

## **Methylation studies in *Peromyscus*: aging, altitude adaptation, and monogamy**

Steve Horvath<sup>1,2#</sup>, Amin Haghani<sup>1#</sup>, Joseph A. Zoller<sup>2</sup>, Asieh Naderi<sup>3</sup>, Elham Soltanmohammadi<sup>3</sup>, Elena Farmaki<sup>3</sup>, Vimala Kaza<sup>4</sup>, Ioulia Chatzistamou<sup>5</sup>, Hippokratis Kiaris<sup>3,4\*</sup>

<sup>1</sup> Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, USA;

<sup>2</sup> Department of Biostatistics, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, California, USA;

<sup>3</sup> Department of Drug Discovery and Biomedical Sciences, College of Pharmacy, University of South Carolina, SC, USA;

<sup>4</sup> *Peromyscus* Genetic Stock Center, University of South Carolina, SC, USA.

<sup>5</sup> Department of Pathology, Microbiology and Immunology, School of Medicine, University of South Carolina, SC, USA.

# Joint first authorship

\* Corresponding authors: SH ([shorvath@mednet.ucla.edu](mailto:shorvath@mednet.ucla.edu)) and HK ([hk@sc.edu](mailto:hk@sc.edu))

**Key words:** deer mouse, aging, epigenetic clock, DNA methylation,

Emails:

[shorvath@mednet.ucla.edu](mailto:shorvath@mednet.ucla.edu)

[ahaghani@g.ucla.edu](mailto:ahaghani@g.ucla.edu)

[jaz18@g.ucla.edu](mailto:jaz18@g.ucla.edu)

[ANADERI@mailbox.sc.edu](mailto:ANADERI@mailbox.sc.edu)

[elhams@email.sc.edu](mailto:elhams@email.sc.edu)

[efarmaki@coh.org](mailto:efarmaki@coh.org)

[kaza@email.sc.edu](mailto:kaza@email.sc.edu)

[chatzistamou@uscmed.sc.edu](mailto:chatzistamou@uscmed.sc.edu)

[hk@sc.edu](mailto:hk@sc.edu)

## ABSTRACT

DNA methylation-based biomarkers of aging have been developed for humans and many other mammals and could be used to assess how stress factors impact aging. Deer mice (*Peromyscus*) are long living rodents that have emerged as an informative model to study aging, adaptation at extreme environments, and monogamous behavior. In the present study we have undertaken an exhaustive, genome-wide analysis of DNA methylation in *Peromyscus*, spanning different species, stocks, sexes, tissues and age cohorts. We describe DNA methylation-based estimators of age for different species of deer mice based on novel DNA methylation data generated on highly conserved mammalian CpGs measured with a custom array. The multi-tissue epigenetic clock for deer mice was trained on 3 tissue sources (tail, liver, brain). Two dual species human-peromyscus clocks accurately measure age and relative age defined as the ratio of chronological age to maximum age. These analyses also allowed us to accurately manifest the increasing impact of age, sex, genetic relatedness, and ultimately tissue identity, in that order, in the acquisition of specific methylation patterns in the genome. Genes that were differentially methylated across different biological variables were determined and their potential impact is discussed. This study describes highly accurate DNA methylation-based estimators of age in deer mice and illustrates how differential methylation may be linked to adaptation at different conditions.

## INTRODUCTION

CpG methylation of DNA is an epigenetic modification that is influenced by multiple variables. Those range from transcriptional programs in response to environmental stimuli and cellular differentiation cues, to fundamental evolutionary processes associated with adaptation at different environments.

Epigenetic markers of aging based on DNA methylation data can accurately estimate chronological age for any tissue across the entire lifespan of mammals<sup>1-7</sup>. These DNAm-based age estimators, also known as epigenetic clocks, profile methylation levels of many CpG loci and apply penalized regression models to predict chronological age based on DNA methylation levels (reviewed in<sup>7</sup>). By examining various mammalian tissues and cell types, a robust correlation between chronological age and DNAm age over the course of entire lifespans has been well established with the human pan tissue DNAm age estimator<sup>8</sup>. Similar pan tissue clocks have been established in mice and other mammals<sup>5,6</sup>.

Although the impact of each of the factors influencing DNA methylation has been documented extensively, their relative contribution in the emergence of specific methylation patterns, across different scales of biological organization, remains poorly understood. In an effort to better understand how methylation profiles change in relation to different biological variables we have undertaken an exhaustive, genome-wide analysis of DNA methylation patterns in *Peromyscus*, spanning different species, stocks, sexes, tissues and age cohorts. Rodents of the genome *Peromyscus* are appealing models for addressing various biological questions in relation to aging because they live up to 8 years in captivity<sup>9</sup>. They are also used to study metabolism, infectious diseases, adaptation at extreme environments such as high altitude and the desert, monogamous behavior as well behavioral responses in response to anxiety stress<sup>9-12</sup>.

Different species of *Peromyscus* are maintained as closed colonies of outbred, genetically diverse stocks at the *Peromyscus* Genetic Stock Center. The present analysis involved 36,000 CpGs that are highly conserved across mammals in DNA and was applied to specimens from two *Peromyscus* subgenera, 5 species and one interspecific hybrid, 3 tissues (tails, brain and liver) and

individuals from ages ranging from 2 months to about 3.6 years old. The specimens analyzed involved both males and females as well as individuals from two different closed colonies of *P. maniculatus*. This experimental set up facilitated a simultaneous, unbiased analysis of DNA methylation signatures in specimens differing across different levels of biological organization. We also developed a highly accurate pan-tissue deer mouse DNAm clock for relating DNAm age with chronological age across multiple tissues. We next assessed the degree to which the deer mouse DNAm clock is conserved, by comparing it to the human DNAm clock to assess the potential translatability of the two. We then turned our focus to determine if epigenetic aging impacts differently monogamous and polygamous animals (across and within specific tissues). Finally, we performed epigenome-wide association studies (EWAS) to assess the degree to which specific genes show differential epigenetic modification as a response to aging and monogamous behavior.

## RESULTS

### **Data sets**

We used a custom methylation array (HorvathMammalMethylChip40) to generate DNA methylation data from six species of peromyscus: *Peromyscus californicus* (n=16), *Peromyscus eremicus* (n=17), *Peromyscus hybrid* between *polionotus* and *maniculatus* (n=6), *Peromyscus leucopus* (n=36), *Peromyscus maniculatus* (n=53), and *Peromyscus polionotus* (n=16). Hierarchical clustering indicated the n=6 samples were technical outliers (Figure 1). These samples were subsequently removed from the analysis.

These DNA samples came from three tissues/organs: whole brain, tail, and liver as detailed in **Table 1**. The ages ranged from 0.083 years to 3.6 years. Additionally, we used DNA methylation profiles from 1207 human samples, from several tissues and with a large age range, to construct two dual species human-peromyscus epigenetic clocks. These human data were generated on the same custom methylation array, which was designed to facilitate cross species comparisons across mammals.

### **Unsupervised hierarchical clustering of deer mouse methylation data**

Unsupervised hierarchical clustering of methylation data was initially performed that identified the tissue identity as the most prominent discriminator of global methylation signatures (Figure 1). Thus, the profile of DNA methylation is primarily guided by function which in turn underscores the impact of this epigenetic modification in regulating gene transcription and therefore guiding cellular differentiation. Within the same tissue, clustering occurred in a manner that overlapped astonishingly with the evolutionary history. Initially, two branches emerged, with the first including *P. californicus* and *P. eremicus*, and the second comprised of *P. leucopus*, *P. maniculatus* and *P. polionotus*, that signify two distinct groups in the evolution of *Peromyscus*<sup>13</sup>. Among them, the highly related *P. maniculatus* and *P. polionotus* clustered together, while *P. leucopus* has diverged earlier. Indeed, *P. maniculatus* and *P. polionotus* can form interspecific F1 hybrids that in relation to their methylation patterns are more closely related to the parental strain, *P. polionotus* (Figure 1) pointing to the impact of the parent of origin effects in guiding methylation signatures<sup>14</sup>. Among the same species, genetic relevance appeared to be of significance. Individuals from two closed *P. maniculatus* colonies were evaluated, *bairdii* and *sonoriensis* (BW and SM2 stocks respectively)<sup>9</sup>. Individuals from each of these colonies clustered accurately together suggesting that genetic relatedness is capable of triggering similar patterns of DNA methylation that surpass those inflicted by the sex and the age of the

individuals from which the DNA samples have been isolated. Within the same species and stocks, clustering occurred according to sex. The fact that the analysis involved tails, whole brain and liver samples that are not major target tissues for sex hormones implies that sex-specific patterns of methylation are inflicted early during development that persist at adulthood.

### **Epigenetic clocks**

To arrive at unbiased estimates of our DNA methylation-based age estimators, we performed a cross-validation study in the training data. The cross-validation study reports unbiased estimates of the age correlation  $R$  (defined as Pearson correlation between the age estimate (DNAm age) and chronological age as well as the median absolute error. Our different clocks can be distinguished along two dimensions (species and measure of age). The pan tissue clock for deer mice applies to multiple tissues from deer mice ( $R = 0.9$  and a median error of 0.24 years=3 months, **Figure 2A**). We also developed tissue specific clocks for brain samples ( $R=0.78$ ,  $MAE=0.49$  years, **Figure 2B**), liver ( $R=0.94$ ,  $MAE=0.20$ , **Figure 2C**), tail ( $R=0.95$ ,  $MAE=0.16$ , **Figure 2D**). We have two different clocks that apply to both humans and peromyscus. The human-peromyscus pan-tissue clock for chronological age can be used to estimate the chronological age of humans and deer mice using the same mathematical formula. The human-peromyscus clock exhibits a high age correlation across both species ( $R = 0.99$ , **Figure 2E**) and when restricted to peromyscus samples ( $R = 0.79$ , **Figure 2F**). The human-peromyscus clock of *relative age*, defined as the ratio of chronological age to maximum lifespan, achieves a similar performance ( $R = 0.97$  and  $R = 0.79$ , **Figure 2G,H**). By definition, the relative age takes values between 0 and 1 and arguably provides a biologically more meaningful comparison between species with different lifespan (deer mouse and human), which is not afforded by mere measurement of absolute age.

### **EWAS of age**

In total, 29,125 probes from the mammalian chip (HorvathMammalMethylChip40) were aligned to specific loci approximate to 5048 genes in the deer mouse genome (*Peromyscus maniculatus bairdii*.HU\_Pman\_2.1.100). These probes have high conservation in mammals, thus comparable among different species. In this study, we first examined the CpG level age related changes in brain, liver, and tail of different *Peromyscus* species. We also added the liver and tail data from B6 mice (*Mus musculus*) for a comparative purpose. In general, DNAm aging in *Peromyscus* species was moderately similar to B6 mice, but at different degrees (Error! Reference source not found.; Error! Reference source not found.). For example, while *P. maniculatus* and *P. eremicus* tail DNAm aging correlation with B6 mice was only 0.1, *P. californicus*, *P. polionotus*, and *P. leucopus* had a correlation between 0.3-0.5. In another analysis, we compared DNAm aging of *P. maniculatus* with other four *Peromyscus* species (**Figure 3**). Similarly, pairwise correlations differed within these species. Such a difference reflects the evolutionary differences of these sibling animals. Despite these differences, there were numerous CpGs with converged aging patterns in *Peromyscus* genus. Some of the shared changes in these animals included hypermethylation in *Bdnf* promoter, *Igsf9b* exon, *Nkx2-9* downstream, *Lhfp14* exon, *Elavl4* intron, and *Trhde* promoter. As expected, age related CpGs were distributed in all genic and intergenic regions relative to transcriptional start site in both directions, but mainly with hypermethylation change in the promoters (Error! Reference source not found.).

*Peromyscus* genus comprises of species that were evolved in diverse range habitats, from deserts to high altitude mountains, in environments with low to high extrinsic risk of predation. Thus,

each one of these species gained several unique features to adapt, reproduce and survive in these environments. In the following sections, we will describe some of the lessons that we learn from a comparative epigenetic analysis of these interesting characteristics. Understanding the epigenetic mechanism of these traits could be a novel roadmap to underpin the gene-to-trait function and expand the findings to other species such as human.

### **Epigenetic profile of North American deer mouse change along the altitude**

The present analysis involved individuals from two colonies of *P. maniculatus*, *sonoriensis* (SM2 stock) and *bairdii* (BW stock). The original founders of these colonies have evolved in low- or high-altitude environments and since then are maintained as closed colonies in our facilities. We identified a total of 1,273 CpGs that differed in their baseline methylation at  $p < 10^{-4}$  (**Figure 4a**). Some of the top DNAm signatures in high-altitude *P. maniculatus* include hypomethylation in *ENSPEMG00000020040* downstream, *Madd* exon, and hypermethylation in *Nsd3* intron, *Pdzd8* exon, and *Stc1* promoter (**Figure 4a**). Functional enrichment studies of these differentially methylated CpGs show that they tend to be located near genes that play a role in the development of rhombomere 3, motor neurons, middle ear, and immune system functioning, antigen processing and presentation (**Error! Reference source not found.**).

Next, we examined the differences in aging patterns, i.e. differences in the age correlations between the two colonies of *P. maniculatus*. Aging effects showed moderately high correlation between the two colonies ( $r = 0.4$ , **Figure 4b,c**) but several CpGs exhibited significant correlations only in one of the colonies. DNAm aging comparison by altitude was done at a nominal  $p < 10^{-3}$  in order to implicate sufficient numbers of CpGs for our subsequent enrichment studies.

For example, CpGs in the promoter of *Nfia*, upstream of *Htatip2*, in an exon of *Mapre3*, and an exon of *Cxx4* correlated significantly with age only in the low altitude colony of *P. maniculatus*. Conversely, CpGs upstream of *Dlx6*, an exon of *Fam160a2*, an intron of *Fam241b* and upstream of *Pax2* correlated significantly with age only in the high altitude colony (**Figure 4d,e**). Interestingly, these altitude-specific age-related changes are adjacent to genes that play a role in brain development, immune system functioning, and T cell development (**Error! Reference source not found.**).

### **North American deer mouse brain has a unique DNAm profile**

Deer mice are long living rodents that exhibit variable lifespan depending on the species. Among them *P. leucopus* lives up to 8 years in captivity as compared to other *Peromyscus* species such as *P. maniculatus* that reportedly have a lifespan of about 4 years<sup>15</sup>. Comparison of brain specimens between older *P. leucopus* and *P. maniculatus* indicated that in the latter, coordination of the unfolded protein response is compromised, and evidence of neurodegenerative pathology was obtained<sup>16</sup>. Therefore, comparative analysis of *Peromyscus* species may be relevant to the study of age-related alterations in the brain. Here, we compared DNAm profile of the *P. maniculatus* brain to *P. leucopus* and recorded notable differences. Interestingly, the brain DNAm profiles of these animals were largely different. A total number of 2,396 CpGs were differentially methylated in these samples at  $p < 10^{-4}$  (**Figure 5a**). Some of the top *P. maniculatus* signatures included hypomethylation in *Cadsp2* exon, *Cas21* exon, and hyper methylation in *Grm8* exon, *Epha3* promoter, and *Lbx1* promoter. These differentially methylated regions enriched circadian rhythm, glucagon secretion, histone H2A ubiquitination, and some abnormal neuron specification in mouse phenotype database. Several brain regions are sensitive to glucagon by induction of cAMP signaling cascades that regulate peripheral

homeostasis<sup>17</sup>. Interestingly, brain insulin resistance is highly related to Alzheimer's disease pathology in human<sup>18</sup>, thus, glucose and glucagon sensing pathway in *P. maniculatus* brain seems an attractive study target for future studies.

Another striking finding in this analysis was a strong inverse correlation ( $r = -0.35$ ) of DNAm aging between *P. maniculatus* and *P. leucopus* brains (**Figure 5b**). This finding requires a larger sample size for validation nevertheless we could only identify two CpGs (adjacent to *Cnot4* exon, *Pax2* intron) that changed with age in the same direction in these species (**Figure 5c**). In contrast, several loci were identified that mainly changed in one of these brains, or even diverged during aging. *Map2k5* intron was an extreme example that was hypermethylated in *P. maniculatus*, but hypomethylated in *P. leucopus* (**Figure 5b,c**). While the unique DNAm aging changes in *P. leucopus* enriched angiogenesis-related processes, in *P. maniculatus* unique brain DNAm aging related to gamma delta T cells (Error! Reference source not found.), which are believed to contribute to neurodegeneration in the brain<sup>19</sup>.

### DNAm relate to pair-bonding behavior in deer mouse species

Monogamy is a perplexing strategy chosen by a relatively small number of mammals to enhance their fitness<sup>20</sup>. Pair bonds based on mating, that are associated with the development of monogamous behavior are estimated to occur in less than 10% of mammals, including humans<sup>21,22</sup>. In *Peromyscus*, monogamous behavior is fairly common and has developed independently at least twice in evolution<sup>21,23</sup>. Our study involved three monogamous (*P. californicus*, *P. polionotus*, and *P. eremicus*) and two polygamous (*P. maniculatus*, *P. leucopus*) species. Strikingly, there was a major baseline methylation difference between these species. In total, 9,411 CpGs were differentially methylated ( $p < 10^{-4}$ , median  $pval = 10^{-7}$ ) by monogamy status in these species (**Figure 6a**). The top changed hits related to monogamy included hypomethylation in *Zeb2* intron, *1700008P02Rik* upstream, *Cadps* intron, and hypermethylation in *Fer* exon, *Rnd3* exon, and *Srsf9* exon (**Figure 6b**). The top differentially methylated CpGs in this analysis related to a wide range of biology such as development, metabolism, and immune system. Interestingly, monogamous related hypomethylation also enriched synaptic dopamine release in mouse phenotype database. Dopamine seems to play a central role in pair bond formation, expression, and maintenance<sup>24</sup>. *Zeb2* gene, the top monogamy related gene in our analysis, is also a key regulator of midbrain dopaminergic neurons development<sup>25</sup>.

## DISCUSSION

The development of deer mouse epigenetic clocks described here was based on novel DNA methylation data that were derived from 3 deer mouse tissue types (brain, tail, and liver). We show that the pure pan-tissue deer mouse clock accurately relates chronological age with estimated DNAm age in different deer mouse tissues. This gives us confidence that these clocks will work on new samples from other tissue types as well. A critical step toward crossing the species barrier was the use of a mammalian DNA methylation array that profiled 36 thousand probes that were highly conserved across numerous mammalian species. The deer mouse DNA methylation profiles reported here represent the most comprehensive dataset thus far of matched single base resolution methylomes across multiple tissues and ages. The two human-*Peromyscus* clocks estimate chronological and relative age, respectively. The dual-species clock for relative age demonstrates the feasibility of building epigenetic clocks for two species based on a single mathematical formula. This dual species clock also effectively demonstrates that epigenetic aging mechanisms are highly

conserved between deer mice and humans. The mathematical operation of generating a ratio also yields a much more biologically meaningful value of aging because it indicates the relative biological age of the organism in relation to its own species. Providing an indicator of biological age empowers the possibility to gauge potential long-term survivability with implications for reproductive fitness potential and individual mate quality. We expect that the availability of these epigenetic clocks will provide a significant boost to the attractiveness of the deer mouse as a biological model in aging research.

Beyond their utility, the epigenetic clocks for deer mice reveal several salient features with regard to the biology of aging. First, the deer mouse pan-tissue clock re-affirms the implication of the human pan-tissue clock that aging might be a coordinated biological process that is harmonized throughout the body. Second, the ability to develop human-peromyscus pan-tissue clock for relative age attests to the high conservation of the aging process across two evolutionary distant species. This increases the likelihood, albeit does not guarantee, that conditions (such as pair bonding status) that alter the epigenetic age of deer mice, as measured using the human-peromyscus clock, will exert a similar effect in humans. Overall, this study provides evidence linking social monogamous life strategies with epigenetic aging in an attractive animal model.

Besides its implications in aging research our results collectively underscore the operation of a robustly maintained hierarchical association of the biological variables that can influence DNA methylation patterns. Processes associated with gene regulation and are reflected to the acquisition of a specific tissue identity operate as the major classifier and emerge as the most important determinant of methylation signatures. These are followed by evolutionary processes, are culminated by genetic relatedness, and range from the distinction to taxa to specific closed populations. Ultimately, the age of the individuals appears to be a factor that is superseded by all others and fine-tunes the profile of DNA methylation in a manner that precisely predicts biological aging.

Our analyses also revealed a series of interesting differentially methylated targets in relation to specific environmentally and biologically relevant conditions. Of interest are the differentially methylated genes that were identified in relation to the altitude at which the original founders of these colonies were captivated. These findings should be interpreted with caution because SM2 and BW stocks are highly diverse and are bred in isolation for extended periods which may have caused the fixation of methylation profiles that are irrelevant to the altitude adaptation of their original ancestors. While however, no apparent differences have been recorded between the SM2 and the BW stocks that could associate with rhombomere 3 and motor neurons development, genes related to immune response and middle ear development may be of relevance. The latter may reflect adaptations associated with the differential atmospheric pressures at high altitudes. As regards to the differential methylation of genes associated with immune system function, this may be relevant to the reported compromise of the immune system at high elevations<sup>26-28</sup>. To that end, high altitude deer mice may have engaged epigenetic strategies to counteract such immune system suppression. Of note is also the unique methylation profile of *P. maniculatus* in comparison with *P. leucopus* that corroborate the recently reported deregulation of stress response genes and the aberrant histological manifestations recorded in aged *P. maniculatus*<sup>16</sup>. Finally, of interest are the findings related to monogamy that identified high differences in overall baseline methylation between monogamous and polygamous species and a series of genes that are specifically methylated differentially between these groups, irrespectively of the species. The fact that these analyses involved tail tissues suggests that inherent differences in bonding behavior instructs specific epigenetic changes in peripheral tissues that may be translated into distinct physiological outcomes. Whether this is due to the differential regulation of

specific neurohormonal circuits in response to hormones and neurotransmitters related to bonding, and which the exact physiological outputs are, remains to be determined.

Collectively our study provided the first epigenetic clock for *Peromyscus* and illustrated the hierarchical association between various biological variables in determining methylation profiles across different scales of biological organization. Finally, it provided hints with regards to global differences and specific gene targets that are epigenetically impacted by biologically and environmentally relevant conditions.



## MATERIALS AND METHODS

Deer mice are maintained as outbred, genetically diverse closed colonies in the Peromyscus Genetic Stock center of the University of South Carolina. The study was approved by the Institutional Animal Care and Use Committee (IACUC) of the UofSC (protocol #: 2356-101506-042720) and were in accordance with the guidelines set forth by the National Institutes of Health. DNA was isolated from live animals by tail snips, or upon sacrifice from livers and brains by using DNeasy DNA isolation kit (Qiagen).

### *Human tissue samples*

To build the human-peromyscus clock, we analyzed previously generated methylation data from  $n=1207$  human tissue samples (adipose, blood, bone marrow, dermis, epidermis, heart, keratinocytes, fibroblasts, kidney, liver, lung, lymph node, muscle, pituitary, skin, spleen) from individuals whose ages ranged from 0 to 93. The tissue samples came from three sources. Tissue and organ samples from the National NeuroAIDS Tissue Consortium<sup>29</sup>. Blood samples from the Cape Town Adolescent Antiretroviral Cohort study<sup>30</sup>. Skin and other primary cells provided by Kenneth Raj<sup>31</sup>. Ethics approval (IRB#15-001454, IRB#16-000471, IRB#18-000315, IRB#16-002028).

### *DNA methylation data*

We generated DNA methylation data using the custom Illumina chip "HorvathMammalMethylChip40". The mammalian methylation array provides high coverage (over thousand-fold) of highly conserved CpGs in mammals, but focuses only on 36k CpGs that are highly conserved across mammals. Out of 37,492 CpGs on the array, 35,988 probes were chosen to assess cytosine DNA methylation levels in mammalian species<sup>32</sup>. The particular subset of species for each probe is provided in the chip manifest file can be found at Gene Expression Omnibus (GEO) at NCBI as platform GPL28271. The SeSaMe normalization method was used to define beta values for each probe<sup>33</sup>.

### *Penalized Regression models*

Details on the clocks (CpGs, genome coordinates) and R software code are provided in the Supplement. Penalized regression models were created with glmnet<sup>34</sup>. We investigated models produced by both "elastic net" regression ( $\alpha=0.5$ ). The optimal penalty parameters in all cases were determined automatically by using a 10-fold internal cross-validation (cv.glmnet) on the training set. By definition, the alpha value for the elastic net regression was set to 0.5 (midpoint between Ridge and Lasso type regression) and was not optimized for model performance.

We performed a cross-validation scheme for arriving at unbiased (or at least less biased) estimates of the accuracy of the different DNAm based age estimators. One type consisted of leaving out a single sample (LOOCV) from the regression, predicting an age for that sample, and iterating over all samples. A critical step is the transformation of chronological age (the dependent variable). While no transformation was used for the pan tissue clock for deer mice, we did use a log linear transformation for the dual species clock of absolute age.

To introduce biological meaning into age estimates of deer mice and humans that have very different lifespan; as well as to overcome the inevitable skewing due to unequal distribution of data points from deer mice and humans across age range, relative age estimation was made using the formula:  $\text{Relative age} = \text{Age} / \text{maxLifespan}$  where the maximum lifespan for the two species was chosen from the an Age data base<sup>35</sup>. We used the following maximum lifespans Peromyscus

californicus (5.5 years), Peromyscus eremicus (7.4 years), Peromyscus leucopus (7.9 years) Peromyscus maniculatus (8.3 years), Peromyscus polionotus (5.5 years), and humans (122.5 years), respectively.

#### *Epigenome wide association studies of age*

EWAS was performed in each tissue separately using the R function "standardScreeningNumericTrait" from the "WGCNA" R package. Next the results were combined across tissues using Stouffer's meta-analysis method.

#### **ACKNOWLEDGEMENTS**

This work was supported by the Paul G. Allen Frontiers Group (PI Steve Horvath) and by NSF (OIA-1736150) (PI H. Kiaris). The PGSC is supported by a grant from NSF (DBI-1755670) (PI H. Kiaris).

#### **CONFLICT OF INTEREST STATEMENT**

SH is a founder of the non-profit Epigenetic Clock Development Foundation which plans to license several patents from his employer UC Regents. These patents list SH as inventor. The other authors declare no conflicts of interest.

#### **DATA AVAILABILITY**

The data will be made publicly available as part of the data release from the Mammalian Methylation Consortium. Genome annotations of these CpGs can be found on Github <https://github.com/shorvath/MammalianMethylationConsortium>

## REFERENCES

- 1 Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol* **14**, R115, doi:10.1186/gb-2013-14-10-r115 (2013).
- 2 Petkovich, D. A. *et al.* Using DNA Methylation Profiling to Evaluate Biological Age and Longevity Interventions. *Cell Metab* **25**, 954-960 e956, doi:10.1016/j.cmet.2017.03.016 (2017).
- 3 Wang, T. *et al.* Epigenetic aging signatures in mice livers are slowed by dwarfism, calorie restriction and rapamycin treatment. *Genome Biol* **18**, 57, doi:10.1186/s13059-017-1186-2 (2017).
- 4 Stubbs, T. M. *et al.* Multi-tissue DNA methylation age predictor in mouse. *Genome Biol* **18**, 68, doi:10.1186/s13059-017-1203-5 (2017).
- 5 Thompson, M. J. *et al.* A multi-tissue full lifespan epigenetic clock for mice. *Aging (Albany NY)* **10**, 2832-2854, doi:10.18632/aging.101590 (2018).
- 6 Meer, M. V., Podolskiy, D. I., Tyshkovskiy, A. & Gladyshev, V. N. A whole lifespan mouse multi-tissue DNA methylation clock. *eLife* **7**, e40675, doi:10.7554/eLife.40675 (2018).
- 7 Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet*, doi:10.1038/s41576-018-0004-3 (2018).
- 8 Horvath, S. Webpage: <http://labs.genetics.ucla.edu/horvath/dnamage>, 2013).
- 9 Havighorst, A., Crossland, J. & Kiaris, H. Peromyscus as a model of human disease. *Semin Cell Dev Biol* **61**, 150-155, doi:10.1016/j.semcdb.2016.06.020 (2017).
- 10 Steinman, M. Q. & Trainor, B. C. Sex differences in the effects of social defeat on brain and behavior in the California mouse: Insights from a monogamous rodent. *Semin Cell Dev Biol* **61**, 92-98, doi:10.1016/j.semcdb.2016.06.021 (2017).
- 11 Bendesky, A. *et al.* The genetic basis of parental care evolution in monogamous mice. *Nature* **544**, 434-439, doi:10.1038/nature22074 (2017).
- 12 Storz, J. F. *et al.* The molecular basis of high-altitude adaptation in deer mice. *PLoS Genet* **3**, e45, doi:10.1371/journal.pgen.0030045 (2007).
- 13 Bedford, N. L. & Hoekstra, H. E. Peromyscus mice as a model for studying natural variation. *eLife* **4**, doi:10.7554/eLife.06813 (2015).
- 14 Vrana, P. B., Guan, X. J., Ingram, R. S. & Tilghman, S. M. Genomic imprinting is disrupted in interspecific Peromyscus hybrids. *Nat Genet* **20**, 362-365, doi:10.1038/3833 (1998).
- 15 Vrana, P. B. *et al.* Peromyscus (deer mice) as developmental models. *Wiley Interdiscip Rev Dev Biol* **3**, 211-230, doi:10.1002/wdev.132 (2014).
- 16 Soltanmohammadi, E. *et al.* Coordination in the unfolded protein response during aging in outbred deer mice. *Exp Gerontol* **144**, 111191, doi:10.1016/j.exger.2020.111191 (2021).
- 17 Abraham, M. A. & Lam, T. K. T. Glucagon action in the brain. *Diabetologia* **59**, 1367-1371, doi:10.1007/s00125-016-3950-3 (2016).
- 18 Kellar, D. & Craft, S. Brain insulin resistance in Alzheimer's disease and related disorders: mechanisms and therapeutic approaches. *Lancet Neurol* **19**, 758-766, doi:10.1016/S1474-4422(20)30231-3 (2020).
- 19 Wo, J. *et al.* The Role of Gamma-Delta T Cells in Diseases of the Central Nervous System. *Front Immunol* **11**, 580304, doi:10.3389/fimmu.2020.580304 (2020).
- 20 Klug, H. Why Monogamy? A Review of Potential Ultimate Drivers. *Frontiers in Ecology and Evolution* **6**, doi:10.3389/fevo.2018.00030 (2018).
- 21 Kleiman, D. G. Monogamy in mammals. *Q Rev Biol* **52**, 39-69, doi:10.1086/409721 (1977).
- 22 Scribner, J. L. *et al.* A neuronal signature for monogamous reunion. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 11076-11084, doi:10.1073/pnas.1917287117 (2020).
- 23 Jašarević, E. *et al.* Evolution of monogamy, paternal investment, and female life history in Peromyscus. *J Comp Psychol* **127**, 91-102, doi:10.1037/a0027936 (2013).
- 24 Curtis, J. T., Liu, Y., Aragona, B. J. & Wang, Z. Dopamine and monogamy. *Brain Res* **1126**, 76-90, doi:10.1016/j.brainres.2006.07.126 (2006).

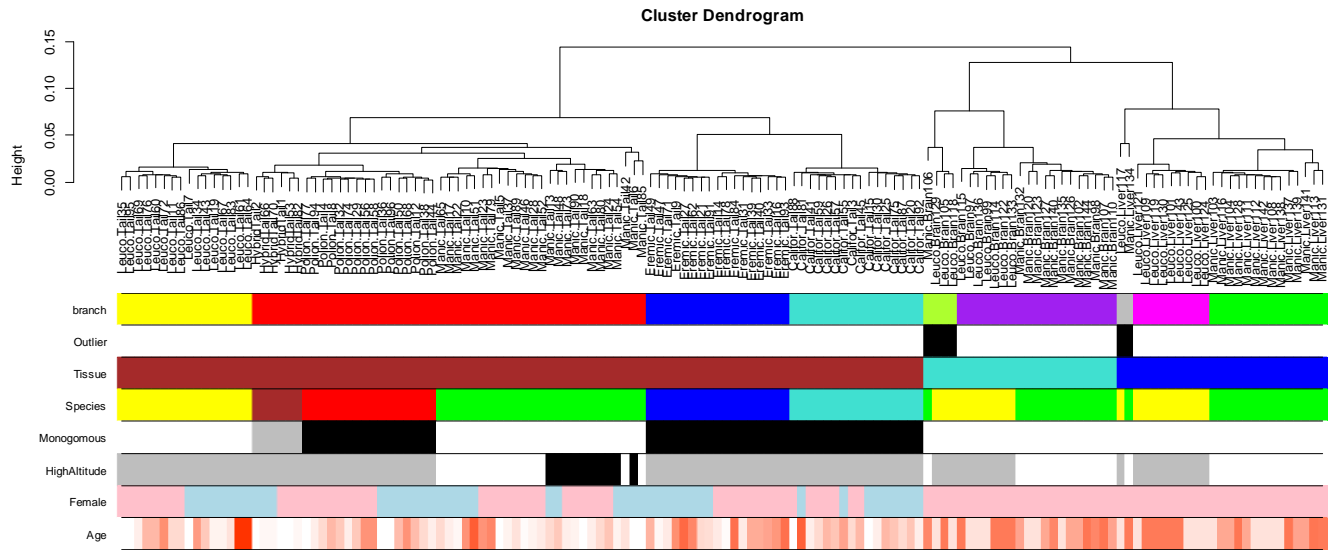
- 25 Hegarty, S. V. *et al.* Zeb2 is a negative regulator of midbrain dopaminergic axon growth and target innervation. *Scientific reports* **7**, 8568, doi:10.1038/s41598-017-08900-3 (2017).
- 26 Mazzeo, R. S. Altitude, exercise and immune function. *Exerc Immunol Rev* **11**, 6-16 (2005).
- 27 Oliver, S. J. *et al.* High Altitude Impairs In Vivo Immunity in Humans. *High Altitude Medicine & Biology* **14**, 144-149, doi:10.1089/ham.2012.1070 (2013).
- 28 Meehan, R. T. Immune suppression at high altitude. *Annals of Emergency Medicine* **16**, 974-979, doi:10.1016/S0196-0644(87)80743-6 (1987).
- 29 Morgello, S. *et al.* The National NeuroAIDS Tissue Consortium: a new paradigm in brain banking with an emphasis on infectious disease. *Neuropathol Appl Neurobiol* **27**, 326-335. (2001).
- 30 Horvath, S. *et al.* Perinatally acquired HIV infection accelerates epigenetic aging in South African adolescents. *AIDS* **32**, 1465-1474, doi:10.1097/QAD.0000000000001854 (2018).
- 31 Kabacik, S., Horvath, S., Cohen, H. & Raj, K. Epigenetic ageing is distinct from senescence-mediated ageing and is not prevented by telomerase expression. *Aging (Albany NY)* **10**, 2800-2815, doi:10.18632/aging.101588 (2018).
- 32 Arneson, A. *et al.* A mammalian methylation array for profiling methylation levels at conserved sequences. *bioRxiv*, 2021.2001.2007.425637, doi:10.1101/2021.01.07.425637 (2021).
- 33 Zhou, W., Triche, T. J., Jr, Laird, P. W. & Shen, H. SeSAmE: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Research* **46**, e123-e123, doi:10.1093/nar/gky691 (2018).
- 34 Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software* **33**, 1-22 (2010).
- 35 de Magalhaes, J. P., Costa, J. & Church, G. M. An analysis of the relationship between metabolism, developmental schedules, and longevity using phylogenetic independent contrasts. *J Gerontol A Biol Sci Med Sci* **62**, 149-160 (2007).

## TABLES

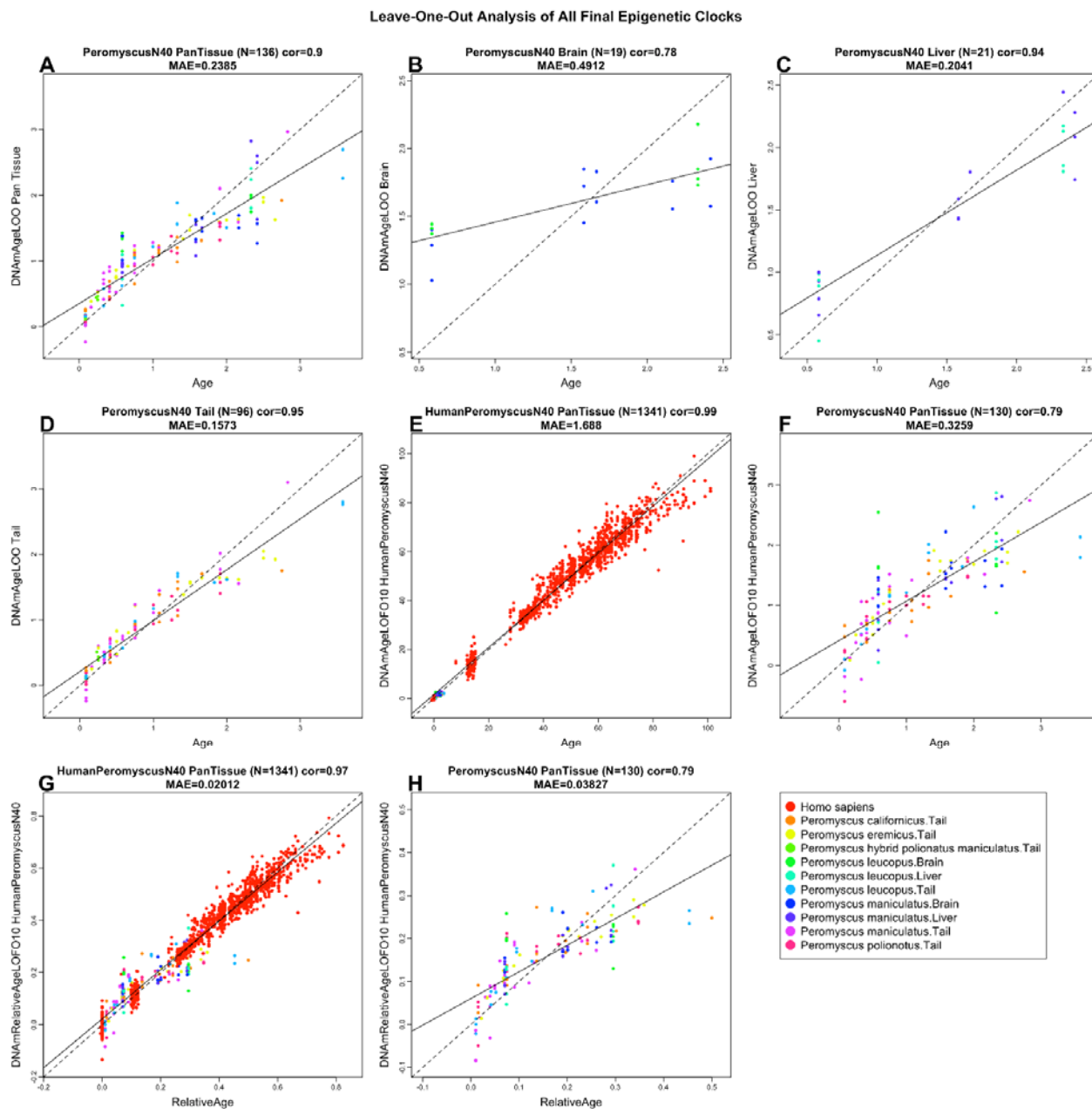
**Table 1. Description of the data.** N=Total number of tissues. Number of females. Age (in units of years): mean, minimum and maximum age.

<b>Species/Tissue</b>	<b>N</b>	<b>No. Female</b>	<b>Mean Age</b>	<b>Min. Age</b>	<b>Max. Age</b>
Peromyscus californicus/Tail	16	7	0.955	0.083	2.75
Peromyscus eremicus/Tail	17	9	1.36	0.16	2.66
Hybrid polionatus+maniculatus/Tail	6	3	0.163	0.083	0.25
Peromyscus leucopus/Brain	7	7	1.58	0.583	2.33
Peromyscus leucopus/Liver	8	8	1.46	0.583	2.33
Peromyscus leucopus/Tail	16	8	1.13	0.083	3.58
Peromyscus maniculatus/Brain	12	12	1.58	0.583	2.42
Peromyscus maniculatus/Liver	13	13	1.46	0.583	2.42
Peromyscus maniculatus/Tail	25	14	0.724	0.083	2.83
Peromyscus polionotus/Tail	16	9	0.909	0.083	1.91

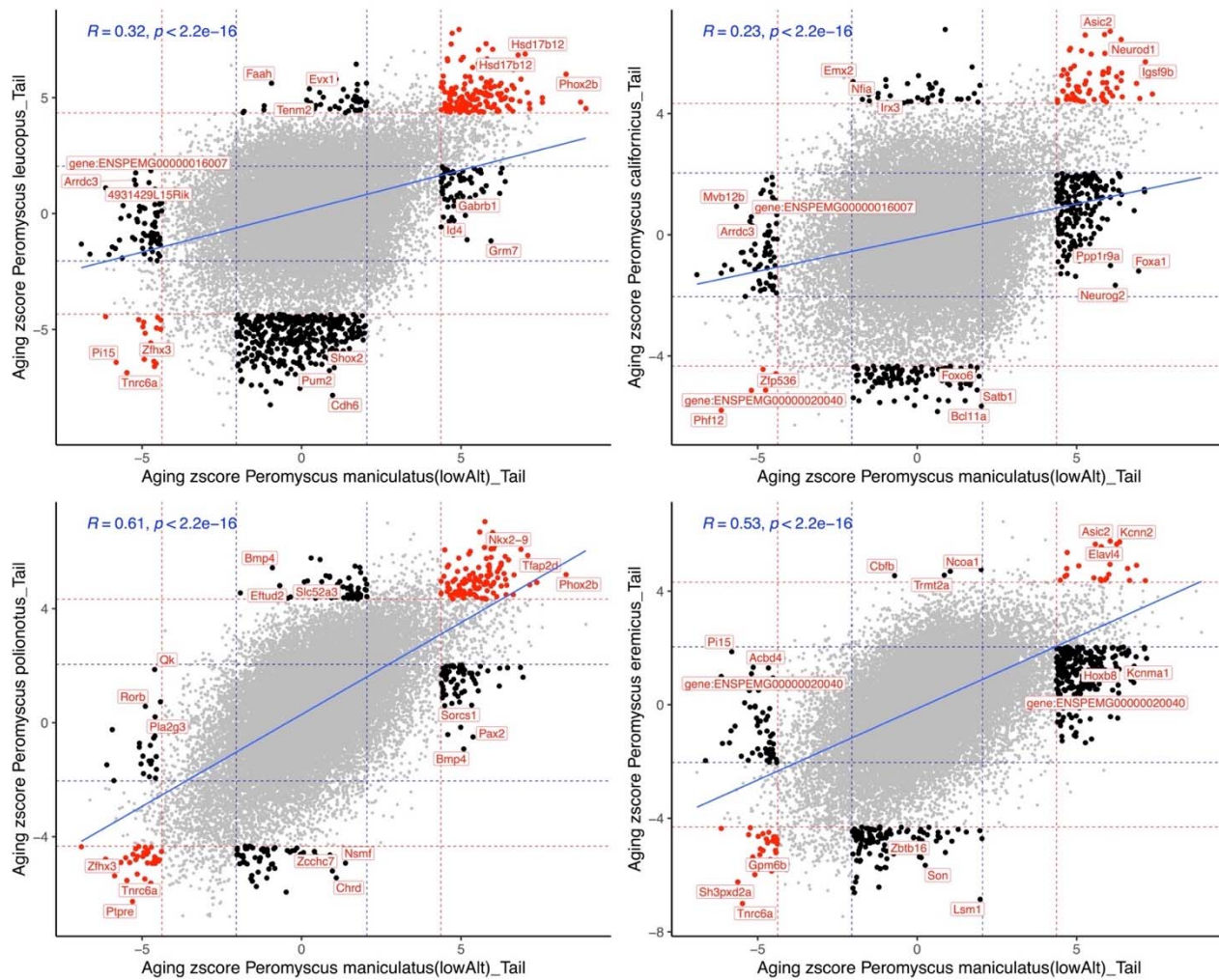
## FIGURE LEGENDS



**Figure 1. Unsupervised hierarchical clustering of tissue samples from deer mice.** Average linkage hierarchical clustering based on the interarray correlation coefficient (Pearson correlation). The first color band is based on cutting the branches. Note that the branch colors correspond to tissue and species.

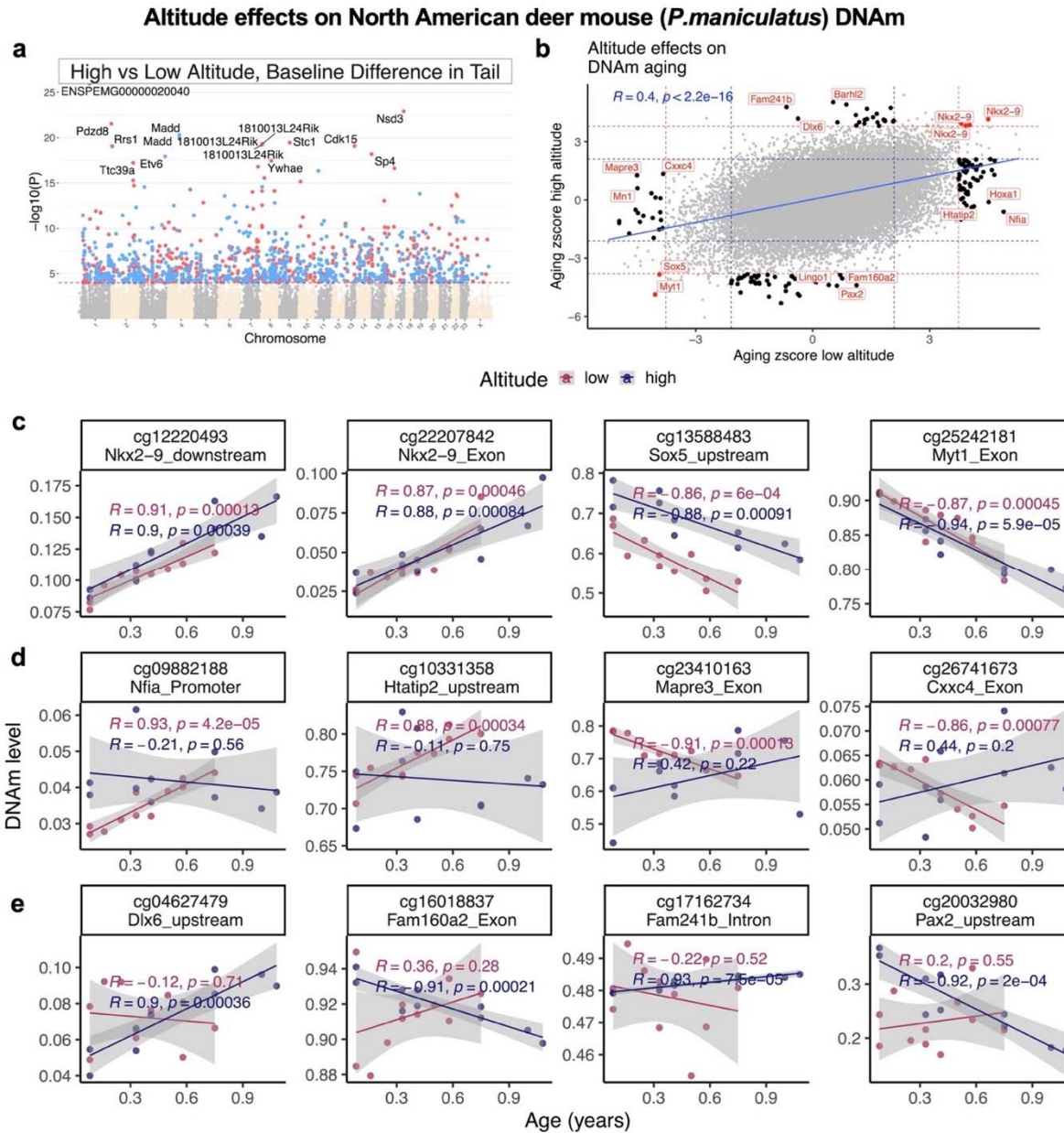


**Figure 2. Cross-validation study of epigenetic clocks for Peromyscus.** A) Epigenetic clock for multiple peromyscus species. Leave-one-sample-out (LOO) estimate of DNA methylation age (y-axis, in units of years) versus chronological age. A) Pan tissue deer mouse clock, B) Brain clock, C) Liver clock, D) Tail clock. E) Human-peromyscus clock for chronological age, F) excerpt of panel (E) restricted to peromyscus samples. G) Human-peromyscus clock for relative age. H) excerpt of panel (G) restricted to peromyscus samples. Dots (DNA samples) are colored by species or tissue sample as outlined in the legend. Each panel reports the sample size, correlation coefficient, median absolute error (MAE).



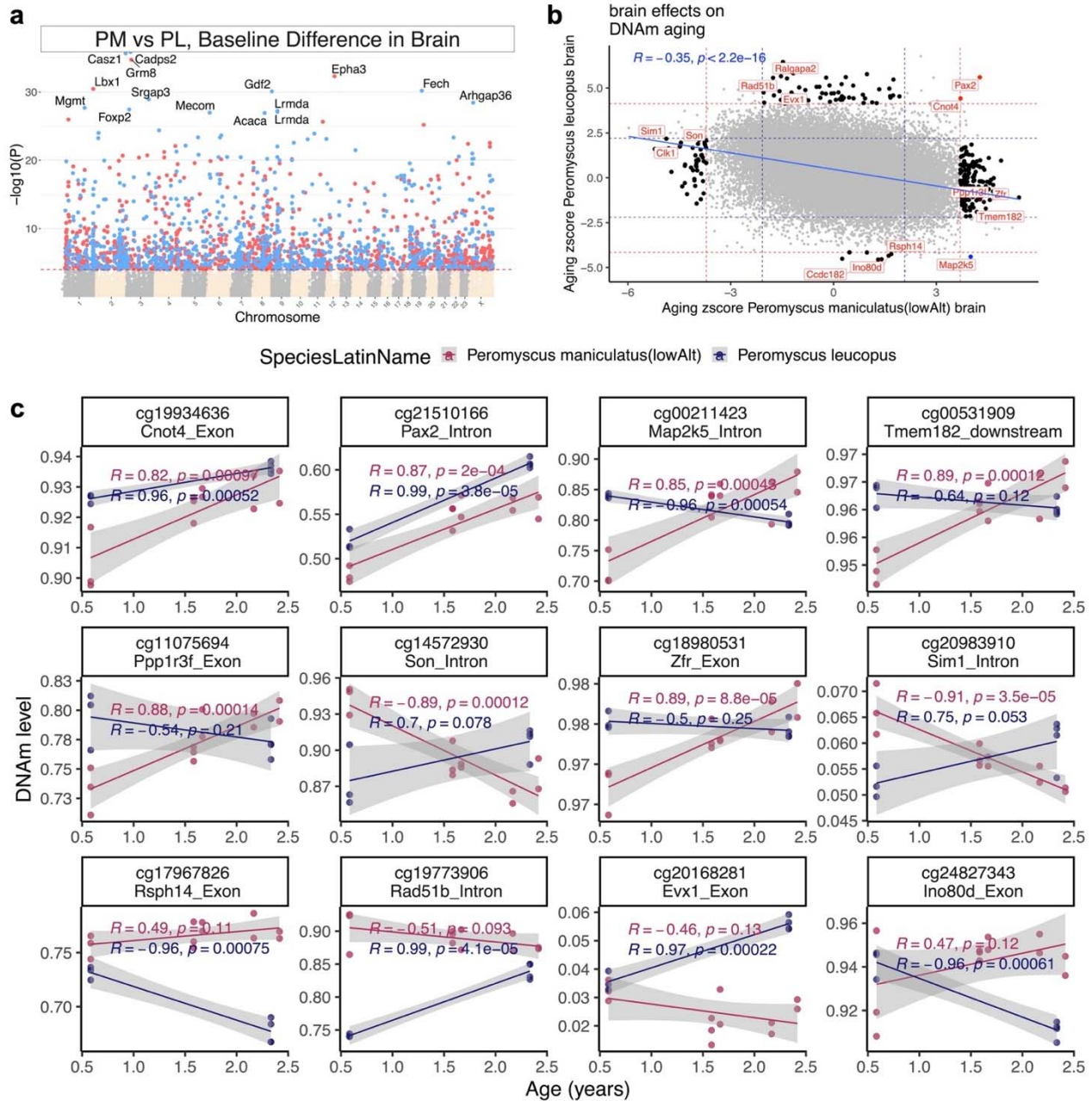
**Figure 3. Comparative DNAm aging between *P. maniculatus* and four other *Peromyscus* species.** Scatter plots represent the EWAS of age in *P. maniculatus* (x-axis) vs other four species. Red dotted line:  $p < 10^{-4}$ ; blue dotted line:  $p > 0.05$ ; Red dots: shared CpGs; black dots: species specific changes. The top CpGs in each sector is labeled by the adjacent gene based on *P. maniculatus* genome.



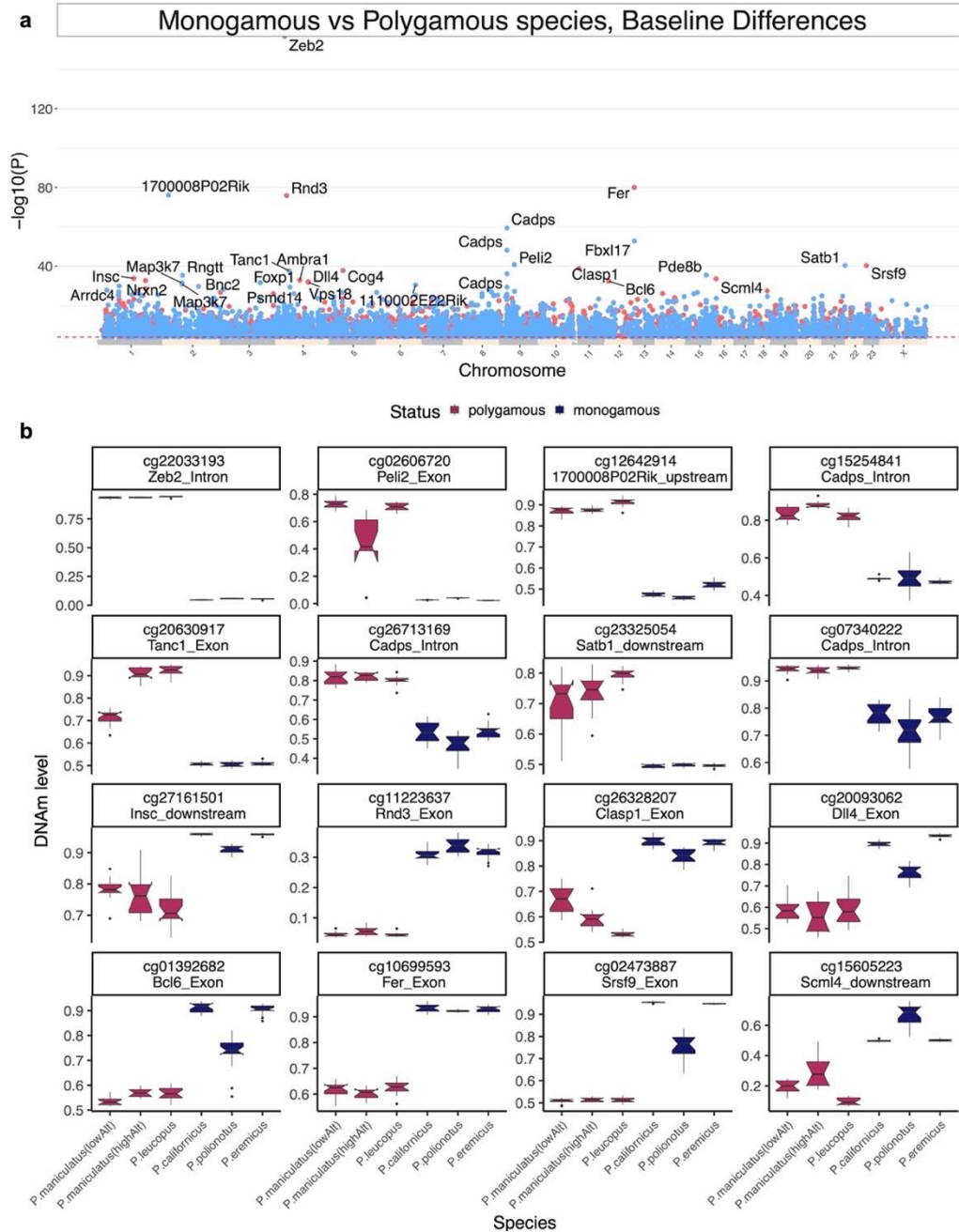


**Figure 4. Altitude difference altered epigenetic profiles of *P. maniculatus*.** This study compares two colonies of *P. maniculatus*, which originated from the founders with different habitats in different altitudes. a) The Manhattan plot of EWAS of altitude in tail of *P. maniculatus* species after adjusting for chronological age. The genome coordinates are estimated based on the alignment of Mammalian array probes to the *Peromyscus maniculatus bairdii*.HU\_Pman\_2.1.100 genome assembly. The direction of associations with  $p < 10^{-4}$  (red dotted line) is colored in red (hypermethylated) and blue (hypomethylated). The 15 most significant 15 CpGs are labeled by adjacent genes. b) Sector plot of DNA methylation aging effects in *P. maniculatus* species from different altitudes. Red dotted line:  $p < 10^{-3}$ ; blue dotted line:  $p > 0.05$ ; Red dots: shared CpGs; black dots: EWAS specific changes. The scatter plots represent the CpGs that change with age in both (c), or uniquely in low (d) or high (e) altitude colonies.

**Brain differences between North American deer mouse (*P. maniculatus*) and white-footed mouse (*P. leucopus*)**

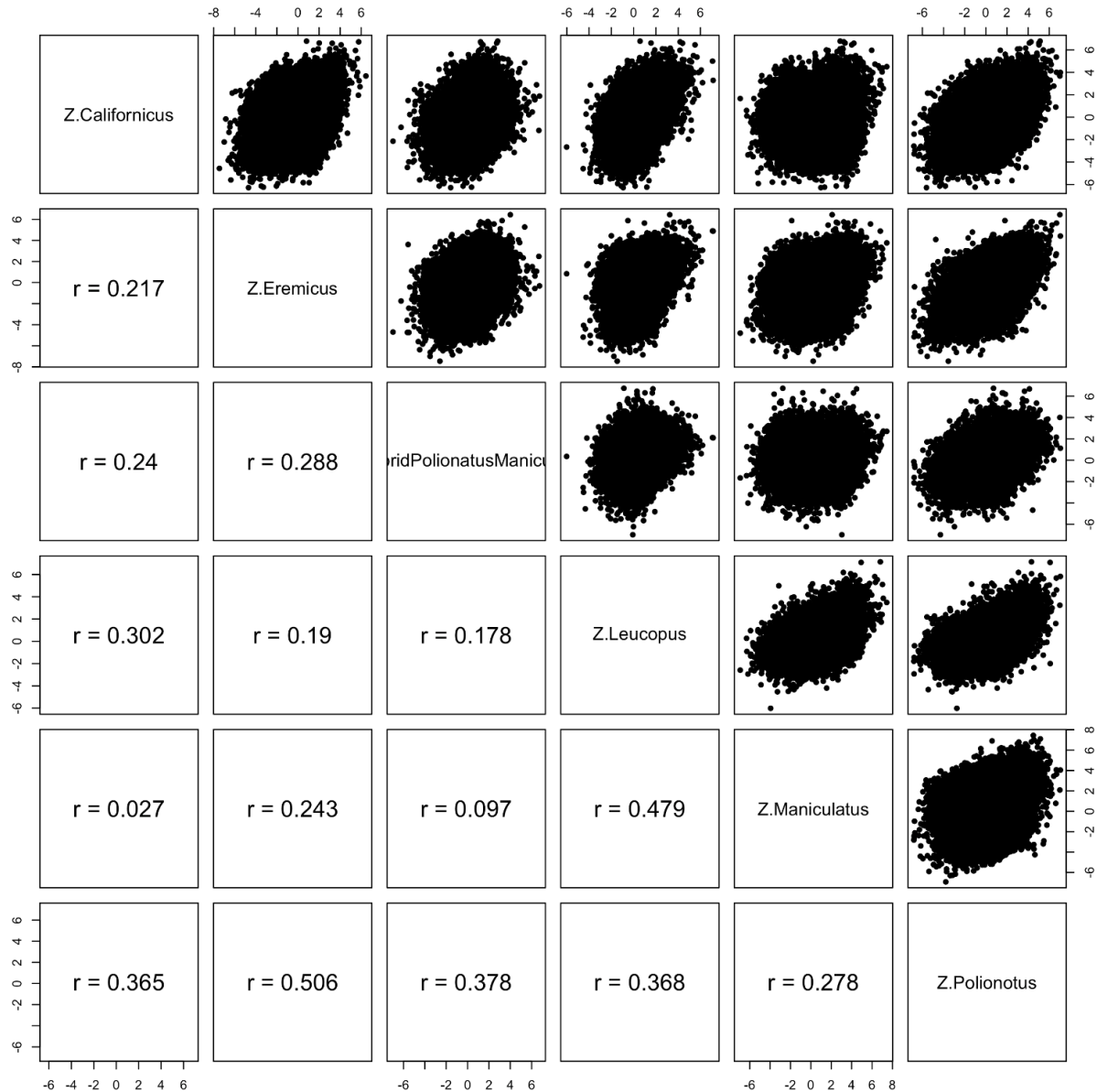


**Figure 5. Brain methylation differences between *P. maniculatus* and *P. leucopus*.** a) The Manhattan plot of mean methylation differences between *P. maniculatus* and *P. leucopus* after adjusting for chronological age. Genome coordinates for the *Peromyscus maniculatus bairdii*.HU.Pman\_2.1.100 genome assembly. The direction of associations with  $p < 10^{-4}$  (red dotted line) is colored in red (hypermethylated) or blue (hypomethylated). Top 15 CpGs are labeled by adjacent genes. b) Sector plot of DNAm aging effects in *P. maniculatus* and *P. leucopus*. Red dotted line:  $p < 10^{-3}$ ; blue dotted line:  $p > 0.05$ ; Red dots: shared CpGs; black dots: species specific changes. c) The scatter plots represent the CpGs that change with age in both, or uniquely in one of these species.

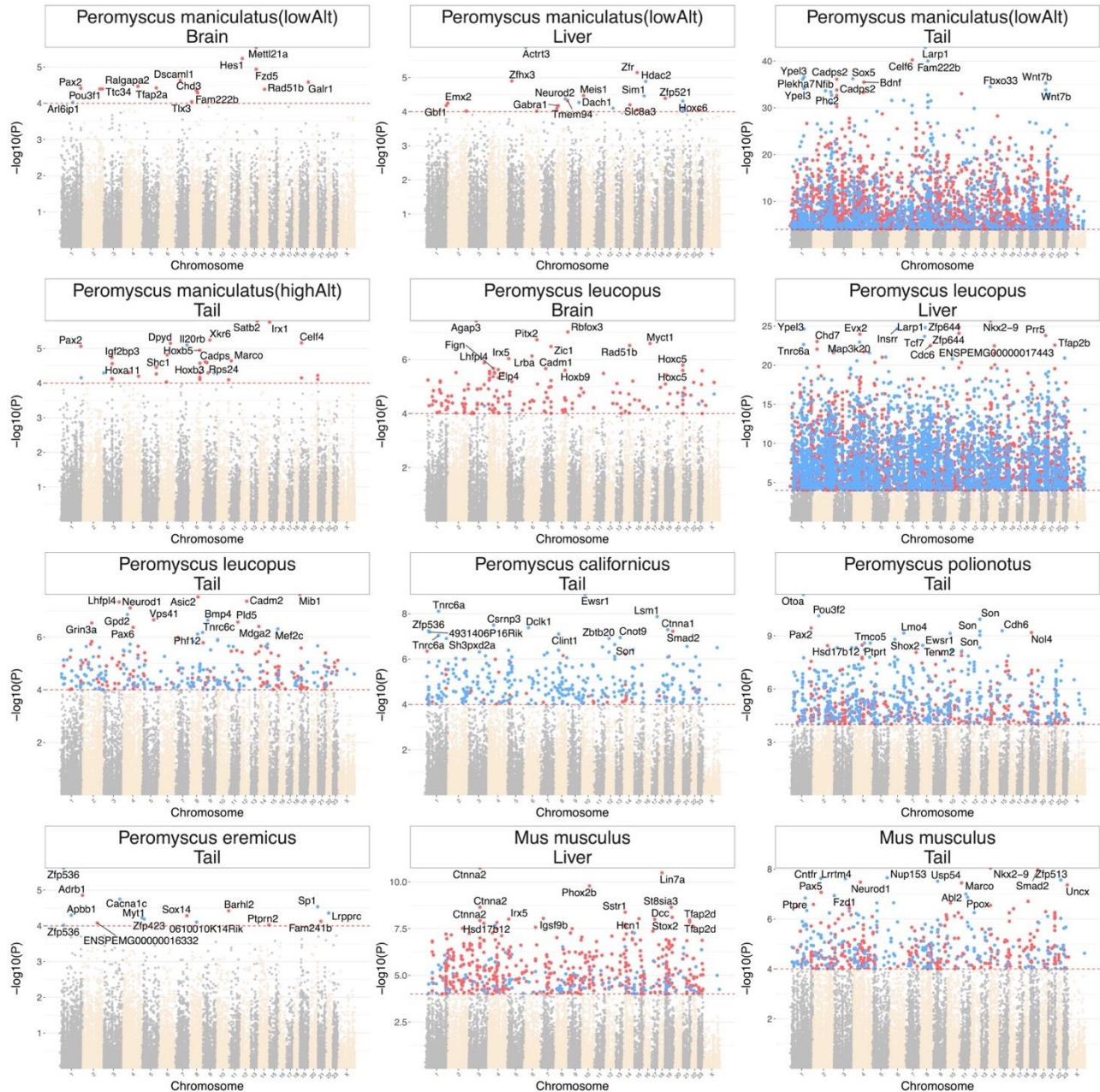


**Figure 6. CpGs that differ between monogamous and polygamous *Peromyscus* species.** a) The Manhattan plot of the mean DNAm difference between monogamous and polygamous *Peromyscus* species after adjusting the analysis for chronological age. Genome coordinates from *Peromyscus maniculatus bairdii*.HU\_Pman\_2.1.100 genome assembly. The direction of associations with  $p < 10^{-3}$  (red dotted line) is colored in red (hypermethylated) and blue (hypomethylated). Top 15 CpGs labeled by their neighboring genes. b) The box plot of the top CpGs that differed by dyadic relationship in *Peromyscus* species.

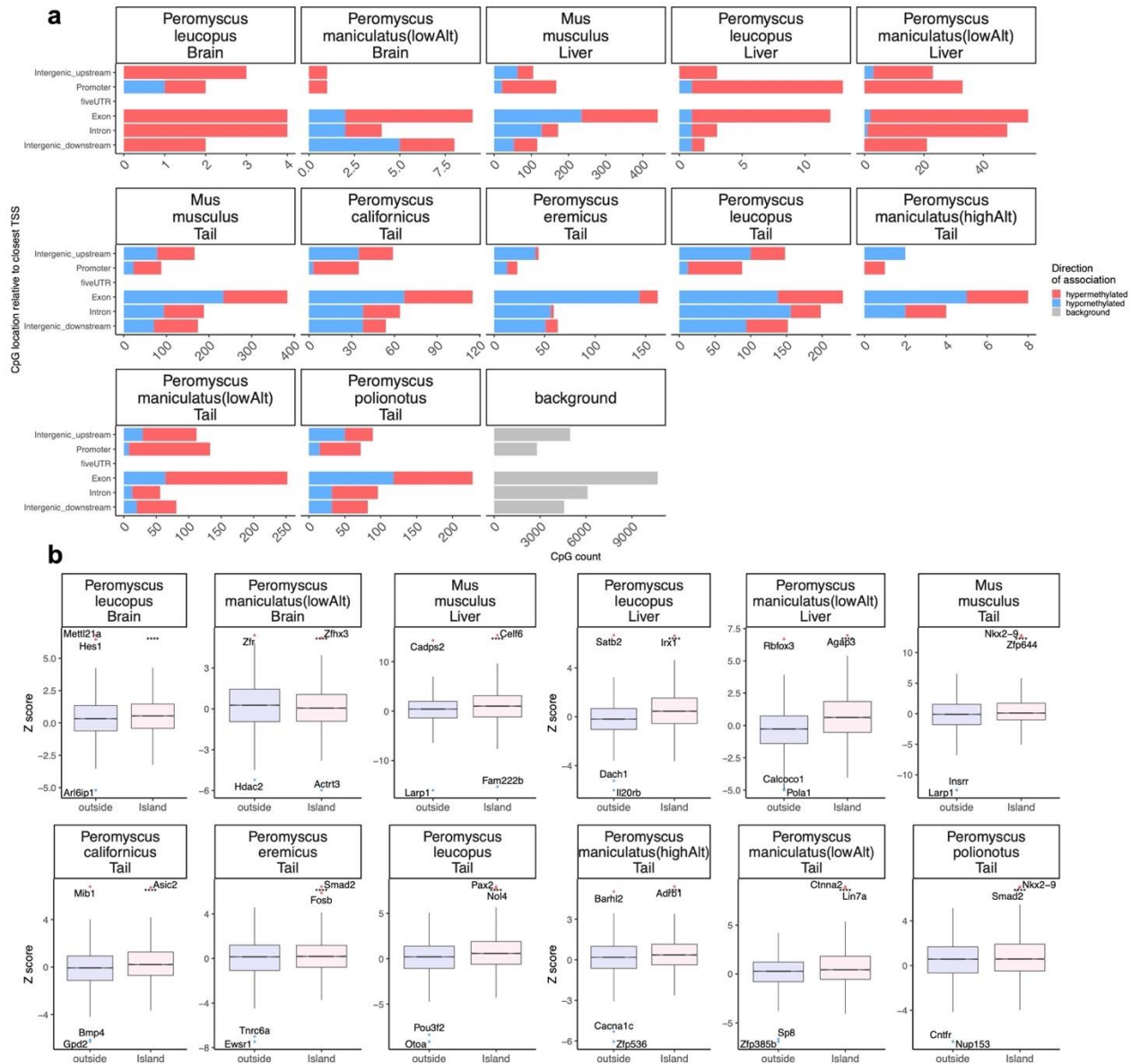
## SUPPLEMENTARY MATERIAL

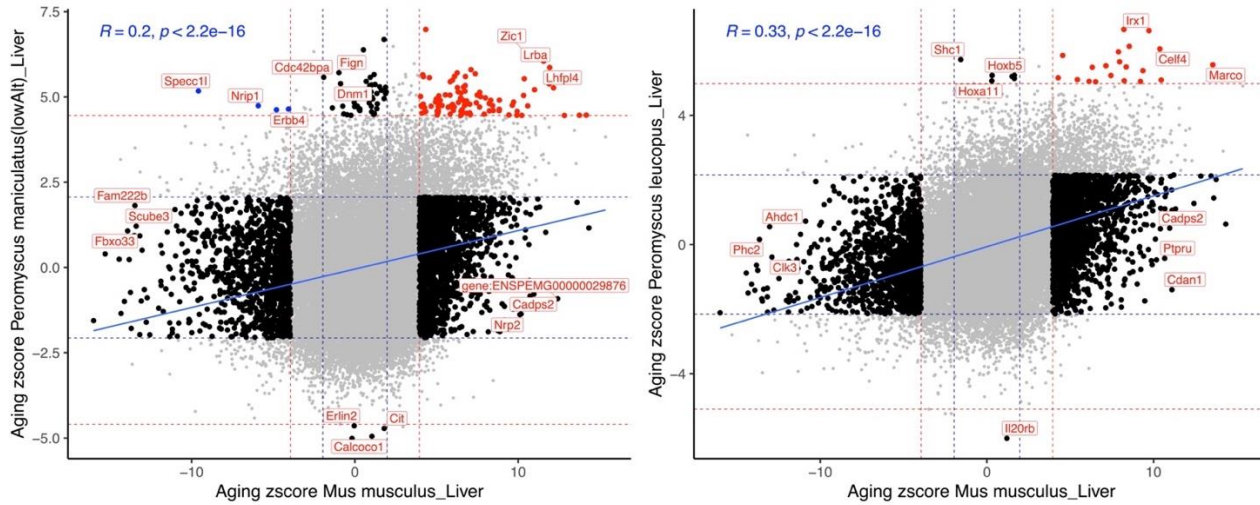


**Supplementary Figure 1. Epigenome wide association study of age correlation in different deer mouse species.** Each dot is a CpG. The Z statistics (which have a standard normal distribution under the null hypothesis of no age effects) results from a correlation test relating CpGs to chronological age in different species.

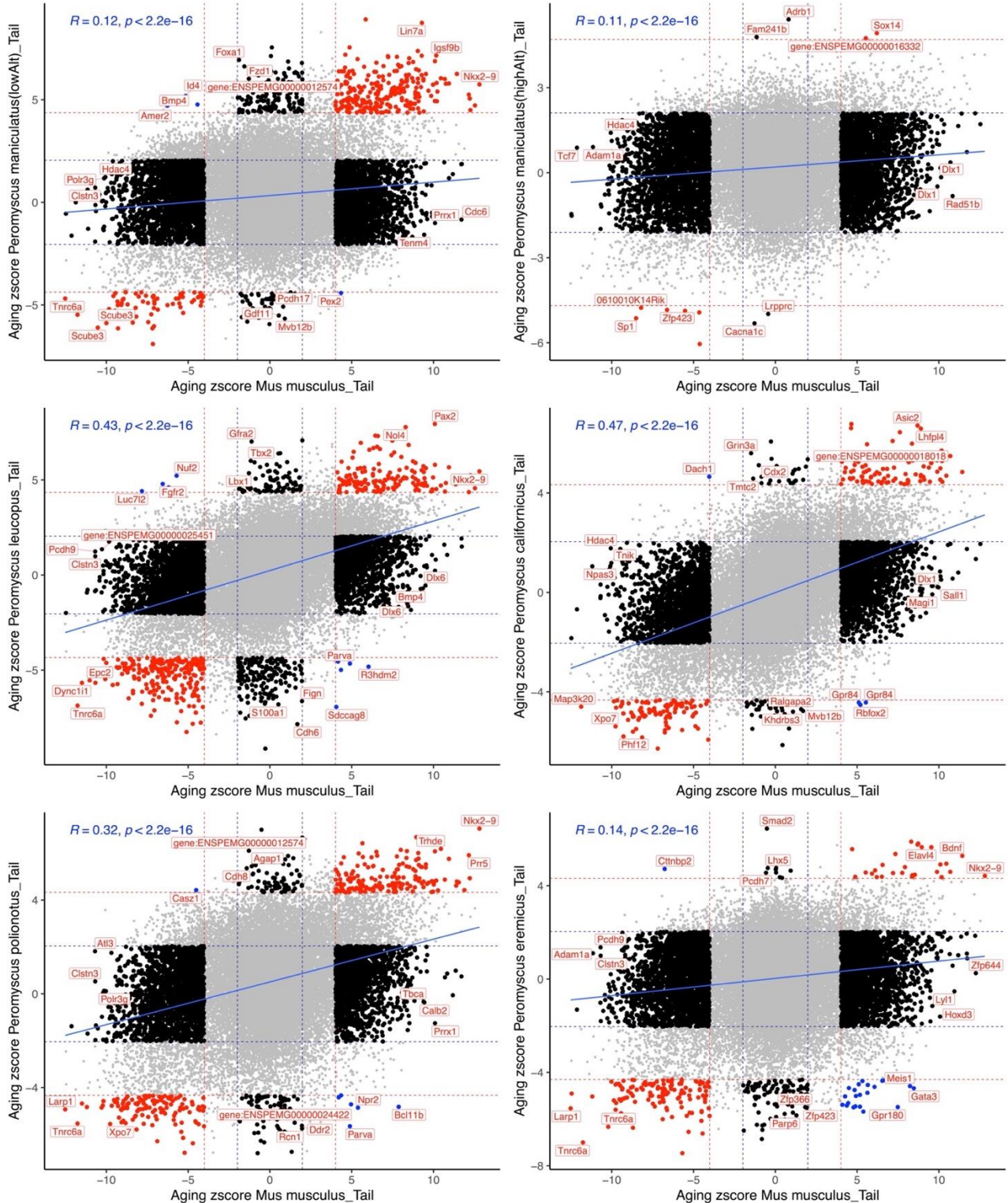


**Supplementary Figure 2. EWAS of age in different tissues of *Peromyscus* genus and B6 mouse.** The coordinates are estimated based on the alignment of Mammalian array probes to *Peromyscus maniculatus bairdii*.HU\_Pman\_2.1.100 genome assembly. The direction of associations with  $p < 10^{-4}$  (red dotted line) is highlighted by red (hypermethylated) and blue (hypomethylated) colors. Top 15 CpGs was labeled by the neighboring genes.



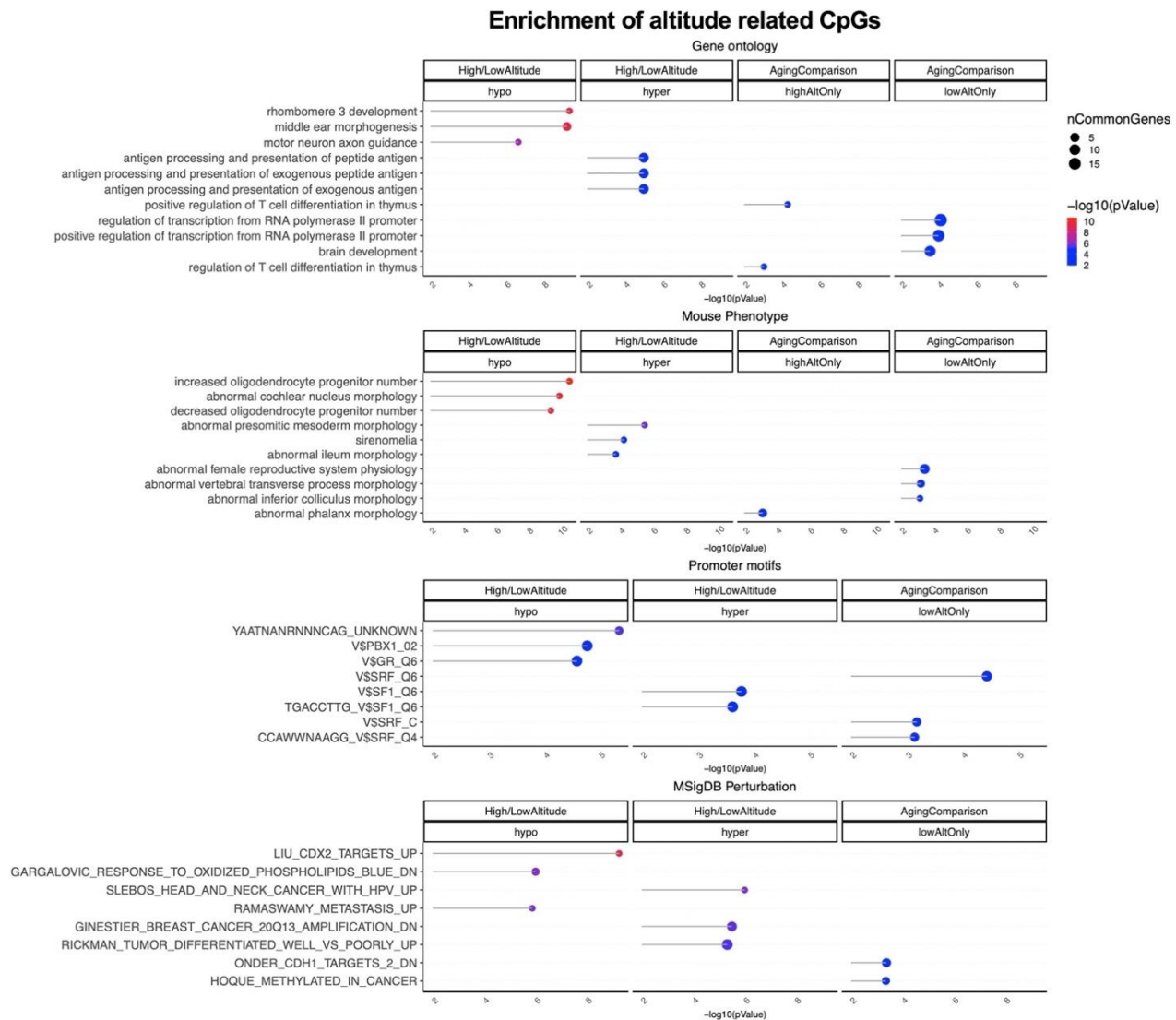


**Supplementary Figure 4. Liver DNAm aging in *Peromyscus* genus and B6 mouse moderately correlate.** Sector plot of DNAm aging in *Peromyscus* species and B6 mouse liver. Red dotted line:  $p < 10^{-4}$ ; blue dotted line:  $p > 0.05$ ; Red dots: shared CpGs; black dots: species specific changes.



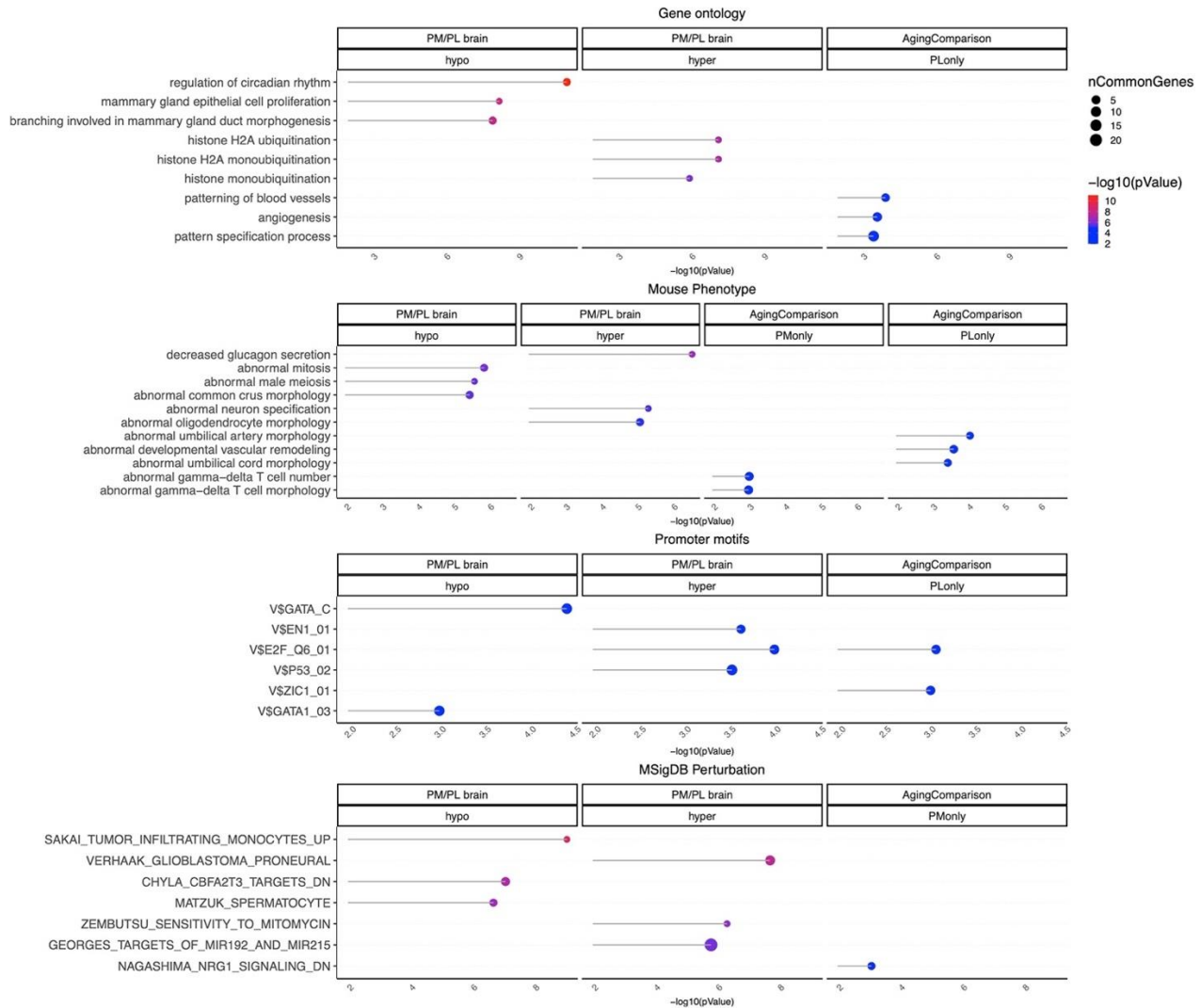
**Supplementary Figure 5. Comparison of DNAm aging in *Peromyscus* genus and B6 mouse tail samples.** Sector plot of DNAm aging in *Peromyscus* species and B6 mouse tails. Red dotted line:  $p < 10^{-4}$ ; blue dotted line:  $p > 0.05$ ; Red dots: shared CpGs; black dots: species specific changes; blue dots: CpGs with divergent aging pattern between species.





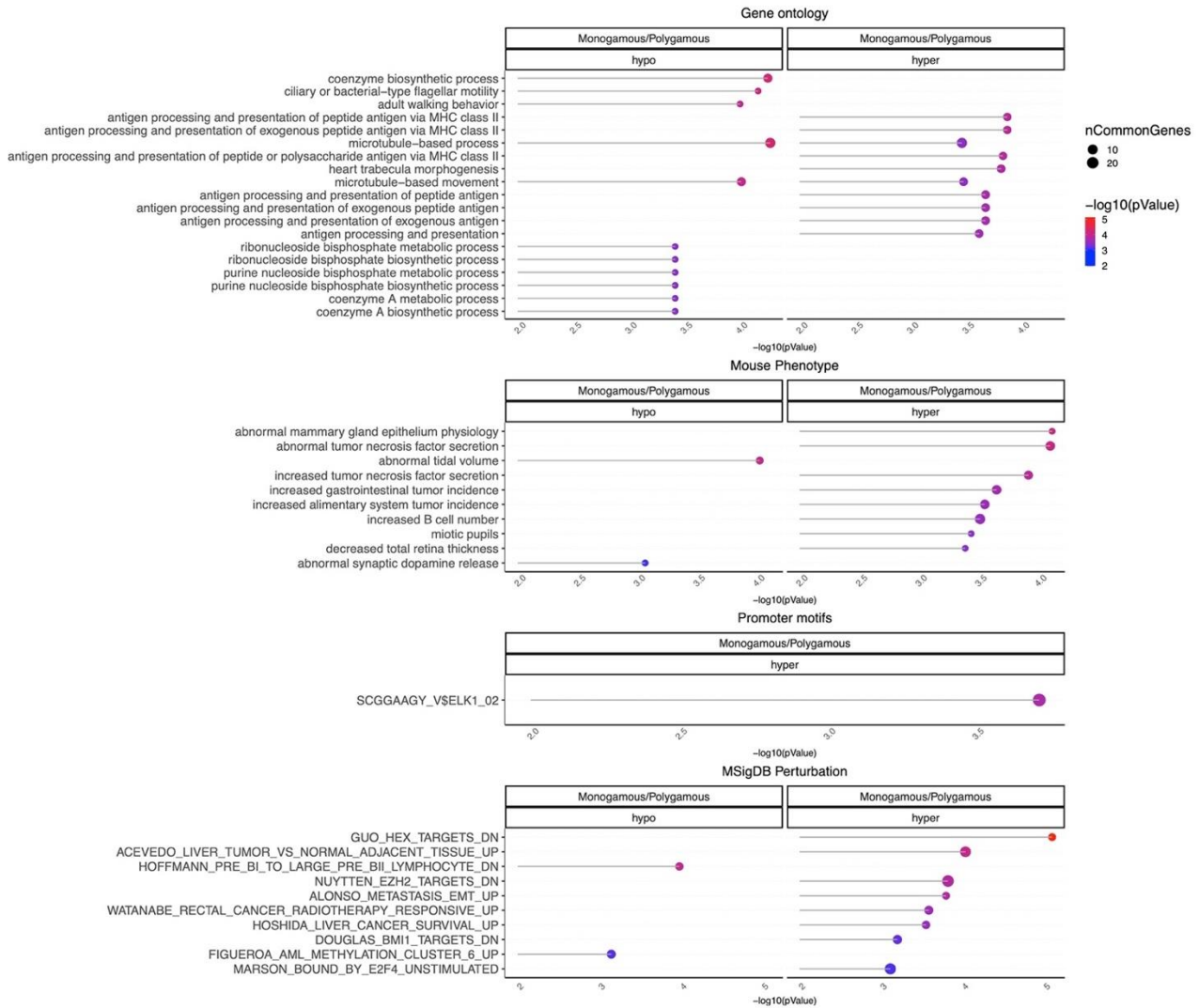
**Supplementary Figure 6. Gene set enrichment analysis of DNA methylation differences by altitude.** The gene level enrichment was done using GREAT analysis and human Hg19 background. Datasets: gene ontology, mouse phenotypes, promoter motifs, and MSigDB Perturbation. The results were filtered for significance at  $p < 10^{-3}$ .

### Enrichment of brain differences between North American deer mouse (*P.maniculatus*) and white-footed mouse (*P.leucopus*)



**Supplementary Figure 7. Gene set enrichment analysis of DNA methylation differences between *P. maniculatus* and *P. leucopus* brains.** The gene level enrichment was done using GREAT analysis and human Hg19 background. Datasets: gene ontology, mouse phenotypes, promoter motifs, and MSigDB Perturbation. The results were filtered for significance at  $p < 10^{-3}$ .

## Enrichment of monogamous related CpGs



**Supplementary Figure 8. Gene set enrichment analysis of DNA methylation differences between monogamous and polygamous *Peromyscus* species.** The gene level enrichment was done using GREAT analysis and human Hg19 background. Datasets: gene ontology, mouse phenotypes, promoter motifs, and MSigDB Perturbation. The results were filtered for significance at  $p < 10^{-3}$ .