HHS Public Access

Author manuscript

Curr Opin Genet Dev. Author manuscript; available in PMC 2021 June 13.

Published in final edited form as:

Curr Opin Genet Dev. 2020 June ; 62: 1–7. doi:10.1016/j.gde.2020.04.007.

Prime time for primate functional genomics

Genevieve Housman[1,*], **Yoav Gilad**[1,2]

[1]Section of Genetic Medicine, Department of Medicine, University of Chicago, 5841 S. Maryland Ave., N417, MC6091, Chicago, IL 60637 USA

[2]Department of Human Genetics, University of Chicago, Cummings Life Science Center, 928 E. 58th St., Chicago, IL 60637 USA

Abstract

Functional genomics research is continually improving our understanding of genotype-phenotype relationships in humans, and comparative genomics perspectives can provide additional insight into the evolutionary histories of such relationships. To specifically identify conservation or species-specific divergence in humans, we must look to our closest extant evolutionary relatives. Primate functional genomics research has been steadily advancing and expanding, in spite of several limitations and challenges that this field faces. New technologies and cheaper sequencing provide a unique opportunity to enhance and expand primate comparative studies, and we outline possible paths going forward. The potential human-specific insights that can be gained from primate functional genomics research are substantial, and we propose that now is a prime time to expand such endeavors.

Keywords

primate; functional genomics; comparative genomics; evolution

Determining how the genome and environment interact and contribute to the development and evolution of human phenotypes is critical for advancing the field of medicine and satisfying our general curiosity of who we are and how we got to be here. Certain human traits, such as sickle cell anemia and lactose tolerance, are associated with genetic variants at individual loci. For such traits, the genetic associations almost always reveal likely causal mechanisms. In the cases of sickle cell anemia and lactose tolerance, the mechanisms of causation consist of alterations in the coded protein structure and changes in enhancer DNA sequences, respectively. However, most phenotypes, like heart disease and height, are complex and influenced by hundreds or thousands of genetic loci. Variants associated with complex traits are often located in regulatory, non-coding regions of the genome. Moreover, the effects of individual loci on such traits are typically very small, and the function of each associated locus is often unclear. Together, these features present a challenge to identifying

[*]Corresponding author: ghousman@uchicago.edu, Address: Cummings Life Science Center, 928 E. 58th St., Chicago, IL 60637, Phone: 773-702-8507.

Disclosures

The authors declare no conflict of interest.

the mechanisms that underlie genotype-phenotype associations for complex traits. Collecting additional information about the relationship between genetic variants and the regulation of gene expression can reveal the regulatory pathways through which loci influence the development and evolution of complex traits. The field of functional genomics leads these efforts by characterizing gene regulatory dynamics, and it has radically advanced our understanding of genotype-phenotype relationships.

Efforts to characterize population-level gene regulation in humans have expanded in recent years, resulting in massive repositories of information. For example, the Genotype-Tissue Expression (GTEx) project [1] classifies regions of the human genome that impact how much a gene is expressed, and the Roadmap Epigenomics Mapping Consortium [2] catalogs human epigenomic data. Additionally, the Encyclopedia of DNA Elements (ENCODE) Consortium [3], the Mouse ENCODE Consortium [4], and the model organism ENCODE (modENCODE) Project [5] categorize functional elements in the human genome, mouse genome, and genomes of other model organisms, respectively. These efforts resulted in a wealth of data and insight into gene regulation and regulatory elements, but they are missing key insights because of their inability to classify genetic variants as either highly conserved or specific to humans. The inference of conservation can indicate that a locus has been associated with a critical function throughout evolution, and mutations in conserved loci are likely to have deleterious consequences. In turn, human-specific genetic variants may underlie human-specific phenotypes, including complex traits and diseases that are observed exclusively or more often in humans.

Comparative genomics provides us with tools to identify features that are shared across species and to detect species-specific features. Comparison across closely related species, such as primates, can provide specific information regarding a feature’s evolutionary history, suggesting either neutrality, conservation, or species-specific divergence. Unfortunately, we do not yet have concerted efforts to collect comparative functional genomics data from closely related species. Indeed, GTEx and the Roadmap Epigenomics Mapping Consortium focus solely on human gene regulation, and do not include samples from other species. Conversely, the ENCODE project catalogs data from multiple species, including humans, and uses these data to evaluate conservation across long evolutionary time scales. However, because the evolutionary distances between the examined species (human, mouse, nematode, and fly) are large, ENCODE data can only be used to classify genetic loci as conserved when they have been unchanged for hundreds of millions of years. Further, ENCODE data cannot be used to identify recent changes that have evolved exclusively in the human lineage. Indeed, in order to infer genetic loci in humans that have been conserved on time scales smaller than hundreds of millions of years (for example, in all primates), as well as isolate human-specific genotype-phenotype relationships, it is imperative to pursue functional genomics investigations in some of our closest living relatives. The field of primate functional genomics is not new, but in this commentary, we propose that now is a prime time to expand research endeavors in this area.

Progress in primate functional genomics has been relatively slow. Most comparative functional studies in primates have been limited to the characterization of bulk gene expression and regulatory patterns in a small number of tissues from a small number of

species [6]. The main reason for the slow progress is that tissue samples from nonhuman primates, especially from apes, are difficult to obtain due to ethical and logistical considerations. Further, when samples are accessible, their preservation is often not optimal for a wide range of molecular assays. Hence, while population studies in humans and other model species have taken full advantage of the technological advances that led to cheaper sequencing-based assays [7], comparative genomic studies have lagged.

Despite the many challenges to obtaining primate samples, recent *in vivo* comparative research in human and nonhuman primates has made efforts to include larger samples of primates both within populations [8,9] and among species [10–12]. Expanded sets of tissue types [10,13–17] and tissue developmental stages [18] have also been examined, as have additional levels of gene regulation [19,20]. Research in this area has advanced our understanding of human diseases [13,21,22] and provided insight into the effects of social status and related environmental factors [23–26].

In particular, cercopithecoids such as baboons and macaques are proving to be beneficial models for evaluating many human traits. Several species within these groups naturally develop disorders and symptoms similar to those observed in humans, and these species can be studied using manipulations and treatments that cannot practically or ethically be applied to humans. One compelling examples is that of cardiovascular disease (CVD) [21]. A prominent contributor to CVD is atherosclerosis, or the build-up of fatty plaques in arterial walls. It is known that levels of low-density lipoprotein cholesterol (LDL-C) in plasma and high-cholesterol, high-fat (HCHF) diets can contribute to the development of atherosclerosis, but the mechanisms by which these factors alter cellular function to initiate atherosclerosis is unclear. It is difficult to perform the appropriate study to address this question in humans because one cannot easily perturb the long-term diet of tissue donors while simultaneously controlling for the many environmental confounding factors that may potentially affect the experiment. However, such an experiment is possible in nonhuman primates. Indeed, one comparative functional genomics study recently identified an interaction between HCHF diets and plasma LDL-C concentrations in captive baboons [22]. This interaction causes a change in miRNA expression in peripheral blood mononuclear cells, which is thought to regulate factors involved in the initiation of atherosclerosis lesions [22].

Another compelling example demonstrating the power of primate comparative functional genomics comes from studies in macaques, which have revealed how environmental factors like social experience influence disease susceptibility [24,25]. Like humans, macaques have complex hierarchical social networks. In macaques it is possible to confidently determine the social ranking of each individual. Moreover, one can influence the social ranks of individual monkeys by manipulating certain variables, such as the order in which animals are introduced to their environment. Importantly, this model system allows researchers to study the impact of social rank while controlling for the effects of other typically associated variables, such as access to food or other resources. In this setting, low-ranking macaques have fewer affiliative interactions than high-ranking macaques. The lack of social integration due to low social status is thought to make individuals more susceptible to stressors, like pathogen infections. After manipulating long-term dominance ranks in macaques,

researchers found that social status does alter immune cell gene expression responses to bacterial and viral infection [24]. Moreover, these effects appear to also depend on the social history of individuals, whether they have increased or decreased social status over time [24]. This observation has clear implications for human health management.

Recent progress in *in vitro* cell culture has also substantially enhanced primate functional genomics research. In particular, induced pluripotent stem cell (iPSC) technology allows us to establish renewable sample resources from different primate species and obtain differentiated cells and tissues that are not easily accessible otherwise. This development is particularly important for comparative studies that involve apes, from which sample collection is becoming increasingly difficult (sample collection from chimpanzees, for instance, is no longer allowed in the USA). The availability of a comparative panel of matched human and chimpanzee iPSC lines may ensure that we are able to continue studying humans alongside their most closely related extant evolutionary relative [27]. This comparative panel of iPSCs has opened the door to a new wave of primate functional genomics research. Primate iPSCs have been used in comparative assessments of transposable element regulation [28], chromatin accessibility [29], and chromatin folding [30] – assays that are difficult to perform on non-renewable sets of frozen tissue samples. Primate iPSCs have also been differentiated into other cell types, such as cardiomyocytes and endoderm cells, to examine static gene expression differences across species [31,32], as well as dynamic gene expression responses to environmental perturbations [33]. Outside of this comparative primate cell culture model, additional cell types [34,35] and organoid models [36] in other primate species [37,38] are also being developed and studied.

The potential insights that can be gained through these and additional lines of primate functional genomics research are substantial. Evaluating gene regulation and genotype-phenotype relationships between humans and nonhuman primates not only reveals human-specific variants but also provides a phylogenetic context and evolutionary timeframe for such changes. Comparative studies can reveal the processes and mechanisms by which aspects of human variation have evolved, adapted, or remained conserved. Because genetic variation is greater between species than within species, features that distinguish species from one another typically have larger effect sizes than features that differ between individuals from the same species. Comparative studies can therefore help expose loci of interest for further exploration within species (Figure 1). Thus, by combining comparative genomics and population genetics, primate functional genomics research presents a unique opportunity to identify human-specific variation that may improve our understanding of human traits.

Despite advances in cell culture and functional genomics technology, there are still substantial barriers to data collection, data processing, and data interpretation that impede progress in primate functional genomics. In particular, sample sizes in comparative studies of primates are often small, even when iPSCs are available. Combining data in meta-analyses could remedy this; however, there is currently no standard for collecting or reporting functional genomics data across species. Further, the poor quality of nonhuman primate genome builds makes data processing difficult. In particular, the predominant challenge in analyzing primate functional genomics data is found at the level of genome

annotation. The human genome is the only primate genome with a robust annotation based on direct functional data, and genome annotations for practically all other primates are pieced together using the human annotation. Recent efforts using long-read sequencing have improved genome annotations in apes and macaques [39–41]. Nevertheless, for nonhuman primate genomes more broadly, the locations of genes and coding regions lack functional characterization, and information on other DNA elements such as noncoding RNAs, transcriptional regulatory sites, and regions associated with specific chromatin structures is extremely limited. The sparse annotations of nonhuman primate genomes make the functional interpretation of differential genetic and regulatory variants almost impossible. Moreover, because the quality of the human genome assembly and annotation is substantially higher than that of other primates, initial data processing steps, such as read alignment to reference genomes, may result in biased data filtering that can lead to biased quantitative estimates in downstream analyses.

To rectify these impediments, we must improve nonhuman primate genome assemblies and annotations. This can be accomplished by either obtaining a large number of genomics assays from a small set of species or collecting a limited set of genomics assays from all available species. Based on our experience (but admittedly, also given our scientific interests), our recommendation is for researchers to invest in acquiring detailed annotation information from a small subset of nonhuman primates. Ideally, this set would span the primate phylogeny and include a subset of species relevant for clinical research. A possible starting sample set could include chimpanzees as representative apes and the closest extant evolutionary relatives of humans, macaques and baboons as representative cercopithecoids, marmosets as representative platyrrhines, and mouse lemurs [42] as representative strepsirrhines. Collecting additional annotation information on chimpanzees will be difficult given the moratorium on chimpanzee research, but previously established cell lines, especially iPSCs, as well as previously collected and preserved tissue samples, are still available for this research.

This endeavor to improve nonhuman primates genome annotations could be similar, in a sense, to the ENCODE Project, which was originally purposed with identifying functional elements encoded in human genome sequences [3]. In the first publication of the ENCODE Project, researchers evaluated transcribed regions (RNA-seq), transcription factor binding sites (ChIP-seq, DNase-seq), chromatin structure (histone ChIP-seq, DNase-seq), and DNA methylation (RRBS assay) in several cell types. We propose that similar datasets be collected from a sample of nonhuman primates using systematic and comparable methods. The ENCODE Project initiated data collections by focusing on a subset of cell lines (K562, GM12878, H1 hESC, HeLA-S3, HepG2, and HUVEC) and then expanding to include additional primary cell types. Similarly, primate functional genomics research would benefit from first focusing on robust nonhuman primate cell lines, such as iPSCs and differentiated cells, and then expanding to include additional primary tissues and cell types. Robust iPSC lines reprogramed from chimpanzee cells exist, and efforts should be taken to develop comparable cell lines from other species.

The potential of iPSC technology for comparative genomics is substantial, but one should also be aware of its limitations. Differentiated cells have been shown to be a decent model of

primary tissues, but these *in vitro* models can inadvertently acquire unintended genetic mutations [43][44] and epigenetic modifications [45]. Cells differentiated *in vitro* also tend to show signatures of early development and rarely achieve full maturation [46]. Thus, we believe that the most effective study design is to perform comparative explorations using iPSCs and differentiated cells, followed by validation experiments that make use of primary tissue samples and cell types. By studying nonhuman primate iPSCs alongside primary cell types, a wealth of reference information can be obtained that would enable even further advances in the field of primate functional genomics.

Overall, while the challenges facing continued research in primate functional genomics are substantial, they are not insurmountable. The field of functional genomics as a whole has expanded well beyond what was initially thought to be possible. Similarly, improvements in nonhuman primate data reporting standards, genome sequence quality, and genome annotations will also expand the breadth of comparative primate research that is possible. Given this path forward, now is a prime time for primate functional genomics research.

Acknowledgements

Funding: This work was supported by the National Institutes of Health [HG010772]. We thank Michelle Ward and Natalia Gonzales for their helpful comments on our manuscript.

References

1. The GTEx Consortium: The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science 2015, 348:648–660. [PubMed: 25954001]
2. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, Kellis M, Marra MA, Beaudet AL, Ecker JR, et al.: The NIH Roadmap Epigenomics Mapping Consortium. Nat Biotechnol 2010, 28:1045–1048. [PubMed: 20944595]
3. The ENCODE Project Consortium: An Integrated Encyclopedia of DNA Elements in the Human Genome. Nature 2012, 489:57–74. [PubMed: 22955616]
4. Stamatoyannopoulos JA, Snyder M, Hardison R, Ren B, Gingeras T, Gilbert DM, Groudine M, Bender M, Kaul R, Canfield T, et al.: An encyclopedia of mouse DNA elements (Mouse ENCODE). Genome Biol 2012, 13:418. [PubMed: 22889292]
5. Celniker SE, Dillon LAL, Gerstein MB, Gunsalus KC, Henikoff S, Karpen GH, Kellis M, Lai EC, Lieb JD, MacAlpine DM, et al.: Unlocking the secrets of the genome. Nature 2009, 459:927–930. [PubMed: 19536255]
6. Pai AA, Gilad Y: Comparative studies of gene regulatory mechanisms. Curr Opin Genet Dev 2014, 29:68–74. [PubMed: 25215415]
7. Guigo R, de Hoon M: Recent advances in functional genome analysis. F1000Research 2018, 7.* The authors review recent methodological advances in functional genomics research, including genomics, transcriptomics, chromatin conformation assays, genomic editing, spatially resolved -omics techniques, and single-cell technologies.
8. Vilgalys TP, Rogers J, Jolly CJ, Analysis BG, Mukherjee S, Tung J: Evolution of DNA Methylation in Papio Baboons. Mol Biol Evol 2019, 36:527–540. [PubMed: 30521003] ** Using reduced representation bisulfite sequencing on blood samples from several baboon species, the authors demonstrate that genome-scale DNA methylation profiles reflect genetic and geographic structure. Given the evolutionary history of these baboon species, this study presents unique insight into DNA methylation divergence at both shallow and deep timescales.
9. Jasinska AJ, Zelaya I, Service SK, Peterson CB, Cantor RM, Choi O-W, DeYoung J, Eskin E, Fairbanks LA, Fears S, et al.: Genetic variation and gene expression across multiple tissues and developmental stages in a non-human primate. Nat Genet 2017, 49:1714–1721. [PubMed: 29083405]

10. Housman G, Quillen EE, Stone AC: Intraspecific and interspecific investigations of skeletal DNA methylation and femur morphology in primates. Am J Phys Anthropol 2020, 1–16. [PubMed: 32058585] ** Using methylation arrays on bone samples from several nonhuman primate species, the authors demonstrate that methylation variation is not strongly associated with skeletal morphology variation within species but that between species methylation variation is present and may contribute to evolutionarily distinct phenotypes. This is the first study to assess gene regulation in skeletal tissues from a phylogenetically diverse sample of primates.
11. Sharma AK, Pafčo B, Vlčková K, Červená B, Kreisinger J, Davison S, Beeri K, Fuh T, Leigh SR, Burns MB, et al.: Mapping gastrointestinal gene expression patterns in wild primates and humans via fecal RNA-seq. BMC Genomics 2019, 20:493. [PubMed: 31200636] * The authors validate a fecal RNA-seq method for evaluating gene regulation in gastrointestinal epithelial cells from wild primates. Such techniques could prove useful for studying intestinal health and aspects of foraging ecology in samples previously inaccessible.
12. Berthelot C, Villar D, Horvath JE, Odom DT, Flicek P: Complexity and conservation of regulatory landscapes underlie evolutionary resilience of mammalian gene expression. Nat Ecol Evol 2018, 2:152–163. [PubMed: 29180706] * The authors examine promotor and enhancer activity in combination with gene expression in liver tissue from 15 mammalian species, including primates. With this phylogenetically diverse sample set, they were able to confirm that the number and conservation of surrounding promoter and enhancer activity correlate with evolutionarily stable gene expression.
13. Housman G, Havill LM, Quillen EE, Comuzzie AG, Stone AC: Assessment of DNA Methylation Patterns in the Bone and Cartilage of a Nonhuman Primate Model of Osteoarthritis. Cartilage 2019, 10:335–345. [PubMed: 29457464] * The authors use cartilage and bone samples from baboons with and without knee osteoarthritis (OA) to identify OA-related methylation changes, some of which are conserved with those previously identified in humans with and without OA. Given the similar progression of OA in baboons and humans, this research presents unique insight into the evolutionarily conserved mechanisms driving this disease.
14. Lim W-J, Kim KH, Kim J-Y, Kim H-J, Park J-L, Kim M, Yoon S, Oh J-H, Cho J-W, Kim YS, et al.: Investigation of Gene Expression and DNA Methylation from Seven Different Brain Regions of a Crab-eating Monkey as Determined by RNA-seq and Whole-Genome Bisulfite Sequencing. Front Genet 2019, 10.* The authors characterize gene expression and DNA methylation variation in several brain regions of a macaque and identify brain-related genes that may be regulated by methylation levels.
15. Blake LE, Roux J, Hernando-Herraez I, Banovich NE, Perez RG, Hsiao CJ, Eres I, Cuevas C, Marques-Bonet T, Gilad Y: A comparison of gene expression and DNA methylation patterns across tissues and species. Genome Res 2020, 30:250–262. [PubMed: 31953346] ** Using a robust study design, the authors assay gene expression and methylation levels between liver, kidney, heart, and lung tissues from humans, chimpanzees, and macaques. To date, this is the most comprehensive catalog of gene regulation similarities and differences in these tissues and species.
16. Babbitt CC, Haygood R, Nielsen WJ, Wray GA: Gene expression and adaptive noncoding changes during human evolution. BMC Genomics 2017, 18:435. [PubMed: 28583075]
17. Swain-Lenz D, Berrio A, Safi A, Crawford GE, Wray GA: Comparative Analyses of Chromatin Landscape in White Adipose Tissue Suggest Humans May Have Less Beigeing Potential than Other Primates. Genome Biol Evol 2019, 11:1997–2008. [PubMed: 31233101] * The authors assess chromatin accessibility in human, chimpanzee, and macaque adipose tissue and find human-specific decreases in accessibility of regulatory regions near genes involved in the beigeing of fat. These results suggest a mechanism by which humans evolved relative increases in body fat.
18. Spradling-Reeves KD, Glenn JP, Lange KJ, Kuhn N, Coalson JJ, Nijland MJ, Li C, Nathanielsz PW, Cox LA: The nonhuman primate kidney transcriptome in fetal development. J Med Primatol 2018, 47:157–171. [PubMed: 29603257] * By collecting transcriptomic data from baboon kidneys at different developmental stages, the authors provide a detailed catalog of expression variation in the developing kidney.
19. Proffitt JM, Glenn J, Cesnik AJ, Jadhav A, Shortreed MR, Smith LM, Kavanagh K, Cox LA, Olivier M: Proteomics in non-human primates: utilizing RNA-Seq data to improve protein identification by mass spectrometry in vervet monkeys. BMC Genomics 2017, 18.

20. Pai AA, Baharian G, Sabourin AP, Brinkworth JF, Nédélec Y, Foley JW, Grenier J-C, Siddle KJ, Dumaine A, Yotova V, et al.: Widespread Shortening of 3' Untranslated Regions and Increased Exon Inclusion Are Evolutionarily Conserved Features of Innate Immune Responses to Infection. PLOS Genet 2016, 12:e1006338. [PubMed: 27690314]

21. Cox LA, Olivier M, Spradling-Reeves K, Karere GM, Comuzzie AG, VandeBerg JL: Nonhuman Primates and Translational Research—Cardiovascular Disease. ILAR J 2017, 58:235–250. [PubMed: 28985395]

22. Karere GM, Glenn JP, Birnbaum S, Garcia R, VandeBerg JL, Cox LA: Identification of coordinately regulated microRNA-gene networks that differ in baboons discordant for LDL-cholesterol. PLoS ONE 2019, 14.** By quantifying gene expression in blood cells from baboons with varying levels of plasma LDL-C, the authors determined that the interaction between high-cholesterol, high-fat diets and plasma LDL-C concentrations impact blood cell miRNA expression which in turn regulates atherosclerosis development. Identification of this mechanism provides insight into the progression of human cardiovascular disease.

23. Snyder-Mackler N, Sanz J, Kohn JN, Brinkworth JF, Morrow S, Shaver AO, Grenier J-C, Pique-Regi R, Johnson ZP, Wilson ME, et al.: Social status alters immune regulation and response to infection in macaques. Science 2016, 354:1041–1045. [PubMed: 27885030]

24. Snyder-Mackler N, Sanz J, Kohn JN, Voyles T, Pique-Regi R, Wilson ME, Barreiro LB, Tung J: Social status alters chromatin accessibility and the gene regulatory response to glucocorticoid stimulation in rhesus macaques. Proc Natl Acad Sci 2019, 116:1219–1228. [PubMed: 30538209]

25. Sanz J, Maurizio PL, Snyder-Mackler N, Simons ND, Voyles T, Kohn J, Michopoulos V, Wilson M, Tung J, Barreiro LB: Social history and exposure to pathogen signals modulate social status effects on gene regulation in rhesus macaques. Proc Natl Acad Sci 2019, doi:10.1073/pnas.1820846116.** By experimentally manipulating dominance ranks in macaques, the authors determined that social history influences how current social status affects immune cell gene expression responses to pathogen infection. This study has a unique design which enables the effects of controlled, long-term social structure changes to be quantified and studied.

26. Lea AJ, Akinyi MY, Nyakundi R, Mareri P, Nyundo F, Kariuki T, Alberts SC, Archie EA, Tung J: Dominance rank-associated gene expression is widespread, sex-specific, and a precursor to high social status in wild male baboons. Proc Natl Acad Sci 2018, 115:E12163–E12171. [PubMed: 30538194] * The authors evaluate the relationship between social status and immune gene expression in wild baboons and detect a stronger relationship in males than females. High status-associated gene expression patterns in males also appear to be precursors, not consequences, of high rank.

27. Romero IG, Pavlovic BJ, Hernando-Herraez I, Zhou X, Ward MC, Banovich NE, Kagan CL, Burnett JE, Huang CH, Mitrano A, et al.: A panel of induced pluripotent stem cells from chimpanzees: a resource for comparative functional genomics. eLife 2015, 4:e07103. [PubMed: 26102527]

28. Ward MC, Zhao S, Luo K, Pavlovic BJ, Karimi MM, Stephens M, Gilad Y: Silencing of transposable elements may not be a major driver of regulatory evolution in primate iPSCs. eLife 2018, 7:e33084. [PubMed: 29648536]

29. Romero IG, Gopalakrishnan S, Gilad Y: Widespread conservation of chromatin accessibility patterns and transcription factor binding in human and chimpanzee induced pluripotent stem cells. bioRxiv 2018, doi:10.1101/466631.* Using a panel of human and chimpanzee iPSCs, the authors collect ATAC-seq data to assess differential genome-wide chromatin accessibility and determine that the high degree of conservation between species suggests that pluripotency is extremely stable and possibly subject to stronger evolutionary constraint than other tissues.

30. Eres IE, Luo K, Hsiao CJ, Blake LE, Gilad Y: Reorganization of 3D genome structure may contribute to gene regulatory evolution in primates. PLOS Genet 2019, 15:e1008278. [PubMed: 31323043] * The authors collect Hi-C data from a panel of human and chimpanzee iPSCs to characterize differential 3D chromatin interactions and determine that while lower-order pairwise genomic interactions are conserved higher-order genomic structure like TADs are less conserved.

31. Pavlovic BJ, Blake LE, Roux J, Chavarria C, Gilad Y: A Comparative Assessment of Human and Chimpanzee iPSC-derived Cardiomyocytes with Primary Heart Tissues. Sci Rep 2018, 8:1–14. [PubMed: 29311619] * The authors compared gene expression patterns in primary adult heart

tissue and iPSC-derived cardiomyocytes from humans and chimpanzees to validate the suitability of this cell culture model for dynamic comparative studies.

32. Blake LE, Thomas SM, Blischak JD, Hsiao CJ, Chavarria C, Myrthil M, Gilad Y, Pavlovic BJ: A comparative study of endoderm differentiation in humans and chimpanzees. Genome Biol 2018, 19:162. [PubMed: 30322406]
33. Ward MC, Gilad Y: A generally conserved response to hypoxia in iPSC-derived cardiomyocytes from humans and chimpanzees. eLife 2019, 8:e42374. [PubMed: 30958265] ** Using a panel of human and chimpanzee iPSC-derived cardiomyocytes, the authors determine that the transcriptional response to hypoxia is conserved across species. This study presents a novel research strategy of inducing controlled environmental perturbations on a comparative cell culture model to better examine how gene-by-environment interactions vary across species and contribute to species-specific phenotypes.
34. Advani AS, Chen AY, Babbitt CC: Human fibroblasts display a differential focal adhesion phenotype relative to chimpanzee. Evol Med Public Health 2016, 2016:110. [PubMed: 26971204]
35. Pizzollo J, Nielsen WJ, Shibata Y, Safi A, Crawford GE, Wray GA, Babbitt CC: Comparative Serum Challenges Show Divergent Patterns of Gene Expression and Open Chromatin in Human and Chimpanzee. Genome Biol Evol 2018, 10:826–839. [PubMed: 29608722] * The authors use a serum challenge that mimics physiological stressors associated with cancer progression on human and chimpanzee fibroblasts. Gene expression and chromatin accessibility pathways in humans show stronger affinities with wound healing and cancer development than such pathways in chimpanzees, suggesting possible evolutionary changes that that may make humans more susceptible to cancer.
36. Pollen AA, Bhaduri A, Andrews MG, Nowakowski TJ, Meyerson OS, Mostajo-Radji MA, Di Lullo E, Alvarado B, Bedolli M, Dougherty ML, et al.: Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution. Cell 2019, 176:743–756.e17. [PubMed: 30735633] ** Using cerebral organoids, the authors are able to evaluate gene expression variation between humans, chimpanzees, and macaques. This cell culture model is unique as it enables the examination of brain development in primate samples like chimpanzees from which tissues are inaccessible.
37. Edsall LE, Berrio A, Majoros WH, Swain-Lenz D, Morrow S, Shibata Y, Safi A, Wray GA, Crawford GE, Allen AS: Evaluating Chromatin Accessibility Differences Across Multiple Primate Species Using a Joint Modeling Approach. Genome Biol Evol 2019, 11:3035–3053. [PubMed: 31599933] * The authors present a joint modeling approach for identifying DNase I hypersensitive sites with differential chromatin accessibility between species.
38. Zhang Y, Li T, Preissl S, Amaral ML, Grinstein JD, Farah EN, Destici E, Qiu Y, Hu R, Lee AY, et al.: Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. Nat Genet 2019, 51:1380–1388. [PubMed: 31427791]
39. Gordon D, Huddleston J, Chaisson MJP, Hill CM, Kronenberg ZN, Munson KM, Malig M, Raja A, Fiddes I, Hillier LW, et al.: Long-read sequence assembly of the gorilla genome. Science 2016, 352.
40. Kronenberg ZN, Fiddes IT, Gordon D, Murali S, Cantsilieris S, Meyerson OS, Underwood JG, Nelson BJ, Chaisson MJP, Dougherty ML, et al.: High-resolution comparative analysis of great ape genomes. Science 2018, 360.
41. He Y, Luo X, Zhou B, Hu T, Meng X, Audano PA, Kronenberg ZN, Eichler EE, Jin J, Guo Y, et al.: Long-read assembly of the Chinese rhesus macaque genome and identification of ape-specific structural variants. Nat Commun 2019, 10:1–14. [PubMed: 30602773]
42. Larsen PA, Harris RA, Liu Y, Murali SC, Campbell CR, Brown AD, Sullivan BA, Shelton J, Brown SJ, Raveendran M, et al.: Hybrid de novo genome assembly and centromere characterization of the gray mouse lemur (Microcebus murinus). BMC Biol 2017, 15.
43. Saha K, Jaenisch R: Technical challenges in using human induced pluripotent stem cells to model disease. Cell Stem Cell 2009, 5:584–595. [PubMed: 19951687]
44. Gore A, Li Z, Fung H-L, Young J, Agarwal S, Antosiewicz-Bourget J, Canto I, Giorgetti A, Israel M, Kiskinis E, et al.: Somatic coding mutations in human induced pluripotent stem cells. Nature 2011, 471:63–67. [PubMed: 21368825]

45. Kim K, Doi A, Wen B, Ng K, Zhao R, Cahan P, Kim J, Aryee MJ, Ji H, Ehrlich LIR, et al.: Epigenetic memory in induced pluripotent stem cells. Nature 2010, 467:285–290. [PubMed: 20644535]

46. Karbassi E, Fenix A, Marchiano S, Muraoka N, Nakamura K, Yang X, Murry CE: Cardiomyocyte maturation: advances in knowledge and implications for regenerative medicine. Nat Rev Cardiol 2020, doi:10.1038/s41569-019-0331-x.

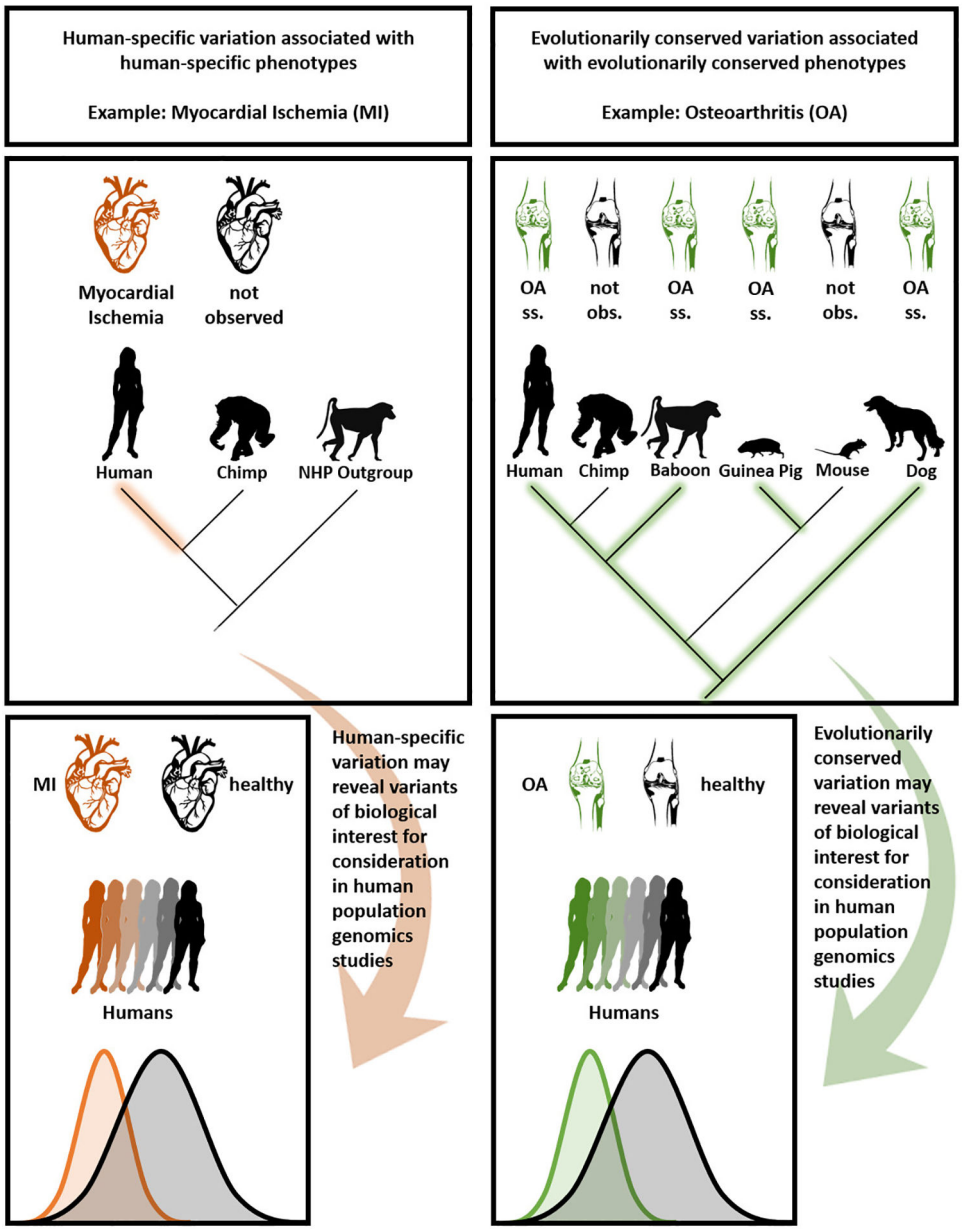

Figure 1. Evolution of genotype-phenotype associations can inform human-specific research.

Comparative research within primates and across large phylogenetic distributions of organisms can inform human-specific research. Because variation between species is greater than variation within species, features that distinguish species from one another typically have larger effect sizes than features that distinguish individuals from one another. Thus, the larger effect sizes between species can expose regions of interest for further exploration within human populations. However, while large phylogenetic comparisons can reveal aspects of conservation within humans, research in primates is necessary to isolate features that uniquely arose in the human lineage. Depicted here are representative examples of diseases that display differential susceptibilities across primates and mammals and that illustrate how comparative research may be used to inform human-specific research. **Left**

Panels: In the case of CVD, humans are prone to myocardial ischemia (MI) and chimpanzees are prone myocardial fibrosis [33]. By comparing human variation with chimpanzee variation, and grounding lineage specificity using a nonhuman primate outgroup, it is possible to identify human-specific genotype-phenotype associations related to MI susceptibility (top left panel). Once isolated, this human-specific variation may reveal variants of biological interest that can be considered further in human population genomics studies, where different individuals have different susceptibilities to MI and the effect sizes of genetic variants associated with trait variation are substantially smaller (bottom left panel). **Right Panels:** In the case of osteoarthritis (OA), humans and several other primates and mammals are susceptible to OA (OA ss.), while other animals have not been observed (not obs.) to naturally develop OA [13]. By comparing human variation with variation observed in other OA-susceptible lineages, it is possible to identify conserved genotype-phenotype associations related to OA susceptibility (top right panel). Once isolated, this conserved variation may reveal variants of biological interest that can be considered further in human population genomics studies, where different individuals have different susceptibilities to OA and the effect sizes of genetic variants associated with trait variation are substantially smaller (bottom right panel). Figure images were adapted from http://phylopic.org/, http://clipart-library.com/clipart/pco5XgRLi.htm, and https://springloadedtechnology.com/guide-to-severe-knee-osteoarthritis/.