Review

Cell PRESS

Human evolutionary genomics: ethical and interpretive issues

Joseph J. Vitti[1,2], Mildred K. Cho[3], Sarah A. Tishkoff[4] and Pardis C. Sabeti[1,2]

[1] Broad Institute of MIT and Harvard, Cambridge, MA, USA
[2] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA
[3] Stanford Center for Biomedical Ethics, Stanford University, Palo Alto, CA, USA
[4] Departments of Genetics and Biology, University of Pennsylvania, Philadelphia, PA, USA

Genome-wide computational studies can now identify targets of natural selection. The unique information about humans these studies reveal, and the media attention they attract, indicate the need for caution and precision in communicating results. This need is exacerbated by ways in which evolutionary and genetic considerations have been misapplied to support discriminatory policies, by persistent misconceptions of these fields and by the social sensitivity surrounding discussions of racial ancestry. We discuss the foundations, accomplishments and future directions of human evolutionary genomics, attending to ways in which the interpretation of good science can go awry, and offer suggestions for researchers to prevent misapplication of their work.

New possibilities for research on natural selection

Access to the genome has revolutionized the way that biology is done. Although the functional concept of the gene long pre-dates the discovery of the molecular structure of DNA, the ability to identify associations between genotypes (see Glossary) and phenotypes has opened up a host of new possibilities for research and medicine. With the development of sequencing technology during the 1970s and 1980s, the goal of unraveling the genome of an organism in its entirety became feasible. As methods continue to come of age, understanding of life, both its molecular basis and its genetic history, will become increasingly fine-grained.

Of course, with the tremendous scientific power that genomic technology affords comes an increased responsibility, and there is a wide range of ethical considerations for researchers who put humans under the microscope. On the one hand, there are immediate issues pertaining to the conduct of research, such as the problem of ensuring fully informed consent when genetic information can often be reused for purposes other than the original study [1]. On the other hand, and more broadly, there are issues concerning the relationship of society to genetics, such as the potential conflicts between personal choice and public health that may arise from the development of genetic interventions or enhancements [2].

The aforementioned issues are by now familiar to the scientific community and appropriately remain the subject of ongoing discussion. Recent advancements in computational methods, however, have opened up a whole new field of inquiry, evolutionary genomics, which poses some unique and previously unaddressed issues.

Glossary

Bottleneck: an event in the history of a population when the population size decreases sharply, resulting in a smaller and less diverse gene pool. This can cause an allele to increase in prevalence within the population as competing variants are wiped out.
Directional selection: a form of natural selection in which a given phenotype is favored (positive selection) or disfavored (negative selection).
Disruptive selection: a form of natural selection in which extreme values of a trait confer greater fitness than do intermediate values. When disruptive selection acts on a population, it results in increased diversity within the population.
Epigenetics: functional changes to the genome that can be inherited, but do not involve changes in the nucleotide sequence.
Epistasis: the process by which certain genes suppress or enhance the expression of other genes.
Eugenics: a social movement in which a community seeks to control its gene pool by regulating reproductive rights, encouraging those with traits seen as desirable to have children, and restricting the reproductive rights (often by exterminating) of those seen as undesirable.
Exaptation: occurs when a trait, originally selected for one function, is co-opted to perform a novel adaptive function.
Fixation: 100% prevalence of an allele within a population.
Genetic determinism: a fallacious misconception, according to which single genes give rise to higher-level traits in a straightforward, law-like fashion.
Genetic drift: an evolutionary process in which the frequency of an allele in a population changes not because of selective pressure, but because of random sampling.
Genetic essentialism: a fallacious misconception, according to which genetic differences between individuals or groups reflect inherent and immutable differences between the individuals or groups themselves.
Genetic reductionism: a fallacious misconception, according to which an individual's genome captures the sum of that individual's traits.
Genetic signatures of selection: observable changes in the DNA pattern that result from selective events. See Box 1 for an explanation of common genetic signatures.
Genotype: alleles that an individual carries, or the genetic makeup of an individual. It is contrasted with phenotype, which refers to the traits expressed in that individual.
Heterozygous: an individual who carries two different alleles (one from each parent) for a given trait.
Homozygous: an individual who carries two of the same allele (one from each parent) for a given trait.
Phenotype: traits observed in an individual, as opposed to the underlying genetic makeup (i.e. genotype).
Pleiotropy: occurs when one gene influences multiple phenotypic traits.
Reduced penetrance: the phenomenon when carriers of an allele fail to express the trait associated with that allele. It is typically caused by a combination of genetic and environmental factors.
Selection on standing variation: occurs when a selective pressure acts on alleles already existing within the population, rather than on alleles that arise by mutation.
Stabilizing selection: a form of natural selection in which intermediate values of a trait confer greater fitness than do extreme values. When stabilizing selection acts on a population, the result is decreased diversity.
Variable expressivity: the fact that individuals with a certain allele may have the traits associated with that allele to a greater or lesser degree.

Corresponding author: Vitti, J.J. (vitti@broadinstitute.org).

0168-9525/$ – see front matter doi:10.1016/j.tig.2011.12.001 Trends in Genetics xx (2012) 1–9

Evolutionary genomics is a powerful field of research, because it unites the disciplines of molecular and evolutionary biology. By applying statistical methods to the genome, one can investigate the history of individual genes and identify loci recently under natural selection [3–8]. This allows one to probe the landscape of evolution at an unprecedented level of efficiency and precision. At the same time, however, evolutionary genomics has the potential to reveal facts about the history of the human species that are too readily misinterpreted by a sensationalistic public, or warped to fit prejudicial agendas. Moreover, the dark history and heavy connotations surrounding the topics of human evolution and genetics heighten the need for an ethically self-aware research agenda, and for an open dialog among biologists, bioethicists and policy-makers.

In this article, we discuss the foundations and early accomplishments of evolutionary genomics. We then describe the ethical issues unique to this field, which inherits a tremendous potential for controversy from its subject matter and surrounding misconceptions. Historically, discoveries in evolution and genetics have been used as an unwitting rationale for discriminatory policies, further fueling the potential controversy and highlighting the need for a self-aware scientific agenda. We discuss the need for deliberate and precise communication of experimental results; in particular, researchers must attend to the way their research could be misconstrued when disseminated through non-scholarly media, and be proactive about preventing the misapplication of bona fide scientific discoveries. To that end, we discuss several concepts, such as ancestry and genetic causation, that are easily misread or assigned normative import by an untutored public. We suggest how researchers and science journalists can avoid these pitfalls, and note that these issues point to a more general need for a thorough discussion of evolution, genetics and surrounding ethical considerations in the public science curriculum.

Human evolutionary genomics: current progress and future directions

Natural selection can act on a population in several ways: it can favor extreme values of a trait over intermediate values (disruptive selection); it can favor intermediate values and so decrease diversity (stabilizing selection); or it can favor or disfavor a given phenotype (directional selection). There has been a recent expansion of computational methods focused on directional selection, specifically, positive selection, in which a beneficial gene is selected for in a population, centered around the selective sweep model. In this model, a favorable allele arises by mutation and then increases in prevalence until it sweeps to high prevalence or even fixation within the population. When it does, it takes with it the linked surrounding region of the genome. This pattern of selection leaves several signatures in the genome that can then be detected computationally (Box 1; reviewed in [9–14]).

The full impact of these new computational methods is still to be realized. Where scientists used to rely on speculative just-so-stories and a patchy archaeological record to support the hypothesis that a trait was adaptive, there are now the resources available to support such claims statistically. Moreover, these computational methods can be applied to the genome in full, allowing for high-throughput analysis, rather than case-by-case examination. Although there are several textbook examples of successful candidate gene studies, such as the discovery of a mutation in the gene encoding hemoglobin (*HBB*) that confers resistance to *Plasmodium falciparum* malaria in heterozygotes, but causes sickle cell anemia in homozygotes [15,16], this approach was typically hit-or-miss, and has now been replaced by whole-genome scans. These scans isolate genetic signatures left by probable selective events, revealing the most probable loci of selection for further investigation (in which the function and phenotypic impact of the gene is identified) [9,11]. In this way, current computational methods reflect a transition in evolutionary genetics from hypothesis-testing to hypothesis-generating science.

The development of statistical methods is only one part of the equation, however; the conceptual foundations of evolutionary genomics were laid some 40 years ago [17]. The central enabling factor for the application of genome-wide selection scans was the cataloguing of extensive maps of human variation. Undertakings such as the International HapMap Consortium and the 1000 Genomes Project have made vast amounts of genetic information publicly available, providing the necessary raw material for human evolutionary genomics to operate [18–22].

Already the ‘genomics revolution’ has given a first glimpse of the landscape of human evolution, verifying pre-genomic hypotheses and identifying new adaptive loci. In the early large-scale studies, scientists identified and examined novel adaptations for lactase persistence in Europe and Africa [23,24], skin pigmentation in Europe and Asia [3,25–29], the production of ear wax [26,30] and hair, sweat and tooth development [3,4,26,27] in Asia, and resistance to infectious disease in Africa [4,31]. Another intriguing discovery is the identification of a family of genes under selection in Tibetan populations that may help the body adjust to high-climate living [32]. These adaptations point to the importance of changes in diet, changes in climate, and infectious disease in driving human evolution, and illustrate the promise of evolutionary genomics in uncovering the history of the human species (Figure 1).

Although these accomplishments inspire confidence, they must also be kept in perspective. There are many genes still to be functionally characterized before ‘a detailed molecular, mechanistic, phenotypic and population genetics characterization of adaptive alleles’ is reached [17]; and even when that point is reached, researchers are still unlikely to have uncovered the whole landscape of human evolution. Current computational methods are best suited for newly arising mutations that undergo strong selection to high prevalence; statistical tests continue to be developed for weaker instances of selection or selection on standing variation. Larger and deeper data sets of human variation will help in this pursuit. Furthermore, most current methods are designed to identify single-locus adaptations, where many complex adaptive traits are likely to result from the interplay of multiple genes, as well as the environment. As methods continue to be refined, it will be helpful to keep in mind ways to identify epistatic interactions among genes. Moreover, there are

Box 1. Genetic signatures of selection

There are several current computational methods to detect genes under selection based on the selective sweep model of positive selection (Figure I). In this model, a beneficial mutation arises in an individual and rapidly increases in prevalence over generations until it reaches high prevalence or fixation (100% prevalence). The adaptive allele also brings with it nearby hitchhiker alleles on the same chromosome, leaving distinctive signatures of selection that stand out against the background of neutral genetic drift and can be detected computationally. Because these signatures are discussed thoroughly elsewhere, [5–14], we provide only a brief review here.

Reduction in genetic diversity
First, selective sweep events tend to reduce diversity in a genetic region. This is because as one allele rises to high prevalence, it also brings carbon copies of the surrounding regions to high prevalence, increasing homogeneity across the population in these regions. Consequently, long strips with unusually few single-nucleotide polymorphisms (SNPs) are a major indicator of a selective sweep.

High-frequency derived alleles
The hitchhiker regions surrounding an adaptive allele contain several other newly arisen alleles at nearby loci. These alleles are referred to as derived (as opposed to ancestral), and are generally rare (at low frequency) in the population, because it takes a long time for new alleles to rise in prevalence in the absence of selection. Under selection, these hitchhiker alleles sweep to high frequency or fixation along with the target of selection; thus, a surplus of derived alleles at high frequency in a region is a second signature of positive selection.

Long haplotype
As an allele proliferates over generations, random recombination events break down the long-range associations of that allele with other alleles (its haplotype). If an allele spreads rapidly (as under selection), not enough time may have passed for these associations to be broken down. Accordingly, regions with long haplotypes (longer than would be expected for their age) are indicative of selection.

Population differences
Relatively large differences in allele frequencies between populations provide further evidence for selection, as this would be expected if the allele were adaptive, and so selected, in one population but not the other.

Figure I. The classic selective sweep model. An advantageous mutation (in red) arises by chance and, over several generations, increases in prevalence. Nearby linked alleles on the chromosome hitchhike along with the mutation, including ancestral alleles (in gray) and derived alleles, or single nucleotide polymorphisms (SNPs; in blue). These hitchhiker regions leave distinctive patterns in the genome that can be detected computationally.

many other evolutionary processes that contribute to the phenotypes observed besides natural selection, including exaptation, genetic drift and demographic events, such as population bottlenecks. Epigenetic factors also play an important role in regulating gene expression and may have conspicuous evolutionary effects. In short, identifying and characterizing alleles that adhere to the simple selective sweep model is only the first step towards the ultimate goal of a complete understanding of the evolutionary history of the human genome. Nonetheless, it is a crucial first step and one that presents a timely opportunity to address associated ethical issues while its methodology is still young.

High ethical stakes for evolutionary genomics

Evolutionary genomics stands at the intersection of two sensitive topics that are widely misunderstood in their own right: evolution and genetics. Regrettably, advances in understanding of these topics have historically been misapplied to provide justification for unethical practices and, even today, scientific advancements are too easily warped to fit prejudicial agendas. Researchers investigating natural selection in the human genome need to be aware of the dark history surrounding these topics, which we describe below, and of the acute social sensitivity that still exists.

On the one hand, public misunderstanding and mistrust of the concept of evolution is widespread, largely because evolution is often mistaken to be progressive and teleological, as if with each generation nature were striving towards some fixed ideal [33]. This fallacy is implicit in misreadings of the term 'survival of the fittest', which Herbert Spencer described in 1864 as the 'preservation of favoured races in the struggle for life' [34]. Although it is now recognized that evolution is a blind process that does not favor certain

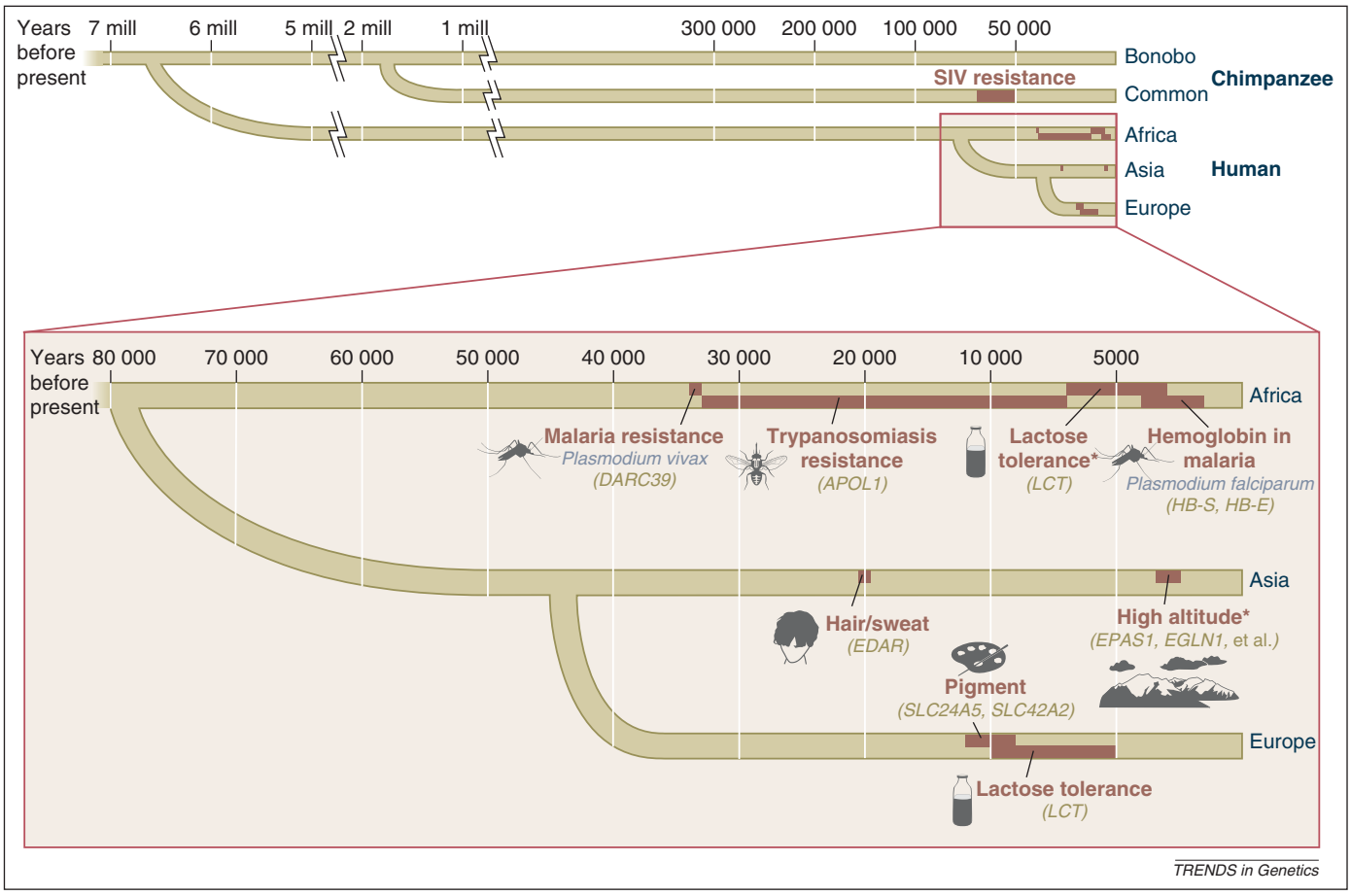

Figure 1. Phylogeny of human adaptive alleles characterized to date. The phylogenetic tree indicates the split between chimpanzee and human lineages (ca 5–7 million years ago), and the subsequent divergence of human populations (ca 50 000–80 000 years ago). This branched structure stands in contrast to misconceptions of evolution as a linear process, in which modern Europeans are the most evolved, with African populations (and then chimps) as their evolutionary predecessors. In actuality, contemporary human and chimpanzee populations are evolutionary cousins, sharing a common ancestor, but separated by distinct evolutionary trajectories. These evolutionary trajectories can be characterized by identifying the genetic adaptations that comprise them. Although recent genome-wide scans have generated many candidates to be investigated, only a handful of genetic adaptations have so far been functionally characterized and dated (error bars reflect current knowledge of when these alleles arose). Resistance to simian immunodeficiency virus (SIV) is conferred by C-C chemokine receptor type 5 (*CCR5*) variants found in common, but not bonobo, chimpanzees. In European populations, genes that affect skin pigmentation [solute carrier family 24 member 5 (*SLC24A5*) and solute carrier family 42 member 2 (*SLC42A2*)] have undergone positive selection. Signatures of selection for hair and sweat production [ectodysplasin A receptor (*EDAR*)] have been detected in Asian populations, and a family of genes [including endothelial PAS domain-containing protein 1 (*EPAS1*) and EGL nine homolog 1 (*EGLN1*)] enabling high-altitude living has been identified as adaptive in Tibetans. In African populations, genes that confer resistance to malaria and trypanosomiasis (sleeping sickness) have been selected [Duffy blood group, chemokine receptor 39 (*DARC39*), hemoglobin S and E (*HB-S* and *HB-E*) and apolipoprotein L, 1 (*APOL1*)]. Distinct variants in lactase (*LCT*), giving rise to lactose tolerance, have been selected in European and African populations. (n.b. alleles marked with a * are specific to subpopulations, not depicted on the phylogenetic tree: The African LCT variant for lactose tolerance arose in pastoralist populations in East Africa, and the Asian family of genes facilitating life at high altitudes arose in inhabitants of the Tibetan plateau.)

groups over others (except inasmuch as they may have differential reproductive success in a specific environment), this admittedly intuitive misreading has been used as justification for ideologies such as social Darwinism, in which welfare policies are suppressed to weed out the socially unfit, or eugenics [35,36]. Additionally, the notion that evolution is progressive has given rise to the false idea that modern African populations are ancestral and so less evolved than other racial groups (Figure 1). For example, in 1854, Nott and Glidden rendered their now-infamous depiction of the heads and skulls of a chimpanzee, an African male and the god Apollo, as an argument that Africans were an intermediate between human and simian (Figure 2 [37]). Misconceptions of this sort, which cut directly along racial lines, exacerbate public confusion and invite political bias, an unfortunate circumstance that surrounds evolutionary research. Researchers focusing on natural selection in humans must bear in mind the moral valence that audiences assign, often unconsciously, to evolution.

On the other hand, evolutionary genomicists must help audiences to avoid the pitfalls of common misconceptions of genetics, such as genetic reductionism, essentialism or determinism. Such oversimplified pictures of genetic causation are still abundant and have been used as justification for dismantling initiatives such as the Head Start Program in the USA [38]. This program provides for educational and social services for children from low-income families, and has been shown to be effective in improving social and cognitive development in participants [39]. However, findings of genetic variation in traits involved in cognitive function have been used to suggest that environmental factors, such as educational interventions, are not influential on phenotype [38]. Nothing that is known about genes supports such fatalistic attitudes, but they remain pervasive and

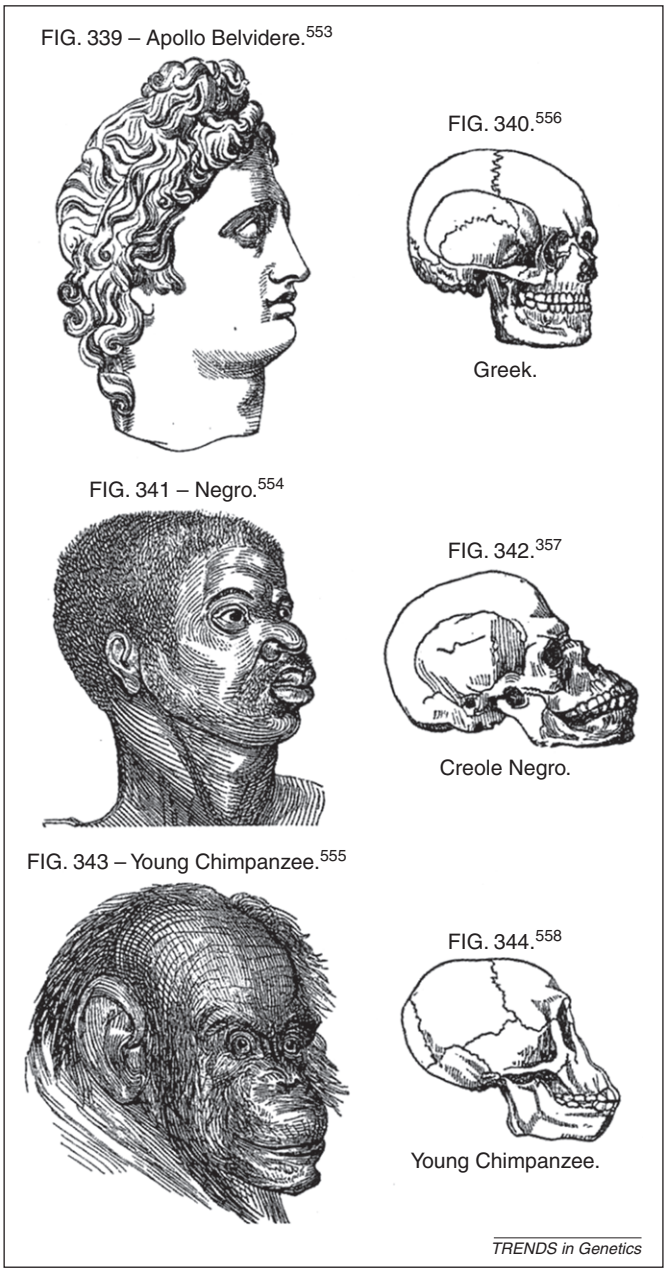

Figure 2. Nott and Gliddon's 'Types of Mankind'. Illustration from Josiah Clark Nott and George Gliddon's book *Types of Mankind,* a treatise that sought to demonstrate that the African race was wholly separate from (and inferior to) the Caucasian race [37].

provide further opportunities for the misapplication of discoveries in evolutionary genomics.

Given these dark histories, it is important for researchers in this field to prevent normative misreadings of their results, and to be sensitive to the potential controversy that may stem from the discovery that a given gene, implicated in the development or manifestation of some trait, underwent selection in a given population. Of concern, real differences might be found between populations with respect to genes that are implicated in the development of psychological–behavioral traits carrying some sort of social or moral valence, such as cognitive ability or tendency to violence. For example, recent publications suggested evidence for natural selection acting at two genes [microcephalin (*MCPH1*) and abnormal spindle-like microcephaly associated (*ASPM*)] found predominantly in Eurasian, but not West African, populations [40,41]. Because mutations in these genes are associated with microcephaly and non-progressive mental retardation, the authors suggested that these genes play a role in brain size and, furthermore, that they may have been recent targets of natural selection, a conjecture that inspired resistance on the one hand and racist conclusions on the other [42]. Neither the claim that these genes underwent recent selection nor the role of selected variants in regulating brain size has been supported by follow-up studies, but the controversy that ensued is illustrative both of the inflammatory nature of evolutionary genomics and of the need for scientific caution [43–46].

Although researchers might in fact discover such population genetic differences as those proposed to be at play with microcephalin and ASPM, the majority of the adaptations found to date are skin-deep, pertaining to metabolic function (e.g. lactase persistence), physical appearance (e.g. skin pigmentation) or disease susceptibility. This is perhaps unsurprising, given that most hominid evolution took place before the migration from Africa somewhere between 125 000 and 60 000 years ago. Recent adaptations within populations such as those current methods are apt to reveal are more likely to be quickly evolved, environment-specific add-ons. Higher-level behavioral tendencies are typically more complex from a genetic standpoint, evolving over a longer span of time, and so less likely to be candidates for recent adaptations. Additionally, the genetic or biological bases for such phenomena are notoriously difficult to tease apart from cultural effects, rendering the interpretation of any such experimental conclusions a delicate issue.

Nevertheless, genes might be discovered that are implicated in behavioral or cognitive traits under selection within certain populations. At present, there are not well-evidenced candidates, but many researchers are investigating the evolution of genes implicated in cognitive development, and one should be prepared for this possibility. Whether such discoveries could be used to justify changes in social policy is a question of enormous scope that we do not broach here, although we do offer some suggestions for discussion; most crucially, we emphasize that the relationship between any single allele and a complex trait is likely to be weak, with environmental factors typically playing a large role (Box 2). We acknowledge the general need for an open dialog about the implications of such findings, a topic that is often stifled because of its sensitive nature; we focus discussion on the prevention of misapplication of discoveries in human evolution more generally (irrespective of whether they concern behavioral or culturally inscribed traits).

Although the controversy of such genetic discoveries may be disheartening for researchers, as exemplified in the resignation of James Watson as chancellor of Cold Spring Harbor Laboratory over comments regarding the intelligence of African individuals, for instance, it is instructive to bear in mind that the facts themselves are not the problem [47]. Discoveries in evolutionary genomics are not normative in nature, and it is up to researchers to determine how they will inform or alter policy (if indeed at

Box 2. Human population differences and ethics

Similar to every other organism, humans have been subject to natural selection, not only in the deep past, but also in the 50 000–80 000 years since branching into continental subpopulations. Researchers may find that genes implicated in cognitive and behavioral tendencies have undergone selection in certain populations. How, if at all, would this impact social and ethical practices?

For example, debate is currently thriving regarding racialized medicine. Proponents hold that race is clinically useful as a proxy for genetic indicators of the efficacy of certain medicines, whereas opponents respond that utilizing racial identity metrics in this way risks legitimizing race as a biological, rather than sociological, phenomenon [61].

Similarly, one could use evolutionary genomics to argue for differential treatment in the form of affirmative action or genetic profiling. Suppose, for example, scientists discovered that certain populations had evolved a gene that was reliably correlated with superior performance on an IQ test. One could argue that individuals in that group should be judged more harshly in admissions programs because of a probable genetic advantage. Just as one is born into a given socioeconomic status that may help or hinder one's educational development, the results of the genetic lottery may put individuals at an advantage or disadvantage, and some might argue that these factors are relevant to an individual's merit. Such principles could also be extended to various social policies beyond institutional admissions.

The question of whether such discoveries could be used to justify changes in social policy is a question of enormous scope that we do not broach here. However, we recapitulate the importance of treating individuals as individuals, recognizing that racial membership is, at best, a sloppy proxy for the presence of genes of interest and, furthermore, that these genes are not foolproof indicators of their correlated traits; in the case of behavioral tendencies in particular, the environment is a major factor.

More generally, it is worth reaffirming that no such scientific finding could justify, much less obligate, discriminatory practices in a democratic society. Even if all men are not created equal in some evolutionary or genetic sense, that does not delegitimize the practice of affording moral worth equally to all persons. Discoveries in evolutionary genomics, as elsewhere in science, are not normative, and so do not commit us to any policy; it remains to us to decide, through conscientious bioethical discourse, how what is learnt will inform ethics.

all). There is a lot of useful information to be found through human evolutionary genetics; what is crucial, and what becomes the major ethical challenge for researchers, is ensuring interpretative rigor and preventing the knowledge acquired from being warped towards harmful ends.

To prevent misapplication of evolutionary genomics results, researchers should be vigilant in three areas: in their own methodology, in the dissemination of results (both through scholarly and lay media) and in their role steering public discourse about science.

Responsible conduct of research: establishing a self-aware scientific agenda

Discoveries in human evolutionary genomics are often high impact, on the one hand because of a common interest in the history of the human species and, on the other, because these discoveries are accessible to a broad audience without technical knowledge of the methods at play. Because this research is so easily sensationalized, it is critical for researchers to ask questions about what counts as evidence, so as to establish methodological rigor (reviewed in [17]).

Computational sweeps of the genome are a powerful tool, and it is tempting to overinvest significance in tentative results. Not all outlier regions are adaptive, however, and the candidate loci generated by computational sweeps must be investigated before a case of natural selection can be confirmed. First, researchers should examine alternate hypotheses for the genetic patterns observed, such as population bottlenecks and other selection-mimicking processes. Moreover, researchers must pursue follow-up study of the genes targeted as adaptive to determine their biological function (e.g. through case and/or control comparison studies, the development of cell lines with the target gene, or *in vivo* transgenic studies), rather than forming conjectures as to the function of the gene based on pathological variants at the locus in question. Once the case for the selection and biological function of the allele has been secured, researchers can go further by examining the ecological history of the population and trying to find evidence of the selective pressure responsible for the adaptation, to yield a comprehensive picture of the adaptive allele [48].

Dissemination of results: problematic concepts

In summarizing experimental results through abstracts and press releases, it is helpful to keep a broad audience in mind, and to use non-technical language wherever possible. Both in the abstract and in the main text, moreover, author interpretations should be acknowledged as such through the use of cautious, non-sensationalistic language. In particular, studies that are likely to attract media attention should include explanations of the limitations of study design in the main text, and conclusions in language that is precise and accessible to a lay audience.

In many cases, however, communication of scientific results goes awry at the stage of the press, perhaps owing to a lack of direct collaboration between journalists and researchers. To prevent this, journalists can reach out to the researchers to supplement their understanding beyond the press release, and researchers and institutional public relations (PR) offices can take care to ensure that those press releases are both comprehensive and cautious, explaining the significance and limitations of the findings in appropriate and non-sensationalistic lay terminology [49–51]. To ensure terminological precision, researchers and journalists must attend to the most easily abused or misinterpreted concepts, and opt to use more specific language wherever possible. Below, we discuss some of the most problematic concepts and their related terminology.

Genetic and evolutionary causation

One well-known fallacy of scientific reporting involves mistaking correlation for causation. More generally, there is a temptation to oversimplify causal stories, by glossing a contributing factor of a phenomenon as its sole cause, for instance. One recent study finds evidence for the hypothesis that evolutionary increases in brain size are related to an increased capacity for physical exercise [52]. Where the

authors are appropriately cautious in the research paper, describing selection for brains that can support exercise as a probable contributing factor (along with several already well-known hypotheses, such as the need for complex social cognition), media reports of the study employ language suggesting that brain size increased exclusively because of the need for more complex locomotion [53,54].

Besides the use of all-or-nothing pictures of causation, media reports of genetic studies often invoke the notion that humans are genetically hardwired, because it is concise and intuitive to a lay audience. However, such wording suggests a law-like, perfect, or inflexible correlation between the presence of a gene and the trait in question, where in fact the majority of phenotypes result from the interaction of multiple genes and the environment. Even in the absence of such explicit language, it is tempting to propagate genetic explanations of social phenomena, such as media reports of the 'warrior gene' as a cause of violence in the Maori, without qualifying that such phenomena are also influenced by social factors, such as economic disparities and discrimination [55]. More generally, although it is intuitive to think of the genome as a blueprint for an organism, written in a code that can be deciphered, this hints at too straightforward a relationship between the genotype and phenotype. A more accurate metaphor depicts the genome as a recipe, in which the end product is more than the sum of its parts.

Even when a trait stems from a single gene, that gene may have reduced penetrance, variable expressivity or pleiotropic effects. Given this intricacy, in most cases it is better to describe genes as associated with traits, or implicated in their development, and the presence of a gene may be an indicator or predictor of that trait (with a certain margin of error; even Mendelian genes are susceptible to environmental influence). In the case of evolutionary genomics, this worry becomes especially acute, because to report that a population group has some trait genetically 'hardwired' implies not only that the trait is Mendelian and inflexible, but also that it has swept to complete fixation within that group, which may or may not be the case.

Lay concepts: innateness and intelligence

In the cases where a robust correlation is found between the presence of a gene and of a trait, there is a temptation to say that trait is innate for carriers of that gene. As with deterministic concepts, the concept of innateness is concise and intuitive, and so tempting to employ, but it is lacking in precision. Scientists use 'innate' in several different senses: present at birth, exhibiting developmental fixity, shared by all members of a certain species or taxonomic group, instinctual or unlearned, and so on [56]. However, because the connotations of inevitability carried by the concept of innateness invoke misconceptions such as genetic determinism, it is better for scientists and writers to use more specific language, especially in headlines and précis. It would be imprecise, for example, to say that high-altitude living is innate for Tibetan individuals; it is more straightforward (and less open to interpretation) to say that these individuals have genetic adaptations that confer the ability to withstand high altitudes.

Similarly, intelligence is a cluster concept that has historically proven dangerous for evolutionary researchers. Studies that seek to investigate general cognitive ability should be explicit about the variables they take to be representative of intelligence (e.g. brain size, response time, performance on an IQ test, etc.), or should specify what form of intelligence they are investigating (logical reasoning, abstract reasoning, emotional intelligence, etc.). More generally, the use and abuse of lay concepts such as intelligence highlight the need for researchers and journalists to invoke clearly defined variables when reporting results, rather than imposing conceptual interpretations.

Biological concepts made normative

Although it should be obvious to researchers that terms such as 'innate' and 'intelligence' are easily misconstrued, there is also a family of concepts indispensable to evolutionary biology that are often assigned a moral valence. Researchers and journalists should be aware of the normative connotations that these words carry, and should take care to deflate them if an audience runs the risk of misinterpretation.

The term 'fitness' for example, and the notion that it is the fittest who survive, is sometimes assigned a value that appears to lend credence to the aforementioned idea that evolution strives towards something better. On top of this, 'fitness' can be misunderstood because it has been used historically to apply to socially valued characteristics (e.g. encouraging fitter families to reproduce). Conversely, the term 'mutation' is sometimes mistaken to represent something undesirable or subideal, whereas in its scientific usage it carries no such connotations [57].

Particularly problematic for evolutionary genomicists is the concept of ancestry, because of its racial implications. The idea that apes were part of an evolutionary spectrum in a direct line with Europeans at the most recently developed end prevailed for decades in popular iconography and still exists [37]. Thus, some may believe that because modern humans originated in Africa, modern Africans are 'ancestral'. In reality, all human populations have had similar time to adapt since separation from a common ancestor within the past approximately 100 000–200 000 years [58]. Indeed, humans and chimpanzees have similarly both been adapting from a common ancestor 5–7 million years ago. For example, a recent comparative study of 200 disease-associated mutations in the human, chimpanzee and rhesus macaque genomes indicated that, in 84 cases, the rhesus variant matched the disease-related human, but not the chimpanzee variant suggesting at these variants that chimpanzees evolved away from the ancestral condition at those sites [59,60]. The notions that chimpanzees are the ancestors of humans and that Africans are the ancestral human populations are false: rather, modern Africans descend from a population that is ancestral to all humans. As such, these modern populations can be considered distant cousins evolving from the same common ancestor.

Science education and public discourse

That these morally neutral concepts of biology are so readily imbued with value suggests that part of the issue with the misapplication of evolutionary genomic discoveries is insufficient public understanding of the topics at play.

Besides contributing to the constant push for improved science education, researchers can make an impact by leveraging public interest in evolutionary genomics to create opportunities to teach. Although the attention that this research draws can engender controversy, it can also provide a forum for researchers to address misconceptions head-on.

Of course, the push to improve public understanding of science involves more than the fundamental concepts: it is also crucial for researchers and educators in the life sciences to engage audiences with the applications of these concepts so that they can appreciate the depth of ethical issues that subsequently arise.

Concluding remarks

Genome-wide scans have generated several promising directions for further investigation, and as we stand on the brink of all these potential breakthroughs in understanding our own evolution, it is more crucial than ever to hold these projects at arm's length and examine the genuine ethical dilemmas they inspire. By cultivating an early awareness of the issues surrounding evolutionary genomics, notably, the gap between the concepts that researchers employ and the public's understanding of these concepts, one can ensure a more productive research agenda and avoid the pitfalls surrounding this promising but easily misunderstood field.

References

1 Caulfield, T. *et al.* (2008) Research ethics recommendations for whole-genome research: consensus statement. *PLoS Biol.* 6, e73
2 Buchanan, A. *et al.* (2000) *From Chance to Choice: Genetics and Justice*, Cambridge University Press
3 Williamson, S.H. *et al.* (2007) Localizing recent adaptive evolution in the human genome. *PLoS Genet.* 3, e90
4 Sabeti, P.C. *et al.* (2007) Genome-wide detection and characterization of positive selection in human populations. *Nature* 449, 913–918
5 Biswas, S. and Akey, J.M. (2006) Genomic insights into positive selection. *Trends Genet.* 22, 437–446
6 Kelley, J.L. *et al.* (2006) Genomic signatures of positive selection in humans and the limits of outlier approaches. *Genome Res.* 16, 980–989
7 Nielsen, R. *et al.* (2007) Recent and ongoing selection in the human genome. *Nat. Rev. Genet.* 8, 857–868
8 Grossman, S.R. *et al.* (2010) A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science* 327, 883–886
9 Bamshad, M. and Wooding, S. (2003) Signatures of natural selection in the human genome. *Nat. Rev. Genet.* 4, 99–111
10 Kreitman, M. (2000) Methods to detect selection in populations with applications to the human. *Annu. Rev. Genomics Hum. Genet.* 1, 539–559
11 Nielsen, R. (2005) Molecular signatures of natural selection. *Annu. Rev. Genet.* 39, 197–218
12 Ronald, J. and Akey, J. (2005) Genome-wide scans for loci under selection in humans. *Hum. Genomics* 2, 113–125
13 Sabeti, P. *et al.* (2006) Positive natural selection in the human lineage. *Science* 312, 1614–1620
14 Vallender, E. and Lahn, B. (2004) Positive selection on the human genome. *Hum. Mol. Genet.* 13, R245–R254
15 Kwiatlowski, D.P. (2005) How malaria has affected the human genome and what human genetics can teach us about malaria. *Am. J. Hum. Genet.* 77, 171–192
16 Cyrklaff, M. *et al.* (2011) Hemoglobins S and C interfere with actin remodeling in *Plasmodium falciparum*-infected erythrocytes. *Science* 334, 1283–1286
17 Akey, J.M. (2009) Constructing genomic maps of positive selection in humans: where do we go from here? *Genome Res.* 19, 711–722
18 The International HapMap Consortium (2005) A haplotype map of the human genome. *Nature* 437, 1299–1320
19 The International HapMap Consortium (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449, 851–861
20 Spencer, G. (2008) *International Consortium Announces the 1000 Genomes Project*, NHGRI
21 The International HapMap Consortium (2010) Integrating common and rare genetic variation in diverse human populations. *Nature* 467, 52–58
22 The 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. *Nature* 467, 1061–1073
23 Bersaglieri, T. *et al.* (2004) Genetic signatures of strong recent positive selection at the lactase gene. *Am. J. Hum. Genet.* 74, 1111–1120
24 Tishkoff, S.A. *et al.* (2007) Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* 39, 31–40 [see comment]
25 Voight, B.F. *et al.* (2006) A map of recent positive selection in the human genome. *PLoS Biol.* 4, e72 [see comment] [erratum appears in *PLoS Biol.* (2006) 4, e154]
26 Kimura, R. *et al.* (2007) A practical genome scan for population-specific strong selective sweeps that have reached fixation. *PLoS ONE* 2, e286
27 Tang, K. *et al.* (2007) A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biol.* 5, e171
28 Lao, O. *et al.* (2007) Signatures of positive selection in genes associated with human skin pigmentation as revealed from analyses of single nucleotide polymorphisms. *Annu. Rev. Hum. Genet.* 71, 354–369
29 Norton, H. *et al.* (2007) Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Mol. Biol. Evol.* 24, 710–722
30 Yoshiura, K. *et al.* (2006) A SNP in the *ABCC11* gene is the determinant of human earwax type. *Nat. Genet.* 38, 324–330
31 Genovese, G. *et al.* (2010) Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* 329, 841–845
32 Yi, X. *et al.* (2010) Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 329, 75–78
33 Bechtel, W. (1986) Teleological functional analyses and the hierarchical organization of nature. In *Current Issues in Teleology* (Rescher, N., ed.), pp. 26–48, University Press of America
34 Spencer, H. (1865) In *Principles of Biology* (vol. 1), Williams and Norgate
35 Dickens, P. (2000) *Social Darwinism: Linking Evolutionary Thought to Social Theory*, Open University Press
36 Kevles, D. (1998) *In the Name of Eugenics: Genetics and the Uses of Human Heredity*, Harvard University Press
37 Nott, J. and Glidden, G. (1854) *Types of Mankind; or, Ethnological Researches Based Upon the Ancient Monuments, Paintings, Sculptures, and Crania of Races*, Lippincott, Grambo & Co
38 Rushton, J. and Jensen, A. (2005) Thirty years of research on race differences in cognitive ability. *Psychol. Public Policy Law* 11, 235–294
39 Love, J. *et al.* (2005) The effectiveness of early Head Start for 3-year-old children and their parents: lessons for policy and programs. *Dev. Psychol.* 41, 883–901
40 Evans *et al.* (2005) Microcephalin, a gene regulating brain size, continues to evolve adaptively in humans. *Science* 309, 1717–1720
41 Mekel-Bobrov, N. *et al.* (2005) Ongoing adaptive evolution of *ASPM*, a brain size determinant in *Homo sapiens*. *Science* 309, 1720–1722
42 Derbyshire, J. (2005) The specter of difference. *National Review Online,* 7 November
43 Mekel-Bobrov, N. *et al.* (2007) The ongoing adaptive evolution of ASPM and Microcephalin is not explained by increased intelligence. *Hum. Mol. Genet.* 16, 600–608
44 Yu, F. *et al.* (2007) Comment on 'Ongoing adaptive evolution of *ASPM*, a brain size determinant in *Homo sapiens*'. *Nature* 316, 370
45 Timpson, N. *et al.* (2007) Comment on papers by Evans *et al.* and Mekel-Bobrov *et al.* on evidence for positive selection of *MCPH1* and *ASPM*. *Science* 24, 1036
46 Rushton, J.P. *et al.* (2007) No evidence that polymorphisms of brain regulator genes *microcephalin* and *ASPM* are associated with general mental ability, head circumference or altruism. *Biol. Lett.* 3, 157–160
47 Cold Spring Harbor Laboratory (2007) *Dr. James D. Watson retires as chancellor of Cold Spring Harbor Laboratory*, Cold Spring Harbor Laboratory
48 MacCallum, C. and Hill, E. (2006) Being positive about selection. *PLoS Biol.* 4, e87

49 Bubela, T.M. and Caulfield, T.A. (2004) Do the print media hype genetic research? A comparison of newspaper stories and peer-reviewed research papers. *Can. Med. Assoc. J.* 170, 1399–1407
50 Brechman, J. *et al.* (2009) Lost in translation? A comparison of cancer-genetics reporting in the press release and its subsequent coverage in the press. *Sci. Commun.* 4, 453–474
51 Ransohoff, D.F. and Ransohoff, R.M. (2001) Sensationalism in the media: when scientists and journalists may be complicit collaborators. *Eff. Clin. Pract.* 4, 185–188
52 Raichlen, D.A. and Gordon, A.D. (2011) Relationship between exercise capacity and brain size in mammals. *PLoS ONE* 6, e20601
53 Mossop, B. (2011) Larger brains may have evolved due to sports, not smarts. *Wired*, 21 July
54 Anon. (2011) Bigger brains evolved due to capacity for exercise. *Physorg*, 4 August
55 Anon. (2006) Warrior gene blamed for Maori violence. *9News MSN*, 8 August
56 Griffiths, P.E. (2001) What is innateness? *Monist* 85, 70–86
57 Wen, P. (2000) Jews fear stigma of genetic studies. *The Boston Globe*, 15 August, F1
58 Tishkoff, S.A. and Verrelli, B.C. (2003) Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annu. Rev. Genomics Hum. Genet.* 4, 293–340
59 Pennisi, E. (2007) Genomicists tackle the primate tree. *Science* 316, 218–221
60 Rhesus Macaque Genome Sequencing and Analysis Consortium *et al.* (2007) Evolutionary and biomedical insights from the rhesus macaque genome. *Science* 316, 222–234
61 Stein, R. (2011) Race reemerges in debate over personalized medicine. *Washington Post*, 31 July