Cell

Human Genetics: Questions, Challenges, and the Future

The Future of Inferring the Past

Sarah Tishkoff
University of Pennsylvania

Since the studies of "mitochondrial Eve" in the 1980s indicating that all modern humans originated in Africa, new and less expensive technologies have been developed, giving us the potential to sequence millions of genomes from individuals across the globe as well as genomes of ancient hominids as old as 400,000 years. Integration of fossil, archeological, and genetic data elucidates the evolutionary processes resulting in uniquely human traits as well as traits that vary across individuals, including risk for disease. The evolutionary history of modern humans is a tangled path. There have been multiple migration events of the genus homo out of Africa and admixture of archaic populations such as Neanderthal and Denisovans with each other and with modern humans. By analyzing larger numbers of ancient genomes from geographically diverse regions (including the tropics, where bones are not well preserved) over multiple time periods, we can directly reconstruct the origins of complex traits including disease. Novel statistical models are needed to distinguish the effects of demographic history and natural selection on patterns of genetic variation in modern populations. Ultimately, we must determine the functional impact of genomic variants on human phenotypic diversity, which requires better understanding of gene regulation as well as gene-by-environment and gene-by-gene interactions. High-throughput genome editing will enable us to systematically distinguish effects of genetic variants on protein activity and gene regulation across cell types. These advances will facilitate novel, and sometimes surprising, discoveries about the history of our species.

From Genomes to Biology

Joshua M. Akey
Department of Ecology and Evolutionary Biology
Lewis-Sigler Institute for Integrative Genomics, Princeton University

An important goal for human genetics is to translate the ever-increasing deluge of sequencing data, which could plausibly reach a billion genomes over the next decade, into quantitative, dynamic, and mechanistic biological knowledge. As is often the case, technological advances will be needed to make the sort of progress necessary to delineate the biology and evolution of genomes. Fortunately, if the previous decade in human genetics is any guide, it seems a safe bet that innovations in genomic technology will not only continue but accelerate. For instance, DNA sequencing will be cheaper, more accurate, and even more pervasive. This will enable single-cell sequencing to become routine, leading to exquisitely precise cell fate maps, elucidation of levels and determinants of somatic mosaicism, and insights into how stochasticity influences biological processes. Genome-editing technology will also mature, creating new hope in the clinic and complex ethical issues to confront for society. Such technology is also poised to revolutionize evolutionary genomics, allowing, for example, mechanisms of adaptive change to be more precisely delimited and determining what parts of our biology was shared and different compared to Neanderthals, Denisovans, and possibly other now-extinct hominins that our ancestors mated with. These, and other, advances in technology will influence the landscape of human genetics over the next decade, as the field markedly shifts from generating to interpreting and manipulating genome.

E Unum Pluribus—Manifold Effects of a Genetic Variant

Tuuli Lappalainen
New York Genome Center
Department of Systems Biology, Columbia University

The effect of genetic variants on human disease risk is one of the central questions in biology and medicine, and interrogation of genetic effects on cellular and molecular traits has become mainstream as well. In both of these domains, ever-increasing datasets have uncovered an astonishing level of complexity of the effects that a single genetic variant can have both at the cellular and organismal levels. Thus, we now know that the mindset of a variant having *a* functional effect is highly inadequate. For example, at the proximal molecular level, a variant's effect on gene-expression level is highly dependent on the cell type and state, and the cellular *trans* effects of that expression change in molecular networks of the cell are even more context dependent. Beyond cellular phenotypes, phenome-wide association studies have shown that the same loci often affect multiple phenotypes, of which only a subset are causally related, such as cholesterol level affecting heart disease risk. Creating biological and medical understanding from this complexity is an exciting challenge. Increasingly complete maps of genetic effects to diverse traits enables sophisticated statistical modeling of these rich data, using genetic variation as a causality anchor. This will allow us to decipher the causal networks of how molecular perturbations contribute to human traits and diseases.

Human Genetics: Emerging Paradigms and the Road Ahead

Elaine A. Ostrander
National Human Genome Research Institute
National Institutes of Health

The notion that access to the genome can change the course of a person's life is exciting and, in the long run, undeniably true. Emerging tools already aid in predicting disease risk, age at onset, and treatment responses. Indeed, while current studies focus only on coding sequences and a modest set of regulatory elements, the potential exists to link every disease phenotype to a set of base pairs. But to fulfill that promise, human genetics must evolve in three ways. First, analysis of sequence data cannot be a "bioinformatics-only" endeavor. Methods are required that more effectively merge sequence analysis with functional data. Second, as we stretch the bounds of the linear genome, evidence from non-traditional models must be exploited to reveal the rich diversity of biological mechanisms. Finally, recall that the men and women who started our Genetics Society relied first, and foremost, on meticulous descriptive science. While most such studies focused on Mendelian traits, the diseases that preoccupy us today are arguably compiled from those simpler disorders, highlighting the same genetic pathways. What's the emerging paradigm? It is time to look to a reinvention of observational science in medicine, alongside mechanistic analysis, with high value placed on detailed deconstruction of disease phenotypes. Indeed, before a process can be understood, it must be observed; without *The Voyage of the Beagle*, there is no *Descent of Man*. We must re-learn the humility of science, recover our past, and remember that hypotheses only have meaning in the light of diligent observation.

Sharing the Wonders

Guillaume Lettre
Université de Montréal

Given its primary goal, the success of genome-wide association studies (GWAS) is undeniable. We know of >10,000 genetic variants associated with hundreds of complex human diseases and traits. Each of these provides a glimpse into the genetic basis of diverse diseases such as myocardial infarction, diabetes, and schizophrenia. Currently, this peep into disease pathophysiology is foggy at best since it remains extremely difficult to move from genetic variants to genes and biological mechanisms. Difficult but not impossible, especially with the advent of tools like CRISPR-Cas9 editing, chromosome conformation capture, organoids, and 3D culture. The coming years will see a plethora of GWAS loci dissected at the molecular level, delivering new clues to predict, prevent, and treat human disease. In parallel to the progress in our labs, we need to embrace strategies to share with the public the implications of our breakthroughs. This must start at school: while Mendel's pea plants experiments remain relevant, there ought to be more exciting ways to teach students the wonders of the genetic code. And it extends everywhere: if we cannot explain to voters and taxpayers why our discoveries matter, how can we expect to make them our allies in convincing our governments that genetic research is worth funding? Translating the excitement of our findings to the broader community will ultimately help to reinforce the idea that, despite our superficial differences, we are pretty much alike.

The Future Lies in Biobanks

Cisca Wijmenga
University of Groningen, the Netherlands

Thousands of SNPs associated with hundreds of traits and diseases have been discovered over the past 10 years of GWAS studies. While this is very exciting, the translation into biology and potential therapies remains challenging. At the same time, the collection of disease cohorts has shifted to longitudinal population-based biobanks. Although it may take one or two generations to fully exploit these biobanks, their value for the next phase of research is evident. The critical ingredients will be the availability of a wide range of phenotypes—including different levels of molecular data—together with lifestyle factors. The opportunities here are enormous. First, they will help in interpreting GWAS signals by linkage to all possible levels of omics data (potentially in differentiated iPSC cells to mimic a specific context). Second, they will help uncover the pleiotropic effects of GWAS SNPs by phenome-wide association studies; such studies will uncover shared biological mechanisms. Third, they will help in understanding why not everyone carrying disease risk goes on to develop disease; such studies may in fact uncover novel ways to prevent disease. At the same time, biobanks engage the general population much more in research. Not only will we be able to give information back but participants will also become true citizen scientists who will provide a lot of the critical information we need through e.g., wearables.