health factors and mortality, to predict lifespan
442 and healthspan (C. G. Bell et al., 2019; Lu et al., 2019). In humans, DNAm GrimAge has been
443 studied in relation to various comorbidities and social factors (Lu et al., 2019; McCrory et al.,
444 2021). Inbreeding has been associated with decreasing longevity and an increase in health concerns
445 in various species (Keller & Waller, 2002; Skotarczak et al., 2020; Yordy et al., 2020) and recently
446 was linked to accelerated DNAm aging in zebras (*Equus quagga*) (Larison et al., 2021). We used
447 these patterns to explore accelerated aging as a component of island-mainland divergence.

448 Masked shrews tend to live on average 14 months; accurately aging small mammals such
449 as shrews can be difficult since they are such small creatures and current methods rely on
450 somewhat subjective classification based on teeth wear, body length measurements and weight
451 (Pruitt, 1954; Rudd, 1955). Our epigenetic clock validated the current aging techniques used in the
452 field, and overall, the clock was highly predictive (R = 0.95; MAE = 0.10) (Figure 1). Further,
453 when we examined the clock model residuals, island shrews appeared to age slightly faster than
454 mainland shrews. Parasites can affect hosts' life-history traits through the energetic investment of
455 maintaining an immune system response (Cooper et al., 2012; Cowan et al., 2009; Morand &
456 Harvey, 2000), possibly resulting in faster biological aging. Group size has been shown to be
457 correlated with prevalence and intensity of parasitism (Coté & Poulin, 1995). The population

density of shrews on BPI is extremely high, at least in some years (Downie, 1986; Stewart et al., 1989; Telfer, 1984) possibly resulting in an increased incidence of various parasites and pathogens. Overall, this study reports divergent phenotypic trends between mainland-island masked shrew populations, and more broadly support the role of epigenetics in shaping phenotypic divergence. Future research should investigate epigenetic aging in insular populations, especially when inbreeding (Larison et al., 2021) or phenotypic divergence (i.e., this study) is present.

Acknowledgments

We want to thank the reviewers for the valuable insights and comments provided. This work was supported by CanSeq150 (CGEn) (ABAS, DTS and MLC); Ontario Graduate Scholarship (MLC), Natural Sciences and Engineering Research Council of Canada (NSERC) Alexander Graham Bell Canada Graduate Scholarship-Master's (MLC) and NSERC Discovery grants (ABAS and DTS); Compute Canada Resources for Research Groups (ABAS); Canadian Foundation for Innovation: John R. Evans Leaders Fund (ABAS); Ontario Early Researcher Awards (ABAS). SH was supported by the Paul G. Allen Frontiers Group. We wish to acknowledge that the work for this study was carried out on the traditional territory of the Mississauga (Michi Saagiig) Anishnaabeg and the Mi'kmaq People. We are grateful to have had the opportunity to work on this land and thank the First Peoples for their care, stewardship, and teaching.

References

Adler, G. H., & Levins, R. (1994). The island syndrome in rodent populations. *The Quaterly Review of Biology*, *69*(4), 18–19.

al Aboud, N., Tupper, C., & Jialal, I. (2021). *Genetics, Epigenetic Mechanism*. StatPearls Publishing.

Altschup, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic Local Alignment Search Tool. In *J. Mol. Biol* (Vol. 215).

Anastasiadi, D., Esteve-Codina, A., & Piferrer, F. (2018). Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. *Epigenetics and Chromatin*, *11*(1). https://doi.org/10.1186/s13072-018-0205-1

Anastasiadi, D., Venney, C. J., Bernatchez, L., & Wellenreuther, M. (2021). Epigenetic inheritance and reproductive mode in plants and animals. *Trends in Ecology and Evolution*, *36*(12), 1124–1140. https://doi.org/10.1016/j.tree.2021.08.006

Anderson, S. J., Côté, S. D., Richard, J. H., & Shafer, A. B. A. (2022). Genomic architecture of phenotypic extremes in a wild cervid. *BMC Genomics*, *23*(1). https://doi.org/10.1186/s12864-022-08333-x

Arneson, A., Haghani, A., Thompson, M. J., Pellegrini, M., Kwon, S. bin, Vu, H., Maciejewski, E., Yao, M., Li, C. Z., Lu, A. T., Morselli, M., Rubbi, L., Barnes, B., Hansen, K. D., Zhou, W., Breeze, C. E., Ernst, J., & Horvath, S. (2022). A mammalian methylation array for profiling methylation levels at conserved sequences. *Nature Communications*, *13*(1), 783. https://doi.org/10.1038/s41467-022-28355-z

Baeckens, S., & van Damme, R. (2020). The island syndrome. *Current Biology*, *30*(8), R338–R339. https://doi.org/10.1016/j.cub.2020.03.029

Bairoch, A., & Apweiler, R. (1998). The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1998. In *Nucleic Acids Research* (Vol. 26, Issue 1). http://www.expasy.ch/

Bell, C. G., Lowe, R., Adams, P. D., Baccarelli, A. A., Beck, S., Bell, J. T., Christensen, B. C., Gladyshev, V. N., Heijmans, B. T., Horvath, S., Ideker, T., Issa, J. P. J., Kelsey, K. T., Marioni, R. E., Reik, W., Relton, C. L., Schalkwyk, L. C., Teschendorff, A. E., Wagner, W., … Rakyan, V. K. (2019). DNA methylation aging clocks: Challenges and recommendations. *Genome Biology*, *20*(1), 1–24. https://doi.org/10.1186/s13059-019-1824-y

Bell, J. T., Pai, A. A., Pickrell, J. K., Gaffney, D. J., Pique-Regi, R., Degner, J. F., Gilad, Y., & Pritchard, J. K. (2011). DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biology*, *12*(1). https://doi.org/10.1186/gb-2011-12-1-r10

Bellocq, M. I., Bendell, J. F., Innes, D. G. L., & Innes', D. G. L. (1994). Diet of Sorex cinereus, the Masked Shrew, in Relation to the Abundance of Lepidoptera Larvae in Northern Ontario. *Source: The American Midland Naturalist*, *132*(1), 68–73.

Benítez-López, A., Santini, L., Gallego-Zamorano, J., Milá, B., Walkden, P., Huijbregts, M. A. J., & Tobias, J. A. (2021). The island rule explains consistent patterns of body size evolution in terrestrial vertebrates. *Nature Ecology and Evolution*, *5*(6), 768–786. https://doi.org/10.1038/s41559-021-01426-y

Bushnell, B. (n.d.). *BBTools*. https://jgi.doe.gov/data-and-tools/bbtools/

Cao, J., Wei, C., Liu, D., Wang, H., Wu, M., Xie, Z., Capellini, T. D., Zhang, L., Zhao, F., Li, L., Zhong, T., Wang, L., Lu, J., Liu, R., Zhang, S., Du, Y., Zhang, H., & Du, L. (2015). DNA methylation Landscape of body size variation in sheep. *Scientific Reports*, *5*. https://doi.org/10.1038/srep13950

Catania, K. C., Hare, J. F., & Campbell, K. L. (2008). Water shrews detect movement, shape, and smell to find prey underwater. In *PNAS* (Vol. 105, Issue 2).

Cervantes, F. A., Arcangeli, J., & Alonso Morales, R. A. (2013). The SRY gene of the y chromosome successfully identifies gender of endemic Mexican shrews (Mammalia: Soricomorpha). *Mammalia*, *77*(3), 335–339. https://doi.org/10.1515/mammalia-2011-0123

Chai, S., Tian, R., Rong, X., Li, G., Chen, B., Ren, W., Xu, S., & Yang, G. (2020). Evidence of echolocation in the common shrew from molecular convergence with other echolocating mammals. *Zoological Studies*, *59*, 1–11. https://doi.org/10.6620/ZS.2020.59-4

Chakraborty, M., Baldwin-Brown, J. G., Long, A. D., & Emerson, J. J. (2016). Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage. *Nucleic Acids Research*, *44*(19), 1–12. https://doi.org/10.1093/nar/gkw654

Chapman, J. A., Ho, I., Sunkara, S., Luo, S., Schroth, G. P., & Rokhsar, D. S. (2011). Meraculous: De novo genome assembly with short paired-end reads. *PLoS ONE*, *6*(8). https://doi.org/10.1371/journal.pone.0023501

Chen, Y., Hu, S., Liu, M., Zhao, B., Yang, N., Li, J., Chen, Q., Zhou, J., Bao, G., & Wu, X. (2021). Analysis of Genome DNA Methylation at Inherited Coat Color Dilutions of Rex Rabbits. *Frontiers in Genetics*, *11*(January), 1–9. https://doi.org/10.3389/fgene.2020.603528

Chikhi, R., & Medvedev, P. (2014). Informed and automated k-mer size selection for genome assembly. *Bioinformatics*, *30*(1), 31–37. https://doi.org/10.1093/bioinformatics/btt310

Choi, W. il, Yoon, J. H., Song, J. Y., Jeon, B. N., Park, J. M., Koh, D. I., Ahn, Y. ho, Kim, K. S., Lee, I. K., & Hur, M. W. (2019). Zbtb7c is a critical gluconeogenic transcription factor that induces glucose-6-phosphatase and phosphoenylpyruvate carboxykinase 1 genes expression during mice fasting. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, *1862*(6), 643–656. https://doi.org/10.1016/j.bbagrm.2019.04.001

Churchfield, S. (1990). *The Natural History of Shrews*. Cornell University Press.

Churchfield, S. (2002). Why are shrews so small? The costs and benefits of small size in northern temperate Sorex species in the context of foraging habits and prey supply. *Acta Theriologica*, *47*, 169–184.

Cooper, N., Kamilar, J. M., & Nunn, C. L. (2012). Host longevity and parasite species richness in mammals. *PLoS ONE*, *7*(8). https://doi.org/10.1371/journal.pone.0042190

Coté, I. M., & Poulin, R. (1995). Parasitism and group size in social animals: a meta-analysis. In *Behavioral Ecology* (Vol. 6). http://beheco.oxfordjournals.org/

Cowan, K. M., Shutler, D., Herman, T. B., & Stewart, D. T. (2009). Splenic mass of masked shrews, sorex cinereus, in relation to body mass, sex, age, day of the year, and bladder nematode, liniscus (=capillaria) maseri, infection. *Journal of Parasitology*, *95*(1), 228–230. https://doi.org/10.1645/GE-1566.1

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes and Development*, *25*(10), 1010–1022. https://doi.org/10.1101/gad.2037511

Dhar, G. A., Saha, S., Mitra, P., & Nag Chaudhuri, R. (2021). DNA methylation and regulation of gene expression: Guardian of our health. In *Nucleus (India)* (Vol. 64, Issue 3, pp. 259–270). Springer. https://doi.org/10.1007/s13237-021-00367-y

Downie, A. B. (1986). *Spatial distribution, movements and age structure in island and mainland populations of the masked shrew (Sorex cinereus Kerr)*.

Etherington, G. J., Heavens, D., Baker, D., Lister, A., McNelly, R., Garcia, G., Clavijo, B., Macaulay, I., Haerty, W., & di Palma, F. (2020). Sequencing smart: De novo sequencing and assembly approaches for a non-model mammal. *GigaScience*, *9*(5). https://doi.org/10.1093/gigascience/giaa045

Farhang-Fallah, J., Randhawa, V. K., Nimnual, A., Klip, A., Bar-Sagi, D., & Rozakis-Adcock, M. (2002). The Pleckstrin Homology (PH) Domain-Interacting Protein Couples the Insulin Receptor Substrate 1 PH Domain to Insulin Signaling Pathways Leading to Mitogenesis and

GLUT4 Translocation. *Molecular and Cellular Biology*, *22*(20), 7325–7336. https://doi.org/10.1128/mcb.22.20.7325-7336.2002

Forsman, K. A., & Malmquist, M. G. (1988). Evidence for echolocation in the common shrew, Sorex araneus. *Journal of Zoology*, *216*(4), 655–662. https://doi.org/10.1111/j.1469-7998.1988.tb02463.x

Foster, B. (1964). Evolution of Mammals on Islands. *Nature*, *202*, 234–235.

Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, *33*(1), 1–22. https://doi.org/10.18637/jss.v033.i01

Gardiner-Garden, M., & Frommer, M. (1987). CpG Islands in Vertebrate Genomes. *Journal of Molecular Biology*, *196*, 261–282. https://doi.org/10.1016/S0040-4020(02)00396-4

George, S. B. (1986). Evolution and Historical Biogeography of Soricine Shrews. *Systematic Biology*, *35*(2), 153–162.

George, S. B. (1988). Systematics, Historical Biogeography, and Evolution of the Genus Sorex. *Journal of Mammalogy*, *69*(3), 443–461.

Goltsman, M., Kruchenkova, E. P., Sergeev, S., Volodin, I., & Macdonald, D. W. (2005). "Island syndrome" in a population of Arctic foxes (Alopex lagopus) from Mednyi Island. *Journal of Zoology*, *267*(4), 405–418. https://doi.org/10.1017/S0952836905007557

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., White, O., Buell, C. R., & Wortman, J. R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biology*, *9*(1), R7. https://doi.org/10.1186/gb-2008-9-1-r7

Haghani, A., Lu, A., Li, C., Robeck, T., Belov, K., Breeze, C., Brooke, R., Clarke, S., Faulkes, C., Fei, Z., Ferguson, S., Finno, C., Gladyshev, V., Gorbunova, V., Goya, R., Hogan, A., Hogg, C., Hore, T., Kiaris, H., … Horvath, S. (2021). DNA Methylation Networks Underlying Mammalian Traits. *BioRxiv*. https://doi.org/10.1101/2021.03.16.435708

Hoff, K. J., Lomsadze, A., Borodovsky, M., & Stanke, M. (2019). *Whole-Genome Annotation with BRAKER* (pp. 65–95). https://doi.org/10.1007/978-1-4939-9173-0_5

Holder, J. L., Zhang, L., Kublaoui, B. M., Dileone, R. J., Oz, O. K., Bair, C. H., Lee, Y.-H., Zinn, A. R., Lloyd, J., & Zinn, R. (2004). Sim1 gene dosage modulates the homeostatic feeding response to increased dietary fat in mice. *Am J Physiol Endocrinol Metab*, *287*, 105–113. https://doi.org/10.1152/ajpendo.00446.2003.-Haploinsufficiency

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, *14*(10), 1–20.

Huggins, J. A., & Kennedy, M. L. (2018). Morphologic Variation in the Masked Shrew (Sorex cinereus) and the Smoky Shrew (S . fumeus). *The American Midland Naturalist*, *122*(1), 11–25.

Humann, J. L., Lee, T., Ficklin, S., & Main, D. (2019). *Structural and Functional Annotation of Eukaryotic Genomes with GenSAS. In: Kollmar M. (eds) Gene Prediction. Methods in Molecular Biology* (Humana, Ed.; Vol. 1962). https://doi.org/10.1007/978-1-4939-9173-0

Iguchi-Ariga, S. M. M., & Schaffner, W. (1989). CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes & Development*, *3*, 612–619.

Illingworth, R. S., & Bird, A. P. (2009). CpG islands - "A rough guide." *FEBS Letters*, *583*(11), 1713–1720. https://doi.org/10.1016/j.febslet.2009.04.012

620 Illingworth, R. S., Gruenewald-Schneider, U., Webb, S., Kerr, A. R. W., James, K. D., Turner,
621 D. J., Smith, C., Harrison, D. J., Andrews, R., & Bird, A. P. (2010). Orphan CpG Islands
622 Identify numerous conserved promoters in the mammalian genome. *PLoS Genetics*, *6*(9).
623 https://doi.org/10.1371/journal.pgen.1001134
624 Jasinska, A. J., Haghani, A., Zoller, J. A., Li, C. Z., Arneson, A., Ernst, J., Kavanagh, K.,
625 Jorgensen, M. J., Mattison, J. A., Wojta, K., Choi, O.-W., DeYoung, J., Li, X., Rao, A. W.,
626 Coppola, G., Freimer, N. B., Woods, R. P., & Horvath, S. (2021). Epigenetic clock and
627 methylation studies in vervet monkeys. *GeroScience*, *0123456789*.
628 https://doi.org/10.1007/s11357-021-00466-3
629 Johnson, R. C., Nelson, G. W., Troyer, J. L., Lautenberger, J. A., Kessing, B. D., Winkler, C. A.,
630 & O'Brien, S. J. (2010). Accounting for multiple comparisons in a genome-wide association
631 study (GWAS). *BMC Genomics*, *11*(1). https://doi.org/10.1186/1471-2164-11-724
632 Jones, P. A. (1999). The DNA methylation paradox. *Trends in Genetics*, *15*(1), 34–37.
633 https://doi.org/10.1016/S0168-9525(98)01636-9
634 Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J.,
635 Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M.,
636 Yong, S. Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein
637 function classification. *Bioinformatics*, *30*(9), 1236–1240.
638 https://doi.org/10.1093/bioinformatics/btu031
639 Juette, T., Garant, D., Jameson, J. W., & Réale, D. (2020). The island syndrome hypothesis is
640 only partially validated in two rodent species in an inland–island system. *Oikos*, *129*(11),
641 1739–1751. https://doi.org/10.1111/oik.07249
642 Karbiener, M., Glantschnig, C., Pisani, D. F., Laurencikiene, J., Dahlman, I., Herzig, S., Amri, E.
643 Z., & Scheideler, M. (2015). Mesoderm-specific transcript (MEST) is a negative regulator
644 of human adipocyte differentiation. *International Journal of Obesity*, *39*(12), 1733–1741.
645 https://doi.org/10.1038/ijo.2015.121
646 Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology*
647 *and Evolution*, *17*(5), 230–241. https://doi.org/10.1016/S0169-5347(02)02489-8
648 Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome
649 alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*,
650 *37*(8), 907–915. https://doi.org/10.1038/s41587-019-0201-4
651 Kita, M., Nakamura, Y., Okumura, Y., Ohdachi, S. D., Oba, Y., Yoshikuni, M., Kido, H., &
652 Uemura, D. (2004). Blarina toxin, a mammalian lethal venom from the short-tailed shrew
653 Blarina brevicauda: Isolation and characterization. *Proceedings of the National Academy of*
654 *Sciences of the United States of America*, *101*(20), 7542–7547.
655 https://doi.org/10.1073/pnas.0402517101
656 Kooke, R., Johannes, F., Wardenaar, R., Becker, F., Etcheverry, M., Colot, V., Vreugdenhil, D.,
657 & Keurentjes, J. J. B. (2015). Epigenetic basis of morphological variation and phenotypic
658 plasticity in Arabidopsis thaliana. *Plant Cell*, *27*(2), 337–348.
659 https://doi.org/10.1105/tpc.114.133025
660 Larison, B., Pinho, G. M., Haghani, A., Zoller, J. A., Li, C. Z., Finno, C. J., Farrell, C., Kaelin,
661 C. B., Barsh, G. S., Wooding, B., Robeck, T. R., Maddox, D., Pellegrini, M., & Horvath, S.
662 (2021). Epigenetic models developed for plains zebras predict age in domestic horses and
663 endangered equids. *Communications Biology*, *4*(1). https://doi.org/10.1038/s42003-021-
664 02935-z

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T., & Carey, V. J. (2013). Software for Computing and Annotating Genomic Ranges. *PLoS Computational Biology*, *9*(8). https://doi.org/10.1371/journal.pcbi.1003118

Lázaro, J., Dechmann, D. K. N., LaPoint, S., Wikelski, M., & Hertel, M. (2017). Profound reversible seasonal changes of individual skull size in a mammal. *Current Biology*, *27*(20), R1106–R1107. https://doi.org/10.1016/j.cub.2017.08.055

Lázaro, J., Hertel, M., Muturi, M., & Dechmann, D. K. N. (2019). Seasonal reversible size changes in the braincase and mass of common shrews are flexibly modified by environmental conditions. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-38884-1

Lázaro, J., Nováková, L., Hertel, M., Taylor, J. R. E., Muturi, M., Zub, K., & Dechmann, D. K. N. (2021). Geographic patterns in seasonal changes of body mass, skull, and brain size of common shrews. In *Ecology and Evolution* (Vol. 11, Issue 6, pp. 2431–2448). John Wiley and Sons Ltd. https://doi.org/10.1002/ece3.7238

Lee, H. S., & Park, T. (2020). The influences of DNA methylation and epigenetic clocks, on metabolic disease, in middle-aged Koreans. *Clinical Epigenetics*, *12*(1). https://doi.org/10.1186/s13148-020-00936-z

Lomolino, M. V., Van der Geer, A. A., Lyras, G. A., Palombo, M. R., Sax, D. F., & Rozzi, R. (2013). Of mice and mammoths: Generality and antiquity of the island rule. *Journal of Biogeography*, *40*(8), 1427–1439. https://doi.org/10.1111/jbi.12096

Lu, A. T., Fei, Z., Haghani, A., Robeck, T. R., Zoller, J. A., Li, C. Z., Zhang, J., Ablaeva, J., Adams, D. M., Almunia, J., Ardehali, R., Arneson, A., Baker, C. S., Belov, K., Blac, P., & Horvath, S. (2021). Universal DNA methylation age across mammalian tissues. *BioRxiv*, 1–28.

Lu, A. T., Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., Hou, L., Baccarelli, A. A., Li, Y., Stewart, J. D., Whitsel, E. A., Assimes, T. L., Ferrucci, L., & Horvath, S. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*, *11*(2), 303–327. https://doi.org/10.18632/aging.101684

Ma, X., & Kang, S. (2019). Functional implications of DNA methylation in adipose biology. *Diabetes*, *68*(5), 871–878. https://doi.org/10.2337/dbi18-0057

MacPherson, I. S., & Stewart, D. T. (2003). Characterisation of a unique dipeptidase allele in an insular population of masked shrews, Sorex cinereus (Insectivora: Soricidae). *Biochemical Systematics and Ecology*, *31*(6), 573–580. https://doi.org/10.1016/S0305-1978(02)00218-1

Matsubara, K., Ishibashi, Y., Ohdachi, S., & Matsuda, Y. (2001). A new primer set for sex identification in the genus Sorex (Soricidae, Insectivora). *Molecular Ecology Notes*, *1*(4), 241–242.

McAlpine, C. S. (2009). *Molecular characterization of the unique Peptidase A Allele within a population of Sorex cinereus on Bon Portage Island, Nova Scotia*. https://scholar.acadiau.ca/islandora/object/theses:571

McCrory, C., Fiorito, G., Hernandez, B., Polidoro, S., O'Halloran, A. M., Hever, A., Ni Cheallaigh, C., Lu, A. T., Horvath, S., Vineis, P., & Kenny, R. A. (2021). GrimAge Outperforms Other Epigenetic Clocks in the Prediction of Age-Related Clinical Phenotypes and All-Cause Mortality. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, *76*(5), 741–749. https://doi.org/10.1093/gerona/glaa286

McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M., & Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology*, *28*(5), 495–501. https://doi.org/10.1038/nbt.1630

Mendes-Soares, H., & Rychlik, L. (2009). Differences in swimming and diving abilities between two sympatric species of water shrews: Neomys anomalus and Neomys fodiens (Soricidae). *Journal of Ethology*, *27*(3), 317–325. https://doi.org/10.1007/s10164-008-0122-z

Meng, G., Li, Y., Yang, C., & Liu, S. (2019). MitoZ: A toolkit for animal mitochondrial genome assembly, annotation and visualization. *Nucleic Acids Research*, *47*(11). https://doi.org/10.1093/nar/gkz173

Meröndun, J., Murray, D. L., & Shafer, A. B. A. (2019). Genome-scale sampling suggests cryptic epigenetic structuring and insular divergence in Canada lynx. In *Molecular Ecology* (Vol. 28, Issue 13). https://doi.org/10.1111/mec.15131

Miller, J. M., Garant, D., Perrier, C., Juette, T., Jameson, J. W., Normandeau, E., Bernatchez, L., & Réale, D. (2021). Linking genetic, morphological, and behavioural divergence between inland island and mainland deer mice. *Heredity*. https://doi.org/10.1038/s41437-021-00492-z

Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology*, *38*(1), 23–38.

Morand, S., & Harvey, P. H. (2000). Mammalian metabolism, longevity and parasite species richness. *Proceedings of the Royal Society B: Biological Sciences*, *267*(1456), 1999–2003. https://doi.org/10.1098/rspb.2000.1241

Nagorsen, D. (1996). *Opossums, Shrews and Moles of British Columbia*. UBC Press.

Nannini, D. R., Joyce, B. T., Zheng, Y., Gao, T., Liu, L., Yoon, G., Huan, T., Ma, J., Jacobs, D. R., Wilkins, J. T., Ren, J., Zhang, K., Khan, S. S., Allen, N. B., Horvath, S., Lloyd-Jones, D. M., Greenland, P., & Hou, L. (2019). Epigenetic age acceleration and metabolic syndrome in the coronary artery risk development in young adults study. *Clinical Epigenetics*, *11*(1). https://doi.org/10.1186/s13148-019-0767-1

Ning, Z., Giordano, F., & Harry, E. (n.d.). *Scaff10X*. https://github.com/wtsi-hpag/Scaff10X

Novosolov, M., Raia, P., & Meiri, S. (2013). The island syndrome in lizards. *Global Ecology and Biogeography*, *22*(2), 184–191. https://doi.org/10.1111/j.1466-8238.2012.00791.x

Ochocińska, D., & Taylor, J. R. E. (2003). Bergmann's rule in shrews: geographical variation of body size in Palearctic Sorex species. *Biological Journal of the Linnean Society*, *78*, 365–381. https://academic.oup.com/biolinnean/article/78/3/365/2639823

Ochocińska, D., & Taylor, J. R. E. (2005). Living at the physiological limits: Field and maximum metabolic rates of the common shrew (Sorex araneus). *Physiological and Biochemical Zoology*, *78*(5), 808–818. https://doi.org/10.1086/431190

Oulion, S., Bertrand, S., & Escriva, H. (2012). Evolution of the FGF Gene Family. *International Journal of Evolutionary Biology*, *2012*, 1–12. https://doi.org/10.1155/2012/298147

Pappalardo, X. G., & Barra, V. (2021). Losing DNA methylation at repetitive elements and breaking bad. In *Epigenetics and Chromatin* (Vol. 14, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s13072-021-00400-z

Pavlidis, P., Jensen, J. D., Stephan, W., & Stamatakis, A. (2012). A critical assessment of storytelling: Gene ontology categories and the importance of validating genomic scans. *Molecular Biology and Evolution*, *29*(10), 3237–3248. https://doi.org/10.1093/molbev/mss136

Pinho, G. M., Martin, J. G. A., Farrell, C., Haghani, A., Zoller, J. A., Zhang, J., Snir, S., Pellegrini, M., Wayne, R. K., Blumstein, D. T., & Horvath, S. (2021). Hibernation slows epigenetic aging in yellow-bellied marmots. *BioRxiv*, 2021.03.07.434299. https://www.biorxiv.org/content/10.1101/2021.03.07.434299v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.03.07.434299v1.abstract

Posbergh, C. J., & Huson, H. J. (2021). All sheeps and sizes: a genetic investigation of mature body size across sheep breeds reveals a polygenic nature. *Animal Genetics*, *52*(1), 99–107. https://doi.org/10.1111/age.13016

Postma, E. (2014). Four decades of estimating heritabilities in wild vertebrate populations: improved methods, more data, better estimates? In Charmantier Anne, Garant Dany, & Kruuk Loeske E. B. (Eds.), *Quantitative Genetics in the Wild* (pp. 16–33). Oxford University Press.

Prado, N. A., Brown, J. L., Zoller, J. A., Haghani, A., Yao, M., Bagryanova, L. R., Campana, M. G., E. Maldonado, J., Raj, K., Schmitt, D., Robeck, T. R., & Horvath, S. (2021). Epigenetic clock and methylation studies in elephants. *Aging Cell*, *20*(7), 1–11. https://doi.org/10.1111/acel.13414

Pruitt, W. O. Jr. (1954). Aging in the Masked Shrew, Sorex cinereus cinereus Kerr. *Journal of Mammalogy*, *35*(1), 35–39.

Quadrana, L., & Colot, V. (2016). Plant Transgenerational Epigenetics. *Annual Review of Genetics*, *50*, 467–491. https://doi.org/10.1146/annurev-genet-120215-035254

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Raj, K., Szladovits, B., Haghani, A., Zoller, J. A., Li, C. Z., Black, P., Maddox, D., Robeck, T. R., & Horvath, S. (2021). Epigenetic clock and methylation studies in cats. *GeroScience*, *43*(5), 2363–2378. https://doi.org/10.1007/s11357-021-00445-8

Razin, A., & Riggs, D. (1980). DNA Methylation and Gene Function. *Science*, *210*(4470), 604–610.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47. https://doi.org/10.1093/nar/gkv007

Robeck, T. R., Fei, Z., Lu, A. T., Haghani, A., Jourdain, E., Zoller, J. A., Li, C. Z., Steinman, K. J., DiRocco, S., Schmitt, T., Osborn, S., van Bonn, B., Katsumata, E., Mergl, J., Almunia, J., Rodriguez, M., Haulena, M., Dold, C., & Horvath, S. (2021). Multi-species and multi-tissue methylation clocks for age estimation in toothed whales and dolphins. *Communications Biology*, *4*(1). https://doi.org/10.1038/s42003-021-02179-x

Roland, A. E. (1982). *Geological background and physiography of Nova Scotia*. Nova Scotian Institute of Science.

Rudd, R. L. (1955). Age, sex, and weight comparison in three species of shrews. *Journal of Mammalogy*, *36*(3), 323–338.

Saino, N., Ambrosini, R., Albetti, B., Caprioli, M., de Giorgio, B., Gatti, E., Liechti, F., Parolini, M., Romano, A., Romano, M., Scandolara, C., Gianfranceschi, L., Bollati, V., & Rubolini, D. (2017). Migration phenology and breeding success are predicted by methylation of a photoperiodic gene in the barn swallow. *Scientific Reports*, *7*(January), 1–10. https://doi.org/10.1038/srep45412

Schachtschneider, K. M., Schook, L. B., Meudt, J. J., Shanmuganayagam, D., Zoller, J. A., Haghani, A., Li, C. Z., Zhang, J., Yang, A., Raj, K., & Horvath, S. (2021). Epigenetic clock

and DNA methylation analysis of porcine models of aging and obesity. *GeroScience*, *43*(5), 2467–2483. https://doi.org/10.1007/s11357-021-00439-6

Schaeffer, P. J., O'Mara, M. T., Breiholz, J., Keicher, L., Lázaro, J., Muturi, M., & Dechmann, D. K. N. (2020). Metabolic rate in common shrews is unaffected by temperature, leading to lower energetic costs through seasonal size reduction. *Royal Society Open Science*, *7*(4). https://doi.org/10.1098/rsos.191989

Schapira, A. H. (2006). Mitochondrial disease. *The Lancet*, *368*(9529), 70–82. https://doi.org/10.1016/S0140-6736(06)68970-8

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. v., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Skotarczak, E., Szwaczkowski, T., & Ćwiertnia, P. (2020). Effects of inbreeding, sex and geographical region on survival in an American bison (Bison bison) population under a captive breeding program. *European Zoological Journal*, *87*(1), 402–411. https://doi.org/10.1080/24750263.2020.1797194

Smit, A., & Hubley, R. (n.d.-a). *RepeatMasker Open-4.0*. Retrieved February 3, 2022, from http://www.repeatmasker.org

Smit, A., & Hubley, R. (n.d.-b). *RepeatModeler Open-1.0*. Retrieved February 3, 2022, from http://www.repeatmasker.org

Smith, A. K., Kilaru, V., Kocak, M., Almli, L. M., Mercer, K. B., Ressler, K. J., Tylavsky, F. A., & Conneely, K. N. (2014). Methylation quantitative trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. *BMC Genomics*, *15*(1). https://doi.org/10.1186/1471-2164-15-145

Stewart, D. T., & Baker, A. J. (1992). Genetic differentiation and biogeography of the masked shrew in Atlantic Canada. *Canadian Journal of Zoology*, *70*(1), 106–114. https://doi.org/10.1139/z92-016

Stewart, D. T., Herman, T. B., & Teferi, T. (1989). Littoral feeding in a high-density insular population of Sorex cinereus. *Canadian Journal of Zoology*, *67*(8), 2074–2077. https://doi.org/10.1139/z89-295

Sugrue, V. J., Zoller, J. A., Narayan, P., Lu, A. T., Ortega-Recalde, O. J., Grant, M. J., Bawden, C. S., Rudiger, S. R., Haghani, A., Bond, D. M., Hore, R. R., Garratt, M., Sears, K. E., Wang, N., Yang, X. W., Snell, R. G., Hore, T. A., & Horvath, S. (2021). Castration delays epigenetic aging and feminizes dna methylation at androgen-regulated loci. *ELife*, *10*, 1–20. https://doi.org/10.7554/eLife.64932

Swanson, E. M., & Dantzer, B. (2014). Insulin-like growth factor-1 is associated with life-history variation across Mammalia. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1782). https://doi.org/10.1098/rspb.2013.2458

Takahashi, M., Kamei, Y., Ezaki, O., Takahashi, Y., & Kamei, O. E. (2005). Mest/Peg1 imprinted gene enlarges adipocytes and is a marker of adipocyte size. *Am J Physiol Endocrinol Metab*, *288*, 117–124. https://doi.org/10.1152/ajpendo

Taylor, J. R. E., Rychlik, L., & Churchfield, S. (2013). Winter reduction in body mass in a very small, nonhibernating mammal: Consequences for heat loss and metabolic rates. *Physiological and Biochemical Zoology*, *86*(1), 9–18. https://doi.org/10.1086/668484

Telfer, K. A. (1984). *Microhabitat associations, morphopology, and demography of shrews (Soricidae) in Nova Scotia*.

Tomasi, T. E. (1979). Echolocation by the Short-Tailed Shrew Blarina brevicauda. In *Source: Journal of Mammalogy* (Vol. 60, Issue 4).

van Egeren, D., Madsen, T., & Michor, F. (2018). Fitness variation in isogenic populations leads to a novel evolutionary mechanism for crossing fitness valleys. *Communications Biology*, *1*(1). https://doi.org/10.1038/s42003-018-0160-1

van Valen, L. (1973). Body Size and Numbers of Plants and Animals. *Evolution*, *27*(1), 27–35.

Vernaz, G., Hudson, A. G., Santos, M. E., Fischer, B., Carruthers, M., Shechonge, A. H., Gabagambi, N. P., Tyers, A. M., Ngatunga, B. P., Malinsky, M., Durbin, R., Turner, G. F., Genner, M. J., & Miska, E. A. (2021). Epigenetic Divergence during Early Stages of Speciation in an African Crater Lake Cichlid Fish. *BioRxiv*, *doi:10.1101/2021.07.30.435319*, 1–12.

Vernaz, G., Malinsky, M., Svardal, H., Du, M., Tyers, A. M., Santos, M. E., Durbin, R., Genner, M. J., Turner, G. F., & Miska, E. A. (2021). Mapping epigenetic divergence in the massive radiation of Lake Malawi cichlid fishes. *Nature Communications*, *12*(1), 1–13. https://doi.org/10.1038/s41467-021-26166-2

Vogt, G. (2017). Facilitation of environmental adaptation and evolution by epigenetic phenotype variation: Insights from clonal, invasive, polyploid, and domesticated animals. *Environmental Epigenetics*, *3*(1), 1–17. https://doi.org/10.1093/eep/dvx002

Wan, J., Oliver, V. F., Wang, G., Zhu, H., Zack, D. J., Merbs, S. L., & Qian, J. (2015). Characterization of tissue-specific differential DNA methylation suggests distinct modes of positive and negative gene expression regulation. *BMC Genomics*, *16*(1). https://doi.org/10.1186/s12864-015-1271-4

Watt, F., & Molloy, P. L. (1988). Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes & Development*, *2*, 1136–1143.

Weisenfeld, N. I., Kumar, V., Shah, P., Church, D. M., & Jaffe, D. B. (2018). Corrigendum: Direct determination of diploid genome sequences. *Genome Research*, *28*(4), 757–767. https://doi.org/10.1101/gr.235812.118

Whitaker, J. O. Jr., & French, T. W. (1984). Foods of six species of sympatric shrews from New Brunswick. *Canadian Journal of Zoology*, *62*, 622–626.

Whitaker John O. Jr., & French, T. W. (1984). Foods of six species of sympatric shrews from New Brunswick. *Canadian Journal of Zoology*, *62*(4), 622–626.

Wilkinson, G. S., Adams, D. M., Haghani, A., Lu, A. T., Zoller, J., Breeze, C. E., Arnold, B. D., Ball, H. C., Carter, G. G., Cooper, L. N., Dechmann, D. K. N., Devanna, P., Fasel, N. J., Galazyuk, A. v., Günther, L., Hurme, E., Jones, G., Knörnschild, M., Lattenkamp, E. Z., … Horvath, S. (2021). DNA methylation predicts age and provides insight into exceptional longevity of bats. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-21900-2

Wolf, J. F., Bowman, J., Keobouasone, S., Taylor, R. S., & Wilson, P. J. (2022). A de novo genome assembly and annotation of the southern flying squirrel (Glaucomys volans). *G3: Genes, Genomes, Genetics*, *12*(1). https://doi.org/10.1093/G3JOURNAL/JKAB373

Wood, D. E., & Salzberg, S. L. (2014). Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*, *15*(3). https://doi.org/10.1186/gb-2014-15-3-r46

Yang, C., Gagnon, D., Vachon, P., Tremblay, A., Levy, E., Massie, B., & Michaud, J. L. (2006). Adenoviral-mediated modulation of Sim1 expression in the paraventricular nucleus affects

food intake. *Journal of Neuroscience*, *26*(26), 7116–7120. https://doi.org/10.1523/JNEUROSCI.0672-06.2006

Yom-Tov, Y., & Yom-Tov, J. (2005). Global warming, Bergmann's rule and body size in the masked shrew Sorex cinereus Kerr in Alaska. *Journal of Animal Ecology*, *74*(5), 803–808. https://doi.org/10.1111/j.1365-2656.2005.00976.x

Yordy, J., Kraus, C., Hayward, J. J., White, M. E., Shannon, L. M., Creevy, K. E., Promislow, D. E. L., & Boyko, A. R. (2020). Body size, inbreeding, and lifespan in domestic dogs. *Conservation Genetics*, *21*(1), 137–148. https://doi.org/10.1007/s10592-019-01240-x

Yu, G., Wang, L. G., & He, Q. Y. (2015). ChIP seeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, *31*(14), 2382–2383. https://doi.org/10.1093/bioinformatics/btv145

Zhou, W., Triche, T. J., Laird, P. W., & Shen, H. (2018). SeSAMe: Reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Research*, *46*(20), 1–15. https://doi.org/10.1093/nar/gky691

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. In *J. R. Statist. Soc. B* (Vol. 67, Issue 2).

Data accessibility statement

Raw sequence reads are deposited in the SRA: PRJNA826195

Genome is deposited in NCBI: Pending

Scripts and normalized β values for each probe and sample uploaded on GitLab:

https://gitlab.com/WiDGeT_TrentU/graduate_theses/-/tree/master/Cossette.

Benefit-sharing statement

Benefits Generated: Benefits from this research accrue from the sharing of our data and results on public databases as described above.

Author contribution:

Marie-Laurence Cossette: Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft preparation, Visualization, Project administration, Funding acquisition. **Don T. Stewart:** Conceptualization, Methodology, Resources, Writing – review and

editing, Supervision, Funding acquisition. **Amin Haghani:** Writing – review and editing, Software, Data curation. **Joseph Zoller:** Writing – review and editing, Software, Data curation. **Aaron B.A. Shafer:** Conceptualization, Methodology, Validation, Investigation, Resources, Writing – original draft preparation, Supervision, Project administration, Funding acquisition. **Steve Horvath:** Conceptualization, Methodology, Software, Validation, Resources, Data curation, Writing – review and editing, Supervision, Project administration, Funding acquisition.

Tables and figures

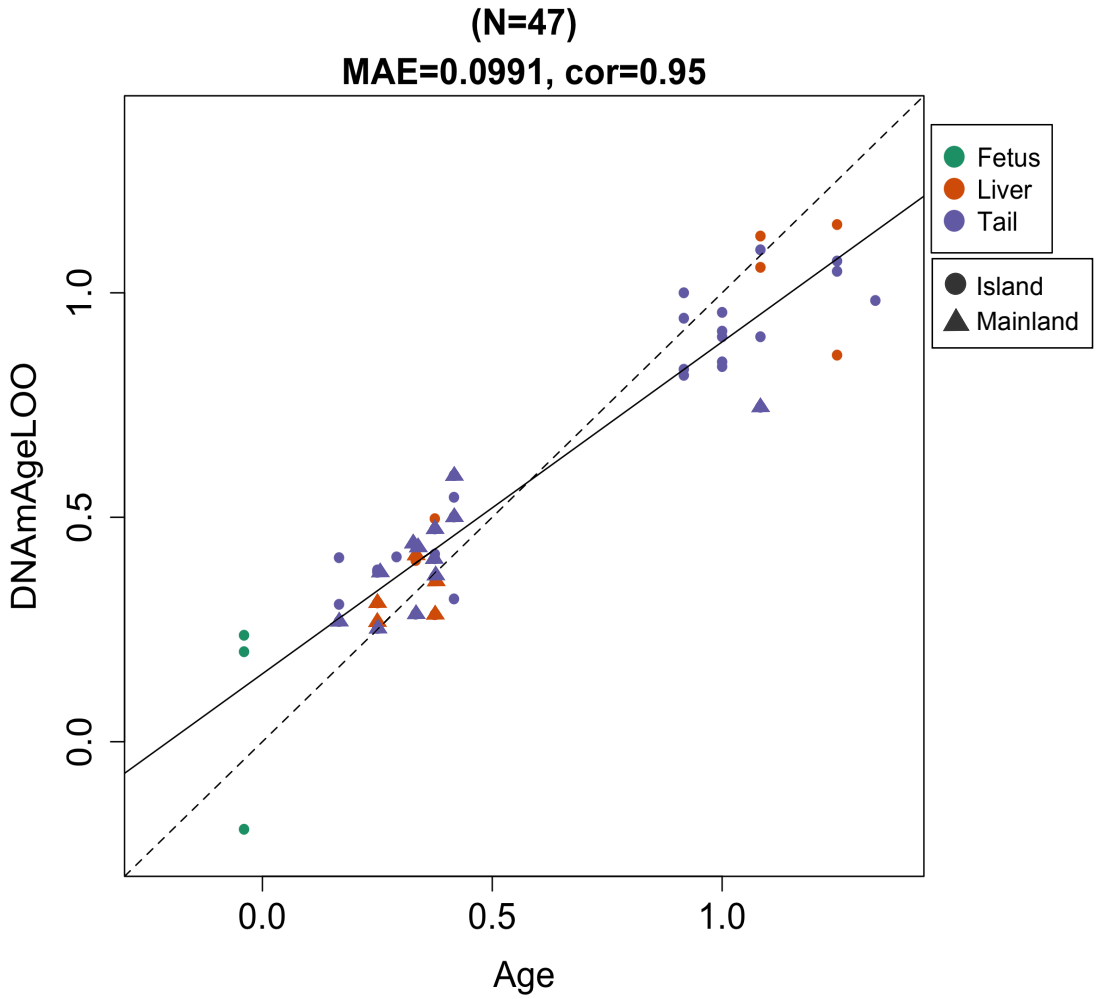

Figure 1. Cross-validation study of pan-tissue epigenetic clock for masked shrews. Leave-one-sample-out (LOO) estimate (y-axis, in units of years) versus chronological age (x-axis, in unit of years). The linear regression of

epigenetic age is indicated by a solid line while the diagonal line (y = x) is depicted by a dashed line. “MAE” represents the mean absolute error and “cor” represents the correlation coefficient

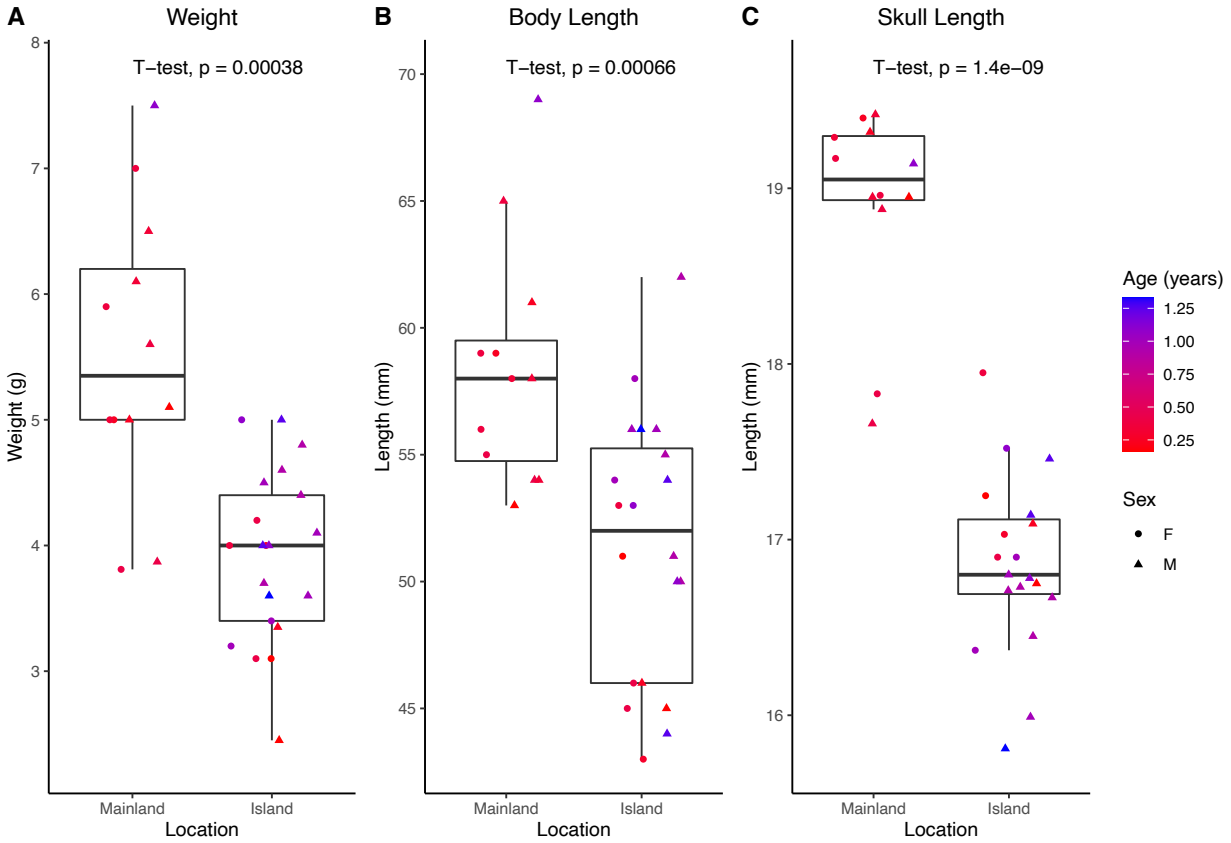

Figure 2. Boxplot of **A** weight (g), **B** body length (mm) and **C** skull length (mm) between mainland and island masked shrew populations.

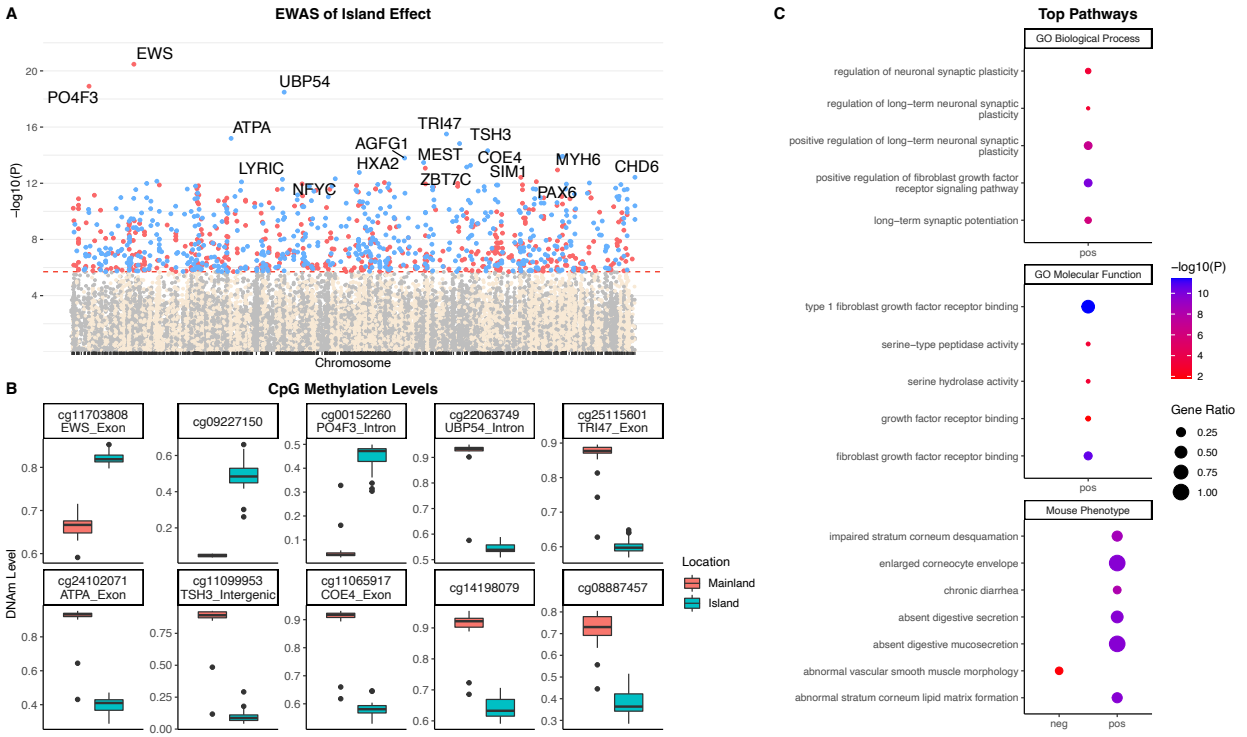

Figure 3. Epigenome-wide association (EWAS) of the multivariate regression model (location, age, sex and tissue) for masked shrew. **A** Manhattan plots of the EWAS for location (island versus mainland). The coordinates are estimated based on the alignment of 29,609 Mammalian array probes to our masked shrew genome assembly. The direction of associations with < 2e-6 (red dotted line) is highlighted by red (hypermethylated) and blue (hypomethylated) colors relative to mainland methylation levels. Top 15 CpGs are indicated by their neighboring genes. **B** DNAm levels of island samples versus mainland for the top 10 significant CpGs (*p* value) associated to location. **C** Enrichment analysis of the top CpGs with positive (hypermethylated) and negative (hypomethylated) correlations to island populations. The gene-level enrichment analysis was carried out using the GREAT software. Background probes were limited to 14,290 probes that had shrew gene annotations. The top ontologies with most significance pathways for island samples were selected based on Bonferroni corrected p-values.

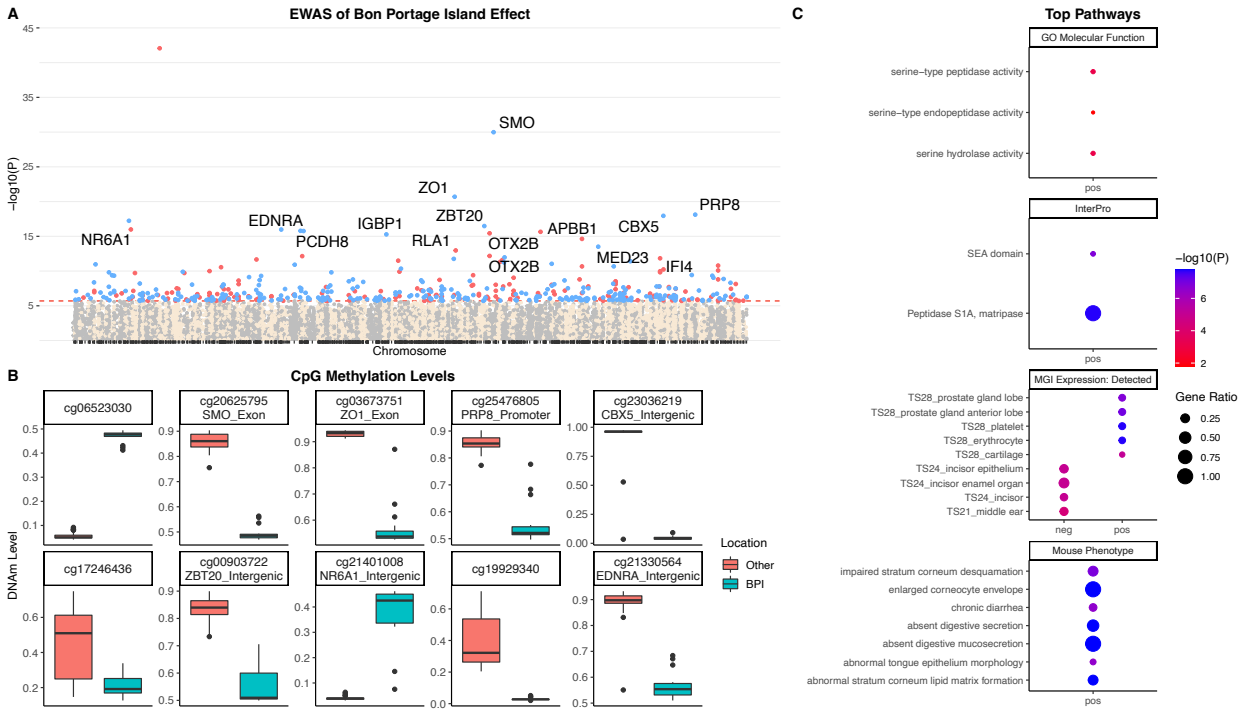

Figure 4. Epigenome-wide association (EWAS) of the multivariate regression model (location, age, sex and tissue) for masked shrew. **A** Manhattan plots of the EWAS for location (Bon Portage Island (BPI) versus others). The coordinates are estimated based on the alignment of 29,609 Mammalian array probes to our masked shrew genome assembly. The direction of associations with < 2e-6 (red dotted line) is highlighted by red (hypermethylated) and blue (hypomethylated) colors relative to non-BPI methylation levels. Top 15 CpGs are indicated by their neighboring genes. **B** DNAm levels of BPI samples versus others for the top 10 significant CpGs (*p* value) associated to location. **C** Enrichment analysis of the top CpGs with positive (hypermethylated) and negative (hypomethylated) correlations to BPI. The gene-level enrichment analysis was carried out using the GREAT software. Background probes were limited to 14,290 probes that had shrew gene annotations. The top 4 ontologies with most significance pathways for BPI were selected based on Bonferroni corrected p-values.

Table 1. Trapping location of samples, N = total number of samples, number of females, minimum and maximum age. Age is represented as the estimated age in months divided by 12. Note, age of 0.00 represents fetus.

Location	Island	N	No. of females	Age (min/max)
Bon Portage Island, NS	Yes	16	6	0.167/1.33
Long Island, NS	Yes	9	4	0.00/1.25
Sandy Cove, NS	No	4	3	0.250/0.375
North Mountain, NS	No	2	1	0.417/0.417
Peterborough County, ON	No	6	1	0.167/1.08

976 **Table 2.** Summary statistics for the linear regression models for weight, body length and skull length of mainland
977 versus island masked shrews. Sex values are relative to females and location relative to mainland populations. The
978 β-coefficient (estimate), 95% confidence interval (CI), p-value and adjusted R^2 are provided.
979

	Estimate	95% CI	P value	Adj. R^2
Model 1: Weight ~ Age + Sex + Location				0.53
Age	1.15	0.21 to 2.08	**0.018**	
Sex	0.01	– 0.61 to 0.64	0.962	
Location	– 2.09	– 2.79 to – 1.39	**<0.001**	
Model 2: Body length ~ Age + Sex + Location				0.49
Age	8.43	3.36 to 13.50	**0.002**	
Sex	– 0.20	– 3.61 to 3.20	0.903	
Location	– 10.35	– 14.15 to – 6.54	**<0.001**	
Model 3: Skull length ~ Age + Sex + Location				0.79
Age	– 0.31	– 0.93 to 0.32	0.321	
Sex	– 0.21	– 0.63 to 0.21	0.319	
Location	– 1.92	– 2.39 to – 1.45	**<0.001**	

980
981
982 **Table 3.** Top 5 genes near outlier CpGs for island versus mainland EWAS and BPI versus all other populations.
983 Functions of each gene are abbreviated from Uniprot.
984

Gene name	CpG	Direction	Location (p)	Putative Function
Island-mainland				
EWS exon	cg11703808	pos	3.35e-21	Transcriptional repressor and might play a role in the tumorigenic process.
PO4F3 intron	cg00152260	pos	1.23e-19	Transcriptional activator. Involved in the auditory system development.
UBP54 intron	cg22063749	neg	3.27e-19	Has no peptidase activity.
TRI47 exon	cg25115601	neg	3.12e-16	Mediates the ubiquitination and proteasomal degradation of CYLD.
ATPA exon	cg24102071	neg	6.37e-16	Produces ATP from ADP
BPI				
SMO exon	cg20625795	neg	1.03e-30	Associates with the patched protein (PTCH) to transduce the hedgehog's proteins signal.
ZO1 exon	cg03673751	neg	2.00e-21	Related to movement of substances through the paracellular space. Plays a role in the regulation of cell migration.
PRP8 promoter	cg25476805	neg	7.60e-19	Plays role in pre-mRNA splicing.
CBX5 intergenic	cg23036219	neg	1.12e-18	Associated with epigenetic repression. Involved in the formation of functional kinetochore.
ZBT20 intergenic	cg00903722	neg	3.32e-17	Transcription factor that may be involved in hematopoiesis, oncogenesis, and immune responses. Plays a role in postnatal myogenesis.

985