The genetics of human migrations

Our ancestors migration out of Africa has left traces in our genomes that explain how they adapted to new environments

Philip Hunter

When our human ancestors began to migrate from the African savannahs across the Alps into Northern Europe and Asia between 200,000 and 60,000 years ago, they encountered a colder climate, different food sources and new predators. Inevitably, evolution began to work on those wandering tribes: their skin became lighter, their metabolism adapted to new food sources, and their immune system had to handle different pathogens. Modern *Homo sapiens* began to evolve into slightly different subspecies or, as they are referred to in the common vernacular, different races.

This shallow notion of race—that differences in our appearance have deep-seated consequences for who and what we are as human beings—is the root of racism, which has led to slavery, discrimination, murder and genocide throughout the millennia. With the emergence of modern genetics and genomics, scientists had hoped that a better understanding of genetic differences and similarities among humans would eventually help to overcome the controversial and toxic notion of race as having anything to do with a person's ability, intelligence or humanity. As J. Craig Venter, then head of the Celera Genomics Corporation in Rockville, MD, USA, put it in 2000, when the first draft of the human genome was published: "Race is a social concept, not a scientific one".

Since the publication of the human genome, further advances in genomic sequencing and statistical analysis are now yielding new insights into the link between past human migrations and today's variety of genetic differences that govern disease resistance, skin pigmentation, tolerance to cold, and ability to digest or metabolize

various nutrients. Such findings are contributing to discussions around the delicate issues of race, ethnicity and identity, but they are also having a significant impact on biomedical research and personalized health care. Furthermore, genetic analysis is shedding light on the migrations themselves, along with the impact on our evolution of facing new pathogens, climates and diets.

Differences in the distribution of alleles and associated phenotypes exist between populations and highlight their origins and migratory history, but the boundaries are rather diffuse, according to Mark Stoneking from the Department of Evolutionary Genetics at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. "I think genetic evidence shows overwhelmingly that there is a strong geographic influence on patterns of human genetic variation, so if you want to equate 'ethnicity' with 'geography', then there is a scientific basis for such concepts", he said. "However, even though we can identify the geographic origin of a person from genetic data with a fair degree of accuracy, there are no distinct boundaries between groups such as would be predicted by 'racial' classifications. Instead, genetic diversity is distributed on a continuum".

The implications of genetic research for the notions of race and ethnicity in the post Human Genome Project era were first set out comprehensively in a seminal paper in 2004 by Francis Collins, who was then at the US National Human Genome Research Institute (NHGRI) and who is now Director of the NIH. He argued that race and ethnicity were poorly defined and over simplistic terms that fail to account for the multiple environmental and genetic factors in disease

causation, including ancestral geographic origins, socio-economic status, education and access to health care [1].

.....
“...scientists had hoped that a better understanding of genetic differences and similarities among humans would eventually help to overcome the controversial and toxic notion of race. . .”
.....

"I think the major thing we have learnt is that human genetic variation does not fall into five distinct racial groups", said Vence Bonham, NHRGI's Associate Investigator for Social and Behavioral Research, in reference to the five racial divisions proposed in 1775 by Johann Friedrich Blumenbach. "There are clearly variations in the prevalence of certain gene variants and areas of the world where there are more clear clusters, but that does not create distinct groups that you can describe and report in a way that is scientifically accurate".

Yet, even if geneticists largely agree that the concept of human races is outdated, they have mixed views over how research on diversity between populations will affect public attitudes. Stoneking suggested that evolution itself has instilled a tendency to draw distinctions on the basis of visual characteristics. "It seems pretty clear that humans are very good at distinguishing subtle differences in facial features etc., including differences in skin, hair, and eye pigmentation. One potential

explanation is that in the past, it was important to be able to identify both other members of your own group and thus friends, as well as members of other groups and thus potential enemies, and so humans became adept at distinguishing others on the basis of appearance", he said. "But that still remains a hypothesis".

.....
"... some geneticists argue that prejudice can only be defeated by moral rather than biological arguments"

For such reasons, some geneticists argue that prejudice can only be defeated by moral rather than biological arguments. "I'm not sure these so-called 'discoveries' and the observation that we are all indeed rather closely related, will be more than an adjunct to a good, liberal, secular education, adequate resources and equal opportunity", said Stephen Oppenheimer, a geneticist at Oxford University in the UK and author of several seminal books on migration, including *Out of Eden: the Peopling of the World*. A different line is taken by Keith Cheng, a human geneticist at the Penn State College of Medicine in the USA. He argues in an upcoming book, *Racism in the 21st Century*, that geneticists should not ignore the issue of race in particular, but should tackle it head on to demystify the scientific aspects of the subject rather than avoiding it.

Meanwhile, the NHRGI has focused on dividing humans into about 1,000 subgroups as a basis for disease research, which, it believes, will be more valuable for studying human disease than attempting to draw lines on racial or ethnic grounds. "We have a number of different studies to improve the diversity of ancestral background in our study population so that we can better understand disease", Bonham said. "We have developed research measures for health services researchers and social scientists to understand how doctors use race in clinical practice". The underlying rationale is that some medical practitioners are making assumptions in diagnosis and prescription on the basis of broad racial or ethnic identities that lack a sound clinical basis.

The 1,000 genomic groups exhibit distinct differences in susceptibility or resistance to certain diseases that reflect their migratory history. One promising avenue of research concerns genetic changes associated with the incidence of auto-immune and inflammatory disorders: there is growing evidence that risk factors for some of these disorders are correlated with resistance to infectious diseases. This so-called yin and yang selection appears when adaptation to a pathogen in a given region is so advantageous that it spreads across the local population even though it has deleterious effects. A growing number of such links between alleles selected for resistance and risks for a range of common diseases have been identified, including auto-immune disorders, inflammation or psychiatric diseases such as schizophrenia and autism, according to a study led by Elinor Karlsson at the Harvard University FAS Centre for Systems Biology [2]. "Recent work in genomics is finding more and more examples of connections between pathogen resistance and common diseases, although in most cases the connection is not yet well understood", Karlsson said.

.....
"One of the best-known adaptations is the ability to digest lactose among Europeans, which enables them to consume milk from domestic animals"

He cited one example in the modern Bangladeshi population, where loci associated with inflammatory bowel disease (IBD) have been subject to selection during the past 5,000–50,000 years for resistance to cholera, which is a major cause of mortality in the region. "We also found signals of selection at many genes connected to a gene called IKBKG", Karlsson said. "Disrupting normal IKBKG function in mice causes an IBD-like disorder, suggesting these genes may act in IBD-related pathways. Our work clearly supports the hypothesis that past natural selection for pathogen resistance influences the incidence of common, immune-mediated diseases like IBD in populations today. We think finding and understanding the genetic changes underlying

these signals of selection could be a powerful tool for tackling a broad spectrum of such diseases. This is ongoing research, and we do not yet know whether selection for pathogen resistance is a cause of IBD". IKBKG codes for a protein known as NF- κ B essential modulator (NEMO), the regulatory subunit of the inhibitor of the I κ B kinase (IKK) complex, which is involved in inflammation, immunity and cell survival.

Another study showed a similar genetic link between different cancer types. A variant of the p53 protein, which binds to DNA to regulate the activity of a large number of genes, increases the risk of testicular cancer up to three times, but also appears to confer increased protection against skin damage and cancer from exposure to ultraviolet radiation in sunlight [3]. The authors, from the University of Oxford in the UK and the NIH, found that nearly 80% of White men carry the variant form of this gene, but only 24% of African men. According to Douglas Bell, one of the NIH authors, p53 responds to ultraviolet light by binding a specific sequence of DNA located in a gene called the KIT ligand oncogene (KITLG). This then stimulates melanocyte production, causing the skin to tan.

It is not yet known exactly how this particular G allele of p53 that is associated with increased risk of testicular cancer boosts tanning, the extent of positive selection among Europeans, or negative selection among Africans. "The question as to whether there is positive selection for the G allele in Europeans, or if there is a relaxation of negative selection and genetic drift in Africans, is still an open question", Bell said. "We hope that follow up studies will resolve this, because current methods do not distinguish between various evolutionary scenarios". The common hypothesis is that the ability to tan in middle latitudes confers resistance to skin cancer during summer months when the sun is strong, while the paling of skin during the winter allows greater absorption of the weaker solar radiation to enable production of vitamin D.

Another good example for adaptation to changing environments is the gene coding for haemoglobin in red blood cells, which seems to blur the line between immunity and non-immunity genes, according to Lluís Quintana-Murci, CNRS Research Director and head of the Human Evolutionary Genetics Unit at the Institut Pasteur in Paris.

“Strictly speaking, haemoglobin (HBB) is not an immunity gene, but as a matter of fact, it is involved in host defence to malaria, so for me, it is an ‘immunity gene’”, he said. Some of the haemoglobin alleles seem to have undergone “yin and yang” selection given the deleterious side effects of sickle cell disease, which causes red blood cells to assume a sickle shape. This shape change does reduce the symptoms of malaria, but can also cause anaemia by retarding the transmission of oxygen in the blood stream. People with just one copy of the sickle cell gene produce both normal and sickle-shaped red blood cells and have some resistance to malaria while being immune to sickle cell disease. However, people who are homozygous for the gene often develop anaemia.

Haemoglobin has also been subject to selective pressure for other functions during human migrations, for example in adaptation to high altitudes where the thinner atmosphere makes it harder to take in sufficient quantities of oxygen. Efficient oxygen transport becomes a considerable selective advantage at high altitude, to the extent that

human evolution has found different ways of reaching the same goal, according to Matteo Fumagalli from the Department of Integrative Biology at the University of California, Berkeley. He highlighted how populations in Tibet, upland areas of Ethiopia and the Andes in South America had all adapted to hypoxia differently, citing a recent paper suggesting that these groups had experienced convergent evolution among various genes on the hypoxia pathway [4], which responds to shortage of oxygen, including increased expression of haemoglobin (Fig 1).

Another selective force that has been working on humankind since its migration from Africa is changes to diet. One of the best-known adaptations is the ability to digest lactose among Europeans, which enables them to consume milk from domestic animals. Mammals have the ability to digest milk from their mothers during infancy through the enzyme lactase, but normally lose this as they are weaned on to solid foods. However, a mutation among Europeans for lactase persistence into adulthood

was a huge selective advantage because it helped them to survive through the winter in colder regions when food crops became scarce and they could drink the milk from livestock instead. In effect, the animals were converting one food, grass, which is largely indigestible to humans, into another, milk, which they could digest and metabolise efficiently, provided they were tolerant to lactose.

As humans migrated out of Africa, they encountered changing food sources that in turn generated adaptations in metabolic pathways, according to a recent study involving institutions from Europe, the USA and Russia [5]. The study identified various alterations associated with two major migration steps, firstly moving to areas rich in roots and tubers and then reaching Polar Regions lacking foods that contain plenty of carbohydrates. “One of our signals is for a specialization in roots and tubers, which are used by some hunter-gatherer groups and may be important fall-back foods”, said study author Anna Rienzo from the Department of Human Genetics at the University of Chicago. “Importantly, we found that the

Figure 1. High-altitude adaptation in humans is an example of convergent evolution.

biological pathways most enriched in signals of specialization for roots and tubers are starch and sucrose metabolism and folate biosynthesis, which makes sense because roots and tubers are rich in starch and poor in folates”.

It is also possible to gain new insights into historical migrations themselves through genetic analysis of current populations. This has been particularly successful in the case of disease resistance, because many pathogens evolved in tandem with the host. It gives the potential for mapping past human migrations by studying evolution of those pathogens and has occasionally led to significant discoveries that modified the prevailing orthodoxy. This happened as a result of looking at the diversity of the bacteria *Helicobacter pylori* [6], which inhabits the gastrointestinal tracts of at least 50% of the world's population. The bacteria evolved in Africa and had infected modern humans before they left the region 60,000 years ago. The study showed that *H. pylori* is about as old as anatomically modern humans, dating back to 116,000 years, and that there must have been a second wave of migration out of Africa around 52,000 years ago. This conclusion was reached because the current European *H. pylori* population is a hybrid between one now found in Asia and one in north-east Africa, with the latter emerging only 36,000–52,000 years ago, well after the original first migration. It suggests that modern Europeans mixed with people originating in north-east Africa, perhaps 10,000 years after the first major migration.

Around the time of these migrations, modern humans are thought to have encountered and interbred with Neanderthals, who were still present in Europe and

Asia, but had left Africa, or become extinct there. It now looks like interbreeding with Neanderthals played a significant role in human evolution and has left distinct traces in our genomes. Modern genomics has shed light on that interbreeding. On the one hand, it introduced some genes for immunity to pathogens that modern humans may not have previously encountered, along with adaptation to colder climates. On the other side, it seems that genomic areas relating to male fertility experienced strong selection against Neanderthal components. In particular, genes that are more expressed in testes than in any other tissue have had Neanderthal components heavily pruned. There is also about a fivefold reduction in Neanderthal ancestry on the X chromosome. These findings make sense because modern humans had previously diverged from Neanderthals about 500,000 years ago, and by the time they met again during the later migration of *Homo sapiens*, the two sub-species were close to reaching biological incompatibility. As a result, it may be that Neanderthal genes have been passed down largely through females, while male Neanderthal/human hybrids were mostly infertile.

.....

“It now looks like interbreeding with Neanderthals played a significant role in human evolution and has left distinct traces in our genomes”

.....

It is not yet clear whether these Neanderthal adaptations have any implications

for modern humans, but they are of great interest for elucidating our migratory past. Many of the other discoveries, though, have both potential therapeutic significance and could help to correct the notion of different human races. There is just one human race, but many clusters of diversity defined by different experiences of natural selection.

References

- Collins FS (2004) What we do and don't know about 'race', 'ethnicity', genetics and health at the dawn of the genome era. *Nat Genet* 36: S13–S15
- Karlsson EK, Kwiatkowski DP, Sabeti PC (2014) Natural selection and infectious disease in human populations. *Nat Rev Genet* 15: 379–393
- Zeron-Medina J, Wang X, Repapi E, Campbell MR, Su D, Castro-Giner F, Davies B, Peterse EFP, Sacilotto N, Walker GJ *et al* (2013) A polymorphic p53 response element in KIT ligand influences cancer risk and has undergone natural selection. *Cell* 155: 410–422
- Huerta-Sánchez E, DeGiorgio M, Pagani L, Tarekegn A, Ekong R, Antao T, Cardona A, Montgomery HE, Cavalleri GL, Robbins PA *et al* (2013) Genetic signatures reveal high-altitude adaptation in a set of Ethiopian populations. *Mol Biol Evol* 30: 1877–1888
- Hanock AM, Witonsky DB, Ehler E, Alkorta-Aranburu G, Beall C, Gebremedhin A, Sukernik R, Utermann G, Pritchard J, Coop G *et al* (2010) Human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. *PNAS* 107(Suppl 2): 8924–8930
- Moodley Y, Linz B, Bond RP, Nieuwoudt M, Soodyall H, Schlebusch CM, Bernhöft S, Hale J, Suerbaum S, Mugisha L *et al* (2012) Age of the Association between *Helicobacter pylori* and Man. *PLoS Pathog* 8: e1002693