Received Date : 27-Jul-2015
Revised Date : 19-Oct-2015
Accepted Date : 22-Oct-2015
Article type : Special Issue

Resource base influences genome-wide DNA methylation levels in wild baboons (*Papio cynocephalus*)

Amanda J. Lea[1], Jeanne Altmann[2,3], Susan C. Alberts[1,3], and Jenny Tung[1,3,4,5*]

1. Department of Biology, Duke University, Box 90338, Durham, NC 27708, USA

2. Department of Ecology and Evolution, Princeton University, 106A Guyot Hall, Princeton, NJ 08544, USA

3. Institute of Primate Research, National Museums of Kenya, P. O. Box 24481, Karen 00502, Nairobi, Kenya

4. Department of Evolutionary Anthropology, Box 90383, Durham, NC 27708, USA

5. Duke University Population Research Institute, Box 90420, Durham, NC 27708, USA

*Corresponding author email: jt5@duke.edu

Other author emails: amanda.lea@duke.edu, alberts@duke.edu, altj@princeton.edu

Abstract: Variation in resource availability commonly exerts strong effects on fitness-related traits in wild animals. However, we know little about the molecular mechanisms that mediate these effects, or about their persistence over time. To address these questions, we profiled genome-wide whole blood DNA methylation levels in two sets of

wild baboons: (i) 'wild-feeding' baboons that foraged naturally in a savanna environment and (ii) 'Lodge' baboons that had ready access to spatially concentrated human food scraps, resulting in high feeding efficiency and low daily travel distances. We identified 1,014 sites (0.20% of sites tested) that were differentially methylated between wild-feeding and Lodge baboons, providing the first evidence that resource availability shapes the epigenome in a wild mammal. Differentially methylated sites tended to occur in contiguous stretches (i.e., in differentially methylated regions or DMRs), in promoters and enhancers, and near metabolism-related genes, supporting their functional importance in gene regulation. In agreement, reporter assay experiments confirmed that methylation at the largest identified DMR, located in the promoter of a key glycolysis-related gene, was sufficient to causally drive changes in gene expression. Intriguingly, all dispersing males carried a consistent epigenetic signature of their membership in a wild-feeding group, regardless of whether males dispersed into or out of this group as adults. Together, our findings support a role for DNA methylation in mediating ecological effects on phenotypic traits in the wild, and emphasize the dynamic environmental sensitivity of DNA methylation levels across the life course.

Introduction

Despite a rich history of studies documenting the relationship between ecological and phenotypic variation in natural populations, we know surprisingly little about the molecular mechanisms that mediate these effects. Insight into these mechanisms is important for understanding how natural phenotypic variation emerges and how organisms cope with environmental change. Genomic approaches can contribute to

these questions by identifying the genes and pathways involved in sensing and responding to selectively relevant ecological variation. For example, studies in eusocial insects have highlighted the impact of nutrient exposure on genome-wide gene regulation, as well as its contribution to the emergence of distinct castes within a hive or colony (Kucharski *et al.* 2008; Foret *et al.* 2012). At the same time, other studies have identified rapid gene expression responses to song in zebra finch (Drnevich *et al.* 2012; Whitney *et al.* 2014); a strong genome-wide signature of social status in hierarchical primates (Tung *et al.* 2012); and widespread transcriptional changes associated with mate choice in fish (Cummings *et al.* 2008). Together, such work points to a fundamental role for gene regulation in mediating physiological responses to environmental inputs. By altering the expression of genes in the genome, gene regulatory mechanisms permit a range of phenotypic values to arise from an otherwise static genome sequence.

Mounting evidence suggests that epigenetic marks, particularly DNA methylation (the best studied to date), play a significant role in mediating the gene regulatory response to environmental conditions. DNA methylation refers to the covalent addition of a methyl group to a cytosine base, and, in mammals, occurs most often at CG dinucleotides (known as 'CpG sites'). CpG sites are strongly enriched in regulatory sequences (e.g., gene promoters, gene bodies, and CpG-dense regions known as 'CpG islands') where changes in methylation can impact the expression of nearby genes. For example, DNA methylation in promoter or enhancer regions can repress gene expression by interfering with transcription factor binding, or by recruiting proteins that induce changes in chromatin accessibility (Klose & Bird 2006; Weber *et al.* 2007).

Meanwhile, gene body methylation is often associated with increased gene expression, and is thought to aid in transcriptional elongation (Jones 2012).

Genome-wide patterns of DNA methylation are first established during development, and once established, are faithfully transmitted across cell divisions throughout the life of the organism. However, environmental conditions can affect this process, either during development itself (when epigenetic patterns are known to be particularly sensitive: Tobi *et al.* 2009; Feil & Fraga 2011; Faulk & Dolinoy 2011) or later in life, when changes in DNA methylation help coordinate the cellular response to new environmental stimuli (Guo *et al.* 2011; Barrès *et al.* 2012; Pacis *et al.* 2015). Thus, changes in DNA methylation are thought to provide an avenue through which environmental inputs can stably alter gene expression levels and, as a consequence, mediate environmental effects on organism-level traits (Meaney & Szyf 2005; Jirtle & Skinner 2007; Feil & Fraga 2011).

The relationship between DNA methylation and diet (both caloric intake and dietary content) is particularly well-studied in this regard (Van den Veyver 2002; Heijmans *et al.* 2008; Carone *et al.* 2010; Moleres *et al.* 2013; Tobi *et al.* 2014). For example, in laboratory mice, maternal diet during pregnancy predicts offspring fur color and susceptibility to diabetes—a relationship mediated by its stable effects on offspring methylation near the *agouti* gene, which in turn changes *agouti* gene expression (Klebig *et al.* 1995; Wolff *et al.* 1998). Long-lasting effects of diet and resource availability (which we refer to in combination as 'resource base') also affect DNA methylation patterns in humans. In Gambian populations that experience dramatic seasonal fluctuations in food availability, season of conception predicts offspring DNA methylation

levels at several metastable epialleles (loci that show consistent, stable epigenetic patterns across tissues: Waterland *et al.* 2010; Dominguez-Salas *et al.* 2014). Similarly, individuals conceived during the Dutch Hunger Winter, a severe war-time famine in the Netherlands, have been shown to exhibit stable differences in DNA methylation levels at both individual growth-related genes (e.g., *IGF2, INSIGF,* and *IL10*) and on a genome-wide scale (Heijmans *et al.* 2008; Tobi *et al.* 2009, 2014). These studies also suggest that diet effects on DNA methylation can be acutely sensitive to timing: long-term changes in DNA methylation levels were only detectable in individuals exposed to the Dutch Hunger winter during the periconception period, but not later in pregnancy (Heijmans *et al.* 2008; Tobi *et al.* 2014, but see Tobi *et al.* 2009).

Variation in resource base, both during development and later in life, is also important in wild mammal populations, where it exerts potent effects on both fertility and mortality components of fitness (Altmann 1991; Gaillard *et al.* 2000; Beehner *et al.* 2006; Nussey *et al.* 2007; Revitali *et al.* 2009; Hamel *et al.* 2009). However, in contrast to human populations or lab model organisms, the role of DNA methylation in mediating these effects has not been investigated, leaving questions about the scope and timing of epigenetic sensitivity to the environment unanswered. To address this gap, we profiled genome-wide DNA methylation levels in a long-term study population of wild baboons in the Amboseli region of Kenya (Alberts & Altmann 2012). Specifically, we compared DNA methylation patterns in 'wild-feeding baboons' to those in 'Lodge group baboons'. Wild-feeding baboons walked 4-6 km per day, foraging in a dry savanna environment on widely distributed foods. In contrast, while Lodge group baboons resided in the same savanna ecosystem, they had access to spatially concentrated

human food scraps. Lodge group baboons were therefore able to feed more efficiently and travel shorter distances each day to meet their caloric requirements (Muruthi *et al.* 1991; Bronikowski & Altmann 1996). In addition, they experienced reduced seasonal and annual variance in resource availability compared to wild-feeding animals. Previous work in Amboseli has documented striking behavioral and physiological differences between these groups. Lodge animals, who expended considerably less energy to achieve the same caloric intake as wild-feeding animals, exhibited higher serum insulin, cholesterol, and body fat levels compared to their wild-feeding counterparts (Muruthi *et al.* 1991; Altmann *et al.* 1993; Kemnitz *et al.* 2002). Further, Lodge juveniles grew faster and matured earlier than wild-feeding animals, suggesting that the combination of more stable resource availability, higher feeding efficiency, and shorter travel distances translated into measurable fitness advantages (Altmann & Alberts 2005).

Here, we investigated whether environmentally induced changes in DNA methylation levels might contribute to the known phenotypic differences between Lodge and wild-feeding animals. To do so, we investigated three sets of questions: (i) do the differences in resource base-associated with the Lodge versus wild-feeding conditions significantly predict DNA methylation levels?; (ii) are sites that are differentially methylated by resource base likely to be functionally important?; and (iii) is the signature of resource base stable or plastic over time, when environmental conditions change? For the third question, we drew on samples from eight male baboons that switched from either the Lodge to wild-feeding condition or from the wild-feeding to Lodge condition as a consequence of natal dispersal. We then asked whether pre-

dispersal (i.e., early life) or post-dispersal (i.e., adult) resource base left a stronger signature on genome-wide DNA methylation patterns.

Materials and methods

Study subjects and sample collection

All study subjects were members of a long-term study population of yellow baboons (*Papio cynocephalus*) that has been monitored by the Amboseli Baboon Research Project (ABRP) for over four decades (Alberts & Altmann 2012). Our study focused on 69 animals from the ABRP study population, including: (i) 39 baboons that resided in a wild-feeding group from birth until the time of sampling (29 males and 11 females); (ii) 22 baboons that either resided in Lodge group from birth until the time of sampling (11 males and 7 females) or gained access to the Lodge resource base early in their lives (4 females born prior to monitoring of Lodge group in 1982); (iii) 3 males that were born in Lodge group and dispersed into a wild-feeding group following reproductive maturation, and were sampled in a wild-feeding group as adults; and (iv) 5 males that were presumed to have been born in a wild-feeding group and that dispersed into Lodge group following reproductive maturation, and were sampled in Lodge (Figure 1). For the 3 males that were born in Lodge group and sampled in a wild-feeding group, their early histories and dispersal events were directly observed. For the 5 males that were presumed born in wild-feeding groups but were sampled in Lodge group, we inferred their early histories based on two pieces of evidence. First, adult male baboons very rarely remain in their natal group to reproduce (Pusey & Packer 1987), and genetic analysis suggested that these adult males were unrelated to members of Lodge group

other than their offspring (Altmann *et al.* 1996). Second, there were very few social groups associated with human food sources in the ecosystem at the time these males matured, suggesting wild-feeding origins for all or most of these animals. Further information about these 8 post-dispersal males, including the timing of dispersal relative to blood sampling, if known, is provided in Table S1.

To investigate epigenetic differences between Lodge and wild-feeding animals, we combined genome-wide DNA methylation data from a previous study on statistical methods for differential DNA methylation analysis (n = 50 individuals; Lea *et al.* 2015b) with additional DNA methylation data generated to study resource base effects in this study (n = 19 individuals; see Figure S7 for quality control comparisons between the two data sets). All data were derived from whole blood samples collected by the ABRP between 1989 and 2011, following well-established procedures (Altmann *et al.* 1996; Tung *et al.* 2009, 2011, 2015). Briefly, animals were immobilized by an anesthetic-bearing dart delivered through a hand-held blow gun, and, following immobilization, were quickly transferred to a processing site for blood sample collection. Following sample collection, study subjects were allowed to regain consciousness in a covered holding cage until they were fully recovered from the effects of the anesthetic. Upon recovery, study subjects were released near their social group and closely monitored. Blood samples were stored at the field site or at an ABRP-affiliated lab at the University of Nairobi until they were transported to the United States.

Generation and processing of genome-wide DNA methylation data

To measure genome-wide DNA methylation levels, we used a cost-effective, high-throughput sequencing approach known as reduced representation bisulfite sequencing (RRBS) (Meissner *et al.* 2008; Gu *et al.* 2011; Boyle *et al.* 2012). RRBS relies on two key steps: (i) digestion of genomic DNA with the enzyme *Msp1*, which produces DNA fragments that begin and/or end with an informative CpG site; and (ii) treatment of *Msp1*-digested DNA with the chemical sodium bisulfite, which leaves methylated cytosines intact but converts unmethylated cytosines to uracil (and ultimately thymine after PCR). Following high-throughput sequencing and mapping of all reads to a reference genome, CpG site-specific DNA methylation levels can be estimated as the ratio of reads read as cytosine (reflecting an originally methylated version of the base) to the total number of mapped reads (reflecting both methylated and unmethylated versions of the base, i.e., reads read as either cytosine or thymine).

To construct RRBS libraries, we followed the protocol of Boyle and colleagues (Boyle *et al.* 2012). For each individual, we created a barcoded library from 180 ng of blood extracted baboon DNA, combined with 1 ng of unmethylated lambda phage DNA to assess the efficiency of the bisulfite conversion (see Supporting Information). Each sample was sequenced to a mean depth (± SD) of 27.25 ± 13.62 million reads on the Illumina HiSeq 2000 platform. Following sequencing, we removed adapter contamination, low-quality bases, and bases artificially introduced during library construction using the program Trim Galore! (Krueger 2015). We then used the program BSMAP (Xi & Li 2009) to map the trimmed reads to the olive baboon genome (*Panu 2.0*), and to extract the methylated read count and total read count for each individual

and CpG site. Before performing differential methylation analyses, we filtered out constitutively hypermethylated and hypomethylated sites from our data set, as well as invariable sites and sites with low levels of mean coverage. Importantly, filtering for hypomethylation should reduce potential biases introduced from mapping yellow baboon RRBS reads to the olive baboon genome (see Supporting Information and Figure S1 for more details).

Testing for differences in DNA methylation levels at individual CpG sites

We first tested for a relationship between resource base and DNA methylation levels using data generated from individuals who had spent most or all of their lives (prior to sampling) in a wild-feeding group (n = 39) or the Lodge group (n = 22). To do so, we used the binomial mixed effects approach implemented in the program MACAU (Lea *et al.*). This approach allowed us to control for kinship in our data set and to work directly with the raw count data – two features that maximize power in bisulfite sequencing data sets. Specifically, for each CpG site, we used MACAU to model DNA methylation levels as a function of the fixed effects of resource base (Lodge or wild-feeding), sex, age of the animal, number of years since blood sample collection, and bisulfite conversion rate (see Table S2 for all covariate values). We also included a random effect that accounts for genetic relatedness among individuals (see Supporting Information and Lea *et al.* 2015b). For each CpG site tested, we extracted the p-value associated with the resource base term and corrected for multiple hypothesis testing using the false discovery rate (FDR) approach implemented in the R package *qvalue* (Storey & Tibshirani 2003; Dabney & Storey 2015). We considered a CpG site to be

differentially methylated by resource base (referred to below as 'resource base-associated') if it passed a 10% FDR threshold.

Because DNA methylation patterns are highly cell type-specific (Reinius *et al.* 2012; Roadmap Epigenomics Consortium *et al.* 2015), we also investigated whether differences in whole blood cell type composition between Lodge and wild-feeding animals could confound our analysis. To do so, we drew on two datasets: (i) cell type proportion data generated from manual counts of Giemsa-stained blood smears (for 15 Lodge and 20 wild-feeding animals); and (ii) genome-wide, cell type-specific DNA methylation data from a previous study of human whole blood (Reinius *et al.* 2012; Jaffe 2015). We used these data to first test whether resource base predicted cell type composition, and then to investigate whether resource base-associated sites were more likely to exhibit cell type-specific DNA methylation patterns, which would suggest a potential confound (Supporting Information).

Enrichment of differentially methylated sites by genomic annotation

We hypothesized that, if shifts in DNA methylation are part of a coordinated regulatory response to resource base, these epigenetic changes should be biased toward regions of the genome that control gene expression, and should be targeted towards genes involved in similar biological processes. To test these hypotheses, we evaluated whether our data were consistent with four predictions. Specifically, we expected resource base-associated sites to be:

1. over-represented in putatively functional gene regulatory elements (i.e., gene bodies, promoters, CpG islands, CpG island shores, or enhancers) and under-represented in regions of the genome with no known regulatory function;

2. over-represented in chromatin states associated with active gene transcription and under-represented in chromatin states associated with gene repression;

3. more likely to fall in or near genes expressed in whole blood, compared to genes not expressed in blood; and

4. enriched near genes involved in coherent biological pathways and processes.

To test prediction (1), we used publicly available annotation tracks for the olive baboon genome to assign each resource base-associated CpG site to one of the following categories: gene body, promoter, CpG island, CpG island shore, H3K4me1-marked enhancer, or unannotated (see Supporting Information and Figure S2 for further information about category definitions). For each category, we used a Fisher's Exact Test to test for significant over- or under-enrichment of resource base-associated sites relative to chance expectations. Importantly, we defined chance expectations based on the CpG sites that we actually profiled in our data set, which are themselves enriched for CpG-rich, putatively functional regions of the genome (Figure S7). Thus, significant over- or under-enrichment of resource base-associated sites would indicate that resource base-associated sites are *even more* likely to fall in a given genomic compartment than other sites captured by the RRBS protocol.

To test prediction (2), we drew on chromatin state annotation data generated by the NIH Roadmap Epigenomics project (Roadmap Epigenomics Consortium *et al.* 2015) for human peripheral blood mononuclear cells. Chromatin states are defined by

combinations of histone marks (acetylation or methylation), and provide information about the transcriptional activity and regulatory element function of the associated DNA. For example, actively expressed gene bodies are associated with a chromatin state defined by H3K36me3 marks, whereas repressed genes are associated with a chromatin state defined by H3K27me3 marks. Importantly, histone marks tend to be highly conserved between humans and closely related primates, such as chimpanzees and rhesus macaques (Zhou *et al.* 2014); the baboon lineage diverged from the human lineage at the same time as rhesus macaques, supporting the overall accuracy of using Roadmap Epigenomics-defined chromatin states here. We therefore assigned each resource base-associated CpG site to one of 15 chromatin states (Figure S3). We then investigated the degree to which resource base-associated sites were over- or under-enriched in each chromatin state. As above, we again used Fisher's Exact Tests against the background of the CpG sites included in our RRBS data set (not compared to the whole genome).

To test predictions (3) and (4), which rely on gene-level information, we first assigned each CpG site to a particular gene if it occurred in the gene body or within 10 kb of the gene transcription start site (TSS) or end site (TES; see also Figure S4 for results based on alternative criteria for CpG assignment to genes). To test prediction (3), we then categorized all genes as either not expressed in whole blood or blood-expressed (based on whether they were included in a whole blood RNA-seq data set also from the Amboseli baboons, Tung *et al.* 2015). We used a Fisher's Exact Test to ask whether CpG sites assigned to blood-expressed genes were more likely to be

differentially methylated by resource base, compared to CpG sites assigned to unexpressed genes.

Finally, to test prediction (4), we performed categorical enrichment analysis using publicly available gene annotations (Kyoto Encyclopedia of Genes and Genomes (KEGG): Ogata *et al.* 1999) and the GeneTrail analysis software (Backes *et al.* 2007). Here, we focused only on genes associated with differentially methylated sites (as defined above), and tested for over-representation of genes that fall within specific pathways, compared to chance expectations. To do so, we used hypergeometric tests followed by FDR correction (Benjamini & Hochberg 1995).

Testing for differences in DNA methylation levels at metabolic pathways

Given the known differences in diet, activity patterns, and physiology between Lodge and wild-feeding animals, we were particularly interested in whether metabolism-related genes showed an epigenetic signature of resource base. To specifically address this question, we focused on CpG sites near (in the gene body or within 10 kb of the gene TSS or TES) genes involved in 36 KEGG pathways related to the metabolism of food or to energy balance (Figure S5) (Ogata *et al.* 1999). These pathways were chosen *a priori* because of their relevance to the phenotypic differences between wild-feeding and Lodge animals. We used the R package 'GlobalTest' (Goeman *et al.* 2004) to ask whether linearly transformed methylation levels (Supporting Information) from CpG sites near genes involved in metabolism-related pathways displayed a signature of resource base. This approach asks whether samples with similar DNA methylation patterns (at predefined sets of sites) also have similar resource base labels (i.e., Lodge

or wild-feeding), using a framework similar to penalized logistic regression. Thus, the level of analysis is shifted from individual CpG sites to sets of CpG sites associated with putatively similar functions, allowing us to specifically test pathway-based predictions (Goeman *et al.* 2004). We corrected all GlobalTest p-values for multiple hypothesis testing using the R package *qvalue* (Storey & Tibshirani 2003; Dabney & Storey 2015).

Identification of differentially methylated regions (DMRs)

Spatially contiguous stretches of differentially methylated sites (often termed 'differentially methylated regions', or DMRs) are more likely to have functional effects on gene expression than differentially methylated sites that occur in isolation (Lister *et al.* 2009; Hansen *et al.* 2011; Jaffe *et al.* 2012). To identify DMRs in our data set, we focused on resource base-associated sites (detected at a 10% FDR) that had at least one other measured CpG site within a 2 kb window centered on the focal site (following the precedent for window size used in Lister *et al.* 2009 and Hansen *et al.* 2012). For sites that met this criterion, we counted the absolute number of nearby sites that also exhibited evidence for differential methylation, at a less conservative 20% FDR threshold. We defined DMRs as a cluster of at least 3 resource base-associated sites. We chose this cutoff because clusters of this size were extremely unlikely to occur by chance in permuted data (Supporting Information and Figure S6). Specifically, despite relaxing the FDR threshold for identifying CpG sites close to the original resource base-associated sites, our criteria for identifying DMRs results in a relatively stringent FDR threshold of 6.5% FDR. Finally, we collapsed any DMRs with overlapping boundaries into a single, longer DMR.

Testing the effects of PFKP promoter methylation on gene expression levels

Our analyses revealed one particularly large DMR at the promoter region of the phosphofructokinase gene (*PFKP*). This DMR stretched across 192 CpG sites, including 30 sites associated with resource base at a 10% FDR (Figure 4A). Because *PFKP* is involved in the rate-limiting step of glycolysis and has been previously implicated in obesity-related traits (Ehrich *et al.* 2005; Scuteri *et al.* 2007), we were interested in understanding whether *PFKP* promoter methylation alone was sufficient to drive differences in gene expression. This relationship is implicitly assumed by arguments linking environmental variation to phenotypic variation via epigenetic mechanisms, but is rarely tested in practice.

To test this hypothesis, we used an experimental reporter gene assay in which we cloned 817 bp of the *PFKP* promoter (containing 72 CpG sites) into a CpG-free vector backbone that contains the luciferase reporter gene (pCpGL, Klug & Rehli 2006). After growing up the *PFKP*-pCpGL construct in competent *E. coli* GT115 cells (InvivoGen), we subjected the purified plasmid to one of three treatments: (i) methylation of all 72 CpGs in the *PFKP* promoter region via treatment with *M.SssI* (a methyltransferase that targets all CG sequence motifs, resulting in a completely methylated *PFKP* promoter); (ii) methylation of 13 CpGs in the *PFKP* promoter via treatment with *HhaI* (a methyltransferase that targets only CGCG sequence motifs, resulting in a partially methylated *PFKP* promoter); and (iii) a mock treatment (water substituted for the methyltransferase enzyme, resulting in a completely unmethylated *PFKP* promoter).

We transfected four replicates of each treatment condition into the human K562 myeloid cell line and incubated the transfected cells for 24 hours (n=12 total transfection experiments). To control for transfection efficiency, a vector containing Renilla luciferase was transfected in parallel. Post-incubation, cells were assayed for luciferase activity using a dual-luciferase reporter assay kit (Promega), and luciferase expression was normalized using measures of co-transfected Renilla activity. Finally, we tested for an effect of DNA methylation at the *PFKP* promoter on luciferase gene expression using pairwise Wilcoxon rank sum tests. Additional details on our experimental procedures are provided in the Supporting Information.

Investigating the stability or plasticity of DNA methylation levels for individuals that switched between resource bases

Finally, we tested two alternative hypotheses about DNA methylation patterns in the 8 males that switched resource base as a consequence of natal dispersal. First, we hypothesized that if resource base exerts stable, long-term effects on DNA methylation patterns, switching individuals should resemble their natal group members (Lodge or wild-feeding) rather than the group to which they belonged at the time of sampling (Figure 5A). Alternatively, if resource base-associated DNA methylation patterns are largely plastic in response to prevailing conditions, we hypothesized that switching individuals should exhibit DNA methylation patterns that reflect their resource base at the time of sampling rather than their natal group (Figure 5B).

To differentiate between these two possibilities, we built a support vector machine (SVM) classifier, a machine-learning approach used for class prediction from

high dimensional data (Cortes & Vapnik 1995). This classifier used DNA methylation data to distinguish between individuals that spent all (or the vast majority) of their lives in either a wild-feeding group or in Lodge group (n = 61 individuals). As predictive features for this model, we included the 334,840 CpG sites that were not associated with age, sex, bisulfite conversion rate, or sample age at a nominal p-value of 0.05. We chose this global approach (rather than using significantly differentially methylated sites only) because it allowed us to include sites that may be truly affected by resource base, but did not pass the genome-wide significance threshold in the site-by-site analysis. Additionally, using all sites ensured that the model classification accuracy was not biased by using features that had already been associated with the response variable in a previous analysis of the same data set (doing so can result in erroneously high classification accuracy even from completely random data: Hastie *et al.* 2009). Because SVMs cannot work on binomially distributed count data, we linearly transformed our data before building the SVM (Supporting Information).

Finally, we used the resulting SVM to ask whether individuals that switched resource base more closely resembled their pre-switch or post-switch conspecifics. To do so, we used the fitted model to predict the resource base of the 8 individuals that dispersed between groups (using DNA methylation data from these 8 individuals, for the same 334,840 CpG sites).

Ethics statement

The data used in this study were generated from wild baboon samples, collected in the Amboseli region of Kenya. This research was conducted under the authority of the

Kenya Wildlife Service (KWS), the Kenyan governmental body that oversees wildlife (current permit numbers NCST/RCD/12B/012/57 to Jenny Tung, NCST/5/002/R/777 to Susan Alberts, and NCST/5/002/R/776 to Jeanne Altmann). As the animals are members of a wild population, KWS requires that we do not interfere with injuries to study subjects inflicted by predators, conspecifics, or through other naturally occurring events. Permission to perform temporary immobilizations (for blood sample collection) was granted by KWS; further, these immobilizations were supervised by a KWS-approved Kenyan veterinarian, who monitored anesthetized animals for hypothermia, hyperthermia, and trauma (no such events occurred during our sample collection efforts). Observational and blood sample collection protocols were approved though IACUC committees at Duke University (current protocol is A020-15-01 to Jenny Tung and Susan C. Alberts).

Results

Genome-wide DNA methylation levels contain a signature of resource base

We found that DNA methylation patterns in our full data set (i.e., after quality control, but before filtering for constitutively hypermethylated, hypomethylated, or invariant sites) recapitulated typical patterns observed in mammalian genomes. Specifically, most of the genome was hypermethylated, with the exception of H3K4me1-marked enhancers, promoters, and CpG islands (Figure S7). Further, DNA methylation levels near the transcription start sites (TSS) of expressed genes were inversely related to their expression levels (Figure S7).

After filtering, we investigated DNA methylation levels at over half a million CpG sites in the baboon genome (n = 535,996 sites). As expected when using RRBS (Gu *et al.* 2011; Boyle *et al.* 2012), many of these sites occurred in CpG-rich regions, particularly gene bodies (224,553 sites), promoters (25,730 sites), CpG islands (57,461 sites), and CpG island shores (117,226 sites; Figure S7). Further, this dataset encompasses many putatively functional regions of the genome, as at least one CpG was measured in 66% of genes, 28% of promoters, 40% of CpG islands, 44% of CpG island shores, and 11% of enhancers. Within our filtered set, we identified 1,014 sites (at a 10% FDR) that were differentially methylated between lifelong wild-feeding and Lodge group animals (Figure 2; see also Figure S8 for a power analysis of our data set). We did not detect significant effects of sex on DNA methylation levels, or significant sex by resource base interaction effects, consistent with previous studies that have identified weak or no sex effects in human blood (Eckhardt *et al.* 2006; Lam *et al.* 2012). Also in line with previous studies (Tobi *et al.* 2014), we did not observe strong directional bias for the 1,014 differentially methylated sites (Figure S9). Importantly, our analyses of cell type composition and cell type-specific DNA methylation data indicate that our results are unlikely to be explained by cell type heterogeneity effects (Supporting Information and Figure S10).

Sites associated with resource base are enriched in functionally important regions of the genome

CpG sites associated with differences in resource base were highly nonrandomly distributed in the genome. Specifically, they were enriched in putative enhancers (p =

9.70 x 10^{-3}) and gene promoters (p = 3.66 x 10^{-3}), and underrepresented in functionally unannotated regions of the genome (p = 8.96 x 10^{-10}; Figure 3). Further, they were more likely to occur near genes expressed in whole blood than near unexpressed genes (odds ratio = 1.51, p = 5.49 x 10^{-7}; see also Supporting Information and Figure S11). More fine-grained analyses of chromatin states also indicated a role for differentially methylated sites in the active regulation of genes: differentially methylated sites were more likely to occur in chromatin states associated with active gene transcription in blood cells, including chromatin states designated as 'active TSS' (p = 2.78 x 10^{-4}), 'flanking active TSS' (p = 1.44 x 10^{-2}), 'strong transcription' (p = 5.24 x 10^{-3}), and 'enhancer' (p = 1.28 x 10^{-3}). In contrast, differentially methylated sites were strongly under-enriched in a chromatin state indicative of gene repression ('repressed polycomb', p = 5.49 x 10^{-3}).

Resource base-associated CpG sites are enriched in specific biological pathways

We observed two pieces of evidence that resource base-associated CpG sites were concentrated in specific biological pathways. First, while no KEGG pathways were enriched at a 10% FDR threshold, a more relaxed 20% FDR threshold revealed that resource base-associated sites were enriched near genes involved in 5 KEGG pathways: the T cell receptor and B cell receptor signaling pathways, axon guidance, phosphatidylinositol signaling, and insulin signaling. Second, our GlobalTest analyses revealed patterns of differential methylation associated with carbohydrate metabolism (galactose: p = 1.29 x 10^{-5}; fructose and mannose: p = 1.79 x 10^{-4}; and the glycolysis pathway: p = 2.94 x 10^{-4}); amino acid metabolism (glycine, serine, and threonine: p =

4.21 x 10^{-3}; tryptophan: p = 0.018); insulin signaling (p = 0.012); and the breakdown of other dietary components (propanoate, p = 0.018; all GlobalTest results reported a 10% FDR threshold; see Figure S5). For pathways that included the *PFKP* gene (specifically, the three carbohydrate metabolism pathways listed above), the observed effect of resource base on DNA methylation levels appears to have been driven almost entirely by differential methylation at *PFKP* (Figure S5).

DMRs occur more often than expected by chance, and near a key metabolic gene

We identified 87 2-kb windows that met our criteria for differentially methylated regions, compared to only 6 such windows observed on average in permuted data (equivalent to a 6.5% FDR; see Figure S6). These 87 windows collapsed into 29 distinct, longer DMRs, the largest of which fell within the promoter region of an insulin sensitive gene that encodes the rate-limiting enzyme in glycolysis (Lo *et al.* 2013; Webb *et al.* 2015). For 90% of the 192 sites we tested in this region, *PFKP* was more highly methylated in wild-feeding individuals than in Lodge group baboons (Figure 4A; no site was significantly more highly methylated in Lodge group animals), suggesting that *PFKP* expression may be down-regulated in less resource-rich environments.

In support of this possibility, our reporter assay experiments confirmed that complete methylation of all CpGs in the *PFKP* promoter region (n = 72 CpGs in the region we tested; Supporting Information) suppressed luciferase expression levels relative to fully unmethylated *PFKP* promoter constructs (Wilcoxon rank sum test, W = 16, p = 0.014). Furthermore, even methylation of a minority of CpGs in the *PFKP* promoter (n = 13 CpGs) produced a graded reduction in gene expression levels,

intermediate between the fully methylated and fully unmethylated versions (Wilcoxon rank sum test, W = 14, p = 0.057 for comparison between fully methylated and partially methylated constructs; Figure 4B).

Individuals that switched between resource bases more closely resembled wild-feeding individuals, regardless of the direction of the switch

Finally, we built a support vector machine (SVM) classifier (trained on 334,840 CpG sites) that discriminated between Lodge and wild-feeding individuals with 82% accuracy. We used this model to classify switching individuals based on their DNA methylation levels to test whether switching individuals would be (i) grouped with their natal conspecifics, suggesting that early life (i.e., pre-dispersal) resource base drives variation in DNA methylation levels; or (ii) grouped with their conspecifics at the time of sampling, suggesting that prevailing conditions are more important than past history (Figure 5).

We did not find evidence in support of either of these hypotheses. Instead, we found that the DNA methylation patterns of most switching males (7 of 8) were consistently predicted to be more like wild-feeding animals. This pattern held regardless of whether we considered males that immigrated from a wild-feeding group to Lodge group, or males that immigrated from Lodge group to wild feeding groups. Only one male—an individual who moved from a wild feeding group into Lodge group—violated this pattern, primarily because our model could not consistently classify him with either wild-feeding or Lodge feeding individuals (Figure 5).

Discussion

Ecological variation, experienced throughout life, can have lasting and dramatic effects on trait variation. However, the molecular mechanisms that mediate these effects remain largely unexplored, especially in natural populations. Here, we present the first evidence that resource base — an environmental variable with known effects on activity patterns, growth rates, insulin levels, and body fat percentages in our study population — influences DNA methylation levels in a wild mammal (Muruthi *et al.* 1991; Altmann *et al.* 1993; Kemnitz *et al.* 2002; Altmann & Alberts 2005). Specifically, we identified over a thousand differentially methylated CpG sites, as well as 29 differentially methylated regions, that differed between wild-feeding and Lodge group baboons. Our results support the importance of DNA methylation in translating signals from the environment into changes in gene regulation within cells.

The functional relevance of differential methylation at resource base-associated sites

Several pieces of evidence suggest that the changes in DNA methylation we observed are targeted, coordinated, and likely to exert downstream effects on gene regulation. Specifically, differentially methylated sites were more likely to occur in promoter and enhancer regions; near genes expressed in blood, the tissue we sampled; and at stretches of DNA marked by transcriptionally active chromatin states. Further, differential methylation consistently occurred near genes involved in metabolism and insulin signaling, one of the known differences between Lodge and wild-feeding baboons in Amboseli (Kemnitz *et al.* 2002).

Because current methods for functional validation (e.g., reporter assay experiments) are not feasible on a genome-wide scale, we focused on validating the functional role of changes in DNA methylation at the largest identified DMR (in the promoter region of *PFKP*). However, several of the additional DMRs we identified fall near genes with relevance to metabolism and energy balance, and may also contribute to organism-level differences (Table S3). For example, we identified a DMR in *KCNIP4*, where genetic variants have been previously associated with obesity-related traits (Comuzzie *et al.* 2012). In addition, we identified a DMR in the 5' UTR of *TPM1*, where genetic variation has been associated with platelet count and volume (Soranzo *et al.* 2009; Gieger *et al.* 2011), both of which are biomarkers of obesity and metabolic syndrome (Coban *et al.* 2005; Jesri *et al.* 2005; Tavil *et al.* 2007). Together, these results points toward a model in which easy access to resources alters metabolic processes in Lodge group animals, at least in part through targeted changes in DNA methylation. Further work is needed to assess the causal effects of changes in DNA methylation on gene expression at these loci, and more generally to improve methods for high-throughput, genome-wide functional validation.

By far, the most striking DMR we identified fell near a gene coding for an isoform of phosphofructokinase*,* which catalyzes the irreversible, committed step of glycolysis. Our reporter assay experiments indicate that lower levels of *PFKP* promoter methylation are sufficient to drive higher levels of *PFKP* expression (Figure 4B). In combination with work in laboratory mice, these results suggest a possible avenue through which regulatory changes at *PFKP* may contribute to organism-level changes – namely, body fat mass accumulation – in Lodge animals. Specifically, several studies have

demonstrated that mice who become obese on high carbohydrate diets exhibit increased levels of the phosphofructokinase enzyme relative to mice that did not become obese, but ingested similar numbers of calories (Yamini *et al.* 1992; Dourmashkin *et al.* 2005). Further, mice with experimentally reduced expression levels of the *PFKM* gene (the muscle isoform of phosphofructokinase) exhibit greatly reduced levels of both lipogenesis (the process by which energy is stored as fat) and overall body fat (Getty-Kaushik *et al.* 2010). Together, these studies argue that phosphofructokinase activity is stimulated by a high carbohydrate diet and consequently favors increased fat accumulation. Previous work in Amboseli has only tested for differences in protein content and overall energy intake between Lodge and wild-feeding animals (revealing that protein intake is higher in wild-feeding animals, while energy intake does not differ: Muruthi *et al.* 1991). However, Lodge animals were likely to have ingested higher levels of simple carbohydrates as well, consistent with the model proposed above.

Stability and plasticity in the epigenetic signature of resource base

DNA methylation marks are largely established during development and subsequently carried across mitotic cell divisions. Consequently, most studies of diet or resource access effects on DNA methylation have focused on exposures during development (Wolff *et al.* 1998; Sinclair & Allegrucci 2007; Khulan *et al.* 2012). These studies have observed strong effects of maternal resource constraint during pregnancy on offspring methylation levels (Heijmans *et al.* 2008; Waterland *et al.* 2010; Tobi *et al.*

2014). Furthermore, they have emphasized the precise timing of maternal resource effects, which are sometimes limited to specific trimesters (Tobi *et al.* 2009, 2014).

The results of such studies have been widely interpreted as support for a 'critical period' model of environmental epigenetic effects, where environmental insults during development are primarily responsible for downstream effects on gene regulation (Meaney & Szyf 2005; Faulk & Dolinoy 2011; Heim & Binder 2012). However, recent work has shown that DNA methylation levels are also affected by environmental conditions later in life, including adult socioeconomic status in humans (McGuinness *et al.* 2012; but see Lam *et al.* 2012), experimentally manipulated social status in captive rhesus macaques (Tung *et al.* 2012), and immune response to infection in humans (Marr *et al.* 2014; McErlean *et al.* 2014; Pacis *et al.* 2015). These observations of epigenetic plasticity also extend to recent studies of energy balance and diet in humans: individuals that subsisted on a high fat diet for one week exhibited epigenetic changes at thousands of CpG sites compared to randomized controls (Jacobsen *et al.* 2012). Similarly, short-term exercise interventions induce widespread changes in DNA methylation levels (Barrès *et al.* 2012; Rönn *et al.* 2013).

The current literature thus indicates that DNA methylation plays two complementary roles. In some cases, it encodes a signature of early life experience, producing stable effects on gene regulation that persist over time. In other cases, it continues to be dynamically regulated, allowing organisms to adjust their phenotypes to prevailing environmental conditions. Whereas the effects of resource availability have been largely studied in the context of the first role (Klebig *et al.* 1995; Wolff *et al.* 1998;

Heijmans *et al.* 2008; Carone *et al.* 2010; Tobi *et al.* 2014), we were able to take advantage of naturally occurring male dispersal to investigate both roles (Figure 1).

Surprisingly, our analyses revealed a lack of support for either long-term stability or global plasticity. Instead, we found that the DNA methylation patterns of switching individuals, whether originating from or immigrating into wild-feeding groups, almost universally resembled those of lifelong wild-feeding individuals (Figure 5). Wild-feeding individuals rely on widely distributed, seasonally available foods, and consequently experience greater seasonal and year-to-year variance in resource access compared to Lodge animals. Our results thus indicate that more challenging environments, in terms of energy balance, may leave a stronger epigenetic signature than more favorable environments. This appears to hold whether exposure occurs earlier in life—consistent with long-term early life effects and developmental constraint models—or later in life (e.g., in post-dispersal males), consistent with a more sustained capacity for plasticity. Our results thus dovetail with recent work on both the evolutionary significance of early life effects (Lindström 1999; Gluckman *et al.* 2005a; Botero *et al.* 2015), including in the Amboseli baboons (Lea *et al.* 2015), as well as the possibility that epigenetic marks mechanistically mediate these effects (Weaver *et al.* 2004; Lam *et al.* 2012; Tobi *et al.* 2014). However, theoretical work is needed to connect the evolution of plasticity to expectations about the epigenetic patterns associated with different levels of adaptive plasticity (Furrow & Feldman 2014).

Finally, our findings indicate an important caveat for studies in ecological epigenetics: that inferences about stability versus plasticity may be contingent on the direction in which an environment changes. If we had only analyzed male baboons that

transitioned from resource abundance to resource limitation (i.e., from the Lodge to wild-feeding resource base), our results would have supported complete plasticity. On the other hand, if we had focused only on males that transitioned from resource limitation to resource abundance, our data would have supported long-term stability. Notably, the latter type of transition is the one that has been highlighted in studies of human health, which have tended to emphasize the special importance of early life exposure (Gluckman & Hanson 2004; Schulz 2010). In contrast, our findings suggest that the specific experience of resource limitation may leave an epigenetic signature that transcends any critical window of exposure, at least in our system. Future research will be needed to assess the generality of these results. Nevertheless, we believe our results as a whole emphasize the importance—already acknowledged in evolutionary and behavioral ecology more generally—of taking a life course approach to ecological epigenetics. Organisms are a product of both their current environment and their past history, and we should expect the epigenetic patterns within their cells to reflect this combination. The more interesting outstanding questions are about the negotiation of this balance, including how it evolves.

Acknowledgments

We thank the Kenya Wildlife Service, Institute of Primate Research, National Museums of Kenya, National Council for Science and Technology, members of the Amboseli-Longido pastoralist communities, Tortilis Camp, and Ker & Downey Safaris for their assistance in Kenya. We also thank M Yuan, T Reddy, C Guo, and the Rehli lab for assistance with *PFKP* experiments; M Akinyi and F Mazurek for their help scoring blood slides; T Vilgalys for assistance with testing for mapping bias; A Meissner, J DeYoung, and J Aman for assistance with RRBS; and RS Mututua, S Sayialel, JK Warutere, M Akinyi, T Wango, and V Oudu for invaluable assistance with sample collection. DR Rubenstein and two anonymous reviewers provided constructive comments and criticism on an earlier draft of this work. Finally, we thank the Baylor College of Medicine Human Genome Sequencing Center for access to the current version of the baboon genome assembly (*Panu 2.0*). Sequencing data generation and analysis, as well as manuscript preparation, were supported by a seed grant from the Duke Population Research Institute (a component of 5R24-HD065563-03), NSF BCS-1455808, and NIA R21-AG049936. Behavioral data collection and blood sample collection was supported by a series of grants over the past several decades, including NSF IBN 9223335, NSF IBN 0322613, NSF DEB 0846286, NSF IOS 0919200, NIA P01 AG031719, and NIA R01 AG034513.

Data accessibility

For this study, we drew on previously published baboon RRBS data (Lea *et al.*), as well as RRBS data we generated for this study. All sequencing data has been deposited in

NCBI (project accession SRP058411). In addition, the following data are available in the Dryad database (doi:10.5061/dryad.2d80m): tables of methylated and total read counts (i.e., the input files for our analyses in MACAU), the output files from MACAU (including p-values and effect size estimates for each site), and cell type proportion data collected from blood smears.

References

Alberts SC, Altmann J (2012) The Amboseli Baboon Research Project: 40 years of continuity and change. In: *Long-Term Field Studies of Primates* (eds Kappeler P, Watts DP), pp. 261–288. Springer, New York.

Altmann S (1991) Diets of yearling female primates (Papio cynocephalus) predict lifetime fitness. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 420–3.

Altmann J, Alberts SC (2005) Growth rates in a wild primate population: ecological influences and maternal effects. *Behavioral Ecology and Sociobiology*, **57**, 490–501.

Altmann J, Alberts S, Haines S *et al.* (1996) Behavior predicts genetic structure in a wild primate group. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 5797–5801.

Altmann J, Schoeller D, Altmann S a, Muruthi P, Sapolsky RM (1993) Body size and fatness of free-living baboons reflect food availability and activity levels. *American Journal of Primatology*, **30**, 148–161.

Backes C, Keller A, Kuentzer J *et al.* (2007) GeneTrail-advanced gene set enrichment analysis. *Nucleic Acids Research*, **35**, 186–192.

Barrès R, Yan J, Egan B *et al.* (2012) Acute exercise remodels promoter methylation in human skeletal muscle. *Cell Metabolism*, **15**, 405–11.

Beehner JC, Onderdonk DA, Alberts SC, Altmann J (2006) The ecology of conception and pregnancy failure in wild baboons. *Behavioral Ecology*, **17**, 741–750.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, **57**, 289–300.

Botero C a, Weissing FJ, Wright J, Rubenstein DR (2015) Evolutionary tipping points in the capacity to adapt to environmental change. *Proceedings of the National Academy of Sciences*, **112**.

Boyle P, Clement K, Gu H, Smith Z (2012) Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biology*, **13**, R92.

Bronikowski AM, Altmann J (1996) Foraging in a variable environment: weather patterns and the behavioral ecology of baboons. *Behavioral Ecology and Sociobiology*, **39**, 11–25.

Carone BR, Fauquier L, Habib N *et al.* (2010) Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. *Cell*, **143**, 1084–96.

Coban E, Ozdogan M, Yazicioglu G, Akcit F (2005) The mean platelet volume in patients with obesity. *International Journal of Clinical Practice*, **59**, 981–982.

Comuzzie AG, Cole S a., Laston SL *et al.* (2012) Novel Genetic Loci Identified for the Pathophysiology of Childhood Obesity in the Hispanic Population. *PLoS ONE*, **7**.

Cortes C, Vapnik V (1995) Support-Vector Networks. *Machine Learning*, 273–297.

Cummings ME, Larkins-Ford J, Reilly CRL *et al.* (2008) Sexual and social stimuli elicit rapid and contrasting genomic responses. *Proceedings of the Royal Society of London, Series B*, **275**, 393–402.

Dabney A, Storey J (2015) qvalue: Q-value estimation for false discovery rate control. R package version 1.43.0.

Dominguez-Salas P, Moore SE, Baker MS *et al.* (2014) Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nature Communications*, **5**, 3746.

Dourmashkin JT, Chang G, Gayles EC *et al.* (2005) Different forms of obesity as a function of diet composition. *International Journal of Obesity*, **29**, 1368–1378.

Drnevich J, Replogle KL, Lovell P *et al.* (2012) Impact of experience-dependent and -independent factors on gene expression in songbird brain. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 17245–17252.

Eckhardt F, Lewin J, Cortese R *et al.* (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. *Nature Genetics*, **38**, 1378–85.

Ehrich TH, Hrbek T, Kenney-hunt JP *et al.* (2005) Fine-Mapping Gene-by-Diet

Interactions on Chromosome 13 in a LG/J x SM/J Murine Model of Obesity. *Diabetes*, **54**, 1863–1872.

Faulk C, Dolinoy DC (2011) Timing is everything: The when and how of environmentally induced changes in the epigenome of animals. *Epigenetics*, **6**, 791–797.

Feil R, Fraga MF (2011) Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics*, **13**, 97–109.

Foret S, Kucharski R, Pellegrini M *et al.* (2012) DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proceedings of the National Academy of Sciences*, **109**, 4968–4973.

Furrow RE, Feldman MW (2014) Genetic Variation and the Evolution of Epigenetic Regulation. *Evolution*, **68**, 673–683.

Gaillard J, Festa-Bianchet M, Yoccoz N, Loison A, Toigo C (2000) Temporal variation in fitness components and population dynamics of large herbivores. *Annual Review of Ecology and Systematics*, **31**, 367–393.

Getty-Kaushik L, Viereck JC, Goodman JM *et al.* (2010) Mice Deficient in Phosphofructokinase-M Have Greatly Decreased Fat Stores. *Obesity*, **18**, 434–440.

Gieger C, Kühnel B, Radhakrishnan a *et al.* (2011) New gene functions in megakaryopoiesis and platelet formation. *Nature*, **480**, 201–208.

Gluckman PD, Hanson MT (2004) The developmental origins of the metabolic syndrome. *Trends in Endocrinology and Metabolism*, **15**, 183–7.

Gluckman PD, Hanson MT, Spencer HG (2005a) Predictive adaptive responses and human evolution. *Trends in Ecology and Evolution*, **20**, 527–33.

Gluckman PD, Hanson M a, Spencer HG, Bateson P (2005b) Environmental influences during development and their later consequences for health and disease: implications for the interpretation of empirical studies. *Proceedings of the Royal Society of London, Series B*, **272**, 671–7.

Goeman JJ, Van de Geer S, De Kort F, van Houwellingen HC (2004) A global test for groups fo genes: Testing association with a clinical outcome. *Bioinformatics*, **20**, 93–99.

Gu H, Smith ZD, Bock C *et al.* (2011) Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nature Protocols*, **6**, 468–81.

Guo JU, Ma DK, Mo H *et al.* (2011) Neuronal activity modifies the DNA methylation

landscape in the adult brain. *Nature Neuroscience*, **14**, 1345–1351.

Hamel S, Gaillard J-M, Festa-Bianchet M, Côté SD (2009) Individual quality, early-life conditions, and reproductive success in contrasted populations of large herbivores. *Ecology*, **90**, 1981–95.

Hansen K, Langmead B, Irizarry R (2012) BSmooth : from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biology*, **13**, R83.

Hansen KD, Timp W, Bravo HC *et al.* (2011) Increased methylation variation in epigenetic domains across cancer types. *Nature Genetics*, **43**, 768–75.

Hastie T, Tibshirani R, Friedman J (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer, New York, NY.

Heijmans BT, Tobi EW, Stein AD *et al.* (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 17046–9.

Heim C, Binder EB (2012) Review of human studies on sensitive periods, gene-environment interactions, and epigenetics. *Experimental Neurology*, **233**, 102–111.

Jacobsen SC, Brøns C, Bork-Jensen J *et al.* (2012) Effects of short-term high-fat overfeeding on genome-wide DNA methylation in the skeletal muscle of healthy young men. *Diabetologia*, **55**, 3341–3349.

Jaffe AE (2015) FlowSorted.Blood.450k: Illumina HumanMethylation data on sorted blood cell populations. R package version 1.5.1.

Jaffe AE, Murakami P, Lee H *et al.* (2012) Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. *International Journal of Epidemiology*, **41**, 200–209.

Jesri A, Okonofua EC, Egan BM (2005) Platelet and white blood cell counts are elevated in patients with the metabolic syndrome. *Journal of Clinical Hypertension*, **7**, 705–711; quiz 712–713.

Jirtle RL, Skinner MK (2007) Environmental epigenomics and disease susceptibility. *Nature reviews. Genetics*, **8**, 253–62.

Jones P (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, **13**, 484–92.

Kemnitz JW, Sapolsky RM, Altmann J *et al.* (2002) Effects of food availability on serum insulin and lipid concentrations in free-ranging baboons. *American Journal of Primatology*, **57**, 13–9.

Khulan B, Cooper WN, Skinner BM *et al.* (2012) Periconceptional maternal micronutrient supplementation is associated with widespread gender related changes in the epigenome: A study of a unique resource in the Gambia. *Human Molecular Genetics*, **21**, 2086–2101.

Klebig ML, Wilkinson JE, Geisler JG, Woychik RP (1995) Ectopic expression of the agouti gene in transgenic mice causes obesity, features of type II diabetes, and yellow fur. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 4728–32.

Klose RJ, Bird AP (2006) Genomic DNA methylation: the mark and its mediators. *Trends in Biochemical Sciences*, **31**, 89–97.

Klug M, Rehli M (2006) Functional analysis of promoter CpG methylation using a CpG-free luciferase reporter vector. *Epigenetics*, **1**, 127–130.

Krueger F (2015) Trim Galore! http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

Kucharski R, Maleszka J, Foret S, Maleszka R (2008) Nutritional control of reproductive status in honeybees via DNA methylation. *Science*, **319**, 1827–1830.

Lam LL, Emberly E, Fraser HB *et al.* (2012) Factors underlying variable DNA methylation in a human community cohort. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 17253–60.

Lea AJ, Altmann J, Alberts SC, Tung J (2015) Developmental constraints in a wild primate. *The American Naturalist*, **185**, 809–821.

Lea A, Tung J, Zhou X A flexible, efficient binomial mixed model for identifying differential DNA methylation in bisulfite sequencing data. *PLoS Genetics*.

Lindström J (1999) Early development and fitness in birds and mammals. *Trends in Ecology & Evolution*, **14**, 343–348.

Lister R, Pelizzola M, Dowen R, Hawkins R (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*.

Lo KA, Labadorf A, Kennedy NJ *et al.* (2013) Analysis of In Vitro Insulin-Resistance Models and Their Physiological Relevance to InVivo Diet-Induced Adipose Insulin Resistance. *Cell Reports*, **5**, 259–270.

Marr AK, MacIsaac JL, Jiang R *et al.* (2014) Leishmania donovani Infection Causes Distinct Epigenetic DNA Methylation Changes in Host Macrophages. *PLoS Pathogens*, **10**, e1004419.

McErlean P, Jr SF, Costa FF *et al.* (2014) Human rhinovirus infection causes different DNA methylation changes in nasal epithelial cells from healthy and asthmatic subjects. *BMC Medical Genomics*, **7**, 37.

McGuinness D, McGlynn LM, Johnson PCD *et al.* (2012) Socio-economic status is associated with epigenetic differences in the pSoBid cohort. *International Journal of Epidemiology*, **41**, 151–60.

Meaney M, Szyf M (2005) Environmental programming of stress responses through DNA methylation: life at the interface between a dynamic environment and a fixed genome. *Dialogues in Clinical Neuroscience*, **3**, 103–123.

Meissner A, Mikkelsen TS, Gu H *et al.* (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, **454**, 766–70.

Moleres A, Campión J, Milagro FI *et al.* (2013) Differential DNA methylation patterns between high and low responders to a weight loss intervention in overweight or obese adolescents: the EVASYON study. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, **27**, 2504–12.

Muruthi P, Altmann J, Altmann S (1991) Resource base, parity, and reproductive condition affect females' feeding time and nutrient intake within and between groups of a baboon population. *Oecologia*, 467–472.

Nussey DH, Kruuk LEB, Morris A, Clutton-Brock TH (2007) Environmental conditions in early life influence ageing rates in a wild population of red deer. *Current Biology*, **17**, R1000–1.

Ogata H, Goto S, Sato K *et al.* (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, **27**, 29–34.

Pacis A, Tailleux L, Lambourne J *et al.* (2015) Bacterial Infection Remodels the DNA Methylation Landscape of Human Dendritic Cells. *Genome Research*. doi: 10.1101/gr.192005.115.

Pusey A, Packer C (1987) Dispersal and philopatry. In: *Primate Societies* (eds Smuts B, Cheney D, Seyfarth RM, Wrangham RW, Struhsaker T), pp. 150–166. University of Chicago Press, Chicago, IL.

Reinius LE, Acevedo N, Joerink M *et al.* (2012) Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PloS One*, **7**, e41361.

Revitali M, Ima M, Eserve P, Elt D (2009) Population dynamics of two sympatric rodents in a variable environmental rainfall resource availability and predation. *Ecology*, **90**,

1996–2006.

Roadmap Epigenomics Consortium, Kundaje A, Meuleman W *et al.* (2015) Integrative analysis of 111 reference human epigenomes. *Nature*, **518**, 317–330.

Rönn T, Volkov P, Davegårdh C *et al.* (2013) A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue. *PLoS Genetics*, **9**, e1003572.

Schulz LC (2010) The Dutch Hunger Winter and the developmental origins of health and disease. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 16757–8.

Scuteri A, Sanna S, Chen WM *et al.* (2007) Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genetics*, **3**, 1200–1210.

Sinclair K, Allegrucci C (2007) DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19351–19356.

Soranzo N, Spector TD, Mangino M *et al.* (2009) A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nature Genetics*, **41**, 1182–1190.

Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 9440–5.

Tavil Y, Sen N, Yazici HU *et al.* (2007) Mean platelet volume in patients with metabolic syndrome and its relationship with coronary artery disease. *Thrombosis Research*, **120**, 245–250.

Tobi EW, Goeman JJ, Monajemi R *et al.* (2014) DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nature Communications*, **5**, 1–13.

Tobi EW, Lumey LH, Talens RP *et al.* (2009) DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Human Molecular Genetics*, **18**, 4046–53.

Tung J, Akinyi MY, Mutura S *et al.* (2011) Allele-specific gene expression in a wild nonhuman primate population. *Molecular Ecology*, **20**, 725–39.

Tung J, Barreiro LB, Johnson ZP *et al.* (2012) Social environment is associated with

gene regulatory variation in the rhesus macaque immune system. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 6490–5.

Tung J, Primus A, Bouley AJ *et al.* (2009) Evolution of a malaria resistance gene in wild primates. *Nature*, **460**, 388–91.

Tung J, Zhou X, Alberts SC, Stephens M, Gilad Y (2015) The genetic architecture of gene expression levels in wild baboons. *eLife*, **4**, 1–22.

Van den Veyver IB (2002) Genetic effects of methylation diets. *Annual Review of Nutrition*, **22**, 255–82.

Waterland R a, Kellermayer R, Laritsky E *et al.* (2010) Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genetics*, **6**, e1001252.

Weaver ICG, Cervoni N, Champagne F a *et al.* (2004) Epigenetic programming by maternal behavior. *Nature Neuroscience*, **7**, 847–54.

Webb BA, Forouhar F, Szu F *et al.* (2015) Structures of human phosphofructokinase-1 and atomic basis of cancer-associatedmutations. *Nature*.

Weber M, Hellmann I, Stadler MB *et al.* (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genetics*, **39**, 457–66.

Whitney O, Pfenning AR, Howard JT *et al.* (2014) Core and region-enriched networks of behaviorally regulated genes and the singing genome. *Science*, **27**.

Wolff GL, Kodell RL, Moore SR, Cooney C (1998) Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. *American Societies for Experimental Biology*, **12**, 949–57.

Xi Y, Li W (2009) BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics*, **10**, 232.

Yamini S, Carswell N, Michaelis OE, Szepesi B (1992) Adaptation in enzyme (metabolic) pathways to obesity, carbohydrate diet and to the occurrence of NIDDM in male and female SHR/N-cp rats. *International Journal of Obesity and Related Metabolic Disorders*, **16**, 765—774.

Zhou X, Cain C, Myrthil M *et al.* (2014) Epigenetic Modifications are Associated with Inter-species Gene Expression Variation in Primates. *Genome Biology*, **15**, 1–19.

Figure 1. Study design. The correspondence between resource base and all study subjects is depicted here. Our main differential methylation analyses focused on individuals that spent their entire life (from birth until the time of sampling) in a wild-feeding social group (n = 39: white circle) or in the Lodge group (n = 22: gray circle). Our analyses of the stability versus plasticity of DNA methylation levels focused on individuals that switched resource base after natal dispersal (n = 8: switching baboons in the center).

Figure 2. **Resource base influenced genome-wide DNA methylation levels**. QQ-plot comparing the cumulative distribution of p-values from a uniform distribution against the cumulative distribution of (i) p-values generated from our main model, which tests for effects of resource base on site-specific DNA methylation levels (plotted in light and dark blue); and (ii) p-values generated from the same model when resource base values (Lodge or wild-feeding) were permuted. If resource base did not affect DNA methylation levels, we would expect the p-values associated with our main model (blue points) to follow a uniform distribution and to fall along the x=y line. The observed deviation from the x=y line suggests a pervasive effect of resource base on DNA methylation levels in the real data. Note that deviation from the x=y line at higher p-values (beginning around p=0.1) suggests that more sites are affected by resource base than we could detect in our sample size of 61 individuals, but do not reach genome-wide significance (thus, our non-significant set of sites is a mix of true positives and true negatives). In contrast, when resource base values were permuted across individuals, the resulting p-values were roughly uniformly distributed, as expected, and fell along the x=y line.

Figure 3. Sites affected by resource base were enriched in functionally important regions of the genome. (A) Sites associated with resource base were more likely to occur within 10 kb of genes expressed in whole blood, in regions homologous to H3K4me1-marked enhancers in humans, and in gene promoters. They were significantly under-represented in regions of the genome with no known functional role. (B) Sites associated with resource base were also non-randomly distributed across chromatin states. “Active TSS” and “Flanking Active TSS” reflect the transcription start sites of actively expressed genes; “Strong transcription” reflects gene bodies of highly expressed genes; “Enhancers” reflects regulatory elements that interact with promoters of expressed genes; and “Repressed Polycomb” reflects the promoter and gene bodies of silenced/unexpressed genes. In both (A) and (B), only regions/states with significant over/underenrichment are shown (see Figure S3 for the full set of regions/states) (C) Histone marks associated with each of the chromatin states presented in panel B (based on Roadmap Epigenomics data (Roadmap Epigenomics Consortium *et al.* 2015)). Each chromatin state is defined by the presence (blue square) or absence (white square) of the histone modifications shown below.

Figure 4. Wild-feeding baboons exhibited consistently higher levels of DNA methylation at the phosphofructokinase (*PFKP*) promoter, where methylation suppresses gene expression in reporter gene assays. (A) The magnitude and

direction of the effect of resource base on DNA methylation levels are plotted for all sites tested within 1.5 kb of the *PFKP* transcription start site. Sites with evidence for an effect of resource base at a nominal p-value of 0.05 are shown as green lines with open green dots, and at a 10% FDR as green lines with filled red dots. All other sites are shown as gray lines. The first exon is denoted by a blue box, and the translation start site is denoted with a black asterisk. (B) Firefly luciferase expression levels (normalized by renilla luciferase expression levels) are plotted for 4 replicates per condition. Results from Wilcoxon signed rank tests are shown as follows: **p=0.014 for comparison between no CpGs methylated and all CpGs methylated; *p=0.057 comparison between no CpGs methylated and some CpGs methylated.

Figure 5. Individuals that switched resource base more closely resembled lifelong wild-feeding individuals, regardless of the direction of the switch. (A-B) Predicted and (C) observed results when an SVM classifier that distinguished between wild-feeding and Lodge individuals was applied to DNA methylation data from individuals that switched resource base at natal dispersal (for all plots, grey background = predicted Lodge group, white background = predicted wild-feeding). We hypothesized that DNA methylation patterns in switching individuals would consistently (A) resemble their early life group mates or (B) resemble their current group mates. However, we observed (C) that regardless of the resource base history of switching individuals, they consistently resembled lifelong wild-feeding individuals (represented below as an SVM prediction value below 0). Results in C are shown as boxplots (distributions of predicted values for each individual) because we randomly subsampled our data (50 separate subsamples) to create balanced training sets before predicting the resource-base of switching individuals.