# Assessment of DNA methylation patterns in the bone and cartilage of a nonhuman primate model of osteoarthritis

Genevieve Housman[1,2*], Lorena M. Havill[3], Ellen E. Quillen[4], Anthony G. Comuzzie[4], Anne C. Stone[1,2]

[1] School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA.

[2] Center for Evolution and Medicine, Arizona State University, Tempe, AZ, USA.

[3] Southwest National Primate Research Center, Texas Biomedical Research Institute, San Antonio, TX, USA.

[4] Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX, USA.

* Corresponding author

Email: ghousman@asu.edu (G.H.)

Address: University of Chicago, 920 E. 58th St., CLSC 317, Chicago, IL 60637, USA

**Abbreviated Title:** Skeletal DNA methylation in primate OA

**Funding:** NIH Grant P01 HL028972 to A.G.C; Max and Minnie Tomerlin Voelcker Foundation to L.M.H.; William and Ella Owens Foundation for Biomedical Research to L.M.H.; SNPRC Internship Funds to G.H.; ASU Chapter of Sigma Xi Grant-in-Aid of Research to G.H.

**Abstract**

***Objective:*** Osteoarthritis (OA) impacts humans and several other animals. Thus, the mechanisms underlying this disorder, such as specific skeletal tissue DNA methylation patterns, may be evolutionary conserved. However, associations between methylation and OA have not been readily studied in nonhuman animals. Baboons serve as important models of disease and develop OA at rates similar to those in humans. Therefore, this study investigated the associations between methylation and OA in baboons to advance the evolutionary understanding of OA.

***Design:*** Trabecular bone and cartilage was collected from the medial condyles of adult female baboon femora, five with and five without knee OA. The Infinium HumanMethylation450 BeadChip (450K array) was used to identify DNA methylation patterns in these tissues.

***Results:*** Approximately 44% of the 450K array probes reliably align to the baboon genome, contain a CpG site of interest, and maintain a wide distribution throughout the genome. Of the two filtering methods tested, both identified significantly differentially methylated positions (DMPs) between healthy and OA individuals in cartilage tissues, and some of these patterns overlap with those previously identified in humans. Conversely, no DMPs were found between tissue types or between disease states in bone tissues.

***Conclusions:*** Overall, the 450K array can be used to measure genome-wide DNA methylation in baboon tissues and identify significant associations with complex traits. The results of this study indicate that some DNA methylation patterns associated with OA are evolutionarily conserved, while others are not. This warrants further investigation in a larger and more phylogenetically diverse sample set.

**Key Words:** osteoarthritis; animal model; articular cartilage; bone; DNA methylation

## Introduction

Osteoarthritis (OA) is a complex degenerative joint disease, and OA of the knee is one of the leading causes of disability across the globe.[1] Thus, research endeavors to describe the molecular mechanisms that contribute to this disorder are underway. Both genetic and environmental factors have some effect.[2,3] However, epigenetic factors, such as DNA methylation which regulates gene expression, are now thought to play a more influential role in the development of degenerative skeletal disorders like OA.[4–10] Animal models, such as mice, rats, rabbits, guinea pigs, dogs, sheep, goats, and horses, have been essential in discerning some of the processes inherent to OA development.[11,12] Nevertheless, all of these animals are limited in their abilities to fully inform our understanding of human OA, so the search to find a gold standard animal model for OA is still ongoing.[13] Lastly, although variation in skeletal tissue DNA methylation patterns are thought to be involved in the development and progression of OA, this epigenetic mechanism has not been readily studied in animal models because assays to assess variation in this regulatory level have not been optimized.

Nonhuman primates can serve as important models of disease for humans because they are phylogenetically close to humans. Baboons (*Papio sp.*) are a particularly good model of disease, especially OA,[14] as they naturally develop OA at rates similar to those observed in humans.[3,14] Additionally, because of their evolutionary proximity to humans, further investigation of the molecular processes innate to OA development and progression in baboons as compared to these mechanisms in humans will advance the evolutionary understanding of this disease. Lastly, the relative genetic conservation between baboons and humans makes the optimization and use of standardized DNA methylation assays possible. Specifically, the Infinium

HumanMethylation450 BeadChip (450K array), which is a cost-efficient application for assessing genome-wide DNA methylation patterns in humans, has been successfully used for some nonhuman primate species. These and other nonhuman primate DNA methylation studies have primarily used DNA extracted from blood or other soft tissues.[15–20] However, this technique has not yet been used to study DNA methylation variation in baboon skeletal tissues or how it relates to the development of OA in a nonhuman primate species.

For this study, we used the 450K array to identify DNA methylation patterns in femur bone and cartilage of age-matched female baboons, five with and five without knee OA, in order to validate that this technique can be used for nonhuman primate skeletal tissue DNA extracts and to determine if DNA methylation variation is associated with OA in baboons and in a manner similar to that observed in humans.

## Methods

### *Ethics Statement*

Nonhuman primate tissue samples included were opportunistically collected at routine necropsy of these animals. No animals were sacrificed for this study, and no living animals were used in this study.

### *Samples*

Baboon (*Papio sp.*) samples come from captive colonies at the Southwest National Primate Research Center in Texas. These samples are ideal because many environmental factors that influence skeletal development and maintenance (e.g., diet and exposure to sunlight which

influences vitamin D production) are controlled and consistent across individuals. Femora were opportunistically collected at routine necropsy of these animals and stored in -20°C freezers at the Texas Biomedical Research Institute after dissection. These preparation and storage conditions ensured the preservation of skeletal DNA methylation patterns. Samples include skeletally healthy adult baboons (n=5) and adult baboons with severe OA (n=5). Age ranges are comparable between each group (19.30±1.70 years and 19.24±1.73 years, respectively), and only females were included in this study.

***Assessment of Osteoarthritis***

Classification of adult baboons as having healthy or OA knees was determined through visual examination of the distal femora and macroscopic inspection of the distal articular surface cartilage. Each specimen was assigned an OA severity score. Briefly, Grade 1 is unaffected, Grade 2 is mild OA as indicated by cartilage fibrillation, Grade 3 is moderate OA as indicated by cartilage lesions, and Grade 4 is advanced OA as indicated by eburnation.[3] From this, binary classifications were made such that all healthy adult baboons have 100% Grade 1 on one or both distal femora, and all OA adult baboons have a variable percentage of Grades 3 or 4 on one or both distal femora (Figure 1).

***DNA Extraction***

From the distal femoral condyles, cartilage scrapings were collected using scalpels and processed with a homogenizer, and trabecular bone samples from the interior of the condyle were obtained using a Dremel and pulverized into bone dust using a BioPulverizer. Tissues were collected from this region because OA legions in both humans and baboons generally appear first on the

cartilage of the medial condyle. Both tissues are included in this project because human skeletal epigenetic studies are based on trabecular bone and cartilage, and it is important to standardize tissue type for comparative purposes. DNA was extracted from these processed tissues using a phenol-chloroform protocol optimized for skeletal tissues[21] and quantified using both Nanodrop and Qubit machines (Table S1).

***Genome-Wide DNA Methylation Profiling***

Genome-wide DNA methylation was assessed using Infinium HumanMethylation450 BeadChips (450K array). These arrays analyze the methylation status of over 485,000 sites throughout the genome, covering 99% of RefSeq genes and 96% of the UCSC-defined CpG islands and their flanking regions. For each sample, approximately 500ng of genomic DNA (Table S1) was bisulfite converted using the EZ DNA Methylation™ Gold Kit, according to the manufacturer's instructions (Zymo Research), with modifications described in the Infinium Methylation Assay Protocol. Following manufacturer guidelines (Illumina), this processed DNA was then whole-genome amplified, enzymatically fragmented, hybridized to the arrays, and imaged using the Illumina iScan system. The array data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE101733.

***Processing of Methylation Data***

Raw fluorescent data were normalized to account for the noise inherent within and between the arrays themselves. Specifically, we performed a normal-exponential out-of-band (Noob) background correction method with dye-bias normalization[22] to adjust for background

fluorescence and dye-based biases and followed this with a between-array normalization method (functional normalization)[23] which removes unwanted variation by regressing out variability explained by the control probes present on the array as implemented in the minfi v1.20.2 package in R v3.3.1[24] which is part of the Bioconductor project.[25] This method has been found to outperform other existing approaches for studies that compare conditions with known large-scale differences,[23] such as those assessed in this study.

After normalization, methylation values (β values) for each site were calculated as the ratio of methylated probe signal intensity to the sum of both methylated and unmethylated probe signal intensities. These β values range from 0 to 1 and represent the average methylation levels at each site across the entire population of cells from which DNA was extracted (0 = completely unmethylated sites, 1 = fully methylated sites). Every β value in the Infinium platform is accompanied by a detection p-value, and those with failed detection levels (p-value > 0.05) in greater than 10% of samples were removed from downstream analyses.

The probes on the arrays were designed to hybridize specifically with human DNA, so our use of nonhuman primate DNA required that probes non-specific to the baboon genome, which could produce biased methylation measurements, be computationally filtered out and excluded from downstream analyses. This was accomplished using two different methods modified from.[16,19]

For both methods, we used blastn[26] to map the 485,512 50bp probes onto the *Papio anubis* genome (Assembly: Panu_2.0, Accession: GCF_000264685.2) using an e-value threshold of $e^{-10}$. We retained probes that successfully mapped to the baboon genome, had only 1 unique BLAST

hit, and targeted CpG sites (Supplemental Probe Annotation File). Then, for the first method, which used criteria based on sequence alignment, we only retained probes that had 0 mismatches in the 5bp closest to and including the CpG site and that had 0-2 mismatches in 45bp not including the CpG site.[16] This filtering retained 133,264 probes. For the second method, which used criteria based on gene symbol similarities, we identified the closest baboon gene to each probe site and checked for corresponding gene name matches between humans and baboons.[19] For baboons, this information was obtained from GFF and Ensembl BioMart data. Only those probes with partial or complete gene matches were retained. This filtering retained 130,307 probes.

Additionally, β values associated with cross-reactive probes,[27] probes containing SNPs at the CpG site, probes detecting SNP information, probes detecting methylation at non-CpG sites, and probes targeting sites within the sex chromosomes were removed using the minfi v1.20.2 package in R v3.3.1.[24] This filtering retained a final set of 120,305 probes for the alignment filter criteria and 112,760 probes for the gene symbol criteria (Figure 2).

***Statistical Analysis of Differential Methylation***

Because β values have high heteroscedasticity, they are not statistically valid for use in differential methylation analyses.[28] Thus, M values were calculated as the log transformed ratio of methylated signal to unmethylated signal and used in these analyses instead.

In order to identify sites that were significantly differentially methylated across comparative groups, we designed and tested generalized linear mixed models (GLMMs) which related the

variables of interest to the DNA methylation patterns for each site, while accounting for latent variables.[29] Sites found to have significant associations were classified as significant differentially methylated positions (DMPs). Specifically, a GLMM was used to estimate differences in methylation levels for each of the following contrasts: (1) between bone and cartilage in OA baboons, (2) between bone and cartilage in healthy baboons, (3) between OA and healthy baboon bone, (4) between OA and healthy baboon cartilage, (5) among all 4 combinations of tissue type and disease state (healthy bone vs. healthy cartilage vs. OA bone vs. OA cartilage).

Additional variables included in this GLMM were unknown latent variables calculated using the iteratively re-weighted least squares approach in the sva v3.22.0 package in R v3.3.1.[30,31] The 4 latent variables estimated were included to help mitigate any unknown batch and cell heterogeneity effects on methylation variation at each site. No predefined batch effects for the arrays were included because these did not appear to have large effects on the data (Figure S1).

Alternative methods to account for cell heterogeneity exist, but they are specific to whole blood,[30,32] require reference epigenetic data, or are reference free methods[33] that are comparable to the sva method.[34] Out of the known cell types in skeletal tissues, only chondrocytes and osteoblasts have reference epigenomes available on the International Human Epigenomics Consortium, and these are only for humans, not nonhuman primates. Thus, because no standard method is available to correct for the heterogenous cell structure in nonhuman primate skeletal tissue, we chose the described sva method.

The GLMM design matrix was fit to the M value array data by generalized least squares using the limma v3.30.13 package in R v3.3.1,[25,35] and the estimated coefficients and standard errors for the defined tissue type and disease status contrasts were computed. Lastly, for each coefficient, an empirical Bayes approach[36] was used to compute moderated t-statistics, log-odds ratios of differential methylation, and associated p-values adjusted for multiple testing.[37] Significant DMPs for the effect of tissue type and disease status contrasts were defined as those having log fold changes in M values corresponding to an adjusted p-value of less than 0.05.

## Results

The aim of this study was to use the 450K array to identify DNA methylation patterns in femur bone and cartilage of age-matched female baboons, five with and five without knee OA. In order to do this, we first assessed the effectiveness of the 450K array in identifying DNA methylation patterns in baboon DNA and of different probe filtering methods.

### *Alignment of 450K Array Probes with the Baboon Genome*

Probes from the 450K array were aligned to the baboon genome (Supplemental Probe Annotation File).[16,19] Out of the 485,512 50bp probes on the array, 213,858 probes (44%) map to the baboon genome with e-values less than $e^{-10}$, have only unique BLAST hits, and target a CpG site (Figure 3A). Out of these reliably mapped probes, 133,264 probes (62%) were retained after the alignment filter criteria (Figure 3B), while 130,307 probes (61%) were retained after the gene symbol filter criteria (Figure 3C). 83,142 probes overlapped between both filtering methods (62% for the alignment filter criteria and 64% for the gene symbol filter criteria, Figure S2).

Probes that reliably mapped to the baboon genome, that met the alignment filter criteria, or that met the gene symbol criteria covered approximately 18,800 genes with an average coverage of 9, 6, or 8 probes per gene, respectively (Table S2). Additionally, the retained probes covered a range of locations with respect to genes and CpG islands (Table S2), indicating that these filtered probes maintain a wide distribution throughout the genome.

***Effectiveness of 450K Array Probes using Baboon DNA***

To determine how effectively the 450K array probes measured DNA methylation in baboon DNA, we performed Spearman correlation tests between the hybridization efficiency of each probe and parameters defining the alignment quality of each probe to the baboon genome. Specifically, both probe alignment bitscores and percent identity were significantly negatively correlated with probe hybridization efficiency, and probe alignment e-values were significantly positively correlated with probe hybridization efficiency, regardless of filtering criteria (Table S3).

However, filtering probes using the alignment filter criteria retained proportionally more successfully hybridized probes than filtering probes using the gene symbol filter criteria (Figure S3). Thus, filtering probes using the alignment filter criteria likely produces more reliable results.

***Differential Methylation and Osteoarthritis***

Significant DMPs were only identified between healthy and OA individuals in cartilage tissues (Table S4). All of these DMPs displayed decreased methylation in OA cartilage samples as compared to healthy cartilage samples, and some of these patterns overlapped with those

previously identified in humans. Conversely, no DMPs were found between tissue types or between disease states in bone tissues.

When filtering probes using the alignment filer criteria, 6 significant DMPs were identified between OA cartilage samples and healthy cartilage samples, while only 2 DMPs were identified when filtering probes using the gene symbol criteria (Table 1). One locus, *RFXAP*, matched between these filtering methods. *RUNX1* has previously been found to be differentially methylated in OA and healthy cartilage in humans, with OA cartilage having lower methylation as compared to healthy cartilage.[5] The other genes associated with these probes (*KLHL26*, *RFXAP*, *MIR497*, *MIR195*, *ELF1*, *ACSL1*, and *CMIP*) have not previously been associated with OA in humans.[4–10]

**Discussion**

Here we used the 450K array to identify DNA methylation variation in bone and cartilage tissues from a baboon model of OA. This was done both to determine the effectiveness of this application for baboon DNA and to assess the evolutionary conservation of epigenetic-OA associations in the primate lineage.

We show that using the 450K array is feasible in baboon tissues. *In silico* probe filtering methods[16,19] indicated that 44% of all human probes could be reliably mapped to the baboon genome and contained an informative CpG. This number was lower than expected since previous researchers were able to reliably use these same methods to map 61% of the human probes to the Cynomologus macaque genome,[19] another Old World monkey that is a close phylogenetic

relative to baboons. This discrepancy in number may be due to the quality of each nonhuman primates' genome assembly. While both are well annotated, the average scaffold length (88,649,475) and contig length (86,040) of the macaque genome (Assembly: Macaca_fascicularis_5.0, Accession: GCF_000264685.2) are higher than those (528,927 and 40,262) of the baboon genome.

Subsequent *in silico* analyses based on sequence alignment criteria[16] and based on gene symbol criteria[19] retained similar numbers of probes (Figure 3) that maintained wide and comparable distributions throughout the genome (Table S2). However, only a little more than half of the resulting probes for each filtering technique overlapped with one another (Figure S2). This discrepancy is likely due to the incomplete nature of the baboon genome annotation. More than half of the probes that fit the alignment filter criteria but not the gene symbol criteria (28,699 out of 50,117) are associated with generic gene symbol identifiers (LOC) to indicate the as of yet unknown functions of these regions. Conversely, all of the probes that fit the gene symbol criteria but not the alignment filter criteria have over 3 mismatches with the baboon genome on average and have a maximum of 9 mismatches with the baboon genome. These high mismatch numbers are a potential concern for proper and accurate probe and baboon DNA hybridization.

Fittingly, after applying the 450K array to measure DNA methylation patterns of genomic material extracted from baboon skeletal tissues, we found that the hybridization efficiency of probes was significantly correlated with the alignment quality of each probe to the baboon genome, and thus, the degree of sequence conservation. The majority of filtered probes for both *in silico* methods passed quality controls and produced robust signals on the array, indicating that

either filtering technique may be appropriate for future research. However, because the filtering method based on the alignment filter criteria retained a larger proportion of successfully hybridized probes than the method based on the gene symbol criteria (Figure S3) and because this method is less influenced by the degree of genome assembly annotation, we recommend that this alignment filter criteria method be preferentially used in subsequent nonhuman primate studies.

This work is an extension of previous work using the 450K array to study DNA methylation patterns in the tissues of nonhuman primates.[16,19] The 450K array is advantageous because it is cost efficient per sample and simultaneously measures a large number of CpG loci with a broad genomic representation.[38] Similar to this study, previous researchers have applied the 450K array to measure DNA methylation patterns in great apes,[16] which are closer to humans evolutionarily than baboons, and in macaques,[19] which are comparable in proximity to humans evolutionarily as compared to baboons. In this study, we used a baboon model of OA to assess the evolutionary conservation of epigenetic-OA associations in the primate lineage, and we identified significant DMPs between healthy and OA individuals in cartilage tissues. All of these loci showed hypomethylation in OA cartilage samples as compared to healthy cartilage samples. Six DMPs were identified when using the alignment filter criteria, and two DMPs were identified when using the gene symbol filter criteria (Table 1). All together these loci are associated with 8 genes that have a variety of functions.

Some of these genes have functions directly related to skeletal development and maintenance. For instance, *RUNX1* (Gene ID: 861), also known as runt related transcription factor 1, is

involved in the regulation of bone and cartilage cell development and differentiation.[39] Additionally, *MIR497* (Gene ID: 574456) and *MIR195* (Gene ID: 406971) are non-coding microRNAs that are involved in post-transcriptional regulation.[40] While both of these microRNAs have roles in the development of cancer,[41,42] they also play important regulatory roles in the differentiation of mesenchymal stromal/stem cells into bone related cells.[43]

Other genes have functions associated with the immune system which may have proximal roles in the development of OA. In particular, *RFXAP* (Gene ID: 5994), also known as regulatory factor X associated protein, codes for a protein that assists in the transcriptional activation of major histocompatibility class II genes which are critical for the development and control of the immune system.[44] Additionally, *CMIP* (Gene ID: 80790), also known as c-Maf inducing protein, codes for a protein that is involved in the T-cell signaling pathway, and SNPs within this gene have been associated with chronic diseases like diabetes.[45]

The remaining genes do not have functions related to skeletal phenotypes which makes their involvement in OA less clear. For example, *KLHL26* (Gene ID: 55295), also known as kelch like family member 26, is part of a family of proteins that may be involved in protein ubiquitination[46]. Additionally, *ACSL1* (Gene ID: 2180), also known as acyl-CoA synthetase long-chain family member 1, codes for a protein that assists in the biosynthesis of lipids and degradation of fatty acids, and SNPs within this gene have been associated with chronic diseases like diabetes.[47] Lastly, *ELF1* (Gene ID: 1997), also known as E74 like E26 transformation-specific related transcription factor 1, is an important positive regulator of the Hox cofactor

Myeloid ectropic viral integration site 1 (*MEIS1*) which is involved in developmental processes.[48]

Out of all of these DMPs and their associated genes, *RUNX1* is the only gene that has previously been associated with OA in humans. Specifically, *RUNX1* was found to be differentially methylated in OA and healthy cartilage in humans, with OA cartilage displaying hypomethylation as compared to healthy cartilage.[5] As of yet, none of the remaining DMPs and their associated genes have been identified as candidate loci in human OA studies.[4–10]

Overall, these findings indicate that some DNA methylation patterns associated with OA are evolutionarily conserved between humans and baboons while others are not. Differences between these two species may be due to general speciation events that took place during the evolution of these taxonomic groups, to slight differences in the development or manifestation of OA in these species, or artifacts of the experimental design itself. For instance, the sample size of this study (n=10) is rather small, and all individuals included were female. The small number of individuals likely reduced our power to detect potentially important OA related variants, and the inclusion of only one sex may have biased our results such that identified OA variants are actually female specific variants. Thus, in order to improve the identification of candidate epigenetic alterations that underlie variation in knee OA, a larger sample set that includes both sexes should be considered. Nevertheless, using baboons as a model of OA in this study has begun to clarify the evolutionary conservation of this disorder, and future research in this animal model will help provide insight into the development and progression of OA in order to begin designing preventative and therapeutic agents.[14]

In conclusion, we determined that the 450K array can be used to measure genome-wide DNA methylation in baboon tissues and identify significant associations with complex traits. This is the first study to specifically assess DNA methylation in skeletal tissues from a nonhuman primate using this method. Some methylation variation is associated with genes that impact skeletal development and maintenance, and this may have direct downstream regulatory and phenotypic effects. Additionally, while some DNA methylation patterns associated with OA in baboons appear to be evolutionarily conserved with humans, others do not. These findings warrant further investigation in a larger and more phylogenetically diverse sample set. Lastly, the work presented here begins to advance areas of research that incorporate an animal model of disease and an evolutionary perspective of diseases across phylogenies.

**Acknowledgements**

We thank members of the Department of Genetics at the Texas Biomedical Research Institute for helpful discussions, including Erin Sybout, Sophia Johnson, Mel Carless, Kara Peterson, Laura Cox, Kenneth Lange, Jerry Glenn, & Clint Christensen.

**Author Contributions**

All authors were involved in drafting or critical revision of this article, and all authors approved the final version to be published. G.H. (ghousman@asu.edu) had full access to the study data and takes responsibility for the integrity of these data and the accuracy of their analysis.

Study conception and design: G.H., L.M.H., E.Q., A.C.S.

Acquisition of data: G.H., L.M.H., A.G.C.

Analysis and interpretation of data: G.H.

**Declaration of Conflicting Interests**

The authors of this manuscript declare that there are no conflicts of interest.

## References


1. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, Bridgett L, Williams S, Guillemin F, Hill CL, Laslett LL, Jones G, Cicuttini F, Osborne R, Vos T, Buchbinder R, Woolf A, March L. The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis*. 2014;73:1323-1330. doi:10.1136/annrheumdis-2013-204763.
2. Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthritis Cartilage*. 2010;18(1):24-33. doi:10.1016/j.joca.2009.08.010.
3. Macrini TE, Coan HB, Levine SM, Lerma T, Saks CD, Araujo DJ, Bredbenner TL, Coutts RD, Nicolella DP, Havill LM. Reproductive status and sex show strong effects on knee OA in a baboon model. *Osteoarthr Cartil OARS Osteoarthr Res Soc*. 2013;21(6):839-848. doi:10.1016/j.joca.2013.03.003.
4. Delgado-Calle J, Fernández AF, Sainz J, Zarrabeitia MT, Sañudo C, García-Renedo R, Pérez-Núñez MI, García-Ibarbia C, Fraga MF, Riancho JA. Genome-wide profiling of bone reveals differentially methylated regions in osteoporosis and osteoarthritis. *Arthritis Rheum*. 2013;65(1):197-205. doi:10.1002/art.37753.
5. Fernández-Tajes J, Soto-Hermida A, Vázquez-Mosquera ME, Cortés-Pereira E, Mosquera A, Fernández-Moreno M, Oreiro N, Fernández-López C, Fernández JL, Rego-Pérez I, Blanco FJ. Genome-wide DNA methylation analysis of articular chondrocytes reveals a cluster of osteoarthritic patients. *Ann Rheum Dis*. 2014;73(4):668-677. doi:10.1136/annrheumdis-2012-202783.

6. Jeffries MA, Donica M, Baker LW, Stevenson ME, Annan AC, Beth Humphrey M, James JA, Sawalha AH. Genome-Wide DNA Methylation Study Identifies Significant Epigenomic Changes in Osteoarthritic Subchondral Bone and Similarity to Overlying Cartilage. *Arthritis Rheumatol*. 2016;68(6):1403-1414. doi:10.1002/art.39555.

7. Moazedi-Fuerst FC, Hofner M, Gruber G, Weinhaeusel A, Stradner MH, Angerer H, Peischler D, Lohberger B, Glehr M, Leithner A, Sonntagbauer M, Graninger WB. Epigenetic differences in human cartilage between mild and severe OA. *J Orthop Res*. 2014;32(12):1636-1645. doi:10.1002/jor.22722.

8. Rushton MD, Reynard LN, Barter MJ, Refaie R, Rankin KS, Young DA, Loughlin J. Characterization of the Cartilage DNA Methylome in Knee and Hip Osteoarthritis: Methylation Profile of OA Cartilage. *Arthritis Rheumatol*. 2014;66(9):2450-2460. doi:10.1002/art.38713.

9. Alvarez-Garcia O, Fisch KM, Wineinger NE, Akagi R, Saito M, Sasho T, Su AI, Lotz MK. Increased DNA methylation and reduced expression of transcription factors in human osteoarthritis cartilage: Differential DNA methylation in OA cartilage. *Arthritis Rheumatol*. February 2016:n/a-n/a. doi:10.1002/art.39643.

10. Aref-Eshghi E, Yuhua Zhang, Ming Liu, Harper PE, Martin G, Furey A, Green R, Guang Sun, Rahman P, Guangju Zhai, Zhang Y, Liu M, Sun G, Zhai G. Genome-wide DNA methylation study of hip and knee cartilage reveals embryonic organ and skeletal system morphogenesis as major pathways involved in osteoarthritis. *BMC Musculoskelet Disord*. 2015;16(1):1-10. doi:10.1186/s12891-015-0745-5.

11. Kuyinu EL, Narayanan G, Nair LS, Laurencin CT. Animal models of osteoarthritis: classification, update, and measurement of outcomes. *J Orthop Surg*. 2016;11. doi:10.1186/s13018-016-0346-5.

12. Cucchiarini M, de Girolamo L, Filardo G, Oliveira JM, Orth P, Pape D, Reboul P. Basic science of osteoarthritis. *J Exp Orthop*. 2016;3(1). doi:10.1186/s40634-016-0060-6.

13. Ameye LG, Young MF. Animal models of osteoarthritis: lessons learned while seekiseeking the "Holy Grail." *Curr Opin Rheumatol*. 2006;18(5):537–547.

14. Cox LA, Comuzzie AG, Havill LM, Karere GM, Spradling KD, Mahaney MC, Nathanielsz PW, Nicolella DP, Shade RE, Voruganti S, VandeBerg JL. Baboons as a Model to Study Genetics and Epigenetics of Human Disease. *ILAR J*. 2013;54(2):106-121. doi:10.1093/ilar/ilt038.

15. Enard W, Fassbender A, Model F, Adorján P, Pääbo S, Olek A. Differences in DNA methylation patterns between humans and chimpanzees. *Curr Biol*. 2004;14(4):R148–R149.

16. Hernando-Herraez I, Prado-Martinez J, Garg P, Fernandez-Callejo M, Heyn H, Hvilsom C, Navarro A, Esteller M, Sharp AJ, Marques-Bonet T. Dynamics of DNA Methylation in Recent Human and Great Ape Evolution. *PLOS Genet*. 2013;9(9):e1003763. doi:10.1371/journal.pgen.1003763.

17. Martin DIK, Singer M, Dhahbi J, Mao G, Zhang L, Schroth GP, Pachter L, Boffelli D. Phyloepigenomic comparison of great apes reveals a correlation between somatic and germline methylation states. *Genome Res*. 2011;21(12):2049-2057. doi:10.1101/gr.122721.111.

18. Molaro A, Hodges E, Fang F, Song Q, McCombie WR, Hannon GJ, Smith AD. Sperm Methylation Profiles Reveal Features of Epigenetic Inheritance and Evolution in Primates. *Cell*. 2011;146(6):1029-1041. doi:10.1016/j.cell.2011.08.016.

19. Ong M-L, Tan PY, MacIsaac JL, Mah SM, Buschdorf JP, Cheong CY, Stunkel W, Chan L, Gluckman PD, Chng K, Kobor MS, Meaney MJ, Holbrook JD. Infinium Monkeys: Infinium 450K Array for the Cynomolgus macaque ( *Macaca fascicularis* ). *G3amp58 GenesGenomesGenetics*. 2014;4(7):1227-1234. doi:10.1534/g3.114.010967.

20. Pai AA, Bell JT, Marioni JC, Pritchard JK, Gilad Y. A Genome-Wide Study of DNA Methylation Patterns and Gene Expression Levels in Multiple Human and Chimpanzee Tissues. *PLOS Genet*. 2011;7(2):e1001316. doi:10.1371/journal.pgen.1001316.

21. Barnett R, Larson G. A Phenol–Chloroform Protocol for Extracting DNA from Ancient Samples. In: Shapiro B, Hofreiter M, eds. *Ancient DNA*. Methods in Molecular Biology. Humana Press; 2012:13-19. doi:10.1007/978-1-61779-516-9_2.

22. Triche TJ, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Res*. 2013;41(7):e90-e90. doi:10.1093/nar/gkt090.

23. Fortin J-P, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, Greenwood CM, Hansen KD. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol*. 2014;15(11). doi:10.1186/s13059-014-0503-2.

24. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30(10):1363-1369. doi:10.1093/bioinformatics/btu049.

25. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T, Gottardo R, Hahne F, Hansen KD, Irizarry RA, Lawrence M, Love MI, MacDonald J, Obenchain V, Oleś AK, Pagès H, Reyes A, Shannon P, Smyth GK, Tenenbaum D, Waldron L, Morgan M. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods*. 2015;12(2):115-121. doi:10.1038/nmeth.3252.

26. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389-3402.

27. Chen Y, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*. 2013;8(2):203-209. doi:10.4161/epi.23470.

28. Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, Lin SM. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*. 2010;11:587. doi:10.1186/1471-2105-11-587.

29. Maksimovic J, Phipson B, Oshlack A. A cross-package Bioconductor workflow for analysing methylation array data. *F1000Research*. 2016;5:1281. doi:10.12688/f1000research.8839.1.

30. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol*. 2014;15(2):R31. doi:10.1186/gb-2014-15-2-r31.

31. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012;28(6):882-883. doi:10.1093/bioinformatics/bts034.

32. Morris TJ, Beck S. Analysis pipelines and packages for Infinium HumanMethylation450 BeadChip (450k) data. *Methods*. 2015;72:3-8. doi:10.1016/j.ymeth.2014.08.011.

33. Houseman EA, Molitor J, Marsit CJ. Reference-free cell mixture adjustments in analysis of DNA methylation data. *Bioinformatics*. 2014;30(10):1431-1439. doi:10.1093/bioinformatics/btu029.

34. Kaushal A, Zhang H, Karmaus WJ, Wang JS. Which methods to choose to correct cell types in genome-scale blood-derived DNA methylation data? *BMC Bioinformatics*. 2015;16(Suppl 15):P7. doi:10.1186/1471-2105-16-S15-P7.

35. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. January 2015:gkv007. doi:10.1093/nar/gkv007.

36. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*. 2004;3:Article3. doi:10.2202/1544-6115.1027.

37. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol*. 1995;57(1):289-300.

38. Michels KB, Binder AM, Dedeurwaerder S, Epstein CB, Greally JM, Gut I, Houseman EA, Izzi B, Kelsey KT, Meissner A, Milosavljevic A, Siegmund KD, Bock C, Irizarry RA. Recommendations for the design and analysis of epigenome-wide association studies. *Nat Methods*. 2013;10(10):949-955. doi:10.1038/nmeth.2632.

39. Stein GS, Lian JB, van Wijnen AJ, Stein JL, Montecino M, Javed A, Zaidi SK, Young DW, Choi J-Y, Pockwinse SM. Runx2 control of organization, assembly and activity of the

regulatory machinery for skeletal gene expression. *Oncogene*. 2004;23(24):4315-4329. doi:10.1038/sj.onc.1207676.

40. Wei C, Luo Q, Sun X, Li D, Song H, Li X, Song J, Hua K, Fang L. microRNA-497 induces cell apoptosis by negatively regulating Bcl-2 protein expression at the posttranscriptional level in human breast cancer. *Int J Clin Exp Pathol*. 2015;8(7):7729-7739.

41. Li D, Zhao Y, Liu C, Chen X, Qi Y, Jiang Y, Zou C, Zhang X, Liu S, Wang X, Zhao D, Sun Q, Zeng Z, Dress A, Lin MC, Kung H-F, Rui H, Liu L-Z, Mao F, Jiang B-H, Lai L. Analysis of MiR-195 and MiR-497 expression, regulation and role in breast cancer. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2011;17(7):1722-1730. doi:10.1158/1078-0432.CCR-10-1800.

42. Liu L, Chen L, Xu Y, Li R, Du X. microRNA-195 promotes apoptosis and suppresses tumorigenicity of human colorectal cancer cells. *Biochem Biophys Res Commun*. 2010;400(2):236-240. doi:10.1016/j.bbrc.2010.08.046.

43. Almeida MI, Silva AM, Vasconcelos DM, Almeida CR, Caires H, Pinto MT, Calin GA, Santos SG, Barbosa MA. miR-195 in human primary mesenchymal stromal/stem cells regulates proliferation, osteogenesis and paracrine effect on angiogenesis. *Oncotarget*. 2016;7(1):7-22. doi:10.18632/oncotarget.6589.

44. Garvie CW, Boss JM. Assembly of the RFX complex on the MHCII promoter: role of RFXAP and RFXB in relieving autoinhibition of RFX5. *Biochim Biophys Acta*. 2008;1779(12):797-804. doi:10.1016/j.bbagrm.2008.07.012.

45. Dastani Z, Hivert M-F, Timpson N, Perry JRB, Yuan X, Scott RA, Henneman P, Heid IM, Kizer JR, Lyytikäinen L-P, Fuchsberger C, Tanaka T, Morris AP, Small K, Isaacs A, Beekman M, Coassin S, Lohman K, Qi L, Kanoni S, Pankow JS, Uh H-W, Wu Y,

Bidulescu A, Rasmussen-Torvik LJ, Greenwood CMT, Ladouceur M, Grimsby J, Manning AK, Liu C-T, Kooner J, Mooser VE, Vollenweider P, Kapur KA, Chambers J, Wareham NJ, Langenberg C, Frants R, Willems-vanDijk K, Oostra BA, Willems SM, Lamina C, Winkler TW, Psaty BM, Tracy RP, Brody J, Chen I, Viikari J, Kähönen M, Pramstaller PP, Evans DM, St. Pourcain B, Sattar N, Wood AR, Bandinelli S, Carlson OD, Egan JM, Böhringer S, van Heemst D, Kedenko L, Kristiansson K, Nuotio M-L, Loo B-M, Harris T, Garcia M, Kanaya A, Haun M, Klopp N, Wichmann H-E, Deloukas P, Katsareli E, Couper DJ, Duncan BB, Kloppenburg M, Adair LS, Borja JB, Wilson JG, Musani S, Guo X, Johnson T, Semple R, Teslovich TM, Allison MA, Redline S, Buxbaum SG, Mohlke KL, Meulenbelt I, Ballantyne CM, Dedoussis GV, Hu FB, Liu Y, Paulweber B, Spector TD, Slagboom PE, Ferrucci L, Jula A, Perola M, Raitakari O, Florez JC, Salomaa V, Eriksson JG, Frayling TM, Hicks AA, Lehtimäki T, Smith GD, Siscovick DS, Kronenberg F, van Duijn C, Loos RJF, Waterworth DM, Meigs JB, Dupuis J, Richards JB. Novel Loci for Adiponectin Levels and Their Influence on Type 2 Diabetes and Metabolic Traits: A Multi-Ethnic Meta-Analysis of 45,891 Individuals. *PLoS Genet*. 2012;8(3). doi:10.1371/journal.pgen.1002607.

46. Dhanoa BS, Cogliati T, Satish AG, Bruford EA, Friedman JS. Update on the Kelch-like (KLHL) gene family. *Hum Genomics*. 2013;7(1):13. doi:10.1186/1479-7364-7-13.

47. Manichaikul A, Wang X-Q, Zhao W, Wojczynski MK, Siebenthall K, Stamatoyannopoulos JA, Saleheen D, Borecki IB, Reilly MP, Rich SS, Bornfeldt KE. Genetic association of long-chain acyl-CoA synthetase 1 variants with fasting glucose, diabetes, and subclinical atherosclerosis. *J Lipid Res*. 2016;57(3):433-442. doi:10.1194/jlr.M064592.

48. Xiang P, Lo C, Argiropoulos B, Lai CB, Rouhi A, Imren S, Jiang X, Mager D, Humphries RK. Identification of E74-like factor 1 (ELF1) as a transcriptional regulator of the Hox cofactor MEIS1. *Exp Hematol*. 2010;38(9):798-798, 808.e1-2. doi:10.1016/j.exphem.2010.06.006.

Table 1. Significant DMPs between OA cartilage and healthy cartilage. Provided information includes the 450K array probe ID for each DMP (Probe ID), the log fold change in M values (ΔM), the p-values adjusted for multiple testing (P-Value), and additional gene details for each DMP. Results are shown for probes filtered using the alignment criteria and probes filtered using the gene symbol criteria. For all of these DMPs, OA cartilage samples are hypomethylated as compared to healthy cartilage samples.

| | Probe ID | ΔM | P-Value | Human Gene Symbol | Baboon Gene Symbol | Baboon Chromosome | Baboon CpG Position |
|---|---|---|---|---|---|---|---|
| Alignment Filter Probes | cg05295841 | -3.6136 | 0.0247 | KLHL26 | CRTC1 | 19 | 17216922 |
| | cg17890983 | -1.9683 | 0.0247 | RFXAP | RFXAP | 17 | 15599705 |
| | cg02329670 | -1.4981 | 0.0411 | MIR497; MIR195 | LOC103878622 | 16 | 6672282 |
| | cg18456803 | -2.0478 | 0.0411 | ELF1 | WBP4; LOC103879193 | 17 | 19630529 |
| | cg13030790 | -3.5437 | 0.0411 | RUNX1 | RUNX1 | 3 | 11500594 |
| | cg24721647 | -2.1838 | 0.0411 | ACSL1 | ACSL1 | 5 | 173909497 |
| Gene Symbol Filter Probes | cg17890983 | -1.9679 | 0.0374 | RFXAP | RFXAP | 17 | 15599705 |
| | cg04759112 | -1.6320 | 0.0381 | CMIP | CMIP | 20 | 63369656 |

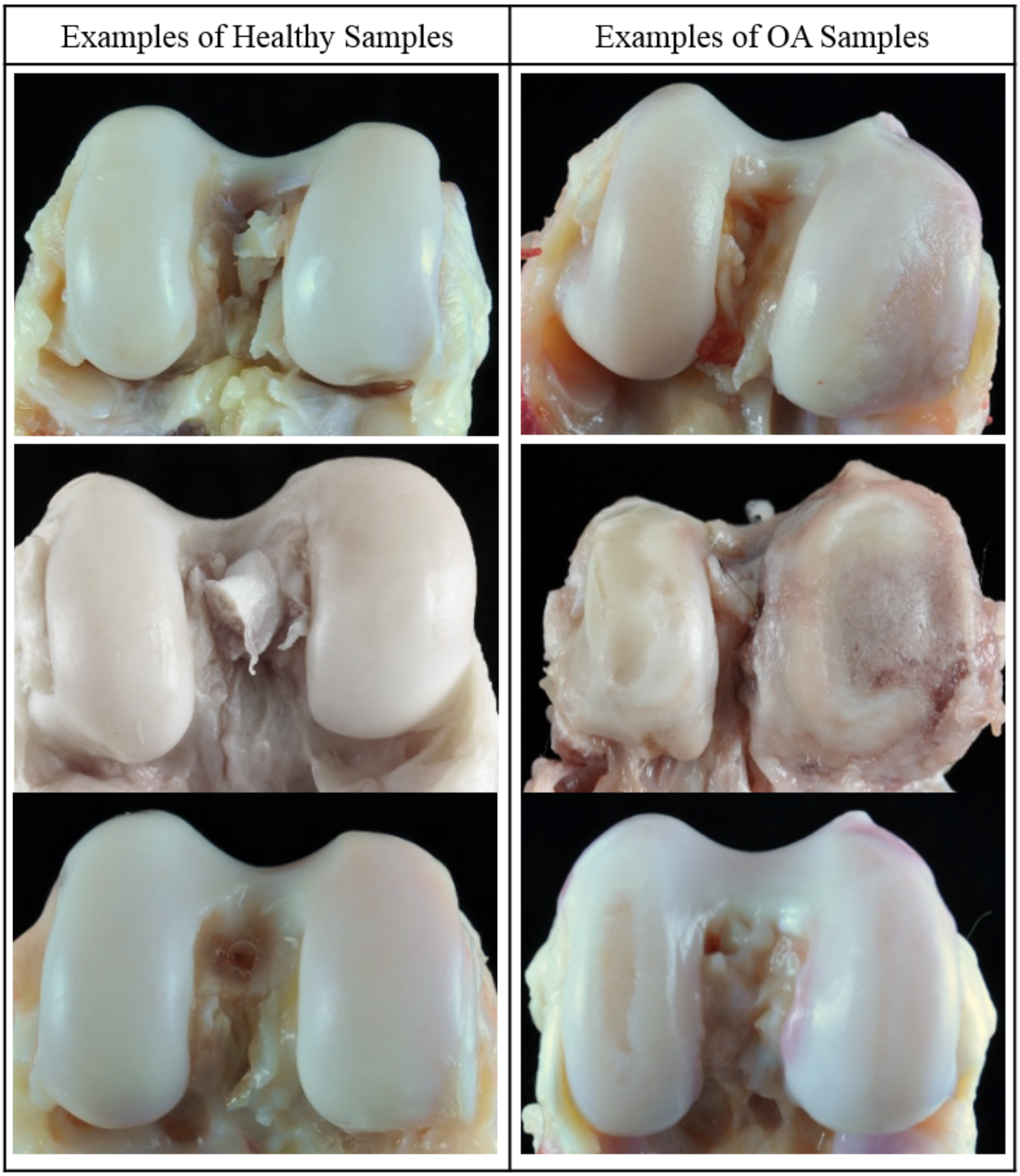


Figure 1. Representative examples of baboon knees (distal femora) that are healthy or have OA.

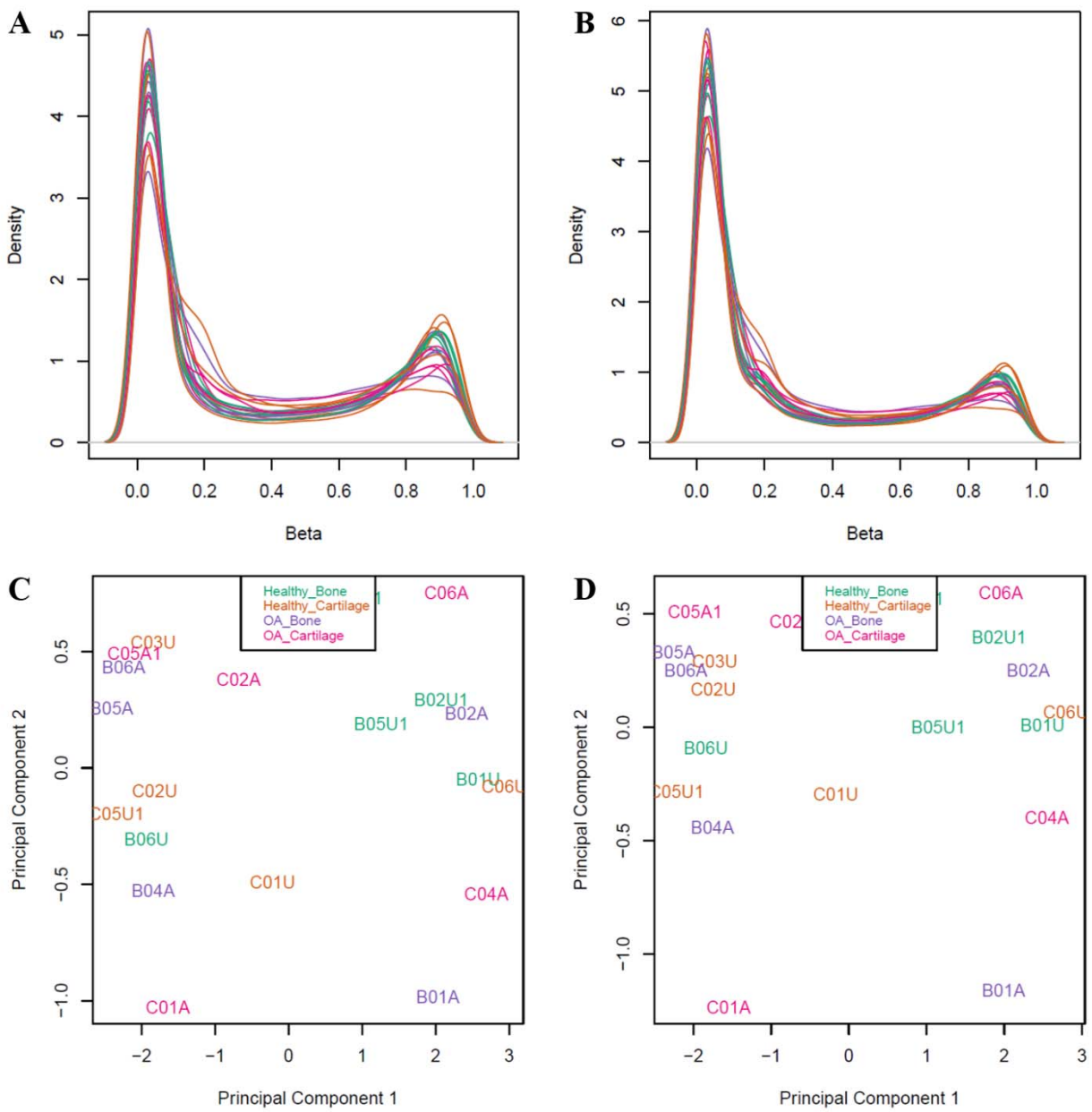


Figure 2. **A** and **B,** Density plots of β values after normalization and probe filtering using the alignment filter criteria **(A)** or the gene symbol filter criteria **(B)**. **C** and **D,** Multidimensional scaling plots showing the first two principle components that describe genome-wide methylation variation after normalization and filtering using the alignment filter criteria **(C)** or the gene symbol filter criteria **(D)**. Each point represents one sample that is either from healthy bone,

healthy cartilage, OA bone, or OA cartilage. In the multidimensional scaling plots, these categories do not form distinct clusters.

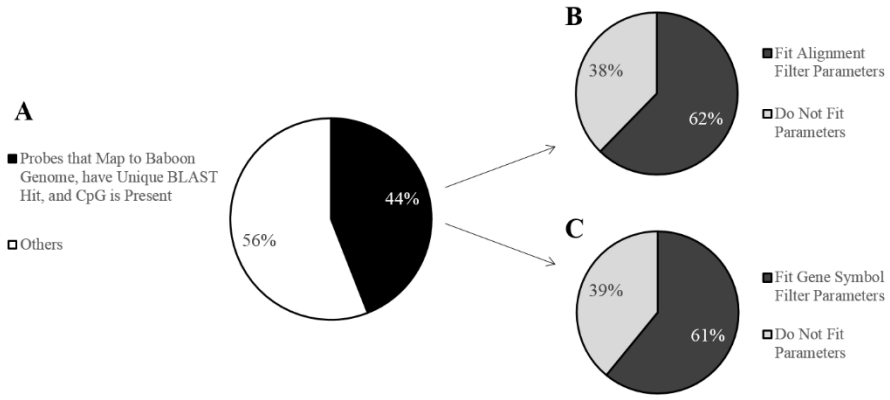


Figure 3. Filtering effects on the 450K array probes for baboons. **A**, Pie chart showing the percent of 450K array probes that map to the baboon (*Papio anubis*) genome with e-values less than $e^{-10}$, have only unique BLAST hits, and target a CpG site. Out of 485,512 probes total, 213,858 probes (44%) meet these criteria. **B,** Pie chart showing the percent of probes, out of those that successfully mapped to the baboon genome, that contain 0 mismatches in 5bp of the probe by and including the targeted CpG site and 0-2 mismatches in 45bp of the probe not including the CpG site. Out of the 213,858 mapped probes, 133,264 probes (62%) meet these criteria. **C,** Pie chart showing the percent of probes, out of those that successfully mapped to the baboon genome, with gene symbol matches to humans. Out of the 213,858 mapped probes, 130,307 probes (61%) meet these criteria.