**Positive selection on the human genome**

Eric J. Vallender\(^1,2\) and Bruce T. Lahn\(^1,\*\)

\(^1\)Department of Human Genetics, Howard Hughes Medical Institute and \(^2\)Committee on Genetics, University of Chicago, Chicago, IL 60637, USA

Received July 6, 2004; Revised and Accepted July 26, 2004

Positive selection has undoubtedly played a critical role in the evolution of *Homo sapiens*. Of the many phenotypic traits that define our species—notably the enormous brain, advanced cognitive abilities, complex vocal organs, bipedalism and opposable thumbs—most (if not all) are likely the product of strong positive selection. Many other aspects of human biology not necessarily related to the ‘branding’ of our species, such as host–pathogen interactions, reproduction, dietary adaptation and physical appearance, have also been the substrate of varying levels of positive selection. Comparative genetics/genomics studies in recent years have uncovered a growing list of genes that might have experienced positive selection during the evolution of human and/or primates. These genes offer valuable inroads into understanding the biological processes specific to humans, and the evolutionary forces that gave rise to them. Here, we present a comprehensive review of these genes, and their implications for human evolution.

Traditionally, studies of human biology have operated under the assumption, either explicitly or implicitly, that much of the molecular processes in humans (and the genes that underlie them) are conserved in other species. The pervasiveness of this sentiment is perhaps best reflected in the wide reliance on model organisms, from microbes all the way up to non-human primates, in the studies of human biology and disease. The assumption of evolutionary conservation, though powerful, has a key deficiency—it fails to address the many aspects of human biology and disease that differ significantly from other species. Although this deficiency has long been recognized, it is only recently that there has been an upsurge of interest in characterizing human-specific traits at the molecular and genetic levels.

In studying human-specific traits, it is necessary to investigate the selective forces that gave rise to them. Evolutionary biologists have typically invoked two types of selective forces that shape the evolution of species. One is purifying selection, which favors the conservation of existing phenotypes. The other is positive selection (also known as Darwinian selection), which promotes the emergence of new phenotypes. Positive selection can leave a set of telltale signatures in the genes under its influence, such as the rapid divergence of functional sites between species and the depression of polymorphism within species (1–3). On the basis of these signatures, investigators are beginning to identify likely target genes of positive selection in the human genome. The identification of these genes represents the first, and necessary, step toward gaining molecular and genetic insights into the evolution of human-specific traits. In this article, we review genes shown to bear evidence of having been the substrate of positive selection in the evolution of humans and/or primates (Table 1). Our emphasis is to provide a comprehensive listing of these genes, and for that reason, we will include genes even if the evidence of positive selection is only suggestive. Given that the number of genes is fairly large, the amount of discussion devoted to each gene will be limited. We therefore encourage readers to refer to the primary literature if they wish to assess the strength of evidence for any individual gene.

For clarity, we will divide genes into several functional domains and consider each domain under a separate heading. We note, however, that these genes can also be coarsely sorted into two categories on the basis of their relevance to the evolution of human-specific traits. One is genes involved in biological functions typically associated with positive selection across a wide range of species, including host–pathogen interactions, reproduction, dietary adaptation and appearance. For these genes, it should come as no surprise that they are also the target of positive selection in humans, and their involvement in human-specific traits may be limited. The other category is genes belonging to biological domains that bear defining differences between humans and other species, and for which positive selection appears to have operated more intensely in the lineage leading to
### Table 1. Genes showing evidence of positive selection in humans and/or primates

<table>
<thead>
<tr>
<th>Category</th>
<th>Gene</th>
<th>Lineage</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host-pathogen interactions</td>
<td>ABO blood group</td>
<td>Human/Primates</td>
<td>28, 29</td>
</tr>
<tr>
<td></td>
<td>APOBEC3G</td>
<td>Primates</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Beta globin (HBB)</td>
<td>Recent human</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>CD45</td>
<td>Primates</td>
<td>14, 15</td>
</tr>
<tr>
<td></td>
<td>CD59</td>
<td>Catarrhines</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Chemokine receptor 5 (CCR5)</td>
<td>Humans</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Cytidine monophospho-N-acetylneuraminic acid hydroylase (CMAH)</td>
<td>Human</td>
<td>24, 131</td>
</tr>
<tr>
<td></td>
<td>Duffy blood group</td>
<td>Recent human</td>
<td>34, 35</td>
</tr>
<tr>
<td></td>
<td>Eosinophil cationic protein (ECP)</td>
<td>Catarrhines</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Glucose-6-phosphate dehydrogenase (G6PD)</td>
<td>Recent human</td>
<td>43– 46</td>
</tr>
<tr>
<td></td>
<td>Glycophorin A</td>
<td>Recent human/Primates</td>
<td>32, 33</td>
</tr>
<tr>
<td></td>
<td>Immunoglobulin A</td>
<td>Primates</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Immunoglobulin heavy chain</td>
<td>Mammals</td>
<td>20, 21</td>
</tr>
<tr>
<td></td>
<td>Interleukin 2 (IL2)</td>
<td>Mammals</td>
<td>17, 18</td>
</tr>
<tr>
<td></td>
<td>Interleukin 4 (IL4)</td>
<td>Human</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Killer cell inhibitory receptors (KIR)</td>
<td>Human</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Major histocompatibility complex (MHC)</td>
<td>Mammals</td>
<td>5– 8</td>
</tr>
<tr>
<td></td>
<td>Pyrin</td>
<td>Recent human</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Rh blood group</td>
<td>Hominids</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Sialic acid-binding immunoglobulin-like lectin 9 (SIGLEC9)</td>
<td>Human</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Sialic acid-binding immunoglobulin-like lectin-like protein 1 (SIGLECL1)</td>
<td>Human</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Tumor necrosis factor ligand superfamily, member 5 (TNFSF5)</td>
<td>Recent human</td>
<td>46</td>
</tr>
<tr>
<td>Reproduction</td>
<td>Chorionic gonadotropin</td>
<td>Primates</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Chromodomain protein Y (CDY)</td>
<td>Primates</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Deleted in azoosperma (DAZ)</td>
<td>Primates</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Fertil</td>
<td>Primates</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>Oviductal glycoprotein 1 (OVGP1)</td>
<td>Mammals</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Protamine 1 (PRM1)</td>
<td>Primates/Mammals</td>
<td>32, 49</td>
</tr>
<tr>
<td></td>
<td>Protamine 2 (PRM2)</td>
<td>Primates/Mammals</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Semenogelin 1 and 2 (SEMG1, SEMG2)</td>
<td>Primates</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Sex-determining region Y (SRY)</td>
<td>Primates/Mammals</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Sperm associated cation channel 1 (CATSPER1)</td>
<td>Primates</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>Sperm protein associated with the nucleus, X-chromosome (SPANX)</td>
<td>Hominids</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Transition protein 2 (TNP2)</td>
<td>Primates/Mammals</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Zona pellucida glycoprotein 2 (ZP2)</td>
<td>Mammals</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Zona pellucida glycoprotein 3 (ZP3)</td>
<td>Mammals</td>
<td>59</td>
</tr>
<tr>
<td>Dietary adaptation</td>
<td>Alanine-glyoxylate aminotransferase</td>
<td>Anthropoids</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Aldehyde dehydrogenase 2 (ALDH2)</td>
<td>Recent human</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Lactase (LCT)</td>
<td>Recent human</td>
<td>72, 73</td>
</tr>
<tr>
<td></td>
<td>Lysozyme (LYZ)</td>
<td>Human/Primates</td>
<td>64, 65</td>
</tr>
<tr>
<td>Appearance</td>
<td>Melanocortin 1 receptor (MC1R)</td>
<td>Recent human</td>
<td>79– 81</td>
</tr>
<tr>
<td>Sensory systems</td>
<td>MAS-related genes (MRG)</td>
<td>Primates</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Olfactory receptors (OR)</td>
<td>Human/Primates</td>
<td>92, 93</td>
</tr>
<tr>
<td></td>
<td>Red and green opsins</td>
<td>Human/Primates</td>
<td>85, 86</td>
</tr>
<tr>
<td></td>
<td>Taste receptor type 1 member 3 (TAS1R3)</td>
<td>Primates</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Taste receptor type 2 (TAS2R)</td>
<td>Primates</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Taste receptor type 2 member 38 (TAS2R38)</td>
<td>Recent human</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Type 1 vomeronasal receptor-like (V1RL)</td>
<td>Primates</td>
<td>94</td>
</tr>
<tr>
<td>Behavior</td>
<td>Dopamine receptor D4 (DRD4)</td>
<td>Primates</td>
<td>97, 98</td>
</tr>
<tr>
<td></td>
<td>Forkhead box P2 (FOXp2)</td>
<td>Human</td>
<td>101, 102</td>
</tr>
<tr>
<td></td>
<td>Monoamine oxidase A (MAOA)</td>
<td>Human</td>
<td>96</td>
</tr>
<tr>
<td>Brain anatomy</td>
<td>Abnormal spindle-like microcephaly associated (ASPM)</td>
<td>Human</td>
<td>107– 109</td>
</tr>
<tr>
<td></td>
<td>Microcephalin (MCPH1)</td>
<td>Human</td>
<td>110, 111</td>
</tr>
<tr>
<td></td>
<td>Myosin heavy chain 16 (MYH16)</td>
<td>Hominine</td>
<td>112</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>Angiogenin (ANG)</td>
<td>Primates</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td>Breast cancer 1 (BRCA1)</td>
<td>Human</td>
<td>113, 114</td>
</tr>
<tr>
<td></td>
<td>Cytochrome c oxidase subunit 4 (COX4)</td>
<td>Human/Primates</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>Cytochrome c oxidase subunit 7a (COX7A)</td>
<td>Primates</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>Cytochrome c oxidase subunit 7c (COX7C)</td>
<td>Primates</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Cytochrome c oxidase subunit 8 (COX8)</td>
<td>Human/Primates</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Forkhead box D4 (FOXD4)</td>
<td>Hominids</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>Morpheus</td>
<td>Hominids</td>
<td>121</td>
</tr>
</tbody>
</table>
GENES LINKED TO POSITIVE SELECTION

Host–pathogen interactions

Van Valen’s Red Queen hypothesis has long served as a theoretical framework for understanding the evolutionary dynamics of host–pathogen interactions (4). It states that the coevolution of two intensely competing species resembles an arms race. Both species evolve continuously to gain advantage over their rival, and yet the long-term outcome is evolutionary stasis, whereby the two species continue to coexist (and do battle) without a clear winner. In the case of host–pathogen coevolution, the pathogen is under strong selection to constantly devise new strategies for gaining access to the host while evading its defense system, whereas the host is under selection to deny access to the pathogen and to neutralize it.

Numerous studies in diverse taxa have indeed shown that genes involved in host–pathogen interactions are a frequent substrate of intense positive selection. The first study of positive selection in humans was on the genes encoding the major histocompatibility complex (MHC), which is a group of genes linked to positive selection in humans (13). These include genes associated with behavior, brain anatomy, and to some extent, the sensory systems. These genes are particularly relevant to understanding the evolution of biological traits that distinguish the human species, such as increased brain size and advanced cognitive abilities.

Among human parasites, those causing malaria, *Plasmodium falciparum* and *P. vivax*, seem to have accorded very recent and strong selective pressure on the cellular phenotype of human erythrocytes. The evolution of a number of erythrocyte genes have now been linked to malaria resistance. The classic example is the mutant allele of β-globin, which causes sickle-cell anemia when homozygous, but imparts some protection against malaria when heterozygous. In malaria-infested regions of the world, particularly sub-Saharan Africa, the mutant allele of β-globin has reached
appreciable frequencies under strong selection, despite its fatal
effect in homozygotes (40). The Duffy antigen is present on
the erythrocyte surface and serves as a receptor for P. vivax.
Null alleles of the Duffy locus, which impart resistance to
P. vivax, appear to have been driven to near fixation by posi-
tive selection in sub-Saharan Africa (41,42). Multiple G6PD-
deficiency alleles, which confer some protection against
malaria, seem to have emerged roughly 10 000 years ago,
and rose to appreciable frequencies under positive selection
in regions with high malaria prevalence (43–46). A promoter
variant of the CD40 ligand gene, TNFSF5, which also confers
some malaria resistance, was similarly shown to have risen to
high frequency quickly under positive selection (46).

It is perhaps not surprising, given the potential for strong
selective pressure, that genetic programs controlling host–
pathogen interactions in humans and other species are littered
with signatures of positive selection (47).

Reproduction

Like genes implicated in host–pathogen interactions, genes
involved in reproduction have also been shown to be under
strong positive selection across taxa (48). Because of
the large reproductive skew in males (i.e. high variance in
reproductive success from male to male), selection on male
reproductive processes is particularly intense. Meta-analysis
of the human sperm proteome has already provided evidence
that these genes may have evolved under positive selection
(32,48), and more in-depth analyses of specific sperm-
related genes are confirming this pattern. Studies of prota-
mines (32,49,50), transition protein 2 (32), semenogelins
(51), sperm ion channels (52) and fertilin (53) have all
shown evidence of positive selection in humans. Many other
genes of lesser known biochemical functions, but which are
expressed predominantly in spermatogenic cells (and are
linked to male infertility in some cases), have also been
suggested to be subject to positive selection (54–58). The
prevalence of positive selection on male reproductive genes
is therefore not unlike genes involved in host–pathogen
interactions.

A number of female reproductive genes have also been
shown to be under positive selection. Two of these encode
the zona pellucida glycoproteins ZP2 and ZP3, which make
up the egg’s protective coat and are intimately involved in
sperm–egg recognition during fertilization (59). Another
gene encodes the oviductal glycoprotein (OGP), which is
also implicated in fertilization (59). It is believed that positive
selection on these fertilization-related proteins may stem from
the ‘arms race’ between sperm and egg, whereby the rapid
evolution of sperm-associated proteins involved in sperm–
egg recognition drives the rapid evolution of corresponding
egg proteins (59).

Another female reproductive protein under positive selec-
tion is chorionic gonadotropin (CG), which is a key
hormone for establishing pregnancy in humans and other
simian primates (60). Why this gene should be under positive
selection is unclear. It is interesting to note, however, that one
of the key reproductive changes in the evolution of humans is
increased gestation time. Although there is no evidence to
suggest that this has any bearing on the evolution of CG or
other female reproductive proteins, the possibility of a link
is tantalizing.

Dietary adaptation

The adaptation to new diet is a major driving force in the evol-
ution of a species. Dietary changes during the evolution of
various primate species, particularly humans, have been well
documented over the years (61,62), and a number of genes
have been implicated in diet-driven positive selection. A
well-studied example can be found in a pancreatic ribonu-
clease in old-world monkey species (63). In this species, the
protein has evolved the ability to better digest bacterial
DNA as a result of the monkey’s changing diet. Another
example is lysozyme, which aids in the degradation of gut
bacteria. This protein has been shown to be under positive
selection in many primate groups including humans (64,65).

The nature of this selection is understood best in the langur
monkey, a species that has evolved a foregut fermentation
method of digestion similar to that of ruminants (64,66). In
humans, the reason for positive selection is less clear, but it
has been speculated that a shift towards a more meat-based
diet, which likely required changes in the ability to digest bac-
teria, might have played a role (67,68). This suggestion is
given credence by studies of alanine-glyoxylate aminotrans-
ferase (AGT), a gene with seemingly divergent roles in herbi-
vores and carnivores, and which also shows evidence of
positive selection across simian primates (69).

Two other examples of positive selection on diet-related
genes have captured the interest of researchers and the
public alike. A common mutation allele of the ALDH2 gene
causes the inability of the body to handle large quantities of
alcohol, and is thought to be responsible for low alcohol toler-
ance in certain East Asian populations. The ALDH2 locus
shows some signature of recent selection in humans (70). It
has been speculated, though far from proven, that selection
might favor the ALDH2-deficient allele in East Asian popu-
lations because alcohol consumption exacerbates the pathol-
ogy of hepatitis B infections (which are prevalent in East
Asia) (71).

In most species the ability to digest dairy ends with weaning
in childhood. However, in some human populations, this ability
persists in many adults. A derived allele of the lactase gene has
been shown to give adults the ability to continue to process
dairy. This allele has been shown to be the substrate of strong
positive selection (72,73). Similar to the malaria-resistance
genes, the selected allele of the lactase gene appears to have
arisen very recently, about 5000–10 000 years ago, coincident
with the emergence of dairy farming in Eurasia (73). This allele
rose quickly in frequency, presumably because individuals with
the allele had a better ability to consume dairy, a clear survival
advantage at that time.

Appearance

Among genes involved in the physical appearance of humans,
the one with the longest story is MC1R, which is involved in
skin and hair coloration. The MC1R gene encodes the melano-
cortin receptor, which regulates the production of eumelanin,
a cause of black pigmentation. Mutations in the gene have
been reported for many taxa in which coat color variation is observed, including rodents (74,75), cats (76), dogs (77), horses (78) and several primate species (79). As coloration has long been recognized as adaptive, it seems likely that changes in this gene are driven by positive selection. Indeed, in many of the non-human species in which the gene has been studied, coloration adaptations have been invoked as the driving selective force. In humans, however, the evolution of MC1R has not been credibly linked to adaptation.

Multiple studies have shown extensive polymorphism at the MC1R locus in human populations, most notably three separate mutations resulting in red hair in the Irish, Dutch and Swedes (80,81). However, there remains some question as to whether this polymorphism represents a relaxation of constraint on the gene outside of African populations or whether there is diversifying selection acting upon the gene. Arguing for the former is the observation that these polymorphisms are located exclusively outside of Africa, whereas the argument for the latter revolves around sexual selection (82). One possibility is that there exists an advantage to novelty in attracting mates and that with humans no longer relying on coloration for camouflage or protection against the sun, diversity was selected for. Although the answer is uncertain, it seems feasible that it lies as some combination of both relaxation of constraint and positive selection.

Sensory systems

Human sensory systems have undergone major changes from their primate ancestors, one of the most notable examples being the greater reliance on visual perception. One aspect of increased visual acuity is the emergence of trichromatic vision in great apes including humans, which is accomplished by three separate opsin genes that each preferentially absorb the red, green or blue wavelength. One of these, absorbing blue light, is autosomal and found ubiquitously among primates. The other two opsins are paralogs located on the X-chromosome, with both copies present in the catarrhine clade (old world monkeys, apes and humans) but not in the more basal platyrhines (new world monkeys). In new world monkeys, the opsin gene on the X-chromosome is a single copy. This gene is polymorphic in some new world monkey species, creating a ‘red’ and a ‘green’ allele. Males and homozygous females are dichromatic, similar to most other mammals, whereas heterozygous females are trichromatic like their simian relatives (83,84). The two X-chromosomal paralogs in catarrhines have been suggested to be the subject of adaptive evolution (85,86).

Taste is another sensory modality for which there are ample examples of adaptive evolution. This is perhaps not surprising given the importance of dietary adaptation as we have discussed. The taste receptor for sweetness exhibits interesting functional differences among primates, with the receptor in simians (including humans) capable of recognizing more sweet compounds than is the case for prosimians (87). That this difference is driven by positive selection is not an unreasonable proposition given the dietary differences among primates, but this is yet to be formally proven. The multiple bitter taste receptors underwent amplifications in diverse mammalian lineages including humans, and the rapid sequence divergence of the ligand-binding domains between amplified paralogs is indicative of positive selection (88). It has been argued that both the diversity and the specificity of the bitter receptors in a species are driven by the need to properly discern poisons commonly encountered by the species (88). The ability to taste the bitter compound phenylthiocarbamide (PTC) is a polymorphic trait across all the major human populations. It has been suggested that this polymorphism is maintained in humans by the balancing selection on two major alleles, the taster and the non-taster, of the PTC taste receptor gene (89). As yet, it is unclear why balancing selection should favor the coexistence of both taster and non-taster phenotypes in humans.

olfactory and pheromone receptor genes have undergone significant evolutionary changes in primates, particularly humans. However, the role of positive selection is much less clear here. In humans, olfaction remains important, but perhaps much less than that in many other mammals including non-human primates. Several studies corroborate this by showing that the human olfactory subgenome has undergone a rapid process of pseudogenization coincident with relaxed functional constraint (90,91). Some studies, however, suggest that a subset of human olfactory genes may have been the subject of positive selection (92,93). For the latter finding, it remains to be seen whether it represents real positive selection on some odorant receptors, or whether it simply reflects the difficulty in distinguishing between positive selection and relaxation of constraint. Indeed, the problem of differentiating positive selection from relaxed constraint, both of which are manifested as the rapid evolution of functional sequences, plagues all studies of positive selection, and is merely exaggerated in the context of olfactory genes. This conundrum is also found in pheromone receptors, where examples of both positive selection and relaxation of functional constraint have been reported (94). It is therefore too soon to state for certain what the overall message is for these genes.

Recent evidence has also shown the effect of positive selection on nociception (i.e. the perception of pain). The MRG nociceptive receptor family has undergone rampant amplification in both the human and murine lineages, and comparisons between paralogs in each lineage showed evidence of strong positive selection in the ligand-binding domains (95). The biological impact of this positive selection remains unclear. However, it seems plausible that selection has favored the functional evolution of this receptor family because of the need to constantly tune the nociceptive properties of a species in response to the changing environment.

Behavior

Many behavioral traits in humans have been shown to have a strong genetic basis, and some of the underlying genes are now beginning to be discovered. Some of these genes have shown evidence of positive selection. One of them was MAO-A, which has been associated with aggressive and impulsive behavior (96). Similarly, a gene associated with novelty-seeking and attention deficit hyperactivity disorder, DRD4, has also been shown to have undergone recent positive selection (97,98). It is tempting to speculate on why these
genes and their associated traits are under selection by invoking a plausible ‘just-so’ story, but the true reasons remain unclear and any speculation is likely premature.

One behavioral trait uniquely associated with humans is language. \textit{FOXP2} is a gene implicated in language abilities in humans, with mutations in this gene leading to a language disorder (99). Corroborating the role of \textit{FOXP2} in language studies showing that the gene may also play a role in song-learning in birds (100). Evolutionary studies showed that \textit{FOXP2} has evolved faster in the human lineage than in several other mammalian species. This, coupled with human polymorphism data, suggests positive selection on this gene during recent human evolution (101,102). This is a tantalizing finding, though additional research is needed to assess whether the gene indeed plays a role in the origin of human language.

**Brain development**

The evolution of human anatomy is marked most prominently by the dramatic expansion of the brain. This is especially true in the last 2–3 million years of hominid evolution, during which the brain more than tripled in size (103,104). Two genes, \textit{ASPM} and \textit{Microcephalin}, have been implicated in the evolution of brain size. Both genes, when mutated, cause primary microcephaly, a disease characterized by a severe reduction in brain size without any other gross abnormalities (105,106). Given the important, and specific, role of these genes in regulating brain size during development, it is enticing to hypothesize that they are also involved in changes of brain size during evolution.

Going on this hunch, several groups investigated the evolution of \textit{ASPM} and \textit{Microcephalin} in primates and other mammals. They found that, indeed, both genes showed robust evidence of positive selection along the primate lineage leading to humans. In particular, this lineage had much higher rates of protein sequence evolution as compared with lineages leading to non-human primates (107–111). For \textit{ASPM}, the intensity of selection is strongest in later portions of the lineage leading to humans, i.e. from ape ancestors to humans. For \textit{Microcephalin}, selection is most pronounced in earlier portions of the lineage, i.e. from simian ancestors to ape ancestors. This suggests that these two genes might have had differential contributions to brain evolution during different periods of the primate lineage leading to \textit{Homo sapiens}.

Another story revolves around a gene expressed in facial musculature. Myosin heavy chain 16 (\textit{MYH16}) is a muscle structural protein found primarily in masticatory muscles, and has been shown to have undergone inactivation in the human lineage, since its divergence from chimpanzees (112). This inactivation was estimated at around 2.5 million years ago, or roughly the time when \textit{Homo erectus} arose, though there are contentions over the accuracy of this estimate. \textit{Homo erectus} has significantly reduced masticatory musculature relative to its predecessors (and a coincident increase in cranial size). The loss of \textit{MYH16} has therefore been speculated to have contributed to smaller jaw muscles, which would perhaps free the constraint on the growth of the brain.

**Miscellaneous**

A number of genes that cannot be readily assigned to the above functional domains have also been linked to positive selection. Two of these, \textit{BRCA1} (113,114) and \textit{angiogenin} (115), are involved in cancer. It is difficult to speculate what adaptive role these genes might have played during evolution. It is interesting to note, however, that \textit{BRCA1} knockout mice show profound defects in nervous system development such as failure of neural tube closure and severely disorganized brain growth (116). This result raises the possibility that positive selection on \textit{BRCA1} was actually directed towards its function in brain development rather than its activity as a tumor suppressor.

Evidence of positive selection has also been found in genes involved in energy production. Four of the cytochrome \textit{c} oxidase subunits exhibit signatures of positive selection in primates (13,117,118). Patterns of mitochondrial DNA (mtDNA) polymorphism in humans are consistent with a role of certain mutations in mtDNA having contributed to human adaptation to northern climates (119,120).

For some genes linked to positive selection, the functions are wholly unknown. One of these is \textit{morphheus}, a gene family found only in humans and African apes, and for which the intensity of positive selection is unrivalled by any other mammalian gene (121). It is worth noting that the \textit{morphheus} genes reside in segmental repeat regions of the genome, where frequent rearrangements (such as duplications, deletions and transpositions) produce extraordinary evolutionary lability (122). Another example is the \textit{FOXD4} gene family which, like \textit{morphheus}, also reside in segmental repeat regions (123). It is possible that repeat regions of the genome, such as subtelomeric and pericentromeric regions, may have a propensity for harboring rapidly evolving gene families.

Lately, large-scale studies have begun to uncover signatures of positive selection on a genomic scale. Two of these studies surveyed the polymorphism patterns across the human genome for signals of positive selective sweeps (124,125). Another pair of studies used human–chimpanzee comparison as a means to detect positive selection (93,126). Finally, genome-wide expression surveys suggested that the transcriptome of the brain might have changed more rapidly during human evolution than during the evolution of other lineages (127). The efficacy of this coarse-grained whole-genome studies for identifying individual positively selected genes remains to be seen. However, the approach can reveal genome-wide patterns of selection not visible to single-gene studies.

**CONCLUSION**

As a species, we are inherently curious about our evolutionary origins. One powerful approach for studying human origins at the molecular and genetic levels is to identify genes that have been the target of positive selection. With the rapid expansion of genomic data and the availability of increasingly sophisticated analytical tools, positively selected genes are being identified at an ever faster pace. Judging from the currently available data, it appears that these genes largely belong to a limited number of functional domains. For some of these,
such as host–pathogen interactions and reproduction, the prevalence of positively selected genes is not surprising. For others, such as the regulation of brain development and behavior, the identification of positively selected genes may offer valuable insights into the evolution of those defining human-specific traits such as enlarged brain and highly advanced cognitive abilities.

A question of particular relevance to the understanding of human origins is whether the selective regimes driving human evolution are of exceptional quality or are more typical. One reason to suspect that selection on humans is exceptional is the remarkable rapidity with which some key traits were acquired. Allometrically scaled brain size, for example, grew by an order of magnitude since the lineage leading to humans diverge from old world monkeys some 20–25 million years ago, with a tripling in size occurring in just the last 2–3 million years of hominid evolution (104). Such a dramatic change within such a short period of time is extraordinary for any tissue system, but is particularly so for the brain, an exceedingly complex organ for which the growth in size is necessarily accompanied by the increase in organizational complexity (128). In the theoretical framework of ‘punctuated equilibrium’ (129), the enlargement of the human brain represents no less a stunning punctuation to an evolutionary equilibrium. The identification of positively selected genes, especially those relating to brain development and cognitive abilities, may offer molecular evidence for the exceptional strength of the selective pressure driving the evolution of our species.

The identification of positively selected genes may also have important implications for human medicine. Many diseases affect humans but not other animals (130). These include infectious diseases such as AIDS, as well as non-infectious conditions such as Alzheimer’s disease. Likewise, many therapeutic strategies developed in animal models fail to work in humans. These dissimilarities between humans and other species are ultimately due to molecular programs that have diverged between lineages. Genes that have experienced positive selection may underlie some of these dissimilarities, and variants of these genes in humans may even participate directly in the pathogenesis of diseases.

Although a fair number of genes have already been identified as targets of positive selection during the evolution of humans and/or primates, these are likely to be the tip of the iceberg. As more genes are added and as alterations in gene sequences are mapped to functional changes, the study of positively selected genes may become a mainstream approach to the dissection of human biology and disease.

REFERENCES


