HHS Public Access

Author manuscript

Curr Opin Genet Dev. Author manuscript; available in PMC 2016 December 01.

Published in final edited form as:

Curr Opin Genet Dev. 2015 December ; 35: 9–15. doi:10.1016/j.gde.2015.08.005.

The Evolution of the Human Genome

Corinne N. Simonti[1] and **John A. Capra**[1,2,*]

[1]Vanderbilt Genetics Institute, Vanderbilt University, Nashville, TN, 37235, USA

[2]Departments of Biological Sciences and Biomedical Informatics, Vanderbilt University, Nashville, TN, 37235, USA

Abstract

Human genomes hold a record of the evolutionary forces that have shaped our species. Advances in DNA sequencing, functional genomics, and population genetic modeling have deepened our understanding of human demographic history, natural selection, and many other long-studied topics. These advances have also revealed many previously underappreciated factors that influence the evolution of the human genome, including functional modifications to DNA and histones, conserved 3D topological chromatin domains, structural variation, and heterogeneous mutation patterns along the genome. Using evolutionary theory as a lens to study these phenomena will lead to significant breakthroughs in understanding what makes us human and why we get sick.

Keywords

human evolution; population genetics; evolution of gene regulation; evolutionary medicine

Introduction

Understanding the evolution of the human genome is a tantalizing goal. Accurately decoding the biological programs encoded in the human genome would reveal molecular answers to fundamental questions about human origins and the genetic basis for human-specific traits. Studying the evolutionary and demographic history of our species also has great promise to reveal how and why modern humans get sick. The human genome has been shaped by evolutionary pressures that, in many cases, no longer reflect the circumstances of most humans, and this mismatch between our genes and our environment can lead to disease [1,2].

In spite of immense progress since the sequencing of the first human genome more than 10 years ago, there is still much we do not understand about the evolution of the human genome. Recent statistical and experimental advances and the sequencing of thousands of human genomes from diverse populations have revealed significant complexity in classical

*Corresponding author: tony.capra@vanderbilt.edu.

topics in human population genetics, including the dynamics of selection across human populations and closely related species [3–6], the determinants of variation in mutation rates [7,8], inference of ancient human population histories [9,10], and how variants, in particular rare variants, contribute to phenotypes [11–13]. Perhaps the most dramatic result in this field over the past five years has been the sequencing of ancient DNA from archaic hominins, like Neanderthals and Denisovans [14,15], and the comprehensive demonstration of admixture between the ancestors of modern humans and several archaic hominin groups [16–20]. Each of these topics has been covered in recent comprehensive reviews (referenced above) and a recent issue of this journal [21], so here we highlight several additional genetic, environmental, and demographic factors influencing human genome evolution that we believe deserve further attention (Figure 1).

How do gene regulatory processes influence human genome evolution?

In the 40 years since King and Wilson hypothesized that phenotypic differences between closely related species were driven by gene regulatory changes, considerable support has been found for the importance of *cis*-regulatory elements (CREs), such as promoters and enhancers, in human evolution [22,23] and disease [24,25]. The recent ability to map gene expression, transcription factor (TF) binding sites, and histone modifications genome-wide in many tissues and species [26–28] has revealed that, while gene expression is generally conserved within similar tissues across species [29], CREs experience rapid turnover [30–32]. For example, a recent study of liver promoters and enhancers across 20 mammalian species found that 25% of a species' enhancers and 10% of its promoters were unique, even when the underlying sequence was deeply conserved [33]; similar results were found for limb CREs across human, macaque, and mouse [34].

There is still much to be learned about the evolution of regulatory sequence, in particular about its dynamics across tissues, species, and different classes of CREs, and how selection acts on these elements. For example, transposable elements (TEs) have helped reprogram gene regulatory networks in tissues relevant to pregnancy in humans and other mammals [35]. In contrast, TEs have made only a modest contribution to the evolution of new CREs in the liver [33]. This suggests that, while fast turnover of CREs and conserved gene expression are common features of mammalian genome evolution, different evolutionary dynamics and pressures act on CREs active in different tissues. It is likely that the regulatory landscapes of some tissues are more conducive to turnover than others; for example, tissues with greater phenotypic diversity across species, like those involved in pregnancy, may be more susceptible to TE-based rewiring. The maintenance and modification of these regulatory processes and their influence on genome evolution requires further investigation.

Integrating genome-wide maps of CREs, TF binding, and expression with recent advances in techniques for determining *in vivo* chromatin conformation of DNA [36] may provide a promising framework for modeling the influence of gene regulation on genome evolution. A recent study of chromatin looping in multiple human and mouse tissues found significant conservation of gene activity within local topological domains across cells and species [37]. These results suggest that, as is true for proteins, the 3D structure of regulatory neighborhoods maybe more deeply conserved and important for function than the sequence-

level conservation of individual CREs. Integrating data about genome structure and CREs across many individuals will likely lead to better models of regulatory sequence evolution and how selection acts on gene expression across evolutionary time and tissues.

How do chemical modifications to DNA and histones constrain human genome evolution?

The human body contains hundreds of different cell types with diverse forms and functions, yet each cell contains (essentially) the same genome. The past decade has seen increasing appreciation for the role of DNA and histone modifications, such as methylation and acetylation, in the diverse gene expression programs observed across different cell types within complex organisms [38–40]. These modifications can be influenced by environmental factors [41] and in some cases inherited across generations, though the extent of trans-generational inheritance in humans is still unclear [42].

In spite of extensive work linking these modifications to nearly all processes of development, aging, and disease [39,43–45], the influence of these modifications on patterns of genome sequence evolution has received comparatively little attention. For example, the extent to which the potential for chemical modification places constraint on DNA sequence patterns, e.g, CpG sites, is not resolved. Several recent studies have explored the degree of conservation of DNA and histone modifications across humans and closely related species [31,33,34,46–48]; changes to the modification status of orthologous regions are common between closely related species and, for DNA methylation, there is a positive correlation between sequence variation and promoter methylation changes. However, even in the presence of deep sequence conservation, many sites show differential modification. Much work remains to model the evolution of these modifications between individuals and species and to identify associated sequence constraints (or lack thereof). Understanding the evolution of these modifications may help resolve debates about whether specific modifications are causal or are the result of other processes like TF binding and transcription [49].

How should interactions between multiple genetic variants and phenotypes be modeled?

The majority of human phenotypes of clinical and evolutionary interest are specified by multiple loci across the human genome. Developing models that account for relationships between multiple genetic variants and phenotypes will be critical to fully dissecting the evolution and complex genetic architecture of most human traits. For example, pleiotropy—when a locus influences multiple independent traits—is found throughout the human genome; however, there is still considerable uncertainty about its prevalence and influence on genome evolution [50–53]. Similarly, epistasis—a non-additive interaction between genetic variants—is common in model organisms, but its influence on human traits has been controversial due to a number of technical and biological factors that can confound current tests for interactions between variants [54,55]. Each of these areas is in need of new

statistical approaches that update existing models to make full use of the wealth of genotype and phenotype data that have become available in the last five years.

What are the causes and effects of mutational biases along the human genome?

There is considerable variation in the rate and pattern of substitution along the human genome. Failure to account for these biases can confound tests for selection, complicate demographic inference, and weaken power in association tests [7]. One of the most potentially influential mutational biases is a recombination-associated process called GC-biased gene conversion (gBGC). gBGC results from a slight preference for G/C alleles in the mismatch repair machinery that has the potential to promote the maintenance of deleterious alleles [57]. The action of gBGC is widespread in human populations and across diverse species [58–60]. Genome-wide modeling of gBGC has demonstrated differences in its strength across the human and chimpanzee lineages [59] and between different human populations [61,62].

The evolution and effects of gBGC are intimately tied to the dynamics of recombination, which vary considerably in rate along the genome within human populations and between closely related species [63]. Recombination patterns influence many drivers of genome evolution, including the efficacy of selection, mutation rates, and the accumulation of deleterious mutations [64,65]. In humans, the fast evolving PRDM9 protein directs recombination to specific hotspots based on the occurrence of a GC-rich motif [66]. Using modern and archaic genome sequences, modeling suggests that gBGC degrades the PRDM9 motif over time and that this may drive the rapid turnover of the recombination landscape in human populations [67].

gBGC is only one of several sources of mutation rate variation that are not well understood [7]. Recent direct estimates from trios indicate that the human germline mutation rate is only half of what is expected from phylogenetic estimates [68], and analysis of whole genome sequence data suggests the evolution of population-specific mutation rates since the divergence of Europeans and Asians [69]. These results underscore the need for further study of the dynamics and causes of human mutation rate variation across evolutionary time and genomic space. We need to develop high-resolution maps of mutation rates in different populations, better models of how it interacts with selection and recombination, and most importantly, a deeper understanding of its effects on organismal fitness.

What are appropriate models for the evolution and functional impact of structural variation?

The initial comparison of the draft human and chimpanzee genomes identified approximately 35 million single nucleotide polymorphisms (SNPs), 5 million small insertions and deletions (indels), and hundreds of larger structural variants (SVs). Indels and large SVs account for far more nucleotide differences between the human and chimpanzee genomes than SNPs [70] and have restructured the genomes of great apes [71]. Recent work on de novo rates of indels and SVs in human populations found that structural changes are

much more rare and occur at lower frequency than SNPs; nonetheless, they influence an average of 4.1 kilobases per generation, which is 91 times more than de novo substitutions [72]. Indels and large SVs (including copy number variants, inversions, and other genomic rearrangements) are more likely to cause disease and have been hypothesized to have a greater influence on recent human evolution than SNPs [73,74]. Indeed, many human- and population-specific SVs have been tied to human-specific phenotypes [75], e.g., human-specific deletion of a conserved enhancer of the androgen receptor gene may be responsible for the lack of penile spines in humans [76]. It has also been suggested that de novo creation of new genes is more common than previously appreciated; tens of new human-specific genes have been detected, with particular enrichment for expression in the brain and testes [77,78].

In spite of the potential importance of indels, large SVs, and new genes to phenotypic differences between human individuals and between closely related species, they have received considerably less attention in evolutionary modeling and testing for association with disease than SNPs. Developing appropriate models for the evolution of indels and SVs faces several challenges including the difficulty of accurately identifying them in short read sequencing data, the diversity of mechanisms that generate them, and their highly heterogeneous mutation rates and distributions along the genome [72,79]. Nevertheless, it is essential to develop evolutionary models akin to those in common use for testing hypotheses about patterns of single nucleotide variant evolution and association with disease for indels and SVs. Sufficiently accurate maps of these events across hundreds of humans are now becoming available [79,80]; these data should facilitate the development of new modeling approaches.

How can we efficiently connect human-specific genomic changes to phenotypes?

Sequencing the genomes of thousands of humans, several archaic humans, and our closest great ape relatives has revealed thousands of loci in the human genome that have experienced accelerated evolution on the human lineage and hundreds more with signatures of recent positive selection [81–83]. These loci hold the promise of explaining much of human-specific biology, and many hypotheses have been proposed about their effects [75]. However, beyond a handful of successes that involved detailed experimental validation [76,84–88], connecting these mutations to effects on human phenotypes has been difficult. The first obstacle comes from the fact that the vast majority of these regions are non-coding and have minimal functional annotation. Furthermore, most human-specific traits have complex genetic architectures in which many coding and non-coding loci influence the phenotype [89]. Finally, appropriate model systems in which to test potential effects of mutations are not available for many phenotypes, and it is challenging to test variants in a high throughput manner in available systems.

Algorithmic and experimental innovations paired with increases in available phenotype and functional genomic data will significantly increase the pace with which human-specific variants can be characterized. For example, algorithms that integrate diverse functional, evolutionary, and DNA sequence data have shown that many human accelerated regions are

developmental gene regulatory enhancers, with particular enrichment for brain activity [81,88,90]. As our understanding of how non-coding mutations influence gene expression and function improves [91], so will the accuracy and specificity of hypotheses about the effects of these regions on human-specific phenotypes.

Over the past ten years, genome-wide association studies (GWAS) have identified hundreds of variants associated with complex diseases [92,93]. These studies provide insight into the functions encoded in specific regions of the genome that can inform evolutionary questions. However, the majority of human- and population-specific variants have not been associated with functions. The recent integration of large databases of electronic health records (EHRs) linked to patient genotypes [94] provides a new approach to this problem. Thousands of phenotypes can be algorithmically derived from EHRs and then simultaneously tested for association with the loci of interest across thousands of individuals in a phenome-wide association study (PheWAS) [95]. As EHR databases grow and sequencing decreases in price, the PheWAS approach will enable efficient testing of hypotheses about the effects of mutations of evolutionary interest.

Finally, new technologies, including directed stem cell differentiation, massively parallel reporter assays [96], and CRISPR gene editing [97], will facilitate faster exploration of the mechanisms driving phenotypic associations in models that closely resemble the *in vivo* human context.

Conclusion

Understanding how evolutionary processes produced the human species and how developmental programs are encoded in the human genome is of great importance to basic and clinical science. The evolutionary history of the human genome is directly relevant to our ability to anticipate and treat human disease [1]. In this review, we have highlighted several research areas that have potential to significantly deepen our knowledge of human genome evolution over the next few years, but our list is not exhaustive. Many other areas, including the evolutionary study of human–microbe interactions [98,99] and experimental evolution [100], are poised for breakthroughs. We are also eager to see how recent technical advances in long-read genome sequencing and single cell analysis will change our understanding of evolutionary processes. Ultimately, continued analysis of the human genome in an evolutionary framework will further reveal the genetic origins of human-specific biology and improve our understanding of the etiology of human disease.

Acknowledgments

C.N.S. was supported by NIH training grant 5T32EY021453. J.A.C. was supported by institutional funds from Vanderbilt University and a March of Dimes Innovation Catalyst award. We are grateful to Seth Bordenstein, Alexandra Fish, Emily Hodges, Dennis Kostka, Katherine Pollard, and Scott Williams for discussions related to topics covered in this review.

BIBLIOGRAPHY AND REFERENCES CITED

1. Rodríguez JA, Marigorta UM, Navarro A. Integrating genomics into evolutionary medicine. Current Opinion in Genetics & Development. 2014; 29:97–102. [PubMed: 25218863]

2. Gluckman, PD.; Beedle, A.; Hanson, MA. Principles of evolutionary medicine. Oxford: Oxford University Press; 2009.

3. Fu W, Akey JM. Selection and Adaptation in the Human Genome. Annual Review of Genomics and Human Genetics. 2013; 14:467–489.

4. Hernandez RD, Kelley JL, Elyashiv E, Melton S, Auton A, McVean G, Sella G, Przeworski M. Classic selective sweeps were rare in recent human evolution. Science. 2011; 331:920–924. [PubMed: 21330547]

5. Jensen JD. On the unfounded enthusiasm for soft selective sweeps. Nat Commun. 2014; 5

6. Prado-Martinez J, Sudmant PH, Kidd JM, Li H, Kelley JL, Lorente-Galdos B, Veeramah KR, Woerner AE, O'Connor TD, Santpere G, et al. Great ape genetic diversity and population history. Nature. 2013; 499:471–475. [PubMed: 23823723] * This study reported the sequencing of 79 wild- and captive-born individuals from all six great ape species. These data revealed radical fluctuations in effective population size and genetic diversity and provide a broader context in which to interpret human evolution.

7. Ségurel L, Wyman MJ, Przeworski M. Determinants of Mutation Rate Variation in the Human Germline. Annual Review of Genomics and Human Genetics. 2014; 15:47–70.

8. Henn BM, Botigue LR, Bustamante CD, Clark AG, Gravel S. Estimating the mutation load in human genomes. Nat Rev Genet. 2015 advance online publication.

9. Gronau I, Hubisz MJ, Gulko B, Danko CG, Siepel A. Bayesian inference of ancient human demography from individual genome sequences. Nature genetics. 2011; 43:1031–1034. [PubMed: 21926973]

10. Li H, Durbin R. Inference of human population history from individual whole-genome sequences. Nature. 2011; 475:493–496. [PubMed: 21753753]

11. Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. Am J Hum Genet. 2014; 95:5–23. [PubMed: 24995866]

12. Gibson G. Rare and common variants: twenty arguments. Nat Rev Genet. 2012; 13:135–145. [PubMed: 22251874]

13. Sulem P, Helgason H, Oddson A, Stefansson H, Gudjonsson SA, Zink F, Hjartarson E, Sigurdsson GT, Jonasdottir A, Jonasdottir A. Identification of a large set of rare complete human knockouts. Nature genetics. 2015; 47:448–452. [PubMed: 25807282] * This study revealed the patterns and prevalence of loss-of-function mutations in the Icelandic population by sequencing the whole genomes of 2,636 Icelanders and imputing the sequence variants identified into more than 100,000 more with genotyping data.

14. Fu Q, Li H, Moorjani P, Jay F, Slepchenko SM, Bondarev AA, Johnson PL, Aximu-Petri A, Prüfer K, de Filippo C. Genome sequence of a 45,000-year-old modern human from western Siberia. Nature. 2014; 514:445–449. [PubMed: 25341783] ** This paper presented the genome sequence of the oldest anatomically modern human sequenced to date. It provides insights into the timing of Neanderthal interbreeding and estimates of the human mutation rate.

15. Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, Heinze A, Renaud G, Sudmant PH, de Filippo C. The complete genome sequence of a Neanderthal from the Altai Mountains. Nature. 2014; 505:43–49. [PubMed: 24352235] ** This study presented a high quality Neanderthal genome sequence and demonstrated several gene flow events between archaic hominins and anatomically modern humans.

16. Kelso J, Prüfer K. Ancient humans and the origin of modern humans. Current opinion in genetics & development. 2014; 29:133–138. [PubMed: 25286439]

17. Pääbo S. The Human Condition—A Molecular Approach. Cell. 2014; 157:216–226. [PubMed: 24679537]

18. Racimo F, Sankararaman S, Nielsen R, Huerta-Sanchez E. Evidence for archaic adaptive introgression in humans. Nat Rev Genet. 2015

19. Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, Pääbo S, Patterson N, Reich D. The genomic landscape of Neanderthal ancestry in present-day humans. Nature. 2014; 507:354–357. [PubMed: 24476815]

20. Vernot B, Akey JM. Resurrecting Surviving Neandertal Lineages from Modern Human Genomes. Science. 2014; 343:1017–1021. [PubMed: 24476670] * This study presented an approach for

identifying introgressed Neanderthal DNA in modern human genomes and mapped the presence and frequency of Neanderthal DNA in the genomes of non-African individuals.

21. Andrés AM, Nowick K. Editorial overview: Genetics of human evolution: The genetics of human origins. Current Opinion in Genetics & Development. 2014; 29:v–vii. [PubMed: 25468516]
22. Siepel A, Arbiza L. Cis-regulatory elements and human evolution. Current opinion in genetics & development. 2014; 29:81–89. [PubMed: 25218861]
23. Villar D, Flicek P, Odom DT. Evolution of transcription factor binding in metazoans [mdash] mechanisms and functional implications. Nature Reviews Genetics. 2014; 15:221–233.
24. Edwards Stacey L, Beesley J, French Juliet D, Dunning Alison M. Beyond GWASs: Illuminating the Dark Road from Association to Function. American Journal of Human Genetics. 2013; 93:779–797. [PubMed: 24210251]
25. Sakabe NJ, Savic D, Nobrega MA. Transcriptional enhancers in development and disease. Genome Biol. 2012; 13:238. [PubMed: 22269347]
26. GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science. 2015; 348:648–660. [PubMed: 25954001] * This study presented pilot RNA sequencing data from 1641 samples across 43 tissues from 175 human individuals. This is the largest analysis of the relationship between genetic variation and gene expression to date
27. Melé M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM, Pervouchine DD, Sullivan TJ, et al. The human transcriptome across tissues and individuals. Science. 2015; 348:660–665. [PubMed: 25954002]
28. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]
29. Brawand D, Soumillon M, Necsulea A, Julien P, Csardi G, Harrigan P, Weier M, Liechti A, Aximu-Petri A, Kircher M, et al. The evolution of gene expression levels in mammalian organs. Nature. 2011; 478:343–348. [PubMed: 22012392]
30. Weirauch MT, Hughes TR. Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same. Trends in genetics. 2010; 26:66–74. [PubMed: 20083321]
31. Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, Ryba T, Sandstrom R, Ma Z, Davis C, Pope BD. A comparative encyclopedia of DNA elements in the mouse genome. Nature. 2014; 515:355–364. [PubMed: 25409824] ** This study determined gene transcription, DNase I hypersensitivity, transcription factor binding, and chromatin modifications across the mouse genome in diverse cell and tissue types. Comparing these profiles to human data they highlighted significant similarities and differences between different levels of the regulatory landscape of mouse and human.
32. Necsulea A, Kaessmann H. Evolutionary dynamics of coding and non-coding transcriptomes. Nature Reviews Genetics. 2014; 15:734–748.
33. Villar D, Berthelot C, Aldridge S, Rayner TF, Lukk M, Pignatelli M, Park TJ, Deaville R, Erichsen JT, Jasinska AJ, et al. Enhancer evolution across-20 mammalian species. Cell. 2015; 160:554–566. [PubMed: 25635462] * This study characterized the conservation of enhancer and promoter activity across 20 diverse mammalian species. The authors demonstrated rapid turnover in enhancer activity and greater stability of promoters across mammals.
34. Cotney J, Leng J, Yin J, Reilly SK, DeMare LE, Emera D, Ayoub AE, Rakic P, Noonan JP. The evolution of lineage-specific regulatory activities in the human embryonic limb. Cell. 2013; 154:185–196. [PubMed: 23827682]
35. Lynch VJ, Nnamani MC, Kapusta A, Brayer K, Plaza SL, Mazur EC, Emera D, Sheikh SZ, Grützner F, Bauersachs S. Ancient Transposable Elements Transformed the Uterine Regulatory Landscape and Transcriptome during the Evolution of Mammalian Pregnancy. Cell reports. 2015; 10:551–561. [PubMed: 25640180]
36. Pombo A, Dillon N. Three-dimensional genome architecture: players and mechanisms. Nat Rev Mol Cell Biol. 2015; 16:245–257. [PubMed: 25757416]
37. Rao Suhas SP, Huntley Miriam H, Durand Neva C, Stamenova Elena K, Bochkov Ivan D, Robinson James T, Sanborn Adrian L, Machol I, Omer Arina D, Lander Eric S, et al. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell. 159:1665–1680. [PubMed: 25497547] ** This paper used in situ Hi-C to determine high

resolution (1–5 kb) maps of the chromatin conformation in nine cell types. These maps demonstrated chromatin contact domains with coherent gene expression and functional patterns that are maintained across cell types and species.

38. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W. Chromatin architecture reorganization during stem cell differentiation. Nature. 2015; 518:331–336. [PubMed: 25693564]
39. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015; 518:317–330. [PubMed: 25693563]
40. Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. Nat Rev Genet. 2011; 12:7–18. [PubMed: 21116306]
41. Feil R, Fraga MF. Epigenetics and the environment: emerging patterns and implications. Nat Rev Genet. 2012; 13:97–109. [PubMed: 22215131]
42. Heard E. Martienssen Robert A. Transgenerational Epigenetic Inheritance: Myths and Mechanisms. Cell. 157:95–109. [PubMed: 24679529]
43. Cantone I, Fisher AG. Epigenetic programming and reprogramming during development. Nat Struct Mol Biol. 2013; 20:282–289. [PubMed: 23463313]
44. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. Nat Rev Genet. 2013; 14:204–220. [PubMed: 23400093]
45. Brunet A, Berger SL. Epigenetics of aging and aging-related disease. J Gerontol A Biol Sci Med Sci. 2014; 69(Suppl 1):S17–S20. [PubMed: 24833581]
46. 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:e1003763. [PubMed: 24039605]
47. Pai AA, Gilad Y. Comparative studies of gene regulatory mechanisms. Current opinion in genetics & development. 2014; 29:68–74. [PubMed: 25215415]
48. Zhou X, Cain CE, Myrthil M, Lewellen N, Michelini K, Davenport ER, Stephens M, Pritchard JK, Gilad Y. Epigenetic modifications are associated with inter-species gene expression variation in primates. Genome Biol. 2014; 15:547. [PubMed: 25468404]
49. Ptashne M. Epigenetics: Core misconcept. Proceedings of the National Academy of Sciences. 2013; 110:7101–7103.
50. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet. 2013; 14:483–495. [PubMed: 23752797]
51. Pavlicev M, Wagner GP. A model of developmental evolution: selection, pleiotropy and compensation. Trends in Ecology & Evolution. 2012; 27:316–322. [PubMed: 22385978]
52. Sivakumaran S, Agakov F, Theodoratou E, Prendergast James G, Zgaga L, Manolio T, Rudan I, McKeigue P, Wilson James F, Campbell H. Abundant Pleiotropy in Human Complex Diseases and Traits. American Journal of Human Genetics. 2011; 89:607–618. [PubMed: 22077970]
53. Wagner GP, Zhang J. The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. Nat Rev Genet. 2011; 12:204–213. [PubMed: 21331091]
54. Hemani G, Shakhbazov K, Westra H-J, Esko T, Henders AK, McRae AF, Yang J, Gibson G, Martin NG, Metspalu A, et al. Detection and replication of epistasis influencing transcription in humans. Nature. 2014; 508:249–253. [PubMed: 24572353]
55. Wood AR, Tuke MA, Nalls MA, Hernandez DG, Bandinelli S, Singleton AB, Melzer D, Ferrucci L, Frayling TM, Weedon MN. Another explanation for apparent epistasis. Nature. 2014; 514:E3–E5. [PubMed: 25279928]
56. Galtier N, Duret L. Adaptation or biased gene conversion? Extending the null hypothesis of molecular evolution. TRENDS in Genetics. 2007; 23:273–277. [PubMed: 17418442]
57. Duret L, Galtier N. Biased gene conversion and the evolution of mammalian genomic landscapes. Annual review of genomics and human genetics. 2009; 10:285–311.
58. Lartillot N. Phylogenetic Patterns of GC-Biased Gene Conversion in Placental Mammals and the Evolutionary Dynamics of Recombination Landscapes. Molecular Biology and Evolution. 2013; 30:489–502. [PubMed: 23079417]

59. Capra JA, Hubisz MJ, Kostka D, Pollard KS, Siepel A. A Model-Based Analysis of GC-Biased Gene Conversion in the Human and Chimpanzee Genomes. PLoS Genet. 2013; 9:e1003684. [PubMed: 23966869]
60. Katzman S, Capra JA, Haussler D, Pollard KS. Ongoing GC-Biased Evolution Is Widespread in the Human Genome and Enriched Near Recombination Hot Spots. Genome Biology and Evolution. 2011; 3:614–626. [PubMed: 21697099]
61. Lachance J, Tishkoff SA. Biased Gene Conversion Skews Allele Frequencies in Human Populations, Increasing the Disease Burden of Recessive Alleles. The American Journal of Human Genetics. 2014; 95:408–420. [PubMed: 25279983]
62. Glemin S, Arndt PF, Messer PW, Petrov D, Galtier N, Duret L. Quantification of GC-biased gene conversion in the human genome. 2014
63. Auton A, Fledel-Alon A, Pfeifer S, Venn O, Ségurel L, Street T, Leffler EM, Bowden R, Aneas I, Broxholme J, et al. A Fine-Scale Chimpanzee Genetic Map from Population Sequencing. Science. 2012; 336:193–198. [PubMed: 22422862]
64. Coop G, Przeworski M. An evolutionary view of human recombination. Nat Rev Genet. 2007; 8:23–34. [PubMed: 17146469]
65. Hussin JG, Hodgkinson A, Idaghdour Y, Grenier J-C, Goulet J-P, Gbeha E, Hip-Ki E, Awadalla P. Recombination affects accumulation of damaging and disease-associated mutations in human populations. Nat Genet. 2015; 47:400–404. [PubMed: 25685891]
66. Baudat F, Imai Y, de Massy B. Meiotic recombination in mammals: localization and regulation. Nat Rev Genet. 2013; 14:794–806. [PubMed: 24136506]
67. Lesecque Y, Glémin S, Lartillot N, Mouchiroud D, Duret L. The Red Queen Model of Recombination Hotspots Evolution in the Light of Archaic and Modern Human Genomes. PLoS Genet. 2014; 10:e1004790. [PubMed: 25393762] * This paper used modern human and archaic hominin genome sequences to provide support for the Red Queen hypothesis that there is a destructive interaction between gBGC and PRDM9 binding preferences that drives the turnover of recombination hotspots.
68. Scally A, Durbin R. Revising the human mutation rate: implications for understanding human evolution. Nature Reviews Genetics. 2012; 13:745–753.
69. Harris K. Evidence for recent, population-specific evolution of the human mutation rate. Proceedings of the National Academy of Sciences. 2015; 112:3439–3444.
70. and Analysis Consortium The Chimpanzee S. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature. 2005; 437:69–87. [PubMed: 16136131]
71. Sudmant PH, Huddleston J, Catacchio CR, Malig M, Hillier LW, Baker C, Mohajeri K, Kondova I, Bontrop RE, Persengiev S. Evolution and diversity of copy number variation in the great ape lineage. Genome research. 2013; 23:1373–1382. [PubMed: 23825009]
72. Kloosterman WP, Francioli LC, Hormozdiari F, Marschall T, Hehir-Kwa JY, Abdellaoui A, Lameijer E-W, Moed MH, Koval V, Renkens I, et al. Characteristics of de novo structural changes in the human genome. Genome Research. 2015 * This study provided a large-scale analysis of de novo rates of indels and structural variation in the human genome. They found that de novo structural changes influence far more nucleotides per generation than de novo substitutions.
73. Stankiewicz P, Lupski JR. Structural Variation in the Human Genome and its Role in Disease. Annual Review of Medicine. 2010; 61:437–455.
74. Chuzhanova NA, Anassis EJ, Ball EV, Krawczak M, Cooper DN. Meta-analysis of indels causing human genetic disease: mechanisms of mutagenesis and the role of local DNA sequence complexity. Human mutation. 2003; 21:28–44. [PubMed: 12497629]
75. O'Bleness M, Searles VB, Varki A, Gagneux P, Sikela JM. Evolution of genetic and genomic features unique to the human lineage. Nat Rev Genet. 2012; 13:853–866. [PubMed: 23154808]
76. McLean CY, Reno PL, Pollen AA, Bassan AI, Capellini TD, Guenther C, Indjeian VB, Lim X, Menke DB, Schaar BT. Human-specific loss of regulatory DNA and the evolution of human-specific traits. Nature. 2011; 471:216–219. [PubMed: 21390129]
77. Knowles DG, McLysaght A. Recent de novo origin of human protein-coding genes. Genome research. 2009; 19:1752–1759. [PubMed: 19726446]

78. Zhang YE, Long M. New genes contribute to genetic and phenotypic novelties in human evolution. Current Opinion in Genetics & Development. 2014; 29:90–96. [PubMed: 25218862]

79. Montgomery SB, Goode DL, Kvikstad E, Albers CA, Zhang ZD, Mu XJ, Ananda G, Howie B, Karczewski KJ, Smith KS, et al. The origin, evolution, and functional impact of short insertion–deletion variants identified in 179 human genomes. Genome Research. 2013; 23:749–761. [PubMed: 23478400]

80. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491:56–65. [PubMed: 23128226]

81. Capra JA, Erwin GD, McKinsey G, Rubenstein JL, Pollard KS. Many human accelerated regions are developmental enhancers. Philosophical Transactions of the Royal Society B. Biological Sciences. 2013; 368 20130025. * This largest study of the functions of human accelerated regions demonstrated that many of them act as developmental gene regulatory enhancers with enrichment for activity in the brain.

82. Grossman SR, Andersen KG, Shlyakhter I, Tabrizi S, Winnicki S, Yen A, Park DJ, Griesemer D, Karlsson EK, Wong SH. Identifying recent adaptations in large-scale genomic data. Cell. 2013; 152:703–713. [PubMed: 23415221]

83. Colonna V, Ayub Q, Chen Y, Pagani L, Luisi P, Pybus M, Garrison E, Xue Y. Human genomic regions with exceptionally high levels of population differentiation identified from 911 whole-genome sequences. 2014

84. Prabhakar S, Visel A, Akiyama JA, Shoukry M, Lewis KD, Holt A, Plajzer-Frick I, Morrison H, Fitzpatrick DR, Afzal V, et al. Human-specific gain of function in a developmental enhancer. Science. 2008; 321:1346–1350. [PubMed: 18772437]

85. Pollard KS, Salama SR, Lambert N, Lambot M-A, Coppens S, Pedersen JS, Katzman S, King B, Onodera C, Siepel A. An RNA gene expressed during cortical development evolved rapidly in humans. Nature. 2006; 443:167–172. [PubMed: 16915236]

86. Kamberov YG, Wang S, Tan J, Gerbault P, Wark A, Tan L, Yang Y, Li S, Tang K, Chen H. Modeling recent human evolution in mice by expression of a selected EDAR variant. Cell. 2013; 152:691–702. [PubMed: 23415220] * This study provided one of the most comprehensive experimental characterizations of a region that has experienced recent positive selection in a human population. Using mouse models, it demonstrated that EDAR variants were likely adaptive in Chinese populations and that they influence relevant hair, mammary, and eccrine gland phenotypes.

87. Boyd JL, Skove Stephanie L, Rouanet Jeremy P, Pilaz L-J, Bepler T, Gordân R, Wray Gregory A, Silver Debra L. Human-Chimpanzee Differences in a FZD8 Enhancer Alter Cell-Cycle Dynamics in the Developing Neocortex. Current Biology. 25:772–779. [PubMed: 25702574]

88. Boyd JL, Skove SL, Rouanet JP, Pilaz L-J, Bepler T, Gordân R, Wray GA, Silver DL. Human-Chimpanzee Differences in a FZD8 Enhancer Alter Cell-Cycle Dynamics in the Developing Neocortex. Current Biology. 2015; 25:772–779. [PubMed: 25702574] ** This study demonstrated that enhancer sequence changes may contribute to unique features of the human brain. It showed that the human version of a human accelerated enhancer region produces a faster progenitor cell cycle and increased neocortical size.

89. Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. Nat Rev Genet. 2015; 16:197–212. [PubMed: 25707927]

90. Erwin GD, Oksenberg N, Truty RM, Kostka D, Murphy KK, Ahituv N, Pollard KS, Capra JA. Integrating diverse datasets improves developmental enhancer prediction. PLoS computational biology. 2014; 10:e1003677. [PubMed: 24967590]

91. Battle A, Khan Z, Wang SH, Mitrano A, Ford MJ, Pritchard JK, Gilad Y. Impact of regulatory variation from RNA to protein. Science. 2014 1260793.

92. Stranger BE, Stahl EA, Raj T. Progress and Promise of Genome-Wide Association Studies for Human Complex Trait Genetics. Genetics. 2011; 187:367–383. [PubMed: 21115973]

93. Bush WS, Moore JH. Chapter 11: Genome-wide association studies. PLoS Comput Biol. 2012; 8:e1002822. [PubMed: 23300413]

94. McCarty CA, Chisholm RL, Chute CG, Kullo IJ, Jarvik GP, Larson EB, Li R, Masys DR, Ritchie MD, Roden DM. The eMERGE Network: a consortium of biorepositories linked to electronic

medical records data for conducting genomic studies. BMC medical genomics. 2011; 4:13. [PubMed: 21269473]

95. Denny JC, Bastarache L, Ritchie MD, Carroll RJ, Zink R, Mosley JD, Field JR, Pulley JM, Ramirez AH, Bowton E, et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. Nat Biotech. 2013; 31:1102–1111. * This study established the feasiblity of the phenome-wide association study (PheWAS) using eletronic health records, by demonstrating that PheWAS replicates the majority of known associations tested and identifies novel gene-phenotype associations.
96. Patwardhan RP, Hiatt JB, Witten DM, Kim MJ, Smith RP, May D, Lee C, Andrie JM, Lee S-I, Cooper GM, et al. Massively parallel functional dissection of mammalian enhancers in vivo. Nat Biotech. 2012; 30:265–270.
97. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotech. 2014; 32:347–355.
98. Brucker RM, Bordenstein SR. The capacious hologenome. Zoology. 2013; 116:260–261. [PubMed: 24035647]
99. Karlsson EK, Kwiatkowski DP, Sabeti PC. Natural selection and infectious disease in human populations. Nature Reviews Genetics. 2014; 15:379–393.
100. Barrick JE, Lenski RE. Genome dynamics during experimental evolution. Nature Reviews Genetics. 2013; 14:827–839.

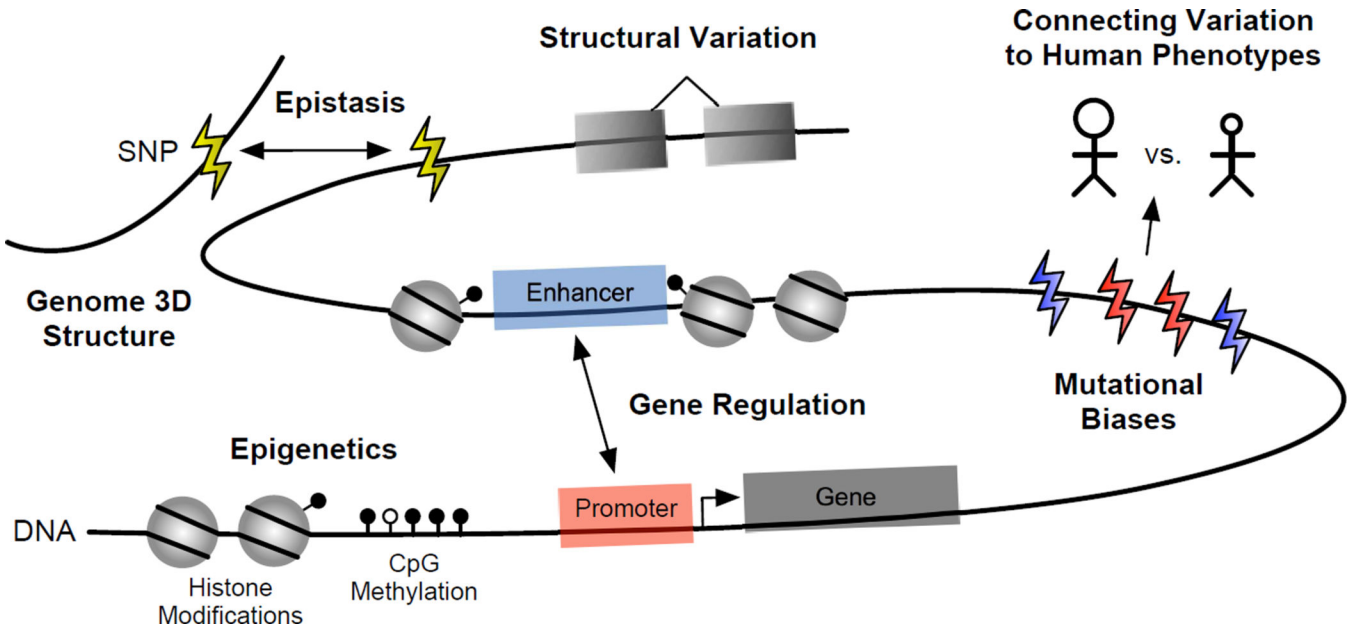

Figure 1.

Many areas of ongoing research in genomics have not been fully integrated into models of genome evolution. In this article, we discuss how study of several emerging topics (bold) in an evolutionary context will enrich our understanding of human evolution.