HACETTEPE JOURNAL OF BIOLOGY AND CHEMISTRY

Review Article

Hacettepe J. Biol. & Chem., 2010, 38 (3) 193-220

Genetic Footprints of Natural Selection and Drift in Human Evolution

Mehmet Somel[1] and Efe Sezgin[2]

[1]CAS-MPG Partner Institute for Computational Biology SIBS, 320 Yue Yang Road, Shanghai, 200031, China

[2]Laboratory of Genomic Diversity, National Cancer Institute, Frederick, MD 21702, USA

Abstract

Humans are genetically little diverged from their closest living relatives, chimpanzees. However, human anatomy, physiology and behavior show substantial divergence from those of chimpanzees and other primates. The evolutionary processes that generated this phenotypic divergence are still debated. Given the incomplete state of the hominid fossil record, genetic information is indispensible for studying human evolution. Recently, data from genome sequencing projects, gene expression profiling, and large-scale genotyping across multiple populations, has significantly improved our understanding of human origins, as well as the forces of selection, adaptation, and demographic forces that shaped human evolution. Recent multispecies comparative genomic studies indicate that random genetic drift, *i.e.* neutral evolution, is a leading evolutionary mechanism in shaping human protein coding sequences, as well as gene expression patterns. Still, there is also compelling evidence that a large number of human genes were affected by adaptive evolution. Genes involved in immunity, sensory perception, reproduction, and apoptosis appear among the most frequently positively selected classes. Likewise, comparative transcriptome studies indicate adaptive expression changes in human brain gene expression. Importantly, variants of both positively and negatively selected genes are frequently found to be responsible for human genetic diseases. Comparison of genetic diversity within and between human populations suggests that humans have been a relatively small, homogeneous species. Accordingly, patterns of genetic diversity are to large extent shaped by neutral, demographic processes. Meanwhile, genome scans and functional studies identify pathogens, diet, and environmental conditions as major selective forces driving genetic diversity in humans. These results portray human evolution as a complex process, simultaneously affected by forces of selection and drift.

Key Words Human evolution, Comparative genomics, Natural selection, Drift, Gene expression evolution

INTRODUCTION

Questions on humankind's place in nature, the history of human species, and the basis of human variation have intrigued people at least since Ancient Greek times [1,2]. But until recently, human evolution was a highly speculative topic. Studies in this field were limited to anatomical and behavioral ape-human comparisons, and analyses of a scant fossil collection. Even with the advent of molecular biology in the second half of the 20th century, human evolutionary genetic studies remained in a primitive state compared to research in model organisms, such as the fruit fly. The publication of first

draft of human genome in 2001 [3] was a turning point in all fields of human genetics, initiating the "genomic" era. Today, armed with high-throughput molecular technologies, we are able to compare the human genome with those dozens of other species, or to characterize human genetic diversity in large-scale, at both the individual and population level.

Why is this exciting? First, from a practical point of view, learning our genome's history is indispensable for studying genetic mechanisms and disease. For instance, the main tool for identifying functional elements in the genome is to measure sequence conservation across different species. To understand the causes of geographical variation in the frequency of a genetic disease, we also need to learn the history of human migration, adaptation and drift.

Second, genetic information is a prolific source for learning history. Given the paucity of the ape and human fossil record, the reconstruction of human history, either ancient or recent, would be very difficult, if not impossible, without molecular data. As we will describe below, studying molecular evolution can provide answers to age-old questions on our origin, our natural niche, and the sources of human diversity.

Analyses of human genetic data also shed light onto fundamental questions regarding general mechanisms of evolution. For example, Darwin and subsequent evolutionary biologists thought of evolution as a purely adaptive process. Today, we know that our evolutionary history is much more complex. Processes including drift, demographic changes, and mutational biases have substantial roles in shaping our genotype. Thus, the current mission is understanding the relative roles of these various processes in evolution.

This paper aims to provide a broad view on human evolution from a genetic point of view. We will focus on human-chimpanzee divergence, review hypotheses on the evolution of human specific traits, and review genetic and transcriptomic evidence on these hypotheses. Finally, we will summarize work on recent human evolution and population genetics. In each section, we will briefly discuss future prospects of that field. The findings will further be evaluated in light of past and ongoing discussions in evolutionary biology, including the selectionist/neutralist debate. For further detailed information on these comprehensive topics, the reader to is encouraged to refer to specific reviews, which will be mentioned in the text.

The human phenotype and hypotheses on human evolution

Today *Homo sapiens* is among the most dominant species on earth, in terms of its geographic distribution, position in the food chain, or impact on global ecology. This owes to our general capacity to systematically modify our environments [4-6] {reviewed in [7]}. Amazingly, our closest relative, the chimpanzee, does not exhibit such a capacity [8]. This has lead to the view that, evolutionary changes that occurred on the human lineage since the human-chimpanzee common ancestor, were more

dramatic than those on the chimpanzee lineage; *i.e.* humans changed more than chimpanzees [9, 10]. But was human evolution really special?

Human-specific characters

Humans' closest relatives are the great apes, that is, chimpanzees, gorillas and orang-utans. Apes and humans are more distantly related to Old World Monkeys {reviewed in [11]}. Compared to primate characteristics, human phenotypes may be classified in three categories: *(i)* traits unique to the human, *(ii)* traits for which the human form is found at the extreme end of a distribution among primates, *(iii)* traits for which humans are only average apes. Below we will describe the first and second categories, although third category traits are also plentiful: for example, human male relative testis size is only average among the great apes [12].

The most conspicuous human-specific characteristics are bipedalism and morphological innovations related to it, such as long legs and changes in backbone and pelvic structure [13]. Similarly unique anatomical features include the opposable thumb [13] (chimpanzees cannot grip objects as firm as humans), small jaws, small teeth, a prominent chin and nose [14], and reduction of terminal hair on the body [15]. In addition, the low position of the human pharynx and vocal communication are traits not observed in the great apes. Other intriguing morphological characters include the wide distribution of eccrine sweat glands [16] the white sclera of the eyes (allowing others to follow a gaze) [17], a rotating birth canal trajectory [18], the loss of the penis bone [19], as well as female reproductive crypsis (chimpanzee females publicize their ovulation by sexual swellings) [20]. Human-specific physiological characters include enhanced T-cell activation [21,22] and the abrupt human menopause [23]. The human diet is also unique, as it diverges from other apes' diets with its high energy and protein content (due to regular meat and tuber eating), as well as a shift towards cooked food [24].

The second category contains characteristics for which the human bears the most extreme form/level among primates. For instance encephalization: human body size is on average twice as large than the chimpanzee, while the human brain is three times larger [25]. Tool production and language may also be considered 'extreme features', given that tool making and non-vocal communication are also observed in apes [26-28]. However, the extent to which these human-ape differences are qualitative or quantitative is still a matter of debate [8]. Similarly, humans show extreme cooperative and altruistic behaviour [29-31], as well as superb social communicative abilities compared to chimpanzees [17] {reviewed in [32]}. To these traits, we may add different life-history patterns, including an extended gestation time, longer juvenile period, and delayed maturation [14,33,34] and a remarkably long lifespan [35-37]. These extended life-history patterns contrast with short weaning and frequent reproduction in humans (chimpanzee weaning can take 5 years) [34], and the immature-helpless state of human neonates, termed 'secondary altriciality' [38].

Hypotheses on human evolution

Under what conditions these characters have evolved? In most cases, the answer is not yet known. Still, there is plenty of room for speculation.

First, some characters, such as life-history traits, appear directly linked to tool use and cultural accumulation, and in general, to niche construction [4]. In fact, human longevity or encephalization may be considered outcomes of an extreme K-strategy – a life-history strategy that is already pronounced in primates relative to many other mammals [14,39,40]. Still other characters may have been adaptations to past human environments, especially the savannah, since this was a novel niche for species on the human lineage [41].

There are numerous, more particular theories on the evolution of human-specific characters. Especially during the 20th century a large number of such hypotheses were proposed. Some of these attempted to account for multiple changes in human evolution through a single ecological or genetic alteration. One example is the once popular “aquatic ape” hypothesis [42], which tried to explain hairlessness, swimming ability, and bidepalism in humans by a hypothetical aquatic life-style of human ancestors. In turn, a whole plethora of morphological, life-history and behavioural characters, from the human chin, small toes, encephalization, to curiosity, were attributed to delayed human development and the similarity of humans to juvenile apes - hence the human neoteny hypotheses [14, 43].

There are theories linking meat eating, encephalization, extended childhood and longevity [33,44]. Similarly, the dual pressure of human encephalization and bipedalism is thought to underlie the early birth of human babies, and further, the complications of human birth is hypothesized to promote social cooperation among females [38]. Meanwhile, human cognitive capacity develops dramatically during childhood, and this is claimed to be due to specific human social abilities enabling infants to acquire concentrated knowledge from society [17,32]. These social abilities include teaching, strong and efficient imitation, and normative behaviour, as well as focusing on processes rather than objects in learning [8].

Others hypotheses have addressed a single character or a related list of characters. The “expensive tissue” hypothesis accounts for increase in brain and decrease in gut size by an increase in human diet quality [45]. The grandmother hypothesis accounts for long post-menopausal lifespan by mother-child food sharing and grandmother contribution to progeny [23]. The “cooling” hypothesis explains sweat glands and hairlessness as an adaptation to the savannah environment {D. Falk’s comment to [45]}, whereas the “parasitic load” hypothesis explains hairlessness by parasite pressures [46]. Sexual selection has been employed to explain a diverse array of characters, from hairlessness to language [1].

What the fossil record tells

Until recently, most evidence available for the study of human evolution derived from comparative morphology {e.*g.* [47-49] and palaeontology {*e.g.* [50,51]}. Intense work in these fields during the 20th century, supplemented by initial genetic studies, provided an outline for evolution on the human lineage {reviewed in [11,20,52]}. Today, we know that human and chimpanzee lineages diverged about ~6 million years ago (mya) in Africa. Fossils assigned to the human lineage are predominantly found in East Africa, whereas chimpanzees of today are found in West Africa, suggesting this geographical separation played role in human speciation. It seems plausible that our common ancestor lived in a forest environment, as extant great apes. Notably, chimpanzee ancestors have left very few fossils; we therefore have little information about their evolution [53].

One of the first species assigned to the human branch is the partially bidepal *Ardipithecus* [54], which lived about 4-6 million years ago, also in a forest environment. In contrast, subsequent *Australopithecine* species occupied the savannah, were small, bipedal, and had brains the same size as current chimpanzees (~500 cm^3). Around 2 million years ago *Homo erectus* species began to appear with two striking novelties: an unprecedented brain size (>1000 cm^3) and a stone tool kit, suggesting a correlation between tool use, brain growth and possibly meat eating [33]. The earliest records of hominins (species on the human and chimpanzee lineage) outside Africa also belong to *H. erectus*.

The first modern *Homo* sapiens remains are found in Africa around 200,000 years ago. Recent work suggests that modern humans had a slower rate of development, and possibly lived longer, compared to Neanderthals and other hominins [51,55,56]. Within this period we also encounter first direct evidence of symbolic human behaviour, such as cave paintings [57]. Archaeological records further indicate that human groups started to spread out of Africa ~50,000 years ago. This period also saw the demise of *Homo neanderthalensis* and possibly other *Homo* species [58]. Indeed, within this 6 million years, many *Australopithecus* and *Homo* species emerged, but all -except for *Homo sapiens*- eventually went extinct, under conditions which are uncertain.

In addition to archaeological work, studies of contemporary and historical pre-industrial societies also suggest contributions to recent human evolution [37,59]. For example, one study found support for the grandmother hypothesis by estimating positive fitness effects on grandchildren of the presence of a grandmother, based on historical records [60].

Nevertheless, reconstructing human evolution using solely archaeological and historical data has its difficulties. There are big gaps in the archaeological record, and fossils provide limited information about physiology and behavioural aspects of extinct ancestors. Moreover, determining which extinct taxa are *Homo sapiens*' direct ancestors is itself an arduous, if not impossible, task. This renders genetic information indispensable for unravelling our past.

Human genetics: an evolutionary overview

Only about 1% of the 3 billion bp long human genome is protein coding, organized in 25,000 genes [61]. The majority of the genome is composed of transposons and simple repeats, although some of this DNA might also be transcribed and have regulatory functions [62]. Indeed, about 5% of the genome is highly conserved (has nearly the same sequence) across mammals, indicating functionality of these loci. Meanwhile, the human and chimpanzee genomes are on the large scale very similar, except for a chromosome fusion in humans (chr. 2) [63]. The nucleotide divergence rate between the two species is 1.2%, and including insertion-deletions, ~4%. Notably, most of these differences are non-coding, and presumably with no phenotypic effect. The average human protein has only 2 amino acid differences compared to the chimpanzee protein. Genomic divergence increases to 7% between humans and macaques [64], suggesting that most differences accumulate linearly with time (human-chimp divergence 6 mya, human-rhesus divergence is 25 mya). Crucially, there is no special acceleration of amino acid changes in the human genome, compared to the chimpanzee and rhesus genomes – on the gross genomic level, humans are just as diverged from rhesus, as are chimpanzees.

What seems special in humans is that we are a very homogeneous species. The average difference between two human genomes is ~0.1% [65,66]. The differences among human populations, compared to populations of chimpanzees or most other species, are minute: Only 10-15% of total genetic variation is seen between human populations, whereas the vast majority of diversity is within populations [67,68]. Strikingly, looking at gene expression variation [69], or even variation in the species composition of microbes living in the human mouth [70], the same proportions appear: 85-90% of variation is seen within populations, and 10-15% is seen between populations. Differentiation among populations, in turn, can be explained largely by the geographic distances among them [71].

Most human diversity is found in Africa, in line with the hypothesis that humans have arisen in Africa, dating back to Charles Darwin [65,72-74]. All modern human mitochondria share a common ancestor that was present in Africa ~200,000 years ago [74], and all Y-chromosomes coalesce back around 100,000 years ago [75]. This evidence argues strongly against earlier “multiregional” theories which claimed that current human populations are descendents of early *Homo erectus* populations who spread out of Africa >1 mya. In other words, humans are a very recent species, and practically ‘all humans are Africans’ [65,73,76,77]. Nevertheless, there have been indications of genetic admixture between putative ancestral *Homo* species, and modern human populations [78,79]. However, this is still a contentious subject, even if true, such events were probably rare [80].

Our low genetic diversity indicates that human effective population size is also very small: ~10,000 [81,82]. That is, modern human ancestors who contributed their genes to today’s pool were a small group, or went through frequent bottlenecks. Human effective population size is much smaller than flies, mice, rhesus macaques [64], and even slightly smaller than common chimpanzees [83,84]. Bottlenecks have been especially strong in populations that moved out of Africa [85]. Importantly, the

smaller the population size, the weaker is the power of natural selection to weed out deleterious alleles, and to choose beneficial alleles [10,86]. This can be already seen at the level of amino acid substitutions and segmental duplications in the human genome, a substantial portion of which are presumably slightly deleterious [64,87,88]. Along the same line, while ~50% of the coding region of the fly genome is estimated to be under positive selection [89], this proportion is ~10% for the human genome [90]. Compared to the chimpanzee, there is also less evidence for positive selection in humans than in chimpanzees [91]. Hence, the human genome appears to have been shaped predominantly by random genetic drift, rather than selection. This clearly complicates the interpretation of genetic differences between humans and other primates.

Selection and drift in human genome

Despite the fact that genome sequences of the human and the chimpanzee are available for a number of years, there are substantial obstacles in pinpointing the genetic causes of phenotypic differences between the two species, and in reconstructing human evolutionary history based on these genetic differences. Associating nucleotide differences between the two species to phenotypic differences, even if we restrict our interest to the ~2 amino acid substitutions per protein [63], is a challenging task. Why is this? First, our current knowledge of protein function generally does not enable inferring function directly from amino acid sequence – we cannot tell what a protein sequence change might phenotypically mean. Second, research on protein evolution suggests that most amino acid differences between humans and chimpanzees have no or very slight phenotypic effects, and are effectively neutral [86,90]. Furthermore, many phenotypic differences might not be determined by amino acid differences, but instead by regulatory differences [9,92]. However, current information about gene regulatory sequences is even sparser than that of amino acids, limiting our understanding of the potential functionality of substitutions in regulatory regions. Finally, even phenotypically influential sequence changes may be driven by neutral drift, and may not be adaptive *per se*. Hence, determining which genes might have undergone adaptive change is a substantial challenge.

Tests for identifying adaptive change

There are a number of approaches for identifying potential adaptive genetic changes between species. The classical approach is to start from a particular phenotype and test substitutions in candidate loci known to be associated with it (see *FOXP2* case below). The second is to scan the genome for signatures of adaptive change and other evolutionary patterns, and combine this with published information on gene function, to infer selection {reviewed in [93,94]}. Third, one can use comparative gene expression {reviewed in [95,96]}.

Scanning species' genomes for loci showing unexpectedly high or low number of amino acid differences is a popular method for studying adaptive evolution and drift. Most mutations changing amino acid sequence of a protein (non-synonymous mutations) are expected to be deleterious, as they may alter the protein function in a random way, and if so, these are less likely to become common or fixed in a species, compared to mutations that have no functional impact on the protein

(synonymous mutations). However, some of these protein alterations may also be beneficial. Beneficial mutations, over long periods of positive selection, may increase in frequency, and eventually get fixed. Comparison of ratios of non-synonymous to synonymous mutations in the coding regions of genome between two species is a common technique used to identify genes under selection. Higher than expected ratios of non-synonymous changes indicate positive, less than expected ratios of non-synonymous changes indicate purifying (conservative/negative) selection [97, 98]. Unfortunately, this strategy works only for genes that have undergone recurrent adaptive changes [93], which severely limits the power to detect selection especially when comparing close species, such as human and chimpanzee.

Combined information across multiple genomes can also help in identifying regions which are presumably functional, and have changed dramatically in the human lineage {*e.g.* [99,100]}.

Another major information source for finding selective signals is genetic polymorphism, *i.e.* variation within a species. Regions of low nucleotide diversity are indicative of selective sweeps, and this information can be combined with high amino acid divergence rates: the McDonald-Kreitman test thus identifies regions which show low diversity but large divergence as candidates for positive selection, and the reverse could indicate negative or balancing selection [98]. Similar tests can also be applied to other functional sites such as non-coding regulatory sequences [101]. The McDonald-Kreitman approach, although very powerful and popular, still may suffer from biases [102].

In addition, there are a plethora of tests using the haplotype structure and the allele frequency spectrum for identifying positive selection, although these are mainly used for comparing across populations {*e.g.* [103]}. Notably, tests based on polymorphism can recognize signatures of selection as far back as ~200,000 years in human history [94, 104], and thus are only partly helpful in studying human-chimpanzee differences.

Negative selection, drift, and gene loss

Before discussing evidence for positive selection in the human genome, we first consider the effects of negative (purifying) selection and neutral evolution. Patterns of negative selection in the human genome are generally similar to those seen in other species: Genes that are critical for development, expressed in multiple tissues and fundemantal for cell structure and function, and in general, those which can severely compromise fitness if they fail in their functions, tend to evolve slowly [105]. For example, human and chimpanzee comparisons show that genes which are of fundamental use in different cell types, such as myosin loci involved in actin binding and cytoskeletal formation, exhibit excess amino acid polymorphism relative to divergence, thus indicate negative or purifying selection. Other examples of genes under strong negative selection include those with roles in ectoderm development and neurogenesis, or vesicle transport [106].

Not surprisingly, human genes showing strong negative selection signatures are overrepresented among genes that cause Mendelian disease if mutated [106]. For example, mutations in conserved skeletal muscle proteins cause muscular dystrophies (*i.e.* Duchene and Becker). Mutations in myosin-VIIA are implicated in deafness and blindness. Alfa- and beta-adducin gene variants are associated with hypertension and cardiovascular disease. The *HD*, one of the genes in the general vesicle transport category that shows negative selection, is the gene that causes Huntington's disease (http://www.ncbi.nlm.nih.gov/omim).

However, the strength of purifying selection on a gene depends not only on the genes' relevance to fitness (s), but also on the species' effective population size (N_e). As mentioned above, amino acid substitution rates and copy number variant distribution patterns indicate relatively high drift in the ape and especially the human genome, compared to many other mammals [64,88]. Similarly, gene regulatory regions in primates were found to be under less purifying selection than in murids [107]. Hence, it is not surprising to find widespread gene losses in the human genome, which could be related to weakening of purifying selection. The pseudogenization of the olfactory receptor gene family in humans is one example, leading to a much weaker smell [108]. Similarly, bitter taste receptors are rapidly being lost in all great apes, including humans [109]. It is of course also plausible that such gene losses happened due to a total loss of reliance on smell and taste, rather than to weak purifying selection.

Meanwhile, some gene losses have been hypothesized to be beneficial in outcome, known as the "less is more" hypothesis [110]. For instance, a myosin gene (*MYH*), expressed in chewing muscles in humans, was lost ~2 million years ago, and this is speculated to have facilitated encephalization by removing growth constraints on the skull [111]. Similarly, the pseudogenization of a keratin gene may be related to hair type differences between humans and chimpanzees [15], and human hairlessness is claimed to have been of adaptive value (see above). Another example is the loss of expression of sialic acid genes in human T-cells, which appears associated with hyper reactivity of these cells in humans (the cause of human vulnerability to HIV) [22]. In some cases, adaptive scenarios for these changes can be supported by polymorphism data. For example, the *CASPASE12* gene, involved in immune response, has been pseudogenized in some humans, and there is a signal of positive selection around the locus, which could be related to the observation that *CASPASE12* loss is beneficial against sepsis [112,113].

Positive selection in the human lineage

A number of recent studies have scanned the human genome for evidence for positive selection in amino acid sequences [63,64,91,106,114-116] {reviewed in [93]}, in regulatory sequences [117], or in potential regulatory non-coding sequences [118-120].

Simultaneous comparison of thousands of annotated genes from humans with their chimpanzee orthologs show evidence of positive selection in genes involved in a variety of functions, including

immunity and defense, chemosensory perception, olfaction, sensory perception, gametogenesis, spermatogenesis and motility, and inhibition of apoptosis [106,121-123]. Immune- and defense-related genes are at the top of a list of genes that show the strongest positive selection, which is a common pattern across mammals [64]. These genes have roles in a wide range of T- and B- cell, antibody- and interferon- mediated immunity processes suggesting a coevolutionary arms race driving selection on the immune system. The four olfactory genes (*OR2WI*, *OR5I1*, *OR2B2*, and *C20orf185*) that show excess numbers of amino acid substitutions (divergence) between chimpanzees and humans also show increased number of polymorphisms in humans. Accordingly, population genetic analyses of *OR5I1* suggest balancing selection (increased polymorphism) as an important evolutionary force that increases the olfactory receptor repertoire [123]. On the other end of the spectrum, some genes show increased divergence between humans and chimpanzees but show no polymorphism in humans suggesting, repeated substitutions and fixations. *SCML1*, a gene with a potential important role in the control of embryonic development, is a good example of this, and may be responsible for some of the developmental differences between humans and chimpanzees.

Although pairwise human-chimpanzee comparisons are informative for detecting genes under selection, analyses including an outgroup to these comparisons are crucial for determining lineage specific evolutionary trends. An analysis of human-chimp-mouse orthologs agreed with accelerated evolution of genes involved in aforementioned biological processes [121]. Furthermore, several genes involved in the amino acid catabolism, particularly the ones involved in branched-chain amino acid catabolism (*ALDH6A1*, *BCKDHA*, and *PCCB*) show faster evolution in human lineage and may represent selective pressures of different dietary habits. Interestingly, another study using rat and dog as outgroup, found ~500 genes positively selected in chimpanzee, compared to <200 in humans, with putatively selected human genes showing enrichment in G-protein coupled receptors and sensory perception-related processes [115].

However, using distant outgroups or fast evolving taxa such as the mouse may be problematic for identifying positive selection, given difficulties in ortholog finding, sequence alignment, and multiple substitutions. Two studies that used the more reliable macaque genome for identifying positively selected genes in the human lineage [64,91], as well as a comprehensive genome scan analyzing six mammalian genomes together [116], consistently identified more positively selected genes in chimpanzees, than in humans. The overlap between these small human-specific gene sets (n<250 in all studies) appears not to be big, although immune response genes and genes related to signal transduction tend to emerge in multiple studies. In turn, across all primates and/or mammals, genes showing positive selection show a clear bias toward being related to either immunity, reproduction, diet, or perception of stimuli [116].

As it is hypothesized that evolutionary changes in gene regulation may be as important as protein coding changes [9], there has also been interest in identifying positively selected regulatory regions. Haygood *et al.* scanned the human and chimpanzee promoters for excess of mutations in either

lineage, which could be indicative of positive selection in gene regulation [117]. Interestingly, promoters that show such human-specific acceleration are associated with neural development and glucose metabolism genes. Similarly, multiple studies focused on the evolution of conserved non-coding sequences, *e.g.* >100 bp sequences with >95% identity across mammals [118]. These regions may function as transcriptional enhancers, although their functions are still not well established, as their deletion may not yield observable phenotypes [124]. A number of studies have identified such conserved regions that show an excess of changes in the human lineage [119,120]. Two interesting examples have been presented by Pollard *et al*., who found that their best human-accelerated candidate regions functioned as enhancer in fetal brain development [99], and Prabhakar et al. identified a similar region regulating limb development [125].

Despite the experimental elegance of this work on non-coding regions, the soundness of the algorithms used in identifying positive selection have been frequently contested. A substantial problem in establishing positive selection in non-coding sequences is that there is no consensus null model, *i.e.* a model of neutral evolution. In other words, there is no equivalent to the use of synonymous substitution rates when analyzing coding regions. Thus, the choice for null models might be contentious {see [126] contesting [117]}. In addition, biases in mutation rate, such as those caused by biased gene conversion, can rapidly increase mutation rate in a single lineage, leaving a signal that may be interpreted as positive selection {see [127] on [99], and [128] on [125]}.

An alternative approach for studying selection is deep investigation of genes with known interesting phenotypes. A well-known example is the *FOXP2* gene. Loss of one copy of *FOXP2* causes severe speech and language disorders in humans. Notably, although *FOXP2* is highly conserved across vertebrates, there are two intriguing amino acid substitutions that occurred in humans since the split from chimpanzees, which is highly unusual [129]. Moreover, the locus shows an unusual pattern of polymorphism in the human lineage, suggestive of a positive selection event in recent human evolution [129,130]. This molecular change in a speech-related gene has attracted much attention: Recent work has shown that transgenic mice carrying the human version of *FOXP2* have alterations in brain development, as well as vocalization [131]. Furthermore, human *FOXP2* appears to influence certain gene expression differences between humans and chimpanzee [132]. An ancient DNA study found the modern human *FOXP2* version also in Neanderthals, implying that this extinct cousin species could in fact have spoken languate [133]. Interestingly, along with *FOXP2*, several genes involved in the development and tuning of hearing also appear in a list of genome-wide scan for positive selection in the human lineage [121]. These studies provide exciting insight into the molecular basis of speech-based communication in humans, one of the most spectacular human traits.

Such work on long-term selection on the human genome not only provides information about human history, but also gives clues regarding the origins of contemporary human disease. For instance, *APOBEC3F*, a host-pathogen interaction gene showing a strong positive selection signal [134], is

associated with anti-HIV activity. Moreover, some of the genes showing the strongest evidence for positive selection are involved in tumor development and control [123], with reasons still not clear.

We expect a number of recent developments to improve the efficiency of genome-wide scans positive selection, as well as work on candidate adaptive loci. One is the increased number of genomes sequenced, especially of taxa closely related to humans, which can enhance the power of tests for positive selection. A related development is the sequencing of extinct human relatives, such as Neanderthals [135,136]. The rapid accumulation of human polymorphism data (http://www.1000genomes.org/page.php) and combining divergence and diversity patterns in tests for positive selection, could be useful for identifying events related to the recent emergence of modern humans. Finally, the increasing feasibility of functional tests on putative adaptive loci {*e.g.* [99,125,131]} is delightful, and could significantly improve the quality of studies of positive selection.

Transcriptome divergence between humans and chimpanzees

Long before the genomes of humans and chimpanzee were available, M.C. King and A. Wilson noted that the two species do not show extraordinary patterns of amino acid sequence divergence relative to other mammals [9]. King and Wilson considered this observation surprising, given their expectation of large phenotypic divergence of the human lineage {see Fig. 5 in [9]}. Hence they proposed that most of phenotypic differences between humans and chimpanzees may be explained by expression level differences, rather than by protein sequence differences.

In the early 2000's, the first studies comparing genome-wide gene expression patterns between humans and chimpanzees began to appear [137-141] {reviewed in [95,96]}. These experiments revealed that for the majority of genes in adult human and chimpanzee transcriptomes, the distribution of expression levels within the two species overlap. In other words, for most genes, there is no indication of an expression level differentiation between the two species, which is compatible with the notion that expression levels are generally under the strong influence of purifying selection [142,143].

On the other hand, even with the modest sample sizes used in these studies, a substantial number of genes do show significant expression level differences between humans and chimpanzees. For example, with 5-6 individuals per species, 5-10% the transcriptome can be identified as differentially expressed in tissues such as the brain or the liver [139,140]. Higher and lower estimates have also been made, depending on the sample size and method employed {*e.g.* [144,145]}. This raised the following questions: could these expression differences underlie phenotypic divergence between humans and chimpanzees? Might they be products of adaptive evolution?

Neutral expression divergence

Khaitovich *et al.* made two interesting observations regarding these questions [146]. First, expression divergence between primate brains appears to increase linearly with species' divergence times.

Second, expression divergence between species is positively correlated with diversity within species. A similar observation was also made in fish [147]. These results lead to the hypothesis that most expression divergence is due to neutral evolution [146]. In contrast, Gilad *et al.* used cDNA arrays, a potentially less biased but noisy technology, and found the same amount of liver expression divergence in all pairwise comparisons among four primate species [140]; the cause of the discrepancy between the studies is not yet clear, but could be related to the technologies or the tissues used.

Another interesting observation regarding the functionality of expression divergence is that differentially expressed genes do not always cluster in particular functional gene groups, as might be expected under an adaptive scenario. In fact, a number of studies have reported overlaps between differential expression and functional groups [138,140,141,148], but in most of these studies, the results may not have withstood correction for multiple hypothesis testing {see [149]}.

It therefore seems plausible that for many, if not most genes showing significant expression difference between humans and chimpanzees, the evolutionary processes that lead to this difference could be neutral, rather than adaptive [105,146,150]. In other words, the effect of most expression differences may either *(a)* be devoid of phenotypic effect, *(b)* be functional but unrelated to fitness, *(c)* have a fitness effect which is too small to be acted upon by natural selection, so that the genetic variants causing expression differences could have been fixed by genetic drift [86,151]. Indeed, as mentioned before, analyses of sequence data indicate that drift has a significant role in ape evolution. We may therefore summarize transcriptome evolution in apes as follows: Gene expression levels are mainly shaped by purifying selection, to a lesser extent by neutral evolution, and to some unknown, possibly even smaller extent, by adaptive changes.

Challenges in identifying adaptive expression divergence

The question therefore becomes how the relative levels of neutral *vs.* adaptive evolution can be distinguished. As with non-coding sequence evolution (mentioned above), we lack a consensus method to make this distinction. That is, we do not know of any transcript type with strictly neutrally evolving expression levels [96,145,146,152]. Mutation accumulation lines in *Drosophila* [143] or crosses between yeast strains [153] have been used to address this problem; however, such methods are not available for use in apes. Therefore, in order to distinguish the relative rates of neutral and adaptive expression divergence, we have to resort to indirect approaches.

One such method is to theoretically estimate neutral expression divergence, by taking into account the species divergence time, ancestral population size, and expression diversity [142,144]. However, such models have not been frequently applied, likely because of the large number of assumptions that they involve. A second, simpler approach is to assume that in every human and chimpanzee comparison, some proportion of expression differences are results of adaptive evolution. Thus, some studies have considered genes which show highest divergence to diversity ratios as candidates for

positive selection [138,140]. A third approach is to use an outgroup, and assign expression changes to human and chimpanzee lineages. If repeated adaptive evolution was restricted to one lineage only, and caused an excess of expression changes on that lineage, such a pattern could be revealed by this strategy. Thus, patterns of asymmetric divergence could be considered indicative of positive selection [137-140]. A fourth approach is to use external evidence for positive selection, or use additional data on the functions of genes that show expression differences. For example, one can test whether genes showing expression differences also overlap with genes known to be positively selected based on polymorphism data or amino acid divergence [154,155], or whether they overlap with genes in specific functional categories [141]. In addition, sequence-based evidence that genes are under purifying selection (*i.e.* they are not under relaxation of constraint) would further support an adaptive explanation. Nevertheless, we note that without a satisfactory neutral model of expression level divergence, how such evidence is interpreted depends largely on the researcher's viewpoint of evolutionary change in general [156-158].

So, is there any indication of adaptive divergence in gene expression patterns between humans and chimpanzees? A number of studies reached affirmative conclusions. Gilad *et al.* investigated gene expression differences in the liver, and reported higher expression levels of transcription factors in the human lineage compared to other primates [140]. Another study comparing humans and chimpanzees reported an extremely high divergence to diversity ratio in the testis, compared to four other tissues [139], even though it is possible that differences in cell type proportions between the two species' testicles [12] may have influenced this result (Somel, unpublished). Approaching this issue from the genome front, Haygood *et al.* identified promoter regions that showed high divergence between humans and chimpanzees [117]. Although these genes did not overlap with previously identified genes showing expression divergence, it is possible that these promoter differences drive expression in tissues or developmental stages not yet investigated.

Adaptive expression divergence in the brain

Another interesting claim for adaptive expression divergence between humans and chimpanzees is related to the brain, and in particular, the cerebral cortex transcriptome {reviewed in [145]}. A number of observations support this notion:

(a) Using an outgroup such as orang-utan or macaque, gene expression levels across species appears to have changed more in the human brain than in the chimpanzee brain, since the time of the two species' common ancestor [137]. In contrast, humans and chimpanzees appear equally diverged in the liver or other tissues.

(b] There is an excess of up-regulated expression levels in the human brain transcriptome compared to the chimpanzee, which again appears specific to the brain [138,139] {but see [144]}.

(c) Genes up-regulated in the human brain are reported to overlap with genes involved in energy metabolism and neural function [138,141].

(d) Brain gene expression changes during postnatal development are generally conserved across mammals, but a distinct group of genes involved in neural development show a human-specific change in their expression profiles: specifically, for these genes, expression changes with age are delayed in humans compared to other primates [159]. This might be related to an extended childhood period in humans, during which human infants develop cognitive abilities and assimilate human culture [8], and also fits the human neoteny hypothesis [14] (see above).

(e) Expression divergence in the human lineage is positively correlated with linkage disequilibrium in humans, a measure of positive or negative selection. This correlation is again specific to the brain [154]. Intriguingly, human-specific metabolites also overlap with metabolites which are affected by psychiatric disease, such as schizophrenia [160].

In addition, gene expression in the central nervous system is in general highly conserved (shows less divergence among species) than other tissues [139]. One reason for this could be the brain being buffered against environmental influences such as diet [96]. Indeed, using a mouse model, a recent study showed that differences in chimpanzee and human diets (which contain raw *vs.* cooked food, and low *vs.* high protein and fat content, respectively) affect gene expression differences between human and chimpanzee livers, but not the brain [155]. Another reason could be that brain function is strongly related to fitness, more than other tissues. Indeed, nervous system genes are also more highly conserved in their amino acid and regulatory sequences compared to other gene classes (see above).

These observations suggest that non-adaptive alternative explanations for the asymmetric divergence of expression changes in the human brain, *i.e.* environmental effects or relaxation of constraints on the human lineage, are not very likely. Instead, the so-called “acceleration” of brain gene expression patterns in human evolution may be driven by adaptive evolution. Because morphological evidence also indicates that the human brain has changed more in size [39,161] and histology [145,162] during the last 6-7 million years than the chimpanzee brain, and as humans show some exceptional cognitive abilities than chimpanzees and orang-utans [8,17,32], the case for functional and adaptive changes in human brain gene expression seems convincing.

However, this is still a preliminary result, for a number of reasons. First, biases caused by microarray technology or by environmental variation have not yet been fully ruled out [96]. Second, because gene expression levels show drastically lower variance in the brain compared to other tissues [139], it is possible that evolutionary divergence patterns are easier to identify in the brain than other tissues – hence the specialty of the brain. Third, because most comparative transcriptome studies have analyzed the brain more often than other organs, it is not easy to assess the uniqueness of

expression patterns in this organ. Fourth, divergent gene expression patterns in humans may be secondary consequences of evolutionary divergence in size and structure of the human brain [145], instead of adaptive changes that were linked to novel cognitive functions and that specifically drove these gene expression levels. Finally, the abovementioned studies that identified the "human-acceleration" predominantly used the prefrontal cortex region of the brain (a structure involved in cognition, planning, social behaviour), and it is unclear if other regions, such as the striatum or cerebellum, with substantially different functions, also show the same pattern or not. In fact, a comparison between humans and chimpanzees across different regions of the brain found very similar patterns of species divergence among cerebral cortex regions [163]. Some of these regions are implicated in human-specific cognitive functions, and some are not, which implies that human-specific expression divergence in the brain might not be directly related to cognitive divergence, although the small sample sizes used here precludes a final conclusion.

To summarize, linking expression differences to phenotypic differences, and distinguishing adaptive changes from neutral changes, remains a complicated task. Our general lack of knowledge on the causes of gene expression differences further aggravates these problems [164,165]. For example, an expression difference between humans and chimpanzees can be caused by *(a)* genetic differences in *cis*, *(b)* gene duplication, deletion, or chromosomal relocation in one of the species, *(c)* epigenetic differences, such as DNA methylation, *(d) trans* effects caused by genotypic differences, such as differences in transcription factor amino acid sequences affecting promoter binding, or the presence/absence of microRNAs, *(e) trans* effects related to differences in the general cellular environment, such as RNA turnover, or protein localization, *(f) trans* effects determined by ontogenetic differences between the species, such as tissue composition, *(g) trans* effects in response to different external environments, or *(h)* a direct reflection of tissue composition differences between the two species, and distinguishing between these modes requires significant experimentation. The pioneering transcriptome studies have revealed this challenge, more than providing definite answers.

How can we face this challenge? A number of developments might considerably improve our understanding the role of gene expression in human evolution. One step already taken is going beyond microarray technology and directly sequencing transcriptomes (RNA-seq) [148]. A second is the analysis of the transcriptomes of cells of similar type, *e.g.* using laser capture microdissection technology [96]. A third path is the systematic collection of transcriptional data from model organisms under controlled and manipulated environments [155], or across different developmental stages [96,159,166]. Comparative data on gene regulation is also accumulating, including species differences in alternative splicing [148,167], microRNA expression [168,169], or epigenetic modifications [170]. Such information could also be combined with other types of data on phenotypic variation, such as protein expression [171] and metabolite levels [160], to build holistic models of gene expression regulation. Through such models, it may be possible to estimate the influence of DNA substitutions on gene expression levels, and expression changes on phenotypic differences

[166]. Finally, experiments on model organisms [131] will be fundamental in directly associating the genotype, gene expression levels, and the phenotype.

Detection and characterization of selection within human populations

We will finally discuss recent human evolution, that is, patterns of within population variation and population divergence that arose since modern humans appeared 200,000 years ago in Africa. As mentioned above, humans are an interesting species compared to other primates such as chimpanzees or rhesus macaques, in that within-species variation is very low: Humans share 99.9% nucleotide sequence identity at the DNA level. Likewise, populations are very little, if any, divergent (~10% differentiation between continental groups). Genetic diversity is largest in Africa, and decreases linearly over geographic distance from East Africa, reaching its lowest levels in the Americas [71,172]. This clearly pinpoints East Africa as the origin of all modern humans.

So, how much did the migrating populations differentiate and adapt to their new local environments? Notably, some of these new environments were fundamentally different, in terms of major factors such as climate, pathogens, or dietary sources [173], which could drive widespread local adaptive evolution. If this were the case, we would expect large differentiation among neighboring groups living in different environments. Instead, as with genetic diversity within groups, genetic divergence between groups also correlates strongly with the geographical distance between them. This implies that patterns of human genetic variation are driven largely by serial founder effects and migration, and less by local adaptation [174]. Even between continental groups, the level of divergence is meager, with only a dozens of loci showing unexpected differentiation across the genome [175]. Signatures of recent selection in the human genome may therefore not to be widespread, this either due to strong migration, to strong drift, or to cultural rather than biological adaptation to new environments [4,7].

The prevalence of neutral processes, of course, does not exclude local adaptation from human history. On the contrary, human populations living in different geographies do show particular morphological, physiological and behavioral differences. Furthermore, observations of geographic patterns for heritable traits like skin pigmentation, body mass, disease risk, pathogen resistance and variable drug response indicates that many of these traits are heritable. Was this variation shaped by local selection events? Population genetic studies have been searching for interesting examples of local adaptation, employing different methods ranging from population differentiation (measured by F_{st}), patterns of excess of rare polymorphisms, to long-range haplotypes, with ever-increasing success {reviewed in [93,176]}.

Already half a century ago researchers had noticed that the geographical distribution of sickle cell disease overlapped with malaria occurrence [177]. Even though early techniques allowed investigation of only a few loci, they were sufficiently powerful to detect spatial correlations and high differentiation (high F_{st}) between populations in this locus [67]. These studies were able to show that malarial selection was strong enough to select for hemoglobin (*HBB*) mutations, which cause the

sickle cell phenotype, in independent populations, as well as to restrict the Duffy (*FY*) null allele in African populations, where *Plasmodium vivax* malaria has been endemic [178-180]. These studies provided the first examples of infectious disease selection shaping the genetic diversity in human populations.

Today, novel techniques allow surveying multiple loci from multiple individuals, representing multiple geographic regions, relatively easily. Moreover, such high-throughput genotyping and sequencing surveys [181] enables the application of more powerful statistical models for detecting selection in human populations [94]. Some of these models try to capture the reduction in variation and/or increase in derived allele frequency near the selected sites due to recent selective sweeps in one population, but not in others [103,182-185].

Such analyses on high-density polymorphism data have already identified interesting examples. For instance, in Africans, patterns of polymorphism in two genes with well-documented biological links to the Lassa fever virus, the *LARGE* and *DMD*, suggest selection in favor of protection against the Lassa [183]. Multiple skin pigmentation genes, including *SLC24A5*, *KITLG* and *MC1R*, show strong signatures of positive selection in Europeans [183,186,187]. Another well-identified locus is the *LCT* (lactase) genes, related to lactose tolerance and milk consumption in Europeans [188]. In Asians, there is evidence of selection in two genes, *EDAR* and *EDA2R*, involved in development of hair, teeth and exocrine glands [183]. The biological basis of this selection is still not clear. In addition to these well-known examples, genome-wide scans have provided a plethora putatively selected genes [103, 184].

However, it is not always straightforward to conclude about positive selection. Most of genome-wide tests rely on choosing extreme loci within the genome, where it is assumed that some part of the genome should be under selection. Other tests use population genetic simulations, which also require assumptions about population history. Can one strictly rule out statistical noise as the reason of these patterns? Functionality of a candidate allele, if it can be shown, is helpful in excluding a neutral explanation. On the other hand, observing genetic change that involves different mutations in different populations, but has the same functional outcome (*i.e.* multiple alleles converging on a similar phenotype) provides strong support for a positive selection scenario. For example, lactase persistence in Europe, Africa and Middle East involves different mutations in *LCT* gene, which arose independently in those populations engaged in dairy farming and might have depended on milk in times of famine [73,189,190]. Similarly, malaria resistance is provided by two different *G6PD* (glucose-6-phosphate dehydrogenase) deficiency alleles: $G6PD^{A-}$ in sub-Saharan Africa and $G6PD^{Med}$ in Europe [191-193]. Likewise, skin pigmentation variants in Europeans and East Asians do not always overlap [194], providing examples of different mutations underlying the same adaptive phenotypes.

Human population genetic studies are expected to accelerate within the next decade, especially with the increasing availability of *de novo* genome sequencing. Such data will be especially favorable for

population genetic analyses, as it will allow overcoming ascertainment biases in single nucleotide polymorphism data, an obstacle for many selection tests [93].

Furthermore, more detailed environmental surveys (such as climate, diet, pathogen diversity and load, physiology) in fine-scale geographic samples may be useful to investigate local adaptations. Already, correlations between such clines and allele frequencies show that candidate susceptibility alleles for hypertension and metabolic syndrome correlate with latitude and climate variables, respectively [195-197]. These gradual changes in allele frequencies may signal adaptations to continuous local environments.

We also expect population genetic studies to diversify in the different types of data used. One obvious direction is identification of gene expression differences between and within populations [69]. Another direction is the study of copy number variation. Indeed, a large proportion of genetic variation among humans is due to copy number differences [198], and it was recently shown that copy number variation in amylase gene (*AMY1*) correlates with amylase enzyme levels, and higher copies of *AMY1* is found in human populations that have a starchy diet [199]. Similarly, initial work on variation in the human microbiome, the bacterial and fungal flora living inside or on the human body, has provided interesting results on the dispersal patterns of the species, as well as human history itself [70,200].

Eventually, the combination of new sequencing methods with more detailed phenotypic and genetic data should greatly increase our understanding of how the forces of selection, adaptation, and demographic history have been shaping genetic and phenotypic diversity in humans.

CONCLUSION

Clearly, deciphering the evolutionary path of the human lineage, in both the long and short terms, is a complex endeavor, and requires multidisciplinary approaches. Today's researchers are generating ever-increasing amounts of genetic and functional data, both at the single gene and at the genome levels, across multiple species or populations. The enlarging scale renders evolution, especially human evolution, an ever-more-appealing field. However, in order to fully exploit the oncoming surge of information and to decipher evolutionary and physiological mechanisms, we also need more advanced mathematical/computational models. Collaborations and multidisciplinary syntheses of empirical and theoretical work are thus expected to abound. Such work will not only provide medically useful information and allow us to learn our species' history, but at the same time enlighten us about humankind's place in nature. More exciting times are imminent for human evolutionary biology.

REFERENCES

1. Darwin, C. (1871/1981) *Descent of man and selection in relation to sex* (Princeton University Press, Princeton).
2. Mayr, E. (1982) *The growth of biological thought* (Harvard University Press, Cambridge, MA).
3. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860-921.
4. Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2001) Cultural niche construction and human evolution. *J Evol Biol* 14: 22-33.
5. Engels, F. (1934) *The part played by labour in the transition from ape to man* (Progress Publishers, Moscow).
6. Smith, J. M. & Szathmary, E. (1998) *The major transitions in evolution* (Oxford University Press, New York).
7. Laland, K. N., Odling-Smee, J., & Myles, S. (2010) How culture shaped the human genome: Bringing genetics and the human sciences together. *Nat Rev Genet* 11: 137-148.
8. Tennie, C., Call, J., & Tomasello, M. (2009) Ratcheting up the ratchet: On the evolution of cumulative culture. *Philos Trans R Soc Lond, B, Biol Sci* 364: 2405-2415.
9. King, M. C. & Wilson, A. C. (1975) Evolution at two levels in humans and chimpanzees. *Science* 188: 107-116.
10. Lynch, M. (2007) *The origins of genome architecture* (Sinauer, Sunderland, MA).
11. Carroll, S. B. (2003) Genetics and the making of homo sapiens. *Nature* 422: 849-857.
12. Harcourt, A. H., Harvey, P. H., Larson, S. G., & Short, R. V. (1981) Testis weight, body weight and breeding system in primates. *Nature* 293: 55-57.
13. Harcourt-Smith, W. E. H. & Aiello, L. C. (2004) Fossils, feet and the evolution of human bipedal locomotion. *J Anat* 204: 403-416.
14. Gould, S. J. (1977) *Ontogeny and phylogeny* (Harvard University Press, Cambridge, MA).
15. Winter, H., Langbein, L., Krawczak, M., Cooper, D., Jave-Suarez, L., Rogers, M., Praetzel, S., Heidt, P., & Schweizer, J. (2001) Human type i hair keratin pseudogene fhhaa has functional orthologs in the chimpanzee and gorilla: Evidence for recent inactivation of the human gene after the pan-homo divergence. *Hum Genet* 108: 37-42.
16. Montagna, W. (1972) The skin of nonhuman primates. *Amer Zool* 12: 109-124.
17. Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B., & Tomasello, M. (2007) Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science* 317: 1360-1366.
18. Weaver, T. D. & Hublin, J. J. (2009) Neandertal birth canal shape and the evolution of human childbirth. *Proc Natl Acad Sci USA*.
19. Gilbert, S. F. & Zevit, Z. (2001) Congenital human baculum deficiency: The generative bone of genesis 2:21-23. *Am J Med Genet* 101: 284-285.
20. Lovejoy, C. O. (2009) Reexamining human origins in light of ardipithecus ramidus. *Science* 326: 74e71-78.
21. Olson, M. V. & Varki, A. (2003) Sequencing the chimpanzee genome: Insights into human evolution and disease. *Nat Rev Genet* 4: 20-28.
22. Nguyen, D. H., Hurtado-Ziola, N., Gagneux, P., & Varki, A. (2006) Loss of siglec expression on t lymphocytes during human evolution. *Proc Natl Acad Sci U S A*: 0510484103.
23. Hawkes, K., O'Connell, J. F., Jones, N. G., Alvarez, H., & Charnov, E. L. (1998) Grandmothering, menopause, and the evolution of human life histories. *Proc Natl Acad Sci U S A* 95: 1336-1339.
24. Wrangham, R. & Conklin-Brittain, N. (2003) 'cooking as a biological trait'. *Comparative Biochemistry and Physiology - Part A: Molecular & Integrative Physiology* 136: 35-46.
25. Finlay, B. L. & Darlington, R. B. (1995) Linked regularities in the development and evolution of mammalian brains. *Science* 268: 1578-1584.
26. Pruetz, J. D. & Bertolani, P. (2007) Savanna chimpanzees, pan troglodytes verus, hunt with tools. *Curr Biol* 17: 412-417.
27. Pollick, A. S. & de Waal, F. B. M. (2007) Ape gestures and language evolution. *Proc Natl Acad Sci U S A* 104: 8184-8189.
28. Taglialatela, J. P., Russell, J. L., Schaeffer, J. A., & Hopkins, W. D. (2008) Communicative signaling activates `broca's' homolog in chimpanzees. *Curr Biol* 18: 343-348.

29. Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007) Egalitarian motives in humans. *Nature* 446: 794-796.
30. Fehr, E. & Gachter, S. (2002) Altruistic punishment in humans. *Nature* 415: 137.
31. Jensen, K., Call, J., & Tomasello, M. (2007) Chimpanzees are rational maximizers in an ultimatum game. *Science* 318: 107-109.
32. Tomasello, M. (2008) *Origins of human communication* (MIT Press, Cambridge, MA).
33. Kaplan, H., Hill, K., Lancaster, J., & Hurtado, A. M. (2000) A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology: Issues, News, and Reviews* 9: 156-185.
34. Bogin, B. & Smith, B. H. (1996) in *Am J Hum Biol*, pp. 703-716.
35. Viegas, J. (2008) in *Discovery News*.
36. Hill, K., Hurtado, A. M., & Walker, R. S. (2007) High adult mortality among hiwi hunter-gatherers: Implications for human evolution. *J Hum Evol* 52: 443-454.
37. Walker, R., Gurven, M., Hill, K., Migliano, A., Chagnon, N., De Souza, R., Djurovic, G., Hames, R., Hurtado, A. M., Kaplan, H.*, et al.* (2006) Growth rates and life histories in twenty-two small-scale societies. *Am J Hum Biol* 18: 295-311.
38. Rosenberg, K. & Trevathan, W. (2002) Birth, obstetrics and human evolution. *BJOG* 109: 1199-1206.
39. Leigh, S. (2004) Brain growth, life history, and cognition in primate and human evolution. *Am J Primatol* 62: 139-164.
40. Walker, R., Burger, O., Wagner, J., & Von Rueden, C. R. (2006) Evolution of brain size and juvenile periods in primates. *J Hum Evol* 51: 480-489.
41. Leonard, W. R. & Robertson, M. L. (1997) Comparative primate energetics and hominid evolution. *Am J Phys Anthropol* 102: 265-281.
42. Hardy, A. (1960) Was man more aquatic in the past? *New Sci* 7: 642-645.
43. Montagu, M. F. A. (1955) Time, morphology, and neoteny in the evolution of man. *American Anthropologist* 57: 13-27.
44. Kaplan, H. S. & Robson, A. J. (2002) The emergence of humans: The coevolution of intelligence and longevity with intergenerational transfers. *Proc Natl Acad Sci USA* 99: 10221-10226.
45. Aiello, L. C. & Wheeler, P. (1995) The expensive-tissue hypothesis: The brain and the digestive system in human and primate evolution. *Curr Anthropol* 36: 199-221.
46. Pagel, M. & Bodmer, W. (2003) A naked ape would have fewer parasites. *Proceedings of the Royal Society B: Biological Sciences* 270: S117-S119.
47. Aiello, L. & Dean, C. (1990) *An introduction to human evolutionary anatomy* (Academic Press, London).
48. Penin, X., Berge, C., & Baylac, M. (2002) Ontogenetic study of the skull in modern humans and the common chimpanzees: Neotenic hypothesis reconsidered with a tridimensional procrustes analysis. *Am J Phys Anthropol* 118: 50-62.
49. Shea, B. T. (1989) Heterochrony in human evolution: The case for neoteny reconsidered. *Am J Phys Anthropol* 32: 69-101.
50. Coqueugniot, H., Hublin, J. J., Veillon, F., Houet, F., & Jacob, T. (2004) Early brain growth in homo erectus and implications for cognitive ability. *Nature* 431: 299-302.
51. Caspari, R. & Lee, S. (2004) Older age becomes common late in human evolution. *Proc Natl Acad Sci U S A* 101: 10895-10900.
52. Tattersall, I. (1998) *Becoming human: Evolution and human uniqueness* (Harcourt Brace and Co, New York).
53. McBrearty, S. & Jablonski, N. G. (2005) First fossil chimpanzee. *Nature* 437: 105-108.
54. White, T. D., Asfaw, B., Beyene, Y., Haile-Selassie, Y., Lovejoy, C. O., Suwa, G., & WoldeGabriel, G. (2009) Ardipithecus ramidus and the paleobiology of early hominids. *Science* 326: 75-86.
55. Smith, T. M., Tafforeau, P., Reid, D. J., Grun, R., Eggins, S., Boutakiout, M., & Hublin, J. (2007) Earliest evidence of modern human life history in north african early homo sapiens. *Proc Natl Acad Sci U S A*: 0700747104.
56. Smith, T. M., Toussaint, M., Reid, D. J., Olejniczak, A. J., & Hublin, J. (2007) Rapid dental development in a middle paleolithic belgian neanderthal. *Proc Natl Acad Sci U S A*: 0707051104.
57. Powell, A., Shennan, S., & Thomas, M. G. (2009) Late pleistocene demography and the appearance of modern human behavior. *Science* 324: 1298-1301.
58. Brown, P., Sutikna, T., Morwood, M. J., Soejono, R. P., Jatmiko, Wayhu Saptomo, E., & Awe Due, R. (2004) A new small-bodied hominin from the late pleistocene of flores, indonesia. *Nature* 431: 1055-1061.

59. Gurven, M., Kaplan, H., & Gutierrez, M. (2006) How long does it take to become a proficient hunter? Implications for the evolution of extended development and long life span. *J Hum Evol* 51: 454-470.
60. Lahdenpera, M., Lummaa, V., Helle, S., Tremblay, M., & Russell, A. F. (2004) Fitness benefits of prolonged post-reproductive lifespan in women. *Nature* 428: 178-181.
61. International_Human_Genome_Sequencing_Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431: 931-945.
62. ENCODE_Project_Consortium, Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET*, et al.* (2007) Identification and analysis of functional elements in 1% of the human genome by the encode pilot project. *Nature* 447: 799-816.
63. Chimpanzee_Sequencing_and_Analysis_Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437: 69-87.
64. Rhesus_Macaque_Genome_Sequencing_and_Analysis_Consortium, Gibbs, R. A., Rogers, J., Katze, M. G., Bumgarner, R., Weinstock, G. M., Mardis, E. R., Remington, K. A., Strausberg, R. L., Venter, J. C.*, et al.* (2007) Evolutionary and biomedical insights from the rhesus macaque genome. *Science* 316: 222-234.
65. Jorde, L. B. & Wooding, S. P. (2004) Genetic variation, classification and 'race'. *Nat Genet* 36: S28-33.
66. Sachidanandam, R., Weissman, D., Schmidt, S. C., Kakol, J. M., Stein, L. D., Marth, G., Sherry, S., Mullikin, J. C., Mortimore, B. J., Willey, D. L.*, et al.* (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409: 928-933.
67. Cavalli-Sforza, L. L., Menozzi, P., & Piazza, A. (1994) *The history and geography of human genes* (Princeton University Press, Princeton, N.J.).
68. Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K., Zhivotovsky, L. A., & Feldman, M. W. (2002) Genetic structure of human populations. *Science* 298: 2381-2385.
69. Storey, J. D., Madeoy, J., Strout, J. L., Wurfel, M., Ronald, J., & Akey, J. M. (2007) Gene-expression variation within and among human populations. *Am J Hum Genet* 80: 502-509.
70. Nasidze, I., Li, J., Quinque, D., Tang, K., & Stoneking, M. (2009) Global diversity in the human salivary microbiome. *Genome Res*.
71. Li, J. Z., Absher, D. M., Tang, H., Southwick, A. M., Casto, A. M., Ramachandran, S., Cann, H. M., Barsh, G. S., Feldman, M., Cavalli-Sforza, L. L.*, et al.* (2008) Worldwide human relationships inferred from genome-wide patterns of variation. *Science* 319: 1100-1104.
72. Darwin, C. (1871) *Descent of man and selection in relation to sex*.
73. Tishkoff, S. A., Reed, F. A., Ranciaro, A., Voight, B. F., Babbitt, C. C., Silverman, J. S., Powell, K., Mortensen, H. M., Hirbo, J. B., Osman, M.*, et al.* (2007) Convergent adaptation of human lactase persistence in africa and europe. *Nat Genet* 39: 31-40.
74. Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K., & Wilson, A. C. (1991) African populations and the evolution of human mitochondrial DNA. *Science* 253: 1503-1507.
75. Underhill, P. A., Shen, P., Lin, A. A., Jin, L., Passarino, G., Yang, W. H., Kauffman, E., Bonné-Tamir, B., Bertranpetit, J., Francalacci, P.*, et al.* (2000) Y chromosome sequence variation and the history of human populations. *Nat Genet* 26: 358-361.
76. Ke, Y., Su, B., Song, X., Lu, D., Chen, L., Li, H. N., Qi, C., Marzuki, S., Deka, R., Underhill, P.*, et al.* (2001) African origin of modern humans in east asia: A tale of 12,000 y chromosomes. *Science* 292: 1151-1153.
77. Paabo, S. (2001) Genomics and society: The human genome and our view of ourselves. *Science* 291: 1219-1220.
78. Evans, P. D., Mekel-Bobrov, N., Vallender, E. J., Hudson, R. R., & Lahn, B. T. (2006) Evidence that the adaptive allele of the brain size gene microcephalin introgressed into homo sapiens from an archaic homo lineage. *Proc Natl Acad Sci U S A*: 0606966103.
79. Plagnol, V. & Wall, J. D. (2006) Possible ancestral structure in human populations. *PLoS Genetics* 2: e105.
80. Serre, D., Langaney, A., Chech, M., Teschler-Nicola, M., Paunovic, M., Mennecier, P., Hofreiter, M., Possnert, G., & Paabo, S. (2004) No evidence of neandertal mtdna contribution to early modern humans. *PLoS Biol* 2: e57.
81. Huff, C., Xing, J., Rogers, A., Witherspoon, D., & Jorde, L. (2010) Mobile elements reveal small population size in the ancient ancestors of homo sapiens. *Proc Natl Acad Sci U S A*.

82. Tenesa, A., Navarro, P., Hayes, B. J., Duffy, D. L., Clarke, G. M., Goddard, M. E., & Visscher, P. M. (2007) Recent human effective population size estimated from linkage disequilibrium. *Genome Res* 17: 520-526.
83. Kaessmann, H., Wiebe, V., Weiss, G., & Pääbo, S. (2001) Great ape DNA sequences reveal a reduced diversity and an expansion in humans. *Nat Genet* 27: 155-156.
84. Won, Y. J. & Hey, J. (2005) Divergence population genetics of chimpanzees. *Mol Biol Evol* 22: 297-307.
85. Keinan, A., Mullikin, J. C., Patterson, N., & Reich, D. (2007) Measurement of the human allele frequency spectrum demonstrates greater genetic drift in east asians than in europeans. *Nat Genet* 39: 1251-1255.
86. Kimura, M. (1968) Evolutionary rate at the molecular level. *Nature* 217: 624-626.
87. Lindblad-Toh, K., Wade, C. M., Mikkelsen, T. S., Karlsson, E. K., Jaffe, D. B., Kamal, M., Clamp, M., Chang, J. L., Kulbokas, E. J., Zody, M. C.*, et al.* (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438: 803.
88. Nguyen, D. Q., Webber, C., Hehir-Kwa, J., Pfundt, R., Veltman, J., & Ponting, C. P. (2008) Reduced purifying selection prevails over positive selection in human copy number variant evolution. *Genome Res* 18: 1711-1723.
89. Smith, N. G. C. & Eyre-Walker, A. (2002) Adaptive protein evolution in drosophila. *Nature* 415: 1022-1024.
90. Boyko, A. R., Williamson, S. H., Indap, A. R., Degenhardt, J. D., Hernandez, R. D., Lohmueller, K. E., Adams, M. D., Schmidt, S., Sninsky, J. J., Sunyaev, S. R.*, et al.* (2008) Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genetics* 4: e1000083.
91. Bakewell, M. A., Shi, P., & Zhang, J. (2007) More genes underwent positive selection in chimpanzee evolution than in human evolution. *Proc Natl Acad Sci U S A* 104: 7489-7494
92. Ohno, S. (1972) An argument for the genetic simplicity of man and other mammals. *J Hum Evol* 1: 651-662.
93. Nielsen, R., Hellmann, I., Hubisz, M., Bustamante, C., & Clark, A. G. (2007) Recent and ongoing selection in the human genome. *Nat Rev Genet* 8: 857-868.
94. Sabeti, P. C., Schaffner, S. F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., Palma, A., Mikkelsen, T. S., Altshuler, D., & Lander, E. S. (2006) Positive natural selection in the human lineage. *Science* 312: 1614-1620.
95. Gilad, Y., Oshlack, A., & Rifkin, S. A. (2006) Natural selection on gene expression. *Trends Genet* 22: 456-461.
96. Khaitovich, P., Enard, W., Lachmann, M., & Paabo, S. (2006) Evolution of primate gene expression. *Nat Rev Genet* 7: 693-702.
97. Li, W. H., Wu, C. I., & Luo, C. C. (1985) A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol Biol Evol* 2: 150-174.
98. McDonald, J. H. & Kreitman, M. (1991) Adaptive protein evolution at the adh locus in drosophila. *Nature* 351: 652-654.
99. Pollard, K. S., Salama, S. R., Lambert, N., Lambot, M., Coppens, S., Pedersen, J. S., Katzman, S., King, B., Onodera, C., Siepel, A.*, et al.* (2006) An rna gene expressed during cortical development evolved rapidly in humans. *Nature* advanced online publication.
100. Kim, S. Y. & Pritchard, J. K. (2007) Adaptive evolution of conserved noncoding elements in mammals. *PLoS Genetics* 3: e147.
101. Andolfatto, P. (2005) Adaptive evolution of non-coding DNA in drosophila. *Nature* 437: 1149-1152.
102. Sella, G., Petrov, D. A., Przeworski, M., & Andolfatto, P. (2009) Pervasive natural selection in the drosophila genome? *PLoS Genetics* 5: e1000495.
103. Voight, B. F., Kudaravalli, S., Wen, X., & Pritchard, J. K. (2006) A map of recent positive selection in the human genome. *PLoS Biol* 4: e72.
104. Cai, J. J., Macpherson, J. M., Sella, G., & Petrov, D. A. (2009) Pervasive hitchhiking at coding and regulatory sites in humans. *PLoS Genetics* 5: e1000336.
105. Nei, M. (2007) The new mutation theory of phenotypic evolution. *Proc Natl Acad Sci U S A* 104: 12235-12242.
106. Bustamante, C. D., Fledel-Alon, A., Williamson, S., Nielsen, R., Hubisz, M. T., Glanowski, S., Tanenbaum, D. M., White, T. J., Sninsky, J. J., Hernandez, R. D.*, et al.* (2005) Natural selection on protein-coding genes in the human genome. *Nature* 437: 1153-1157.
107. Keightley, P. D., Lercher, M. J., & Eyre-Walker, A. (2005) Evidence for widespread degradation of gene control regions in hominid genomes. *PLoS Biol* 3: e42.
108. Gilad, Y., Man, O., Paabo, S., & Lancet, D. (2003) Human specific loss of olfactory receptor genes. *Proc Natl Acad Sci U S A* 100: 3324-3327.

109. Fischer, A., Gilad, Y., Man, O., & Paabo, S. (2005) Evolution of bitter taste receptors in humans and apes. *Mol Biol Evol* 22: 432-436.
110. Olson, M. V. (1999) When less is more: Gene loss as an engine of evolutionary change. *Am J Hum Genet* 64: 18-23.
111. Stedman, H. H., Kozyak, B. W., Nelson, A., Thesier, D. M., Su, L. T., Low, D. W., Bridges, C. R., Shrager, J. B., Minugh-Purvis, N., & Mitchell, M. A. (2004) Myosin gene mutation correlates with anatomical changes in the human lineage. *Nature* 428: 415-418.
112. Wang, X., Grus, W. E., & Zhang, J. (2006) Gene losses during human origins. *PLoS Biol* 4: e52.
113. Xue, Y., Daly, A., Yngvadottir, B., Liu, M., Coop, G., Kim, Y., Sabeti, P., Chen, Y., Stalker, J., Huckle, E.*, et al.* (2006) Spread of an inactive form of caspase-12 in humans is due to recent positive selection. *Am J Hum Genet* 78: 659-670.
114. Clark, A. G., Glanowski, S., Nielsen, R., Thomas, P. D., Kejariwal, A., Todd, M. A., Tanenbaum, D. M., Civello, D., Lu, F., Murphy, B.*, et al.* (2003) Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science* 302: 1960-1963.
115. Arbiza, L., Dopazo, J., & Dopazo, H. (2006) Positive selection, relaxation, and acceleration in the evolution of the human and chimp genome. *PLoS Computational Biology* 2: e38.
116. Kosiol, C., Vinar, T., da Fonseca, R. R., Hubisz, M. J., Bustamante, C. D., Nielsen, R., & Siepel, A. (2008) Patterns of positive selection in six mammalian genomes. *PLoS Genetics* 4: e1000144.
117. Haygood, R., Fedrigo, O., Hanson, B., Yokoyama, K., & Wray, G. A. (2007) Promoter regions of many neural- and nutrition-related genes have experienced positive selection during human evolution. *Nat Genet* 39: 1140-1144.
118. Pollard, K. S., Salama, S. R., King, B., Kern, A. D., Dreszer, T., Katzman, S., Siepel, A., Pedersen, J. S., Bejerano, G., Baertsch, R.*, et al.* (2006) Forces shaping the fastest evolving regions in the human genome. *PLoS Genetics* 2: e168.
119. Bird, C., Stranger, B., Liu, M., Thomas, D., Ingle, C., Beazley, C., Miller, W., Hurles, M., & Dermitzakis, E. (2007) Fast-evolving noncoding sequences in the human genome. *Genome Biology* 8: R118.
120. Prabhakar, S., Noonan, J. P., Paabo, S., & Rubin, E. M. (2006) Accelerated evolution of conserved noncoding sequences in humans. *Science* 314: 786-.
121. Clark, A. G., Glanowski, S., Nielsen, R., Thomas, P. D., Kejariwal, A., Todd, M. A., Tanenbaum, D. M., Civello, D., Lu, F., Murphy, B.*, et al.* (2003) Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science* 302: 1960-1963.
122. Gilad, Y., Segre, D., Skorecki, K., Nachman, M. W., Lancet, D., & Sharon, D. (2000) Dichotomy of single-nucleotide polymorphism haplotypes in olfactory receptor genes and pseudogenes. *Nat Genet* 26: 221-224.
123. Nielsen, R., Bustamante, C., Clark, A. G., Glanowski, S., Sackton, T. B., Hubisz, M. J., Fledel-Alon, A., Tanenbaum, D. M., Civello, D., White, T. J.*, et al.* (2005) A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol* 3: e170.
124. Ahituv, N., Zhu, Y., Visel, A., Holt, A., Afzal, V., Pennacchio, L. A., & Rubin, E. M. (2007) Deletion of ultraconserved elements yields viable mice. *PLoS Biol* 5: e234.
125. Prabhakar, S., Visel, A., Akiyama, J. A., Shoukry, M., Lewis, K. D., Holt, A., Plajzer-Frick, I., Morrison, H., FitzPatrick, D. R., Afzal, V.*, et al.* (2008) Human-specific gain of function in a developmental enhancer. *Science* 321: 1346-1350.
126. Taylor, M. S., Massingham, T., Hayashizaki, Y., Carninci, P., Goldman, N., & Semple, C. A. (2008) Rapidly evolving human promoter regions. *Nat Genet* 40: 1262-1263; author reply 1263-1264.
127. Galtier, N. & Duret, L. (2007) Adaptation or biased gene conversion? Extending the null hypothesis of molecular evolution. *Trends Genet* 23: 273-277.
128. Duret, L. & Galtier, N. (2009) Comment on "Human-specific gain of function in a developmental enhancer". *Science* 323: 714; author reply 714.
129. Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S., Wiebe, V., Kitano, T., Monaco, A. P., & Paabo, S. (2002) Molecular evolution of foxp2, a gene involved in speech and language. *Nature* 418: 869-872.
130. Coop, G., Bullaughey, K., Luca, F., & Przeworski, M. (2008) The timing of selection at the human foxp2 gene. *Mol Biol Evol*: msn091.
131. Enard, W., Gehre, S., Hammerschmidt, K., Hölter, S., Blass, T., Somel, M., Brückner, M., Schreiweis, C., Winter, C., & Sohr, R. (2009) A humanized version of foxp2 affects cortico-basal ganglia circuits in mice. *Cell* 137: 961-971.

132. Konopka, G., Bomar, J. M., Winden, K., Coppola, G., Jonsson, Z. O., Gao, F., Peng, S., Preuss, T. M., Wohlschlegel, J. A., & Geschwind, D. H. (2009) Human-specific transcriptional regulation of cns development genes by foxp2. *Nature* 462: 213-217.
133. Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R. E., Burbano, H. A., Hublin, J. J., Hänni, C., Fortea, J., de la Rasilla, M.*, et al.* (2007) The derived foxp2 variant of modern humans was shared with neandertals. *Curr Biol* 17: 1908-1912.
134. Sawyer, S. L., Emerman, M., & Malik, H. S. (2004) Ancient adaptive evolution of the primate antiviral DNA-editing enzyme apobec3g. *PLoS Biol* 2: E275.
135. Noonan, J. P., Coop, G., Kudaravalli, S., Smith, D., Krause, J., Alessi, J., Chen, F., Platt, D., Paabo, S., Pritchard, J. K.*, et al.* (2006) Sequencing and analysis of neanderthal genomic DNA. *Science* 314: 1113-1118.
136. Green, R. E., Krause, J., Ptak, S. E., Briggs, A. W., Ronan, M. T., Simons, J. F., Du, L., Egholm, M., Rothberg, J. M., Paunovic, M.*, et al.* (2006) Analysis of one million base pairs of neanderthal DNA. *Nature* 444: 330-336.
137. Enard, W., Khaitovich, P., Klose, J., Zollner, S., Heissig, F., Giavalisco, P., Nieselt-Struwe, K., Muchmore, E., Varki, A., Ravid, R.*, et al.* (2002) Intra- and interspecific variation in primate gene expression patterns. *Science* 296: 340-343.
138. Caceres, M., Lachuer, J., Zapala, M. A., Redmond, J. C., Kudo, L., Geschwind, D. H., Lockhart, D. J., Preuss, T. M., & Barlow, C. (2003) Elevated gene expression levels distinguish human from non-human primate brains. *Proc Natl Acad Sci U S A* 100: 13030-13035.
139. Khaitovich, P., Hellmann, I., Enard, W., Nowick, K., Leinweber, M., Franz, H., Weiss, G., Lachmann, M., & Paabo, S. (2005) Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* 309: 1850-1854.
140. Gilad, Y., Oshlack, A., Smyth, G. K., Speed, T. P., & White, K. P. (2006) Expression profiling in primates reveals a rapid evolution of human transcription factors. *Nature* 440: 242-245.
141. Uddin, M., Wildman, D. E., Liu, G., Xu, W., Johnson, R. M., Hof, P. R., Kapatos, G., Grossman, L. I., & Goodman, M. (2004) Sister grouping of chimpanzees and humans as revealed by genome-wide phylogenetic analysis of brain gene expression profiles. *Proc Natl Acad Sci U S A* 101: 2957-2962.
142. Lemos, B., Meiklejohn, C. D., Caceres, M., & Hartl, D. L. (2005) Rates of divergence in gene expression profiles of primates, mice, and flies: Stabilizing selection and variability among functional categories. *Evolution Int J Org Evolution* 59: 126-137.
143. Rifkin, S. A., Houle, D., Kim, J., & White, K. P. (2005) A mutation accumulation assay reveals a broad capacity for rapid evolution of gene expression. *Nature* 438: 220.
144. Hsieh, W., Chu, T., Wolfinger, R. D., & Gibson, G. (2003) Mixed-model reanalysis of primate data suggests tissue and species biases in oligonucleotide-based gene expression profiles. *Genetics* 165: 747-757.
145. Preuss, T. M., Caceres, M., Oldham, M. C., & Geschwind, D. H. (2004) Human brain evolution: Insights from microarrays. *Nat Rev Genet* 5: 850-860.
146. Khaitovich, P., Weiss, G., Lachmann, M., Hellmann, I., Enard, W., Muetzel, B., Wirkner, U., Ansorge, W., & bo, S. (2004) A neutral model of transcriptome evolution. *PLoS Biol* 2: e132.
147. Oleksiak, M. F., Churchill, G. A., & Crawford, D. L. (2002) Variation in gene expression within and among natural populations. *Nat Genet* 32: 261.
148. Blekhman, R., Marioni, J. C., Zumbo, P., Stephens, M., & Gilad, Y. (2010) Sex-specific and lineage-specific alternative splicing in primates. *Genome Res* 20: 180-189.
149. Prüfer, K., Muetzel, B., Do, H., Weiss, G., Khaitovich, P. R., Pääbo, S., Lachmann, M., & Enard, W. (2007) Func: A package for detecting significant associations between gene sets and ontological annotations *BMC Bioinformatics* 8: 41.
150. Whitehead, A. & Crawford, D. L. (2006) Neutral and adaptive variation in gene expression. *Proc Natl Acad Sci U S A*: 0507648103.
151. Lynch, M. (2006) The origins of eukaryotic gene structure. *Mol Biol Evol* 23: 450-468.
152. Fay, J. C. & Wittkopp, P. J. (2007) Evaluating the role of natural selection in the evolution of gene regulation. *Heredity* 100: 191-199.
153. Fraser, H. B., Moses, A. M., & Schadt, E. E. (2010) Evidence for widespread adaptive evolution of gene expression in budding yeast. *Proc Natl Acad Sci USA*.

154. Khaitovich, P., Tang, K., Franz, H., Kelso, J., Hellmann, I., Enard, W., Lachmann, M., & Paabo, S. (2006) Positive selection on gene expression in the human brain. *Curr Biol* 16: R356-R358.
155. Somel, M., Creely, H., Franz, H., Mueller, U., Lachmann, M., Khaitovich, P., & Paabo, S. (2008) Human and chimpanzee gene expression differences replicated in mice fed different diets. *PLoS ONE* 3: e1504.
156. Gould, S. J. & Lewontin, R. C. (1979) The spandrels of san marco and the panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London Series B, Biological Sciences* 205: 581-598.
157. Hahn, M. W. (2008) Toward a selection theory of molecular evolution. *Evolution Int J Org Evolution* 62: 255-265.
158. Nei, M. (2005) Selectionism and neutralism in molecular evolution. *Mol Biol Evol* 22: 2318-2342.
159. Somel, M., Franz, H., Yan, Z., Lorenc, A., Guo, S., Giger, T., Kelso, J., Nickel, B., Dannemann, M., Bahn, S*., et al.* (2009) Transcriptional neoteny in the human brain. *Proc Natl Acad Sci USA*.
160. Khaitovich, P., Lockstone, H. E., Wayland, M. T., Tsang, T. M., Jayatilaka, S. D., Guo, A. J., Zhou, J., Somel, M., Harris, L. W., Holmes, E*., et al.* (2008) Metabolic changes in schizophrenia and human brain evolution. *Genome Biology* 9: R124.
161. Deacon, T. W. (1997) What makes the human brain different? *Annu Rev Anthropol* 26: 337-357.
162. Sherwood, C. C., Stimpson, C. D., Raghanti, M. A., Wildman, D. E., Uddin, M., Grossman, L. I., Goodman, M., Redmond, J. C., Bonar, C. J., Erwin, J. M*., et al.* (2006) Evolution of increased glia-neuron ratios in the human frontal cortex. *Proc Natl Acad Sci U S A*: 0605843103.
163. Khaitovich, P., Muetzel, B., She, X., Lachmann, M., Hellmann, I., Dietzsch, J., Steigele, S., Do, H., Weiss, G., Enard, W*., et al.* (2004) Regional patterns of gene expression in human and chimpanzee brains. *Genome Res* 14: 1462-1473.
164. Heissig, F., Krause, J., Bryk, J., Khaitovich, P., Enard, W., & Paabo, S. (2005) Functional analysis of human and chimpanzee promoters. *Genome Biology* 6: R57.
165. Hsieh, W., Passador-Gurgel, G., Stone, E., & Gibson, G. (2007) Mixture modeling of transcript abundance classes in natural populations. *Genome Biology* 8: R98.
166. Freimer, N. & Sabatti, C. (2003) The human phenome project. *Nat Genet* 34: 15-21.
167. Calarco, J. A., Xing, Y., Caceres, M., Calarco, J. P., Xiao, X., Pan, Q., Lee, C., Preuss, T. M., & Blencowe, B. J. (2007) Global analysis of alternative splicing differences between humans and chimpanzees. *Genes Dev* 21: 2963-2975.
168. Berezikov, E., Thuemmler, F., van Laake, L. W., Kondova, I., Bontrop, R., Cuppen, E., & Plasterk, R. H. A. (2006) Diversity of micrornas in human and chimpanzee brain. *Nat Genet* 38: 1375-1377.
169. Hu, H., Yan, Z., Xu, Y., Hu, H., Menzel, C., Zhou, Y. H., Chen, W., & Khaitovich, P. (2009) Sequence features associated with microrna strand selection in humans and flies. *BMC Genomics* 10: 413.
170. Enard, W., Fassbender, A., Model, F., Adorján, P., Pääbo, S., & Olek, A. (2004) Differences in DNA methylation patterns between humans and chimpanzees. *Curr Biol* 14: R148-R149.
171. Fu, N., Drinnenberg, I., Kelso, J., Wu, J. R., Paabo, S., Zeng, R., & Khaitovich, P. (2007) Comparison of protein and mrna expression evolution in humans and chimpanzees. *PLoS ONE* 2: e216.
172. Conrad, D. F., Jakobsson, M., Coop, G., Wen, X., Wall, J. D., Rosenberg, N. A., & Pritchard, J. K. (2006) A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nat Genet* 38: 1251-1260.
173. Pickrell, J. K., Coop, G., Novembre, J., Kudaravalli, S., Li, J. Z., Absher, D., Srinivasan, B. S., Barsh, G. S., Myers, R. M., Feldman, M. W*., et al.* (2009) Signals of recent positive selection in a worldwide sample of human populations. *Genome Res*.
174. Ramachandran, S., Deshpande, O., Roseman, C. C., Rosenberg, N. A., Feldman, M. W., & Cavalli-Sforza, L. L. (2005) Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in africa. *Proc Natl Acad Sci USA* 102: 15942-15947.
175. Coop, G., Pickrell, J. K., Novembre, J., Kudaravalli, S., Li, J., Absher, D. M., Myers, R. M., Cavalli-Sforza, L. L., Feldman, M. W., & Pritchard, J. K. (2009) The role of geography in human adaptation. *PLoS Genetics* 5: e1000500.
176. Sabeti, P. C., Schaffner, S. F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O., Palma, A., Mikkelsen, T. S., Altshuler, D., & Lander, E. S. (2006) Positive natural selection in the human lineage. *Science* 312: 1614-1620.
177. Allison, A. C. (1954) Protection afforded by sickle-cell trait against subtertian malareal infection. *Br Med J* 1: 290-294.
178. Currat, M., Trabuchet, G., Rees, D., Perrin, P., Harding, R. M., Clegg, J. B., Langaney, A., & Excoffier, L. (2002) Molecular analysis of the beta-globin gene cluster in the niokholo mandenka population reveals a recent origin of the beta(s) senegal mutation. *Am J Hum Genet* 70: 207-223.

179. Hamblin, M. T. & Di Rienzo, A. (2000) Detection of the signature of natural selection in humans: Evidence from the duffy blood group locus. *Am J Hum Genet* 66: 1669-1679.

180. Ohashi, J., Naka, I., Patarapotikul, J., Hananantachai, H., Brittenham, G., Looareesuwan, S., Clark, A. G., & Tokunaga, K. (2004) Extended linkage disequilibrium surrounding the hemoglobin e variant due to malarial selection. *Am J Hum Genet* 74: 1198-1208.

181. Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A., Belmont, J. W., Boudreau, A., Hardenbol, P., Leal, S. M.*, et al.* (2007) A second generation human haplotype map of over 3.1 million snps. *Nature* 449: 851-861.

182. Marshall, J. M. & Weiss, R. E. (2006) A bayesian heterogeneous analysis of variance approach to inferring recent selective sweeps. *Genetics* 173: 2357-2370.

183. Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E. H., McCarroll, S. A., Gaudet, R.*, et al.* (2007) Genome-wide detection and characterization of positive selection in human populations. *Nature* 449: 913-918.

184. Tang, K., Thornton, K. R., & Stoneking, M. (2007) A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biol* 5: e171.

185. Grossman, S. R., Shylakhter, I., Karlsson, E. K., Byrne, E. H., Morales, S., Frieden, G., Hostetter, E., Angelino, E., Garber, M., Zuk, O.*, et al.* (2010) A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science*: science.1183863v1183861.

186. Lamason, R. L., Mohideen, M. A., Mest, J. R., Wong, A. C., Norton, H. L., Aros, M. C., Jurynec, M. J., Mao, X., Humphreville, V. R., Humbert, J. E.*, et al.* (2005) Slc24a5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science* 310: 1782-1786.

187. Williamson, S. H., Hubisz, M. J., Clark, A. G., Payseur, B. A., Bustamante, C. D., & Nielsen, R. (2007) Localizing recent adaptive evolution in the human genome. *PLoS Genet* 3: e90.

188. Bersaglieri, T., Sabeti, P. C., Patterson, N., Vanderploeg, T., Schaffner, S. F., Drake, J. A., Rhodes, M., Reich, D. E., & Hirschhorn, J. N. (2004) Genetic signatures of strong recent positive selection at the lactase gene. *Am J Hum Genet* 74: 1111-1120.

189. Enattah, N. S., Jensen, T. G., Nielsen, M., Lewinski, R., Kuokkanen, M., Rasinpera, H., El-Shanti, H., Seo, J. K., Alifrangis, M., Khalil, I. F.*, et al.* (2008) Independent introduction of two lactase-persistence alleles into human populations reflects different history of adaptation to milk culture. *Am J Hum Genet* 82: 57-72.

190. Enattah, N. S., Sahi, T., Savilahti, E., Terwilliger, J. D., Peltonen, L., & Jarvela, I. (2002) Identification of a variant associated with adult-type hypolactasia. *Nat Genet* 30: 233-237.

191. Cappellini, M. D. & Fiorelli, G. (2008) Glucose-6-phosphate dehydrogenase deficiency. *Lancet* 371: 64-74.

192. Flint, J., Harding, R. M., Boyce, A. J., & Clegg, J. B. (1998) The population genetics of the haemoglobinopathies. *Baillieres Clin Haematol* 11: 1-51.

193. Hill, A. V. (1992) Molecular epidemiology of the thalassaemias (including haemoglobin e). *Baillieres Clin Haematol* 5: 209-238.

194. Norton, H. L., Kittles, R. A., Parra, E., McKeigue, P., Mao, X., Cheng, K., Canfield, V. A., Bradley, D. G., McEvoy, B., & Shriver, M. D. (2007) Genetic evidence for the convergent evolution of light skin in europeans and east asians. *Mol Biol Evol* 24: 710-722.

195. Hancock, A. M., Witonsky, D. B., Gordon, A. S., Eshel, G., Pritchard, J. K., Coop, G., & Di Rienzo, A. (2008) Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genet* 4: e32.

196. Thompson, E. E., Kuttab-Boulos, H., Witonsky, D., Yang, L., Roe, B. A., & Di Rienzo, A. (2004) Cyp3a variation and the evolution of salt-sensitivity variants. *Am J Hum Genet* 75: 1059-1069.

197. Young, J. H., Chang, Y. P., Kim, J. D., Chretien, J. P., Klag, M. J., Levine, M. A., Ruff, C. B., Wang, N. Y., & Chakravarti, A. (2005) Differential susceptibility to hypertension is due to selection during the out-of-africa expansion. *PLoS Genet* 1: e82.

198. Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., Fiegler, H., Shapero, M. H., Carson, A. R., Chen, W.*, et al.* (2006) Global variation in copy number in the human genome. *Nature* 444: 444-454.

199. Perry, G. H., Dominy, N. J., Claw, K. G., Lee, A. S., Fiegler, H., Redon, R., Werner, J., Villanea, F. A., Mountain, J. L., Misra, R.*, et al.* (2007) Diet and the evolution of human amylase gene copy number variation. *Nat Genet* 39: 1256-1260.

200. Linz, B., Balloux, F., Moodley, Y., Manica, A., Liu, H., Roumagnac, P., Falush, D., Stamer, C., Prugnolle, F., van der Merwe, S. W.*, et al.* (2007) An african origin for the intimate association between humans and helicobacter pylori. *Nature* 445: 915-918.