

# Genetics and the making of *Homo sapiens*

Sean B. Carroll

Howard Hughes Medical Institute and Laboratory of Molecular Biology, University of Wisconsin, 1525 Linden Drive, Madison, Wisconsin 53706, USA

**Understanding the genetic basis of the physical and behavioural traits that distinguish humans from other primates presents one of the great new challenges in biology. Of the millions of base-pair differences between humans and chimpanzees, which particular changes contributed to the evolution of human features after the separation of the *Pan* and *Homo* lineages 5–7 million years ago? How can we identify the ‘smoking guns’ of human genetic evolution from neutral ticks of the molecular evolutionary clock? The magnitude and rate of morphological evolution in hominids suggests that many independent and incremental developmental changes have occurred that, on the basis of recent findings in model animals, are expected to be polygenic and regulatory in nature. Comparative genomics, population genetics, gene-expression analyses and medical genetics have begun to make complementary inroads into the complex genetic architecture of human evolution.**

What is a man,  
If his chief good and market of his time  
Be but to sleep and feed? a beast, no more.  
Sure, he that made us with such large discourse,  
Looking before and after, gave us not  
That capability and god-like reason  
To fust in us unused.

W. Shakespeare, *Hamlet* IV:iv

**W**hat makes modern humans different from the great apes and earlier hominids? In what hominids and when in evolution did important physical traits and behaviours appear? Where in our larger brains do human-specific capabilities reside? These have been long-standing questions in palaeoanthropology and comparative anatomy, since the discovery of Neanderthal skulls and the first studies of great apes in the nineteenth century. Now, the mystery of human origins is expanding beyond the description and history of human traits, towards the genetic mechanisms underlying their formation and evolution. With the characterization of the human genome, and that of our chimpanzee cousin on the way, the quest to discover the genetic basis of the physical and behavioural traits that distinguish us from other apes is rapidly gaining momentum.

Genomes diverge as a function of time, and most of the sequence changes that accumulate between any two related species are selectively neutral or nearly neutral in that they do not contribute to functional or phenotypic differences. The great challenge is to elucidate the number, identity and functions of genes, and the specific changes within them, that have shaped the evolution of traits. This has been accomplished for only a few traits in model systems, so it is a difficult task for human features about which we know little, and an enormous prospect to consider the whole arc of human evolution.

In this article, I will examine both the physical and genetic scope of human evolution and the approaches being used to try to understand it. I will first review the current state of our understanding of human evolution from the viewpoint of the fossil record, comparative anatomy and development. These disciplines point to many key traits to be considered, and define the magnitude and nature of evolutionary change in the human lineage. I will preview the picture of human evolution that we might expect to emerge in view of our current knowledge of the genetic architecture of trait evolution in model systems. I will then examine the variety of methods being used on a genome-wide scale and at the level of individual loci to identify genes that may have contributed to the evolution of key traits. I will discuss some of the crucial methodological challenges in distinguishing causative from potentially large

numbers of candidate loci. Finally, I will address some of the disciplines in which future advances are likely to have a central role in furthering our knowledge of the genetic and developmental basis of human evolution.

## Hominid evolution

### The hominid tree

To approach the origins of human traits at the genetic level, it is essential to have as a framework a history of our lineage and the characters that distinguish it. It is inadequate and misleading to consider just the comparative anatomy and development (or genomes) of extant humans, chimpanzees and other apes, and then to attempt to infer how existing differences might be encoded and realized. Each of these species has an independent lineage that reaches back as far or further than hominins (‘hominins’ refers to humans and our evolutionary ancestors back to the separation of the human and ape lineages; ‘hominids’ to humans and the African apes) (Fig. 1). The evolution of ‘modern’ traits was not a linear, additive process, and ideas about the tempo, pattern and magnitude of change can only be tested through fossil evidence, which is always subject to revision by new finds. The fossil record continues to shape views of three crucial issues in hominid evolution. First, what distinguishes hominins from the apes? Second, what distinguishes modern humans (*Homo sapiens*) from earlier hominins? And third, what was the nature of the last common ancestor of hominins and the *Pan* lineage?

Inferences about the chronological order and magnitude of the evolution of hominin characters depend on a model of the hominin evolutionary tree. This has always been a contentious issue and remains unsettled, in part because of the pace of exciting new fossil finds over the past two decades. An emerging view portrays hominin evolution as a series of adaptive radiations in which many different branches of the hominin lineage were formed, but died out<sup>1,2</sup> (Fig. 1). One prediction of this model is that various anatomical features would be found in different combinations in hominins through their independent acquisition, modification and loss in different species. For example, the recent, stunning discovery of a 6–7 million year (Myr)-old fossil cranium of *Sahelanthropus tchadensis* that had a chimpanzee-sized brain but hominin-like facial and dental features<sup>3</sup> is the sort of morphology that would be consistent with a radiation of ape-like animals from which the stem of the hominin lineage emerged (although the interpretation of this fossil’s affinities is controversial<sup>4</sup>).

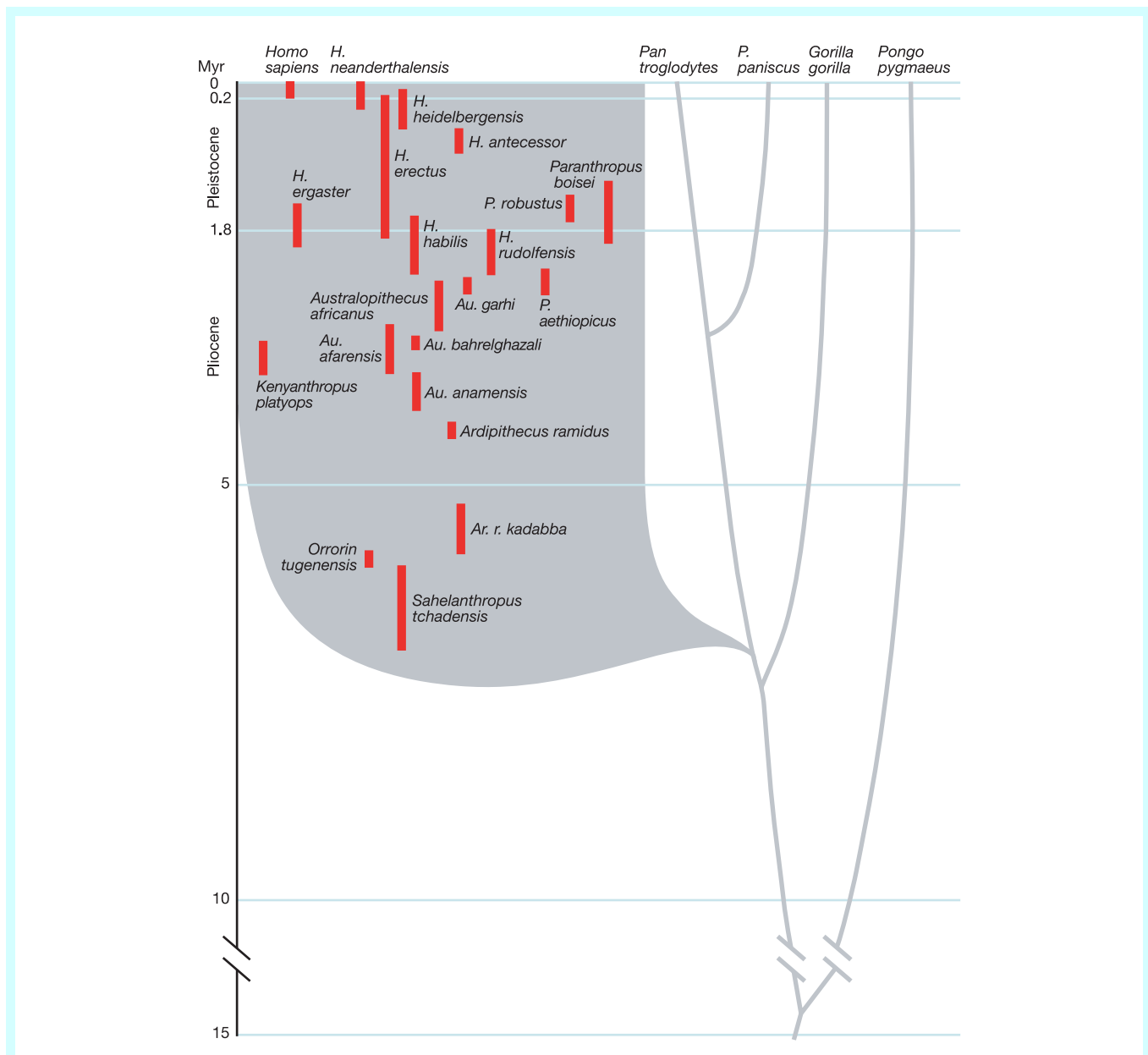
### What makes a human?

**Evolutionary trends in fossil hominins.** Ideally, if every possible fossil human and ape species were identified and many fairly

complete specimens were available, one could reconstruct the emergence of human and chimpanzee features through time. But that is not the case for most lineages; in fact, there are no identified archaic chimpanzee fossils. We must make do with a partial and often confusing picture of human trait evolution. From a rather extensive list of qualitative and quantitative features that distinguish humans from other apes<sup>5,6</sup> (Box 1), our large brain, bipedalism, small canine teeth, language and advanced tool-making capabilities<sup>7,8</sup> have been the focus of palaeoanthropology. The major physical traits are generally not singular elements, but entail concomitant changes in skeletal features involved in locomotion (for example, in the vertebral column, pelvis and feet, and in limb proportions), grasping (hand morphology and an opposable,

elongated thumb) and chewing of food (the mandible and dentition), as well as life-history traits such as lifespan. It is fortunate that most of the skeletal features lend themselves to detailed quantitative studies of the fossil record.

Some trends in the evolution of body size, brain size and dentition are evident within the hominins. More recent species are characterized by larger body mass, relatively larger brains, longer legs relative to the trunk and small teeth, whereas earlier species had, in general, smaller brains and bodies (Table 1), shorter legs relative to the trunk and large teeth<sup>7,9,10</sup>. I highlight these traits to focus attention on the magnitude and timescale of character evolution, and on the (increasing) number of generally recognized hominin taxa. Whatever the branching pattern in the hominin tree, sub-



**Figure 1** The timescale and phylogeny of hominids. Ape relationships are shown in grey for the chimpanzee (*Pan troglodytes*), bonobo (*P. paniscus*), gorilla and orangutan (*Pongo pygmaeus*). The approximate times of divergence are derived from molecular data (summarized in ref. 89). The phylogenetic relationships among hominins (shaded) are uncertain. The solid red bars denote the time span of the fossil species and/or the uncertainty of fossil ages. The identity of the last common ancestor of chimpanzees and

humans (LCA) is not known. Note that the estimated age of *Sahelanthropus tchadensis* predates molecular estimates of the time of the chimpanzee–human divergence. This species could pre- or postdate the LCA. Also note that *Homo sapiens* represent only the last 3% of the time span of hominin evolution. Hominin distributions and nomenclature are based primarily on refs 1, 90.

stantial relative changes occurred over an extended time span and a significant number of speciation events. There was a marked increase in absolute brain size by the Early Pleistocene and again in the Middle Pleistocene, with a long interval of perhaps 1 Myr during which brain size did not change significantly<sup>9,11</sup> (Table 1). With regard to modern *H. sapiens*, it is interesting to note that body and brain size were even greater in *H. neanderthalensis*; there is no obvious physical explanation for the success of *H. sapiens* and the demise of *H. neanderthalensis*<sup>11</sup>.

**A beautiful mind: insights from comparative neuroanatomy.** The relative increase in brain size, although marked, is only a crude index of a potential increase in cognitive abilities. Because it has long been appreciated that there are discrete areas of the brain that process various cognitive, motor and sensory functions, comparative neuroanatomists have sought to identify areas that might be central to the evolution of human capacities. There is a long-standing notion that the frontal cortex (involved in planning, organization, personality, behaviour and other 'higher' cognitive functions) is disproportionately larger in humans<sup>12</sup>, but this now seems not to be the case (it is larger, but not disproportionately so<sup>13</sup>). As gross anatomical differences do not account for cognitive capabilities, relative differences in the size, cellular composition, detailed cytoarchitecture and/or connectivity of human and great ape brain areas have been sought to explain the emergence of human capabilities.

Of paramount interest is the production and understanding of speech. Two areas in particular have commanded the greatest attention. One is Broca's area in the frontal lobe of the neocortex (Fig. 2). This region is larger in the left hemisphere of the brain than in the right, an asymmetry that has been correlated with language ability. From magnetic resonance images of chimpanzees, bonobos and gorillas, a similar left–right asymmetry has been found in these great apes<sup>14</sup> (Fig. 2). This indicates that the neuroanatomical substrate of left-hemisphere dominance in speech production preceded the origin of hominins. The left hemisphere also usually controls right-handedness, so it is interesting to note that in captive apes, manual gestures are right-hand biased, and this bias is increased when vocalization is combined with gesturing<sup>15</sup>, indicating a left-hemisphere-controlled communication process.

A second area of interest is Wernicke's posterior receptive language area in the temporal lobe (Fig. 2). A site within this area, the planum temporale, is implicated in human communication (both spoken and gestural) and musical talent, and also shows a left-hemisphere dominance. In most humans, the Sylvian fissure associated with the left planum temporale extends more posteriorly. Evidence for this asymmetry has been found in fossil

endocasts in *H. habilis*, *H. erectus* and *H. neanderthalensis*<sup>16</sup>. More importantly, an asymmetrical planum temporale pattern has recently been demonstrated in chimpanzees<sup>17,18</sup> (Fig. 2).

Several hypotheses have been forwarded to explain the presence of these two human-communication-associated neuroanatomical landmarks in great apes<sup>14,17</sup>. Although it is possible that they arose and acquired their functions independently in each lineage, the most parsimonious explanation is that the common ancestor of great apes and humans had asymmetrical centres that were involved in communication, and that these structures underwent independent evolutionary modifications in chimpanzees and hominins. If this is the case, then the challenge to comparative neuroanatomy is to identify more subtle differences in suborganization (that is, 'microanatomy') that affect the interconnections of cortical regions, in local circuitry and/or cytoarchitecture<sup>19,20</sup> that might be unique to human brains. Recently, it has been found that the dimensions of the vertical columns of neurons in the cortex, known as 'mini-columns', differ between humans and chimpanzees in the planum temporale<sup>21</sup>. In addition, area 10 of the prefrontal cortex, which is involved in higher cognitive functions, has been shown to be enlarged and specialized in humans relative to apes<sup>22</sup>. These observations suggest that human capacities are more a product of quantitative changes in specialized areas than of neuroanatomical novelties.

#### Development of hominid features

The morphological differences between modern humans, earlier hominins and the great apes are, of course, the product of changes during development. Comparative studies of human and chimpanzee skull development<sup>23</sup>, remarkably detailed examinations of Neanderthal craniofacial ontogeny<sup>22,24–26</sup>, and early hominin dentition formation<sup>27</sup> have yielded crucial insights into a variety of developmental shifts that underlie modern human cranial size and morphology.

One of the long-appreciated, fundamental differences in chimpanzee and human development is the relative rate of skull growth and maturation. Human neonates have less mature skulls in terms of shape in comparison to young chimpanzees, but much larger skulls (and brains)<sup>23,28</sup>. These are classically described as heterochronic changes, which produce neotenic features in which maturation is retarded, size increases and shape resembles juvenile forms of ancestors<sup>29</sup>. Chimpanzee and human skulls eventually grow to the same size, reflecting further relative shifts in juvenile and adolescent growth periods, and marked differences in face size and cerebral volume. Importantly, all of the skeletal changes associated with bipedalism are structural innovations independent of neoteny. These observations suggest that the human brain is not a product of simple shifts in growth relationships, but of multiple, independent and superimposed modifications.

#### Box 1

##### Selected traits that distinguish humans from other apes<sup>5–7</sup>

- Body shape and thorax
- Cranial properties (brain case and face)
- Relative brain size
- Relative limb length
- Long ontogeny and lifespan
- Small canine teeth
- Skull balanced upright on vertebral column
- Reduced hair cover
- Elongated thumb and shortened fingers
- Dimensions of the pelvis
- Presence of a chin
- S-shaped spine
- Language
- Advanced tool making
- Brain topology

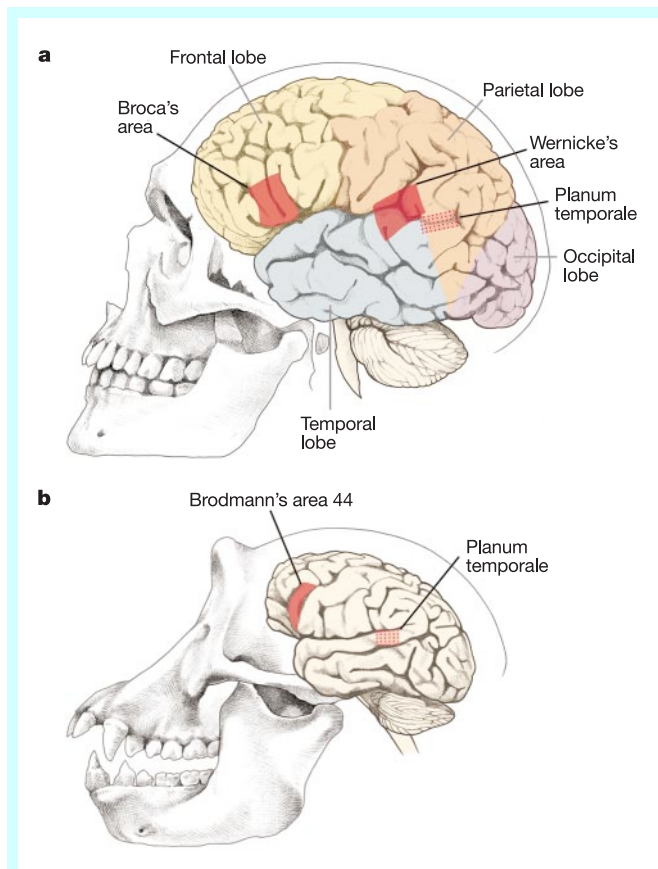
Table 1 Evolution of brain and body size in hominids

Species	Estimated age <sup>1,3,7,86–88</sup> (Myr ago)	Body size <sup>7</sup> (kg)	Brain size <sup>3,7</sup> (cm <sup>3</sup> )
<i>Homo sapiens</i>	0–0.2	53	1,355
<i>H. neanderthalensis</i>	0.03–0.25	76	1,512
<i>H. heidelbergensis</i>	0.3–1	62	1,198
<i>H. erectus</i>	0.2–1.9	57	1,016
<i>H. ergaster</i>	1.5–1.9	58	854
<i>H. rudolfensis</i>	1.8–2.4	-	752
<i>H. habilis</i>	1.6–2.3	34	552
<i>Paranthropus boisei</i>	1.2–2.2	44	510
<i>Australopithecus africanus</i>	2.6–3	36	457
<i>Au. afarensis</i>	3–3.6	-	-
<i>Au. anamensis</i>	3.5–4.1	-	-
<i>Ardipithecis ramidus kadabba</i>	5.2–5.8	-	-
<i>Sahelanthropus tchadensis</i>	6–7	-	~320–380

This list does not include all recognized species.

With respect to more recent hominin species, modern human craniofacial form appears to have been shaped by changes in elements that influence the spatial position of the face, neurocranium and cranial base. Modern humans are marked by greater roundedness of the cranial vault and facial retraction (the antero-posterior position of the face relative to the cranial base and neurocranium)<sup>25</sup>. Comparisons with Neanderthal skull development inferred from fossils of different aged individuals suggest that the characteristic cranial differences between Neanderthal and modern human features arise early in ontogeny<sup>24,26</sup>.

Our prolonged childhood, delayed sexual maturation and long lifespan are life-history traits that shape important aspects of human society. Insights into the evolution of these development shifts in evolution are offered by detailed comparative analysis of dental development, which is correlated with stages of primate growth and development. Rates of enamel formation in fossil hominins suggest that tooth-formation times were shorter in australopiths and early members of the genus *Homo* than they are in modern humans<sup>27</sup>. This indicates that the modern pattern of dental formation and correlated developmental traits appeared late in human evolution. When considered in the context of other traits, such as brain size and body proportions, a mosaic pattern of evolution emerges with different traits appearing at different times and perhaps in different combinations in hominin history<sup>30</sup>.



**Figure 2** Comparative neuroanatomy of humans and chimpanzees. Lateral views of the left hemispheres of a modern human and a chimpanzee brain. Although the overall skull sizes are roughly comparable, the human cranial capacity and brain are much larger. **a**, Two areas of the human brain that are associated with communication are shown: Broca's area in the frontal lobe and Wernicke's area, which overlaps the posterior temporal lobe and parts of the parietal lobe. In the left hemisphere, Broca's area is larger, as is the planum temporale, which lies below the surface in Wernicke's area. **b**, These asymmetries have been found in corresponding regions of chimpanzee brains<sup>15,17</sup>, suggesting that the areas in humans might be elaborations of a pre-existing communication centre in a common ancestor of apes and humans.

### Was human evolution special?

The magnitude, rate and pattern of change during hominin evolution, inferred from the fossil record, comparative neuroanatomy and embryology, provide the essential foundation for approaching the genetics of human evolution. From the studies discussed above, five key points emerge that have a bearing on attempts to reconstruct the genetic events that underlie the origin and modification of human traits.

First, trait evolution was nonlinear. The  $\sim 1,000\text{-cm}^3$  increase in brain size over 5–7 Myr did not occur at the same relative rate in hominin phylogeny: it was static at times, faster in some intervals, and reversed slightly more recently. Second, most trait evolution can be characterized as simple quantitative changes (that is, traits are continuous). Third, evolutionary rates were not at all exceptional with respect to mammalian evolution. For example, fossil horse lineages in the late Pliocene–Pleistocene show similar rates of body-size and other character changes as do those of hominids<sup>31</sup>. Fourth, much evolutionary change preceded the origin of the *Homo* genus and of *H. sapiens*: the history of our species represents just the last 3% of the time span of hominin evolution (Fig. 1). And fifth, many characters are present not only in humans, but also in apes. This suggests that modification of existing structures and developmental pathways, rather than the invention of new features, underlies much of human evolution.

These observations indicate that morphological evolution in hominins was not special, but the product of genetic and developmental changes typical of other mammals and animals.

## Genetics of human evolution

### Genetic architecture of trait evolution

Given the dimensions of hominin evolution, inferred from the fossil record and comparative anatomy, what can we expect in terms of the genetic complexity underlying trait evolution? For example, there is a long-standing tendency for events that are perceived to be relatively 'rapid' in the fossil record to be ascribed to perhaps one or a few radical mutations<sup>32</sup>, including recent human evolution<sup>33,34</sup>. Could the relative increase in brain size over 5 Myr, or its expanded cognitive function, be due to just one or a few genetic changes? The best (and, at present, the only available) guides for this question are detailed genetic studies in model organisms, which have achieved success in dissecting the genetics of complex trait formation, variation and evolution. Six essential general concepts have been established in model systems that pertain to the potential genetic architecture of human trait evolution:

- (1) Variation in continuous, quantitative traits is usually polygenic. Studies of variation in model species<sup>35–38</sup> reveal that many genes of small effect, and sometimes one or a few genes of large effect, control trait parameters. In humans, a study of variation in 20 anthropometric variables in two different ethnic human populations suggested that more than 50% of variation was polygenic<sup>39</sup>.
- (2) The rate of trait evolution tells us nothing about the number of genes involved. Studies of artificial selection<sup>35,38,40</sup> and of interspecific divergence<sup>41</sup> indicate that the intensity of selection and heritability are more important determinants of evolutionary rate than is the genetic complexity of the traits under selection. There is considerable standing variation in traits, including characters that might be thought of as highly constrained, such as limb morphology in tetrapods<sup>42</sup>. In general, the observed rates of evolution under natural selection are far slower than is potentially possible<sup>43,44</sup>. Genetic variation or genetic complexity is not the limiting factor<sup>32</sup>; indeed, considerable genetic variation underlies even phenotypically invariant traits<sup>45–48</sup>. Because the rate at which a trait emerges in the fossil record tells us nothing about genetic architecture, the temptation to invoke macromutational models for 'rapid change'<sup>33,49</sup> must be resisted in the absence of genetic evidence.
- (3) Morphological variation and divergence are associated with

genes that regulate development. Comparisons of the developmental basis of body-pattern evolution in animals suggest that morphological evolution is a product of changes in the spatiotemporal deployment of regulatory genes and the evolution of genetic regulatory networks<sup>50–52</sup>. Developmental changes in the human lineage are expected to be associated with genes that affect developmental parameters, such as those that encode transcription factors and members of signal-transduction pathways.

(4) Mutations responsible for trait variation are often in non-coding, regulatory regions. When it has been possible to localize variation in genes that underlie phenotypic variation or protein-level differences, insertions or substitutions in regulatory regions and non-coding regions are often responsible<sup>53–57</sup>.

(5) Multiple nucleotide replacements often differentiate alleles. Fine-scale analysis of quantitative trait loci has often revealed that functional differences between alleles are due to multiple nucleotide differences<sup>55,56</sup>. It also indicates that non-additive interactions between sites within a locus may be key to the differentiation of alleles, and that the contribution of any individual site may be modest (and difficult to detect).

(6) There is some concordance between genes responsible for intraspecific variation and interspecies divergence. Genetic analyses of interspecies divergence is only possible under certain circumstances, when laboratory breeding can overcome species barriers and traits can be mapped. In some cases, it has been found that some of the same loci are involved in both within-species variation and between-species divergence<sup>35,58</sup>. This raises some hope that studies of intraspecific variation in humans could lead to genes that have been important in human history.

Since human trait evolution has followed a similar, incremental course as traits studied in model systems, these six concepts suggest that we should expect a highly polygenic basis for complex traits such as brain size, craniofacial morphology and development, cortical speech and language areas, hand and digit morphology, dentition and post-cranial skeletal morphology. We should also anticipate that multiple changes in non-coding regulatory regions and in regulatory genes are of great importance. But how can we find them?

#### The arithmetic of human sequence evolution

All genetic approaches to human origins are fundamentally comparative, and seek to identify genetic changes that occurred specifically in the human lineage and contributed to the differentiation of humans from our last common ancestor with either apes or other species of *Homo*. Our primary comparative reference is the genome of the chimpanzee (*Pan troglodytes*), our closest living relative, with whom we share a common ancestor that lived 5–7 Myr ago. The arithmetic that sets the problem for human evolutionary genetics is as follows: first, the most extensive comparison of chimpanzee and human genomic sequences indicates an average substitution level of ~1.2% in single-copy DNA<sup>59</sup>; second, the human genome comprises  $\sim 3 \times 10^9$  base pairs; third, it is reasonable to assume that one-half of the total divergence between chimpanzees and humans occurred in the human lineage (~0.6%); and fourth, this amounts to  $\sim 18 \times 10^6$  base-pair changes. In addition, there are an unknown number of gene duplications and pseudogene, transposon and repetitive element changes in each lineage. A recent small-scale survey indicated that insertions and deletions (indels) might account for another 3.4% of differences between chimpanzee and human genomes, with the bulk of that figure contributed by larger indels<sup>60</sup>. A good deal of genomic change might be the noise of neutral substitutions and the gain and loss of repetitive elements over long time spans (more than 46% of human DNA is composed of interspersed repeats), but some small fraction of the changes in genomic sequence is responsible for the hereditary differences between species. The crux of the challenge is how to identify specific changes that are biologically meaningful from the many that are not.

In the case of human evolution, there are three basic genetic issues that we would like to grasp. First, how many genes were directly involved in the origin of human anatomy, physiology and behaviour (a few, dozens, hundreds or thousands)? Second, which specific genes contributed to the emergence of particular human traits? And third, what types of change in these genes contributed to evolution (for example, gene duplications, amino-acid replacements or regulatory sequence evolution)? In the few pioneering studies that are directly addressing the genetic basis of human–chimpanzee divergence, different but somewhat complementary strategies are being pursued that are beginning to reveal the scope of human genetic evolution and, in some cases, specific genes that might have been under selection in the course of recent human evolution.

#### Comparative genomics

The most readily detected differences between animal genomes are expansions or contractions of gene families. Although the full chimpanzee genome is not yet available, a partial comparative map indicates that there are regions of the human genome that might not be represented in chimpanzees or other apes<sup>61</sup>. Such regions could be due to duplications or insertions that occurred in the hominin lineage or to deletions in the chimpanzee lineage. One gene family, dubbed *morpheus*, underwent expansion as part of a segmental duplication on human chromosome 16 (ref. 62). This expansion is shared by other great apes, but it seems that there were human lineage-specific duplications as well.

On the basis of comparisons with other genomes, particularly the recently reported draft mouse sequence<sup>63</sup>, such lineage-specific duplications are expected. In the 75 Myr or more since the divergence of the common ancestor of mice and humans, several dozen clusters of mouse-specific genes arose that are generally represented by a single gene in the human genome<sup>63</sup>. The shorter divergence time between humans and apes suggests that the human-specific gene set will be smaller. It is interesting to note that a significant fraction of the mouse gene clusters encode proteins with roles in reproduction, immunity and olfaction. This indicates that sexual selection, pathogens and ecology can shape the main differences in coding content between mammals. It should also be noted that 80% of mouse genes have a 1:1 orthologue in the human genome, and that more than 99% have some homologue<sup>63</sup>. These figures and synteny data suggest that there is a gene repertoire that is qualitatively nearly identical among mammals. The presence or absence of particular gene duplicates might reflect adaptively driven change, but further evidence will be necessary to determine whether positive selection has acted on genes.

#### Thousands of adaptive changes in the human proteome?

The first place that adaptive genetic changes have been looked for is in the coding sequences for proteins. If the  $18 \times 10^6$  substitutions in the human lineage are evenly distributed throughout the genome, only a small fraction will be expected to fall within coding regions. Assuming that the average protein is ~400 amino acids in length, and that there are ~30,000 protein-coding genes, only  $\sim 3.5 \times 10^7$  base pairs (or a little more than 1.5% of the genome) consists of coding regions<sup>64,65</sup>. So, assuming neutrality and ignoring the selective removal of deleterious changes in protein sequences, ~1.5% of these  $18 \times 10^6$  substitutions (or 270,000 sites) may contribute to protein evolution. A fraction of these (roughly one-quarter) are synonymous substitutions, so the total number of amino-acid replacements in the human lineage could be of the order of ~200,000. This figure is in good agreement with observed average rates of amino-acid replacement in mammals<sup>66</sup>.

Various methods have been developed to detect whether amino-acid replacements could be the result of positive selection—that is, adaptive evolution<sup>67,68</sup>. To estimate the extent of positive selection in human protein evolution, Fay *et al.*<sup>69</sup> surveyed sequence-divergence data for 182 human and Old World monkey genes, and polymorph-

ism data for a similar number of human genes. Taking into consideration the frequency of common polymorphisms (ignoring rare alleles), a greater-than-expected degree of amino-acid replacements was observed, which is evidence of selection. When extrapolated to the entire proteome, 35% of amino-acid substitutions between human and Old World monkeys were estimated to have been driven by positive selection. Applied to human–chimpanzee divergence, this would extrapolate to ~70,000 adaptive substitutions in the human lineage. This figure is substantially larger than would be expected if most mutations were neutral or nearly neutral<sup>70</sup>. If it is even the correct order of magnitude, it forecasts a nightmare for the identification of key genes under selection, because this figure suggests that, on average, two or more adaptive substitutions have occurred in every human protein in the last 5 Myr.

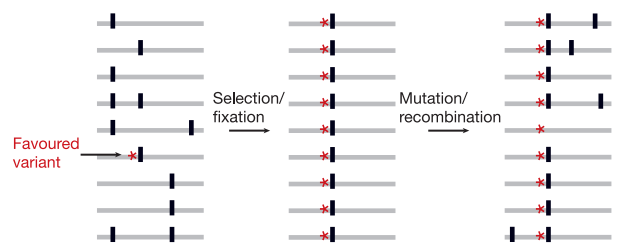
It is possible that the figure, based on the study of less than 0.5% of the human proteome, is an overestimate of the fraction or distribution of adaptive replacements. It is clear that some proteins

are under strong pressure to remain constant, whereas others, especially those involved in so-called ‘molecular arms races’, are under pressure to change. For example, major histocompatibility complex proteins, which interact with diverse and changing foreign substances, show clear signatures of selection<sup>71</sup>. Proteins involved in reproduction that play a part in sperm competition or gamete recognition also appear to evolve faster and under some degree of positive selection<sup>72</sup>. A host of human male reproductive proteins have greater-than-average ratios of amino-acid replacements<sup>73</sup>. Although accelerated protein evolution can also be the consequence of relaxed constraints, the correlation of higher levels of amino-acid replacements in proteins that have a role in reproduction and immunity seems to be biologically and selectively driven.

The population genetics- and protein-sequence-based statistical estimates of adaptive evolution require three caveats regarding how much they tell us about human evolution. First, there are generally no direct functional data that either test or demonstrate whether a human protein is indeed functionally diverged from an ape orthologue. Second, the proteins for which signatures of selection have been detected generally do not affect development. And third, the proteome is just part of the whole picture of genome evolution. Non-coding sequences, including transcriptional *cis*-regulatory elements, the untranslated regions of messenger RNAs, and RNA-splicing signals, contribute considerably to evolution by affecting the time, place and level of gene expression (see above). Ever since the pioneering comparative analysis of ape and human protein-sequence divergence nearly three decades ago<sup>74</sup>, it has generally been anticipated that changes in gene regulation are a more important force than coding-sequence evolution in the morphological and behavioural evolution of hominins.

**Box 2**  
**Selective sweeps**

If a change in a gene is favoured, then selection may drive the allele bearing that change to fixation (left and centre of the figure). In the process, neutral variation at linked sites ‘hitchhikes’ along with the selected site; this is known as a ‘selective sweep’. The physical limits of the sweep depend on the strength of the linkage between selected and adjacent sites. After a sweep, variation may again begin to build up, and initially there will be a relatively high frequency of rare polymorphisms (right of the figure). Tajima<sup>91</sup> proposed a statistic (*D*) that tests for selective neutrality. If the frequencies of polymorphisms are skewed, with an excess of rare types, this gives a negative value (neutral value = 0) and can be indicative of a recent selective sweep. Tajima’s *D* is sensitive to other factors apart from selection that can also yield a negative value. A recent expansion in population size from a relatively small population will produce similar patterns of genetic variation and *D* values. In human populations, population history (for example, drift and expansion) and population structure (ethnicity, migration and immigration) will affect *D* values at all loci, whereas selection will affect *D* values at selected and linked loci. The mean *D* values for 437 loci range from –0.69 to –1.25, depending on sampling methods<sup>92</sup>, indicating that population structure and history has had an effect. These negative values underscore a challenge in human evolutionary genetics to distinguish selective sweeps at loci from population-based effects. The *D* value obtained for the human *FOXP2* locus was –2.20, the second largest negative value among all human genes surveyed so far<sup>84</sup>. Other methods have been developed to detect positive selection: for example, by identifying areas of extended haplotype homozygosity<sup>81</sup>. It is important to emphasize that all of these methodologies detect signs of recent selection in the *Homo sapiens* lineage. The preceding 5–6 Myr of hominin genetic history, a period when we know from the fossil record that many human features arose, is not addressed by these methods.



**Evolution of human gene expression**

How large is the functional compartment of non-coding sequences—the other 98% of the genome? A recent estimate suggests that perhaps twice as much non-coding DNA is under selection than coding DNA<sup>63</sup>. So, we would also expect a large number of substitutions in the human lineage, of the order of several hundred thousand, with potential functional consequences in non-coding DNA. Even if one applies a much smaller, more conservative estimate of the fraction of adaptive substitutions in non-coding DNA, such as 2% (ref. 75), one still reaches a figure of more than 10,000 adaptive substitutions in human genes and their regulatory regions. The problem is that regulatory sequences are more difficult to analyse: we have no algorithms that can infer biological function from tracts of intergenic or intronic sequence, let alone to decipher how base-pair changes affect function. It is therefore perhaps understandable why non-coding regions have received little attention at the level of population genetics. However, a growing body of work in quantitative genetics and on the evolution of development has shown that regulatory sequences are central to changes in gene expression and morphology. New methodologies have been required to detect the evolution of gene expression and regulatory sequences.

A first step towards the identification of human-specific gene-expression patterns was recently taken by Enard *et al.*<sup>76</sup>, who used genome microarrays to analyse within- and between-species differences in primate gene expression. Analysis of RNA expression profiles from the left prefrontal lobe (Brodmann area 9, which is thought to be involved in cognitive functions) of adult male humans, chimpanzees and an orangutan, and from the neocortex of humans, chimpanzees and macaques, indicated an apparent acceleration of gene-expression differences in the human brain relative to other primates and to other tissues. Protein-expression analyses were also consistent with the idea that relative changes in protein-expression levels were accelerated in the evolution of the human brain, and could be detected for ~30% of the proteins surveyed. The concordance between RNA and protein-level data

indicates that regulatory changes have occurred in a substantial fraction of genes. Indeed, a recent survey of humans heterozygous at 13 loci revealed allelic variation in gene-expression levels at 6 loci<sup>77</sup>. Both intraspecific variations and interspecific divergence in gene expression are probably due to substitutions in non-coding regions that influence transcript or protein abundance through transcriptional or post-transcriptional mechanisms. These data further suggest that quantitative changes in gene expression should be expected as a general feature that accompanies species divergence, and that the raw material for evolutionary changes in gene expression appears to be widely available in non-coding DNA.

The microarray experiments raise many challenges for future progress. Specifically, how can changes that contribute to human anatomy, physiology or behaviour be sorted out from those that don't? Gene-expression data are correlative, not definitive in terms of identifying cause and effect. Many developmental and genetic mechanisms could contribute to the overall pattern observed. For example, a change in the composition of a tissue (for example, the relative proportions of cell types) will be accompanied by altered expression profiles, but many of these changes will be an indirect consequence of a developmental change, not the cause. Similarly, changes in levels or activities of regulatory proteins may affect batteries of downstream genes, but again are indirect and do not necessarily involve substitutions at the loci whose expression changes. Therefore, different approaches have to be taken to identify primary changes in regulatory pathways.

#### Candidate genes in human evolution

The ultimate goal of microarray analyses, quantitative trait genetics, population genetics or other comparative genetic methods is the identification of genes that are candidates for being causally associated with phenotypic divergence. Although genome-wide, large-scale surveys provide an overview, rigorous tests of causality demand a gene-by-gene approach. In choosing genes to be studied in greater detail, molecular geneticists will be opportunistic, focusing on those loci for which additional information from human or model-animal biology suggests an association with a trait of greater evolutionary interest, such as craniodental development<sup>78</sup>. Therefore, it is unlikely that all traits will be pursued with equal vigour or success.

To implicate a gene in human evolution, two types of data need to be assessed. First, functional evidence that a gene is involved in a developmental, behavioural or physiological trait is required to formulate hypotheses about the role of an individual gene. This may come from analysis of human mutations at a locus (see below). Second, the molecular evolution and population genetics of the locus need to be analysed for evidence of natural selection. Comparison of orthologues from chimpanzees and other primates and mammals, and analyses of intraspecific variation in humans, can reveal signs of positive selection at the sequence level or of a recent 'selective sweep' through a locus (Box 2). Evidence of positive selection has been found at several human loci<sup>73,79–81</sup>. Although these might be physiologically important (for example, in immunity or reproduction), most genes studied so far are not expected to contribute to the divergence of morphological or behavioural traits. More recently, attention has turned to candidate genes identified from human mutations that do affect such traits.

**The evolution of a gene affecting speech.** Human medical genetics has made substantial progress, and sophisticated mapping techniques for polymorphisms are accelerating the characterization of genes involved in complex traits, particularly those of medical interest. One of the most provocative reports of late was the identification of the gene *FOXP2* (forkhead box P2), mutations of which are associated with a speech and language disorder<sup>82</sup>. The gene encodes a transcription factor and is therefore expected to control the expression of other genes. The excitement surrounding *FOXP2* stems from the observation that affected individuals appear

to have not an overt impairment, but a lesion in the neural circuitry that affects language processes<sup>62,82,83</sup>.

Is this a novel human 'language' gene? No, the gene is found in other species. In fact, the human *FOXP2* protein differs from the gorilla and chimpanzee sequence at just two residues, and from the orangutan and mouse sequences at three and four residues, respectively<sup>76</sup>. This history is typical of human and other species' genes, in that most genes have orthologues in other mammals and animals. However, there is the possibility that the two replacements in the *FOXP2* protein that evolved in the human lineage are of functional significance to the origin of language.

To examine whether the *FOXP2* gene has been the target of selection during human evolution, a detailed analysis was undertaken of nucleotide variation over a 14-kilobase (kb) subregion of the large *FOXP2* locus, of amino-acid polymorphism in a segment of the protein, and of chimpanzee and orangutan sequences<sup>84</sup>. An unusual excess was found of rare alleles at the human *FOXP2* locus, and of high-frequency alleles. Reduced genetic variation in neutral linked regions is a predicted consequence of a selective sweep (Box 2), so these observations are consistent with natural selection acting on the *FOXP2* locus. Estimates of the time of fixation of the two amino-acid replacements place them within the last 200,000 yr of human evolution, an intriguing correlation with the estimated age of *H. sapiens*.

However, it should be noted that there are no biological data to support the hypothesis that these amino-acid replacements are functionally important. In the 14-kb region surveyed, more than 100 fixed differences exist; the entire *FOXP2* locus is large (267 kb), and more than 2,000 differences would be expected to exist between the *FOXP2* genes of humans and chimpanzees. No assessment has been made of potential non-coding regulatory sequences that might have contributed to a divergence in the role of *FOXP2* in hominids. Trait differences are often due to changes in regulatory networks that govern development, and need not be in coding regions (although that would be much more convenient, given just two changes in the human *FOXP2* protein). Because *FOXP2* is a transcription factor, changes in *FOXP2* expression could be of functional and evolutionary significance.

The typical genetic architecture that underlies complex traits makes it extremely unlikely that *FOXP2* was the only gene under selection in the evolution of our language capabilities. However, we have no means of assessing the relative contribution of *FOXP2* and other candidate genes. The encouraging lesson of *FOXP2* is that medical genetics has provided an interesting lead into a regulatory network that affects the development of speech ability. Further study of *FOXP2* should lead, at a minimum, to a better understanding of the neurodevelopmental biology of speech and language, and perhaps to more genes with interesting evolutionary histories.

#### The functions of selected genes

The three genome-scale approaches highlighted here—population genetics, comparative genomics and gene-expression profiling—have all succeeded in finding what each sought: thousands of potential adaptive coding substitutions, regulatory differences in gene expression, and gene duplications and rearrangements. Each has yielded many candidates through which to sift and, interestingly, because of the different search regimens used, there is virtually no overlap in the sets of genetic changes that have been surveyed. It is almost certain that, as in other lineages, all of these types of genetic mechanism have contributed to hominid evolution. The crucial challenge now is to obtain functional data for individual genes and to scrutinize the molecular evolution of candidates for signatures of selection.

To place any candidate gene into a functional context of human trait evolution, advances in primate and human developmental neurobiology will be essential. Non-primates are limited as models

of the development and function of primate and hominid neocortex, and thus as models of the function of proteins such as FOXP2 in the development and elaboration of neural networks. Direct empirical work on developing primates, which faces serious methodological constraints as well as bona fide ethical questions, will be necessary to advance beyond associations and correlations. Testing the functional role of what may be subtle changes in human orthologues of primate genes, a daunting task in the most technically developed model species, will be even more difficult.

There are two immediate avenues to increasing the power of human evolutionary genetics. First, we would increase the value of chimpanzee–human comparative genomics by sequencing the gorilla genome, which is the next earliest branching ape to humans and chimpanzees. This would help us to determine the polarity of genetic changes by distinguishing those changes in the human lineage from those in the chimpanzee lineage. Second,  $6 \times 10^9$  interbreeding humans is a very large resource for identifying rare mutations (for example, in *FOXP2*) with subtle behavioural or developmental effects, and for mapping genetic variation that underlies morphological variation, both of which could lead to genes that govern the formation of human traits and that might have played a part in hominid evolution.

### The fine print below the headlines

It is easy to foresee the media headlines that will announce the completion of the chimpanzee genome. One aim of this article has been to anticipate both the excitement that accomplishment warrants and the more sobering aspects of complex trait genetics and genome-scale evolution. Despite our enhanced understanding of functional genetic architecture, there remains a tendency to associate the development, function or evolution of a trait with single genes (genes ‘for’ speech, cancer and so on). The ghost of ‘hopeful monsters’ still haunts biology and is, unfortunately, a prevalent misconception in the scientific and general press. Perhaps wishful thinking is also an intrinsic part of human nature, but it seems unlikely that the traits that interest us most—bipedalism, skeletal morphology, craniofacial morphology, brain size and speech—were the products of selection of just a few major genes. Just as palaeoanthropology now recognizes a complex pattern of hominin phylogeny and the uncertainties in identifying long-sought common ancestors, and comparative neurobiology now searches for more subtle explanations of human capabilities, the lessons of model-system genetics and comparative genomics should prepare us for the finding that the genetics of hominid trait evolution are, in fact, subtle and complicated.

I underscore this point not just for its scientific relevance, but also because of the larger issues at stake—the meaning of the pursuit of the material basis of human evolution. Evolutionary biology has always faced public resistance. It has been difficult enough to gain acceptance of fundamental ideas using humble finches or fruitflies as examples. We can anticipate even more hostile challenges to human evolutionary genetics. Opponents will be sure to exploit any instances where claims or hypotheses are founded on weak or contradictory data. Witness how the recent scrutiny of data supporting the classic paradigm of industrial melanism has been hijacked by the anti-evolution agenda<sup>85</sup>. The sequencing of the chimpanzee genome will reveal no more directly about the origin of human traits than the sequence of the human genome tells us about how to construct a healthy baby. Headlines may claim more, but we would be well advised to describe this as just the beginning of a large, complex and profoundly important story. □

doi:10.1038/nature01495.

1. Wood, B. Hominid revelations from Chad. *Nature* **418**, 133–135 (2002).
2. Leakey, M. *et al.* New hominin genus from eastern Africa shows diverse middle Pliocene lineages. *Nature* **410**, 433–440 (2001).
3. Brunet, M. *et al.* A new hominid from the Upper Miocene of Chad, Central Africa. *Nature* **418**, 145–151 (2002).

4. Wolpoff, M., Senut, B., Pickford, M. & Hawks, J. Palaeoanthropology: *Sahelanthropus* or ‘*Sahelipithecus*? *Nature* **419**, 581–582 (2002).
5. Groves, C. P. in *Comparative Primate Biology* Vol. 1 (eds Swindler, D. R. & Erwin, J.) 187–218 (Alan R. Liss, New York, 1986).
6. Klein, J. & Takahata, N. *Where Do We Come From? The Molecular Evidence for Human Descent* (Springer, New York, 2002).
7. Wood, B. & Collard, M. The human genus. *Science* **284**, 65–71 (1999).
8. Relethford, J. H. *Genetics and the Search for Modern Human Origins* (Wiley-Liss, New York, 2001).
9. Ruff, C. B., Trinkhaus, E. & Holliday, T. W. Body mass and encephalization in Pleistocene *Homo*. *Nature* **387**, 173–176 (1997).
10. Conroy, G. C. *et al.* Endocranial capacity in an early hominid cranium from Sterkfontein, South Africa. *Science* **280**, 1730–1731 (1998).
11. Conroy, G. C., Weber, G. W., Seidler, H., Recheis, W. & Zur Nedden, E. Endocranial capacity of the Bodo cranium determined from three-dimensional computed tomography. *Am. J. Phys. Anthropol.* **113**, 111–118 (2000).
12. Brodmann, K. Neue Ergebnisse über die vergleichende histologische Lokalisation der Grosshirnrinde mit besonderer Berücksichtigung des Stirnhirns. *Anat. Anz.* **41**, 157–216 (1912).
13. Semendeferi, K., Lu, A., Schenker, N. & Damasio, H. Humans and great apes share a large frontal cortex. *Nature Neurosci.* **5**, 272–276 (2002).
14. Cantalupo, C. & Hopkins, W. D. Asymmetric Broca’s area in great apes. *Nature* **414**, 505 (2001).
15. Hopkins, W. D. & Leavens, D. A. The whole-hand point: The structure and function of pointing from a comparative perspective. *J. Comp. Psychol.* **112**, 95–99 (1998).
16. Holloway, R. L. Indonesian ‘Solo’ (Ngandong) endocranial reconstructions: Some preliminary observations and comparisons with Neandertal and *Homo erectus* groups. *Am. J. Phys. Anthropol.* **53**, 285–295 (1980).
17. Gannon, P. J., Holloway, R. L., Broadfield, D. C. & Braun, A. R. Asymmetry of chimpanzee planum temporale: Humanlike pattern of Wernicke’s brain language area homolog. *Science* **279**, 220–222 (1998).
18. Hopkins, W. D., Marino, L., Rilling, J. K. & MacGregor, L. Planum temporale asymmetries in great apes as revealed by magnetic resonance imaging (MRI). *NeuroReport* **9**, 2913–2918 (1998).
19. Hof, P. R., Nimchinsky, E. A., Perl, D. P. & Erwin, J. M. An unusual population of pyramidal neurons in the anterior cingulate cortex of hominids contains the calcium-binding protein calretinin. *Neurosci. Lett.* **307**, 139–142 (2001).
20. Nimchinsky, E. A. *et al.* A neuronal morphological type unique to humans and great apes. *Proc. Natl Acad. Sci. USA* **96**, 5268–5273 (1999).
21. Buxhoeveden, D., Switala, A., Litaker, M., Roy, E. & Casanova, M. Lateralization of minicolumns in human planum temporale is absent in nonhuman primate cortex. *Brain Behav. Evol.* **57**, 349–358 (2001).
22. Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K. & Van Hoesen, G. W. Prefrontal cortex in humans and apes: A comparative study of area 10. *Am. J. Phys. Anthropol.* **114**, 224–241 (2001).
23. Penin, X., Berge, C. & Baylac, M. 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 (2002).
24. Ponce de León, M. S. & Zollikofer, C. P. E. Neandertal cranial ontogeny and its implications for late hominid diversity. *Nature* **412**, 534–538 (2001).
25. Lieberman, D. E., McBratney, B. M. & Krovitz, G. The evolution and development of cranial form in *Homo sapiens*. *Proc. Natl Acad. Sci. USA* **99**, 1134–1139 (2002).
26. Williams, F. L., Godfrey, L. R. & Sutherland, M. R. in *Human Evolution through Developmental Change* (eds Minugh-Purvis, N. & McNamara, K. J.) 405–441 (Johns Hopkins Univ. Press, Baltimore, 2002).
27. Dean, C. *et al.* Growth processes in teeth distinguish modern humans from *Homo erectus* and earlier hominins. *Nature* **414**, 628–631 (2001).
28. Rice, S. H. in *Human Evolution through Developmental Change* (eds Minugh-Purvis, N. & McNamara, K. J.) 154–170 (Johns Hopkins Univ. Press, Baltimore, 2002).
29. Gould, S. J. *Ontogeny and Phylogeny* (Belknap, Cambridge, Massachusetts, 1977).
30. Moggi-Cecchi, J. Questions of growth. *Nature* **414**, 595–597 (2001).
31. MacFadden, B. J. Fossil horses from ‘Eohippus’ (*Hyracotherium*) to *Equus*: Scaling, Cope’s Law, and the evolution of body size. *Paleobiology* **12**, 355–369 (1986).
32. Charlesworth, B., Lande, R. & Slatkin, M. A neo-Darwinian commentary on macroevolution. *Evolution* **36**, 474–498 (1982).
33. Schwartz, J. H. Homeobox genes, fossils, and the origin of species. *Anat. Rec.* **257**, 15–31 (1999).
34. Klein, R. G. Archeology and the evolution of human behavior. *Evol. Anthropol.* **9**, 17–36 (2000).
35. Mackay, T. F. C. Quantitative trait loci in *Drosophila*. *Nature Rev. Genet.* **2**, 11–20 (2001).
36. Atchley, W. R., Plummer, A. A. & Riska, B. Genetic analysis of size-scaling patterns in the mouse mandible. *Genetics* **111**, 579–595 (1985).
37. Atchley, W. R. & Zhu, J. Developmental quantitative genetics, conditional epigenetic variability and growth in mice. *Genetics* **147**, 765–776 (1997).
38. Doebley, J. & Stec, A. Inheritance of the morphological differences between maize and teosinte: Comparison of results for two F2 populations. *Genetics* **134**, 559–570 (1993).
39. Livshits, G., Roset, A., Yakovenko, K., Trofimov, S. & Kobylansky, E. Genetics of human body size and shape: Body proportions and indices. *Ann. Hum. Biol.* **29**, 271–289 (2002).
40. Brakefield, P. *et al.* Development, plasticity and evolution of butterfly eyespot patterns. *Nature* **384**, 236–242 (1996).
41. True, J. R., Liu, J., Stam, L. F., Zeng, Z. B. & Laurie, C. C. Quantitative genetic analysis of divergence in male secondary sexual traits between *Drosophila simulans* and *Drosophila mauritiana*. *Evolution* **51**, 816–832 (1997).
42. Shubin, N., Wake, D. B. & Crawford, A. J. Morphological variation in the limbs of *Taricha granulosa* (Caudata: Salamandridae): Evolutionary and phylogenetic implications. *Evolution* **49**, 874–884 (1995).
43. Gingerich, P. D. Rates of evolution on the time scale of the evolutionary process. *Genetics* **112**, 113, 127–144 (2001).
44. Gingerich, P. D. Rate of evolution: Effects of time and temporal scaling. *Science* **222**, 159–161 (1983).
45. Lauter, N. & Doebley, J. Genetic variation for phenotypically invariant traits detected in teosinte: Implications for the evolution of novel forms. *Genetics* **160**, 333–342 (2002).
46. Polaczyk, P. J., Gasperini, R. & Gibson, G. Naturally occurring genetic variation affects *Drosophila* photoreceptor determination. *Dev. Genes. Evol.* **207**, 462–470 (1998).

47. Rutherford, S. L. & Lindquist, S. Hsp90 as a capacitor for morphological evolution. *Nature* **396**, 336–342 (1998).
48. Gibson, G., Wempe, M. & van Helden, S. Potential variance affecting homeotic *Ultrabithorax* and *Antennapedia* phenotypes in *Drosophila melanogaster*. *Genetics* **151**, 1081–1091 (1999).
49. Goldschmidt, R. *The Material Basis of Evolution* (Yale Univ. Press, New Haven, Connecticut, 1940).
50. Carroll, S. B., Grenier, J. K. & Weatherbee, S. D. *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design* (Blackwell Scientific, Malden, Massachusetts, 2001).
51. Davidson, E. H. *Genomic Regulatory Systems: Development and Evolution* (Academic, San Diego, 2001).
52. Wilkins, A. S. *The Evolution of Developmental Pathways* (Sinauer Associates, Sunderland, Massachusetts, 2002).
53. Laurie, C. C. & Stam, L. F. Molecular dissection of a major gene effect on a quantitative trait: The level of alcohol dehydrogenase expression in *Drosophila melanogaster*. *Genetics* **144**, 1559–1564 (1996).
54. Long, A. D., Lyman, R. F., Morgan, A. H., Langley, C. H. & Mackay, T. F. C. Both naturally occurring insertions of transposable elements and intermediate frequency polymorphisms at the *achaete-scute* complex are associated with variation in bristle number in *Drosophila melanogaster*. *Genetics* **154**, 1255–1269 (2000).
55. Wang, R.-L., Stec, A., Hey, J., Lukens, L. & Doebley, J. The limits of selection during maize domestication. *Nature* **398**, 236–239 (1999).
56. Long, A. D., Lyman, R. F., Langley, C. H. & Mackay, T. F. C. Two sites in the *Delta* gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics* **149**, 999–1017 (1998).
57. Lai, C., Lyman, R. F., Long, A. D., Langley, C. H. & Mackay, T. F. C. Naturally occurring variation in bristle number and DNA polymorphisms at the *scabrous* locus of *Drosophila melanogaster*. *Science* **266**, 1697–1702 (1994).
58. Nuzhdin, S. V. & Reiwitich, S. G. Are the same genes responsible for intra- and interspecific variability for sex comb tooth number in *Drosophila*. *Heredity* **84**, 97–102 (2000).
59. Chen, F.-C. & Li, W.-H. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *Am. J. Hum. Genet.* **68**, 444–456 (2001).
60. Britten, R. J. Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels. *Proc. Natl Acad. Sci. USA* **99**, 13633–13634 (2002).
61. Fujiyama, A. *et al.* Construction and analysis of a human–chimpanzee comparative clone map. *Science* **295**, 131–134 (2002).
62. Johnson, M. E. *et al.* Positive selection of a gene family during the emergence of humans and African apes. *Nature* **413**, 514–518 (2001).
63. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
64. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1323 (2001).
65. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
66. Li, W.-H. *Molecular Evolution* (Sinauer Associates, Sunderland, Massachusetts, 1997).
67. Kreitman, M. Methods to detect selection in populations with applications to the human. *Ann. Rev. Genomics Hum. Genet.* **1**, 539–559 (2000).
68. McDonald, J. H. & Kreitman, M. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**, 652–654 (1991).
69. Fay, J. C., Wyckoff, G. J. & Wu, C.-I. Positive and negative selection on the human genome. *Genetics* **158**, 1227–1234 (2001).
70. Ohta, T. Near-neutrality in evolution of genes and gene regulation. *Proc. Natl Acad. Sci. USA* **99**, 16134–16137 (2002).
71. Hughes, A. L. *Adaptive Evolution of Genes and Genomes* (Oxford Univ. Press, New York, 1999).
72. Swanson, W. J. & Vacquier, V. D. The rapid evolution of reproductive proteins. *Nature Rev. Genet.* **3**, 137–144 (2002).
73. Wyckoff, G. J., Wang, W. & Wu, C.-I. Rapid evolution of male reproductive genes in the descent of man. *Nature* **403**, 304–309 (2000).
74. King, M.-C. & Wilson, A. C. Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116 (1975).
75. Nachman, M. W. Single nucleotide polymorphisms and recombination rate in humans. *Trends Genet.* **17**, 481–485 (2001).
76. Enard, W. *et al.* Intra- and interspecific variation in primate gene expression patterns. *Science* **296**, 340–343 (2002).
77. Yan, H., Yuan, W., Velculescu, V. E., Vogelstein, B. & Kinzler, K. W. Allelic variation in human gene expression. *Science* **297**, 1143 (2002).
78. McCollum, M. A. & Sharpe, P. T. Developmental genetics and early hominid craniodental evolution. *BioEssays* **23**, 481–493 (2001).
79. Andolfatto, P. Adaptive hitchhiking effects on genome variability. *Curr. Opin. Genet. Dev.* **11**, 635–641 (2001).
80. Diller, K. C., Gilbert, W. A. & Kocher, T. D. Selective sweeps in the human genome: A starting point for identifying genetic differences between modern humans and chimpanzees. *Mol. Biol. Evol.* **19**, 2342–2345 (2002).
81. Sabeti, P. C. *et al.* Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**, 832–837 (2002).
82. Lai, C., S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413**, 519–523 (2001).
83. Pinker, S. Talk of genetics and vice versa. *Nature* **413**, 465–466 (2001).
84. Enard, W. *et al.* Molecular evolution of *FOXP2*, a gene involved in speech and language. *Nature* **418**, 869–872 (2002).
85. Coyne, J. A. Evolution under pressure. *Nature* **418**, 19–20 (2002).
86. Kimbel, W. H., Johanson, D. C. & Rak, Y. The first skull and other new discoveries of *Australopithecus afarensis* at Hadar, Ethiopia. *Nature* **368**, 449–451 (1994).
87. Haile-Selassie, Y. Late Miocene hominids from the Middle Awash, Ethiopia. *Nature* **412**, 178–181 (2001).
88. Asfaw, B. *et al.* Remains of *Homo erectus* from Bouri, Middle Awash, Ethiopia. *Nature* **416**, 317–320 (2002).
89. Hacia, J. G. Genome of the apes. *Trends Genet.* **17**, 637–645 (2001).
90. Richmond, B. G., Aiello, L. C. & Wood, B. A. Early hominid limb proportions. *J. Hum. Evol.* **43**, 529–548 (2002).
91. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–596 (1989).
92. Ptak, S. E. & Przeworski, M. Evidence for population growth in humans is confounded by fine-scale population structure. *Trends Genet.* **18**, 1–5 (2002).

**Acknowledgements** Thanks to B. Hopkins and C. Cantalupo for guidance on Fig. 2, and to L. Olds for illustrations; to B. Williams, A. Kopp, S. Paddock, A. Rokas, D. Bownds, J. Doebley, N. Shubin and J. Crow for comments on the manuscript; to P. N. and J. Carroll for inspiration, and to J. Carroll for preparation of the manuscript. S.B.C. is an Investigator of the Howard Hughes Medical Institute.

**Correspondence** and requests for materials should be addressed to the author (e-mail: sbcarroll@facstaff.wisc.edu).