Humans vs apes

1. Structural DNA diversity (chromosomes)
2. Rough diversity estimates (DNA-DNA Hybridization)
3. DNA sequencing (autosomes, mtDNA and Y)
4. **Proteins**
5. Comparison of genetic diversity in apes and humans

Proteins

ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Gene 346 (2005) 215–219

GENE
SECTION
EVOLUTIONARY GENOMICS

www.elsevier.com/locate/gene

Eighty percent of proteins are different between humans and chimpanzees

Galina Glazko, Vamsi Veeramachaneni, Masatoshi Nei, Wojciech Makałowski*

Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, PA 16802, USA
Department of Biology, Pennsylvania State University, University Park, PA 16802, USA

Received 12 August 2004; received in revised form 1 October 2004; accepted 5 November 2004
Available online 29 January 2005
Received by T. Gojobori

Proteins

A large portion (about 98%) of the human genome is known to be non-protein-coding DNA, and **the estimate of 1–2% nucleotide difference is largely based on** the comparison of **non-protein-coding** DNA, which has little effect on phenotypic characters.

Functional genes or proteins as the **units of comparison**, because these are the genetic units that **control phenotypic characters**.

We compiled **127 human and chimp orthologous** proteins (44,000 amino acid residues) from GenBank. **Only 25 (20%) of these proteins showed the identical amino acid sequence between humans and chimpanzees**. In other words, the **proportion of different proteins was 80%,** in contrast to the 1–2% difference at the nucleotide level.

Typical human and chimp homologs of proteins differ in only an average of two amino acids.

Table 1
Number of orthologous gene pairs used in this study

	Human	Chimpanzee	Gorilla	Orangutan
Chimpanzee	127			
Gorilla	60	56		
Orangutan	56	52	44	
Gibbon	31	26	22	24

Table 4
Percentage of proteins showing 100%, 99%, and 98% sequence identity between humans and chimps for different functional categories

Functional categories	Level of amino acid identity			
	100%	99%	98%	Less than 98%
Enzymes	22.7	30.6	26.1	20.0
Signal transduction	45.5	52.8	56.5	45.0
Transporters	22.7	8.3	4.3	0.0
Others	9.1	8.3	13.0	35.0

Signal transduction proteins (e.g. *growth factors, cytokines and neurotransmitters. steroid hormone) are* highly conserved, whereas **transporter proteins** are least conserved: *s*

Table 2
List of identical proteins in humans and chimpanzees

Gene name	Protein length (amino acids)
Beta-2-microglobulin	119
Opsin 1	348
Poly(A) binding protein	382
5-Hydroxytryptamine (serotonin) receptor 1B	390
5-Hydroxytryptamine (serotonin) receptor 1E	365
A-gamma globin	147
Alpha 2 globin	142
Beta 1,3-galactosyltransferase polypeptide 1	326
Beta defensin 1	68
Epsilon globin	147
G-gamma globin	147
Histamine receptor H2	359
Lysozyme precursor	148
Superoxide dismutase 1	154
Hemochromatosis protein isoform 1 precursor	348
Dopamine receptor D2	443
G protein-coupled receptor 15	360
Renin precursor	406
Mitogen-activated protein kinase 14,	360
Chemokine (C-X-C motif) receptor 4 (fusin)	352
Epididymal secretory protein	151
CD81 antigen	236
Triosephosphate isomerase 1	249
DEAD-box protein	428
Ubiquitin B	229

Beta-2-microglobulin has the identical nucleotide sequence as well.

Table 5
Genes showing $d_N > d_S$ in the comparison of human and chimpanzee genes

Gene name	Modified Nei–Gojobori		
	d_N	d_S	Z-boot
Protamine 1	0.088	0.000	2.642*
Glycophorin A	0.047	0.012	2.522*
Protamine 2	0.053	0.023	1.370
SRY	0.016	0.006	1.588
EP2_variantC	0.009	0.000	1.471
Rhesus-like	0.037	0.029	0.791
EP2_variantE	0.006	0.000	1.008
TGIF-like	0.010	0.005	0.784
EP2_variantD	0.004	0.000	0.964
Rh50	0.009	0.006	0.687
BRCA1	0.008	0.005	1.593
Apolipoprotein_AII	0.015	0.011	0.239
Glycoprotein alpha-2	0.009	0.006	0.605
Interleukin-4	0.003	0.000	1.068
STRL33	0.003	0.000	1.455
Interleukin-8 receptor	0.006	0.003	0.584

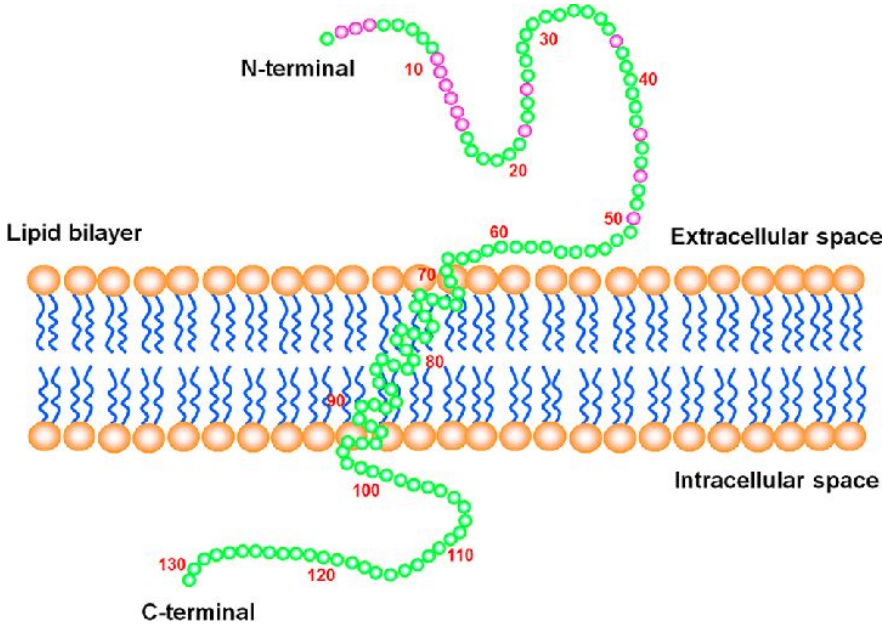
Z-boot: the standard error for this test was obtained by the bootstrapping. Z-a

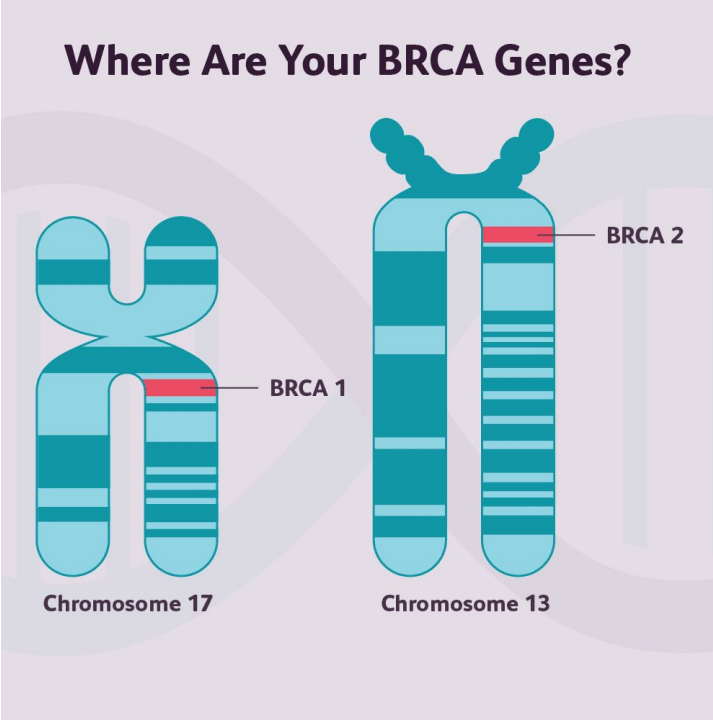
* Significant at the 5% level.

Protein with a significant dN/dS ratio

Glicophorin A
caused by the interaction of this receptor protein with the malaria parasites.

BRCA1 (BREAST CANCER 1 GENE): some mutations at this locus are known to cause breast cancer, but the normal function does not appear to be well understood (Narod and Foulkes, 2004). Therefore, it is difficult to relate the difference of this gene to the phenotypic differences between humans and chimpanzees.





Espressione genica: Anche se molti dei geni sono simili o identici tra umani e scimpanzé, l'espressione di questi geni può variare notevolmente; produzione di proteine in quantità diverse o in contesti cellulari diversi. Le differenze nell'espressione genica sono state osservate in molti tessuti, inclusi il cervello, dove possono influenzare lo sviluppo neuronale e la funzione cerebrale.

Modificazioni post-traduzionali: Possono influenzare l'attività delle proteine, la loro interazione con altre molecole e persino il loro destino all'interno della cellula.

Splicing alternativo: Il processo di splicing alternativo permette a un singolo gene di produrre più forme di RNA messaggero (e quindi più proteine) mediante l'inclusione o l'esclusione di certe sequenze. Le variazioni nello splicing alternativo tra umani e scimpanzé possono portare alla produzione di varianti proteiche con funzioni leggermente diverse.

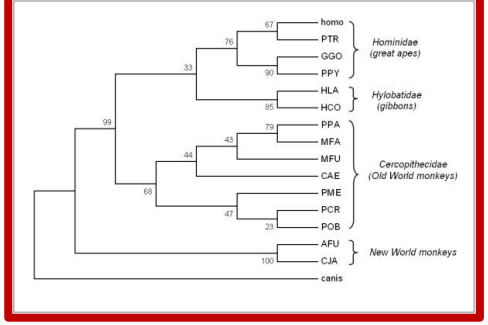
Duplicazioni geniche e perdite: Nel corso dell'evoluzione, alcuni geni possono essere duplicati o persi. Le duplicazioni possono portare a nuove funzioni geniche o aumentare la complessità di certi tratti, mentre le perdite di geni possono riflettere l'adattamento a nuovi stili di vita o ambienti. Ad esempio, i geni coinvolti nel sistema olfattivo mostrano un numero significativo di perdite nell'antenato comune di umani e scimpanzé, ma il pattern di queste perdite differisce tra le due specie, riflettendo differenze nelle capacità olfattive.

Humans vs apes

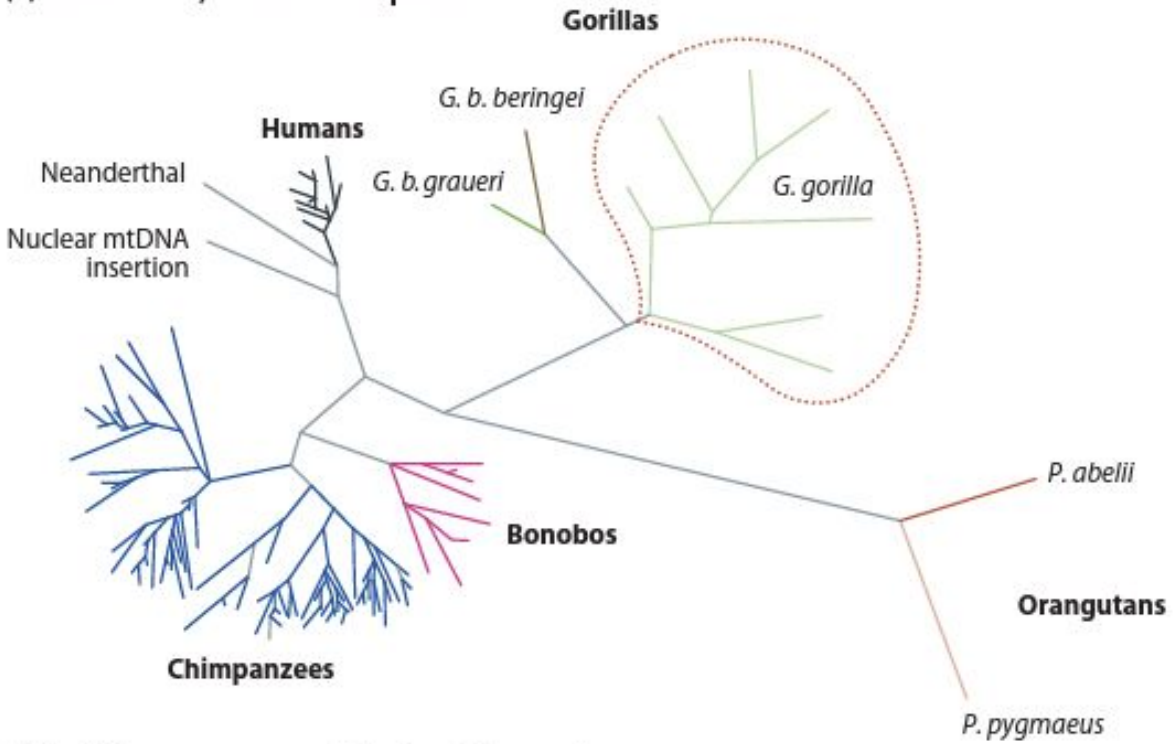
1. Structural DNA diversity (chromosomes)
2. Rough diversity estimates (DNA-DNA Hybridization)
3. DNA sequencing (autosomes, mtDNA and Y)
4. Proteins
5. Comparison of genetic diversity in apes and humans

Within species diversity
population size

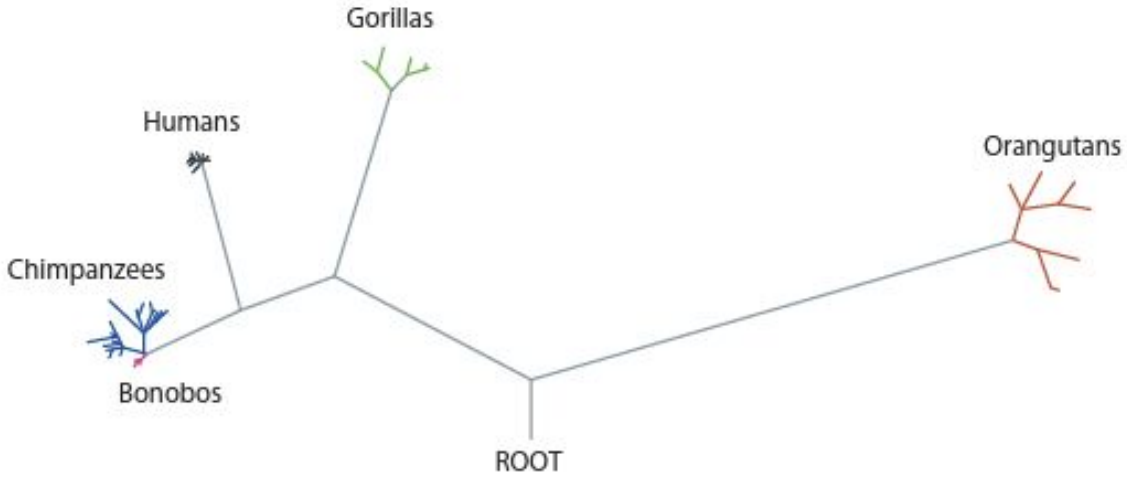
Human	*Homo sapiens*	**7,658,000,000**
Common chimpanzee	*Pan troglodytes*	**172,700–299,700**
Bonobo	*Pan paniscus*	**29,500–50,000**
Western gorilla	*Gorilla gorilla*	**150,000–250,000**
Eastern gorilla	*Gorilla beringei*	**5,880**
Bornean orangutan	*Pongo pygmaeus*	**47,000–73,000**



(a) mtDNA HVSI; unrooted and pruned

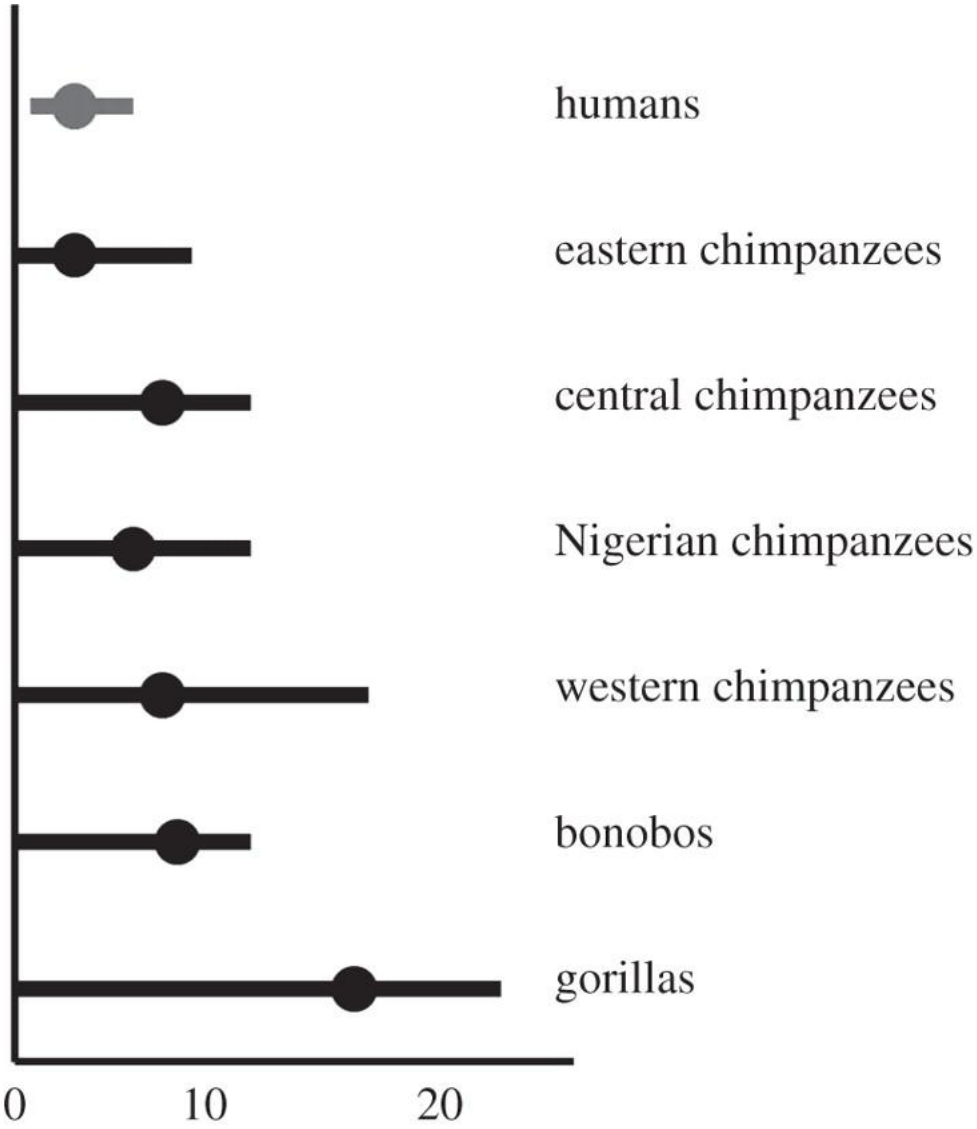


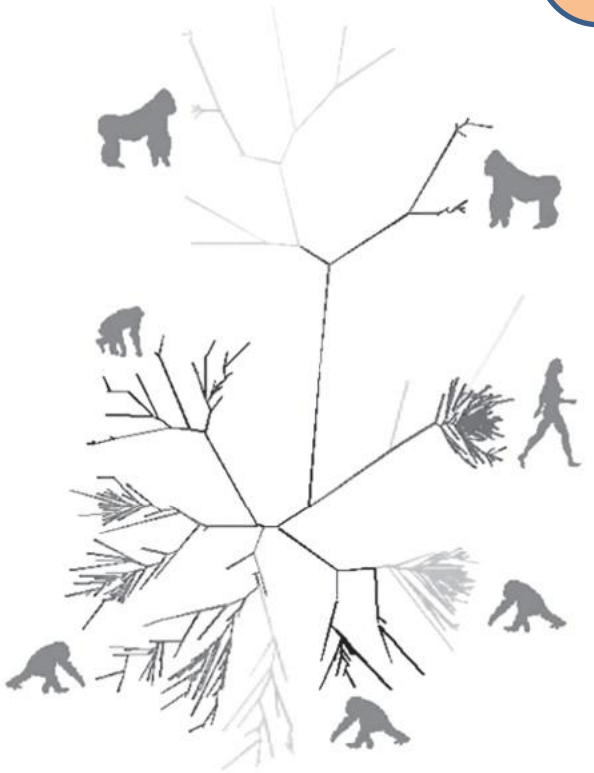
(b) Xq13.3 sequences; rooted using gibbon outgroup

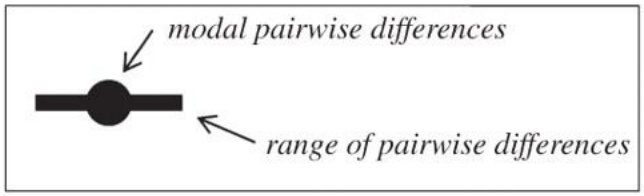


Comparative genetic diversity of apes and humans

mtDNA







Foley R A , Mirazón Lahr M Phil. Trans. R. Soc. B 2011;366:1080-1089

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

Within species diversity

TABLE 7. 3: RELATIVE DIVERSITIES OF VARIOUS LOCI IN HUM

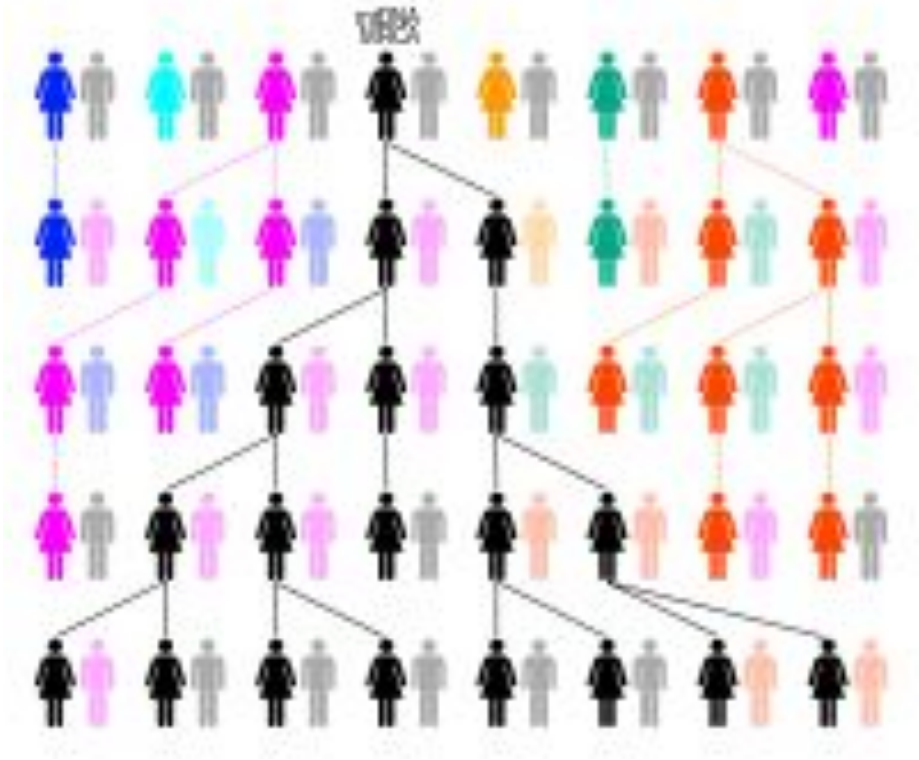
Locus	Chimpanzee vs. human	Bonobo vs. human
Xq13.3	3-fold greater	n.d.
mtDNA	3–4-fold greater	Greater
Y chromosome	Greater	n.d.
MHC class I genes	Greater, but less in HLA-A comparison	n.d.
ABO blood group genes	2–3-fold less	4–7-fold less
Microsatellites	Less	n.d.
Minisatellites	Less	n.d.

TABLE 7. 3: RELATIVEANS AND GREAT APES.

Locus	Gorilla vs. human	Orangutan vs. human
Xq13.3	2-fold greater	3.5-fold greater
mtDNA	Greater	Greater
Y chromosome	Less	Greater
MHC class I genes	n.d.	n.d.
ABO blood group genes	n.d.	n.d.
Microsatellites	n.d.	n.d.
Minisatellites	n.d.	n.d.

Compiled from data in Ely *et* agneux *et al.* (1999); Adams *et al.* (2000); Sumiyama *et al.* (2000); Gibbons (2001);lo not give quantitative measures, and comparison between them is complicated es.

Why greater diversity in Apes than humans?



- ★ Greater age of Apes (TMRCA)
- ★ Fragmentation and isolation in Apes
- ★ Population expansion in humans

What made us humans

Human-lineage-specific (HLS) traits are phenotypes of the human lineage that arose a**fter the split from the *Pan lineage***. A substantial number of forces are likely to have contributed to the development and maintenance of these traits

Plausible forces commonly discussed are macro- and micro-level climate changes that occurred frequently over the course of human evolution.

Box 1 | **Examples of human-lineage-specific traits and potential forces shaping them**

Example of phenotypic feature	Human-lineage-specific trait	Possible evolutionary advantages
Brain growth trajectory	Prolonged postnatal brain growth and delayed myelinization period; enhanced cognition	Allowed creation of novel solutions to survival threats; increased the critical period for learning new skills; facilitated emergence of uniquely human cognitive skills
Brain size	Increased brain/body size ratio; enhanced cognition	Allowed creation of novel solutions to survival threats; improved social cognition
Descended larynx	Portion of tongue resides in throat at level of pharynx; larynx descended into throat	Helped to develop spoken language
Eccrine sweat gland density	Higher density of eccrine glands; enhanced sweating capacity	Enhanced cooling ability; allowed protection of heat-sensitive tissues (for example, the brain) against thermal stress; facilitated endurance running
Endurance running	Improved energy use during periods of high energy demand; increased capacity to transfer energy (in the form of glycerol) from fat stores to muscle; anatomical changes relating to running ability	Allowed persistence hunting to emerge as a viable strategy for accessing the benefits of increased meat consumption; increased range of food sources; improved diet may have facilitated brain evolution
Labour	Earlier onset and longer duration of labour	Partially protected the child and mother from damage due to increased head circumference
Lacrimation	Emotional lacrimation (crying)	Enhanced emotional communication within social groups; increased affective communication
T cell function	Relative T cell hyper-reactivity	Enhanced immune function
Thumb	Increased length; more distally placed; larger associated muscles	Allowed creation of more detailed tools; allowed manipulation of objects on a finer scale

Perspective

Comparing the human and chimpanzee genomes: Searching for needles in a haystack

Ajit Varki[1] and Tasha K. Altheide

Evolution of genetic and genomic features unique to the human lineage

Majesta O'Bleness[1], Veronica B. Searles[1], Ajit Varki[2,3,4], Pascal Gagneux[3,4] and James M. Sikela[1,4]

NATURE REVIEWS | GENETICS VOLUME 13 | DECEMBER 2012 | 853

Biomedical Differences Between Human and Nonhuman Hominids: Potential Roles for Uniquely Human Aspects of Sialic Acid Biology

Annu. Rev. Pathol. Mech. Dis. 2011.6:365-393.

Nissi M. Varki,[1] Elizabeth Strobert,[2] Edward J. Dick Jr.,[3] Kurt Benirschke,[1] and Ajit Varki[1]

http://www.sciencedirect.com/science/article/pii/S0168952516301548

Genome Res. 2005 15: 1746-1758

- Sequencing of the chimpanzee genome signals not an end, but rather a **beginning** for researchers across diverse fields.

- understanding what makes us evolutionarily, biomedically, and cognitively different from chimpanzees will require extensive comparative **phenomics** to complement the comparative genomics now possible using the chimpanzee genome.

- studies of **intra-specific variation among great apes** are in their infancy, and biomedical and physiological data are few.

- Genomic data alone cannot predict **epistatic** interactions between various loci, nor can it reveal the **pleiotropic** effects of changes that have occurred in a single gene

- studies on captive apes should try to financially contribute toward their **conservation in the wild**, e.g., via a proposed Great Apes Conservation Trust

Genome Res. 2005 15: 1746-1758

Perspective

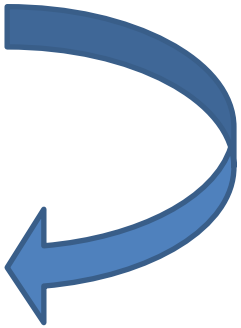
Comparing the human and chimpanzee genomes: Searching for needles in a haystack

Ajit Varki[1] and Tasha K. Altheide

1 **Narrowing the search to important differences**

2 **Identification of sites of sequence difference.**

3 **Comparative expression analysis in human and chimpanzee.**



1a

Using outgroups to define human-specific changes

additional primate **outgroups** will be important for detecting selection over longer time periods and for eliminating false positives

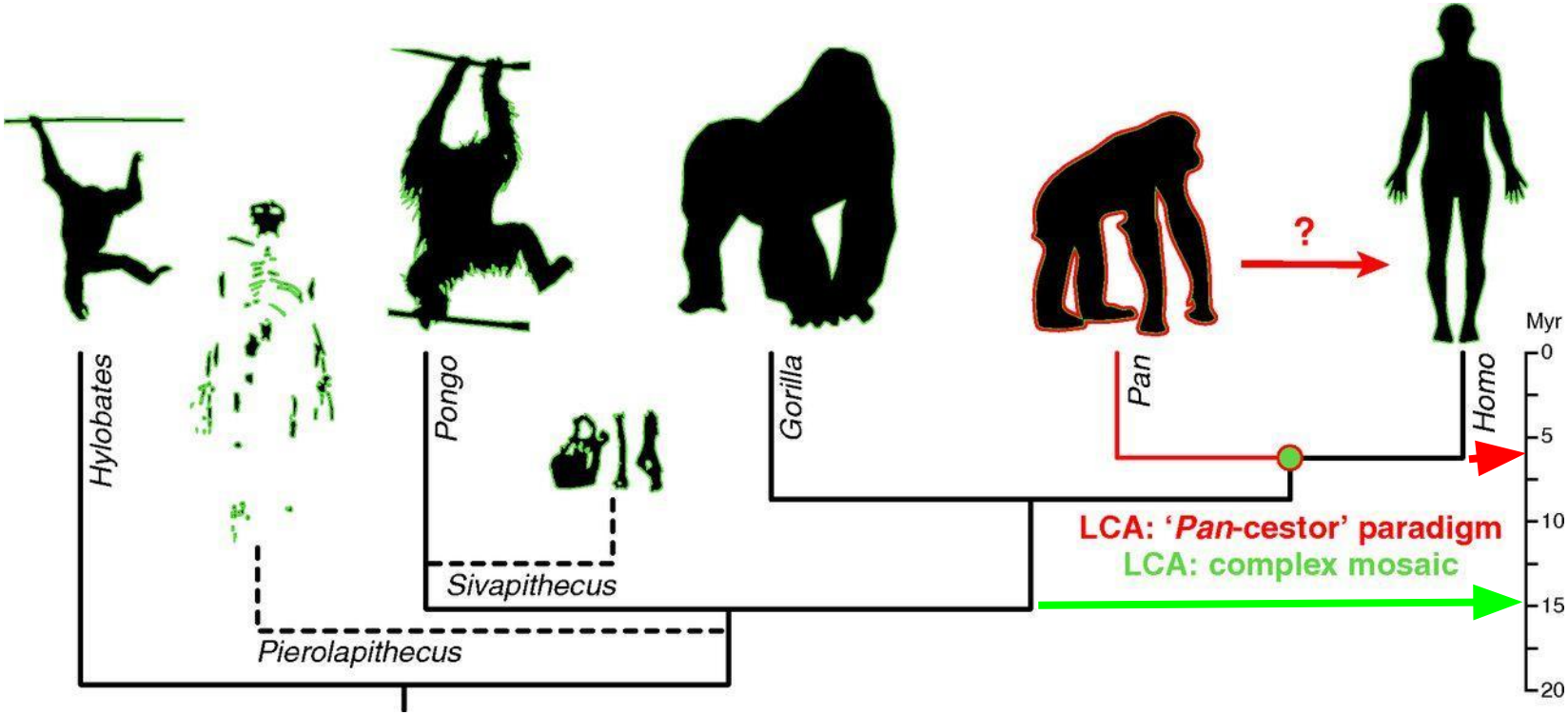
"**outgroups**" to determine the ancestral state (nucleotide or otherwise) at any given locus, and thus establish the subset of human specific changes.

Sequencing additional primate genomes is thus important, as is the careful choice of an appropriate outgroup species.

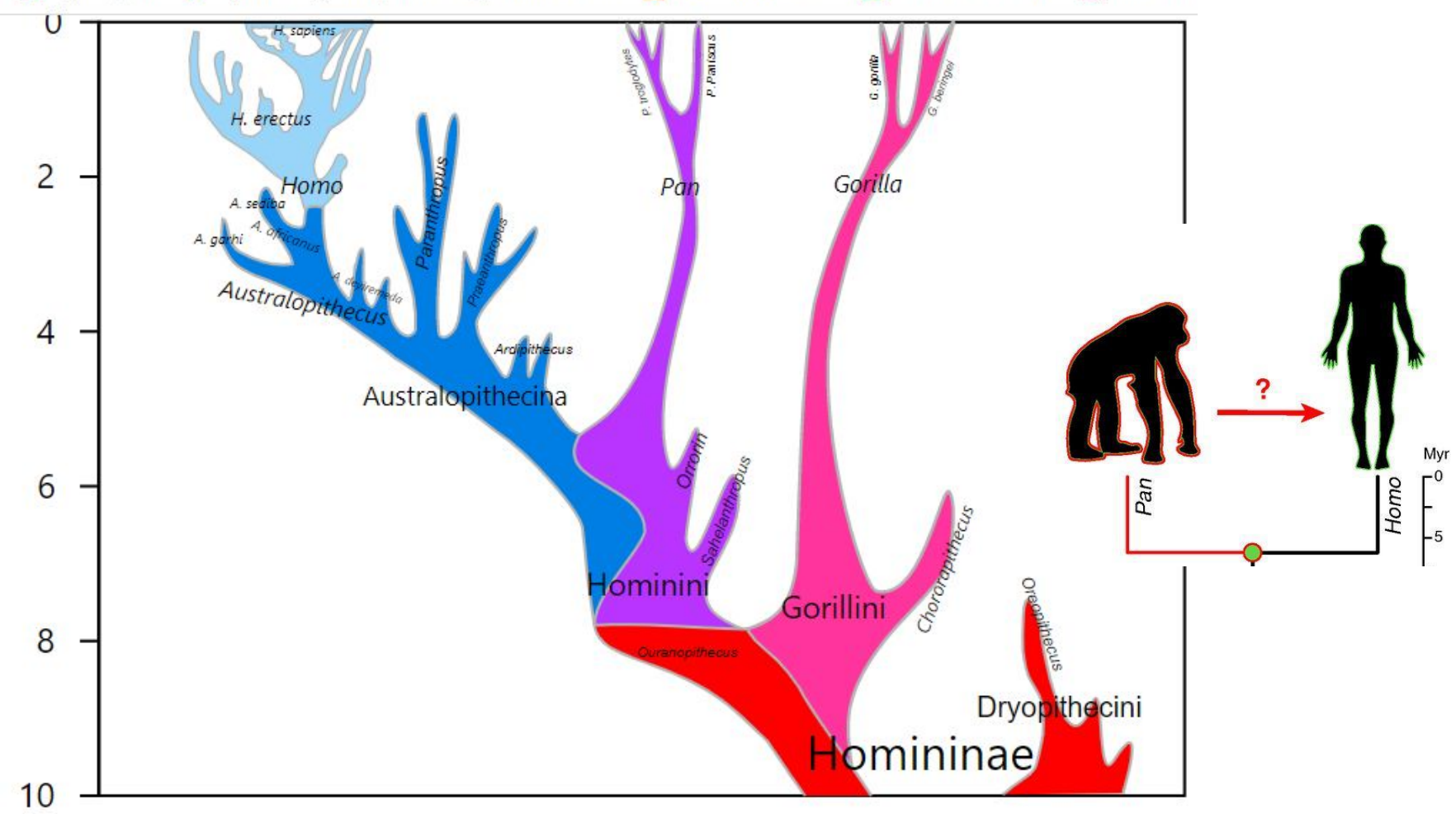
The **large divergence time between primates and rodents** (>60 Myr) means that some mutational events and/or orthology between loci will be obscured when using rodents as outgroups.

The orangutan (***Pongo***) *provides* the appropriate level of sequence divergence to define regions of functional sequence conservation.

Time-calibrated hominoid phylogeny highlighting the position of the LCA of Pan and Homo.



Sergio Almécija PNAS 2016;113:8:E943-E944



HYBRID SPECIATION "original divergence" between populations may have occurred as early as 13 million years ago (Miocene), hybridization may have been ongoing until as recently as 4 million years ago (Pliocene).

but… looking at transitions at CpG sites in genome sequences, which exhibit a more clocklike behavior than other substitutions, arriving at an estimate for human and chimpanzee divergence time of 12.1 million years. doi:10.1073/pnas.1600374113

1b

Excluding intra-species polymorphisms

- One must also ensure that an apparent genomic difference between humans and chimpanzees is not simply due to a polymorphism in one of the species

- Sequences from multiple individuals of both species are needed

- Surveying sequence variation in a minimum of 10 globally distributed humans has also been suggested to ensure a high probability that a given sequence is fixed

- This minimum number will be higher for chimpanzees because of their greater intra-specific diversity

2 Identification of sites of sequence difference.

A. Newly created genes in one of the species. e.g. by duplication or retroposition: neofunctional, non functional or sub-functional

B. Deletions in one of the species, leading to gene loss

C. Pseudogene in one of the two species. Pseudogenes can arise in one of the two species by recent mutation.

D. Non-conservative amino acid changes, especially at positions that are highly conserved in evolution with more distantly related mammals or at known functional sites.

E.Accellerated evolution

F. Changes in regulatory sequences. As regulatory regions are identified, it will then be possible to recognize differences that alter the sequences in functionally significant ways.

2

A. Newly created genes in one of the species

may allow for the evolution of new functions unique to the species

33% of human duplications are human specific (Cheng et al. 2005)

estimated 200–300 species-specific retroposed gene copies in humans and chimpanzees; ample landscape to explore.

possible role of Alu insertions (exonization)

Table 3. Some candidate genes and gene families that may contribute to phenotypic differences between humans and apes[a]

Gene(s)	Gene product(s)	Unusual hominid or human-specific features	Potential relevance to the human condition
SPANX (Xq27.1)	Sperm proteins associated with nucleus—genes on X chromosome	*SPANX-C* is specific to humans. *SPANX-B* has duplicated in humans	Rapidly evolving in all hominids. Expressed in normal testis, and in some cancers (Kouprina et al. 2004a)
PCDH11Y	Protocadherin XY	Duplicated onto Y in Yp11.2/Xq21.3 pseudoautosomal region only in humans	Expressed from Y and escapes X-inactivation? Y copy has undergone structural changes. Selectively expressed in brain. Probable adhesion molecule. Significance unknown, hypothesized to be involved in brain development, lateralization and schizophrenia risk (Ross et al. 2003; Blanco-Arias et al. 2004)

subfamily of cell adhesion molecules
predominantly expressed in the nervous system
believed to play a role in establishing the complex neural network during development

Online Mendelian Inheritance in Man (OMIM) is a database that catalogues all the known diseases with a genetic component, and—when possible—links them to the relevant genes in the human genome and provides references for further research and tools for genomic analysis of a catalogued gene.[1] OMIM is one of the databases housed in the U.S. National Center for Biotechnology Information (NCBI) and included in its search menus

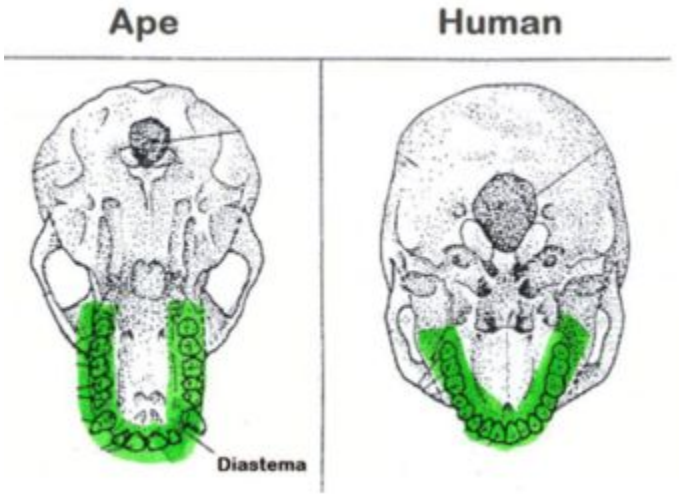
2 B. deletion

Table 3. Some candidate genes and gene families that may contribute to phenotypic differences between humans and apes[a]

Gene(s)	Gene product(s)	Unusual hominid or human-specific features	Potential relevance to the human condition
Individual genes			
MYH16	Myosin heavy chain 16	Human-specific 2-bp deletion causing frameshift—predicted 76-kD unstable head domain	Claimed to be cause of reduction in the type II fibres of human jaw muscle. (Stedman et al. 2004; Perry et al. 2005)

- Myosin heavy chain 16 is a specialized muscle protein found only in the temporalis and masseter muscles of the jaw of Primates
- If they are missing, the muscles will be smaller.
- In non-human primates, MYH16 is functional and the animals have powerful jaw muscles. In humans, the MYH16 gene has a mutation that causes the protein not to function.
- Mutation dated between 2.4 and 5.3 MYA

It caused narrower jaws, smaller teeth and space for the 3rd molar became less. 3rd molar is almost vestigeal now on account of the MYH-16



2

C. Pseudogene in one of the two species

Table 3. Some candidate genes and gene families that may contribute to phenotypic differences between humans and apes[a]

Gene(s)	Gene product(s)	Unusual hominid or human-specific features	Potential relevance to the human condition
CMAH	CMP-Neu5Ac hydroxylase	92-bp deletion of exon 6 causing frameshift and inactive enzyme. Fixed in modern humans	Absence of sialic acid Neu5Gc. Change in resistance or susceptibility to pathogens. Loss of ligand for some Siglecs. Dated ~2.5–3 Mya. Dietary Neu5Gc in meat became foreign antigen (Chou et al. 1998, 2002; Irie et al. 1998; Havakawa et al. 2001)

pseudogenized by an **Alu-mediated deletion event**, which seems to have occurred ~2-3 million years ago in our ancestors

***N*- or *O*-substituted derivatives of neuraminic acid, a monosaccharide with a nine-carbon backbone**

Sialic acid-rich oligosaccharides on the glycoconjugates (glycolipids, glycoproteins, proteoglycans) found on surface membranes help keep water at the surface of cells

can "hide" mannose antigens on the surface of host cells or bacteria from mannose-binding lectin preventing the activation of complement

Sialic acid also plays an important role in human influenza infections, interacting with viral hemagglutinins

Siglecs, *sialic acid binding Ig-like lectins* are cell surface receptors and members of the immunoglobulin superfamily (IgSF) that recognize sialic acids

2 D. Non-conservative amino acid changes

Table 3. Some candidate genes and gene families that may contribute to phenotypic differences between humans and apes[a]

Gene(s)	Gene product(s)	Unusual hominid or human-specific features	Potential relevance to the human condition
MAOA	Monoamine oxidase A	Human-specific nonconservative change Glu151Lys in active site	Substitution affects protein dimerization according to a 3D structural model and predicts functional change (Andres et al. 2004)

MAOA is a mitochondrial enzyme, encoded by nuclear genes that are located on the long arm of the X chromosome (Xp11.4- p11.3)

Violence gene? Mice lacking MAOA show enhanced aggression and altered emotional learning, but we know that no simple explanation apply to humans.

Deletion of a genomic region including MAOA causes congenital blindness and in many cases deafness and mental retardation, psychoses and other behavioral, cognitive, and sleep disorders

The specific depletion of MAOA functional protein by a point mutation in humans s reported to be associated with intellectual and behavioral alteration.

Only one human exclusive (absent in chimpanzee, and gorilla(non-conservative change is present in the gene: Glu151Lys. This human substitution affects protein dimerization according to a three-dimensional structural model that predicts a non-negligible functional shift.

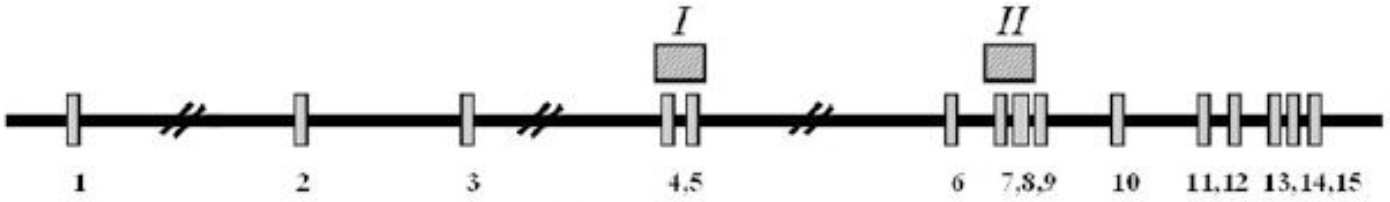


Fig. 1 Graphical representation of MAOA gene. *Vertical boxes*, the 15 exons of the gene analyzed for interspecific variation between humans, chimpanzees, and gorillas, *boxes above* regions I (4–5) and II (7–8), analyzed for intraspecific variability in primates

- 112 humans and 1056 human samples from **HGDP-CEPH** Human Genome Diversity Panel; Cann et al. 2002)
- 35 **chimpanzee** (Pan troglodytes subspecies troglodytes, verus, schweinfurthii, or unknown), eight **bonobo** (P. paniscus), ten **gorilla** (Gorilla gorilla gorilla), ten **orangutan** (Pongo pygmaeus subspecies pygmaeus and abelii), and five **gibbon** (Hylobates syndactylus, H. lar, and H. klossii) unrelated X chromosomes

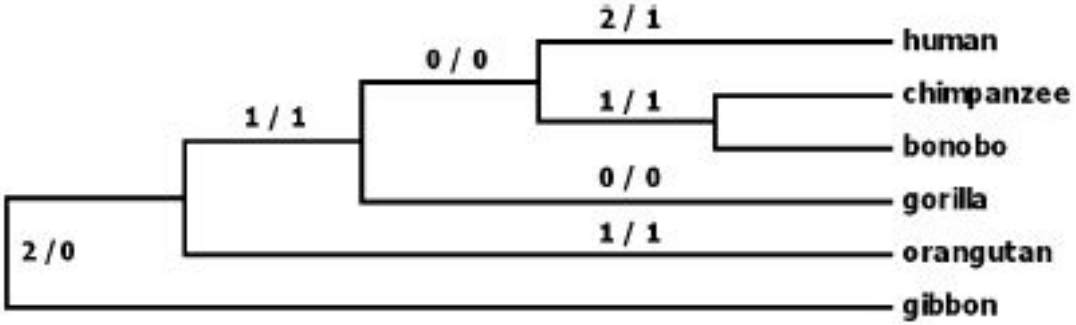


Fig. 2 Cladogram for the species analyzed for 4–5 (I), 7–8 (II) segments of the MAOA gene (note that this figure does not show differences for the whole coding region), with the number of non-synonymous/synonymous substitutions that occurred in every branch. One synonymous change, common to human and gorilla

- **Glu151Lys (G451A)** only in humans (changing a negatively charged, hydrophilic amino acid to a positively charged),
- Strong conservation of amino acid 151 among species and between paralogous genes-
- 3D models suggest that could a**Glu151Lys a**ffect MAO activity by altering the dimerization of the protein

2

D. Non-conservative amino acid changes

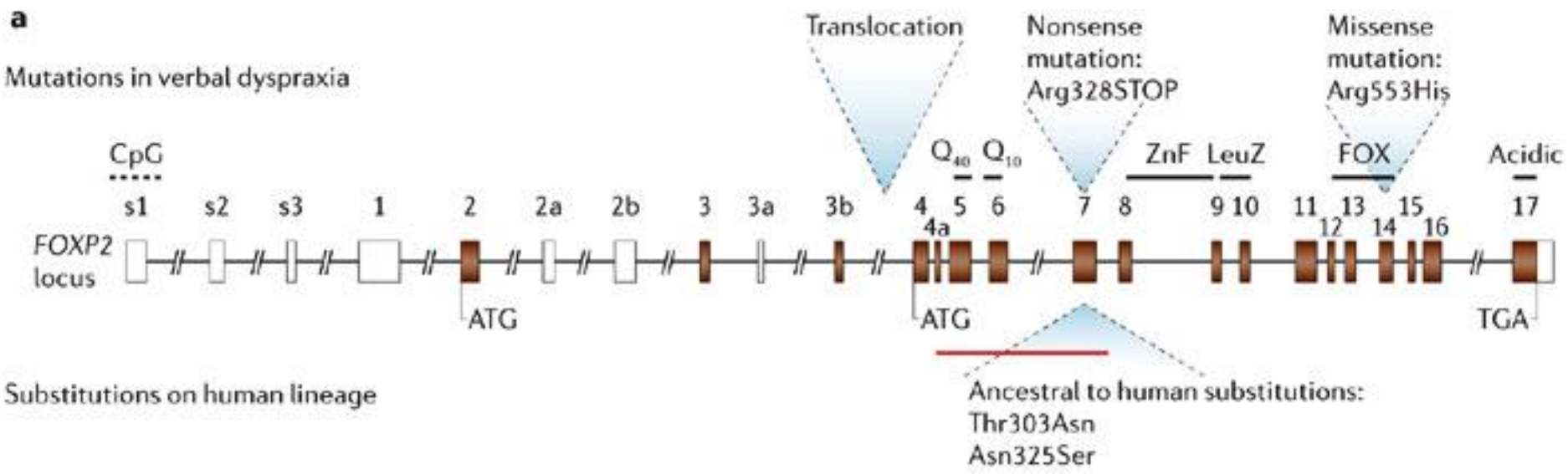
Table 3. Some candidate genes and gene families that may contribute to phenotypic differences between humans and apes[a]

Gene(s)	Gene product(s)	Unusual hominid or human-specific features	Potential relevance to the human condition
Individual genes			
FOXP2	Putative transcription factor with polyglutamine tract and forkhead DNA binding domain	Two human-specific amino acid changes	Mutant humans have motoric speech disorder (developmental verbal dyspraxia). Region positively selected and fixed in humans <200,000 years ago (Enard et al. 2002b; Zhang et al. 2002)

- ***FOXP2**, a transcription factor associated* with an inherited human disorder of speech production (Enard et al. 2002b; Zhang et al. 2002), highly conserved in vertebrates

- found to have two human-specific amino acid changes, and the genomic region in question appears to have been positively selected and fixed in humans <200,000 years ago (Enard et al. 2002b).

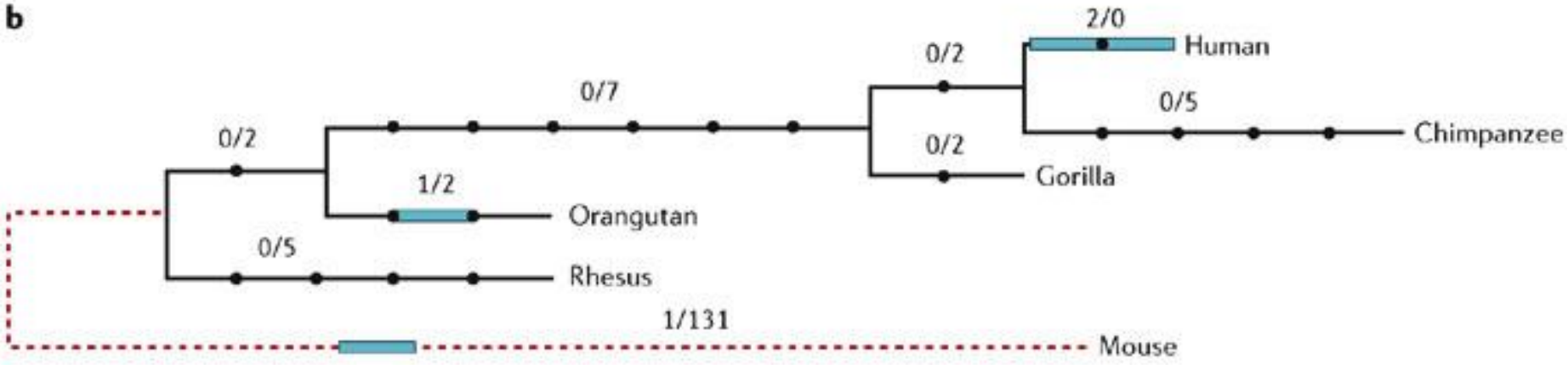
FOXP2 on human **chromosome 7q31**, encodes a **protein of 715 amino acids** belonging to the forkhead class of transcription factors2



A multidisciplinary perspective on language evolution. a | Genetics — the genomic structure of human

forkhead box P2 (FOXP2), showing the location of mutations that cause verbal dyspraxia, which are **distinct from sites of evolutionary substitution in the human lineage** (filled rectangles, coding exons; white rectangles, non-coding exons). The red bar indicates genomic regions that show evidence of a selective sweep113,114. Exons encode polyglutamine tracts (Q40 and Q10), a zinc-finger motif (ZnF), a leucine zipper (LeuZ), the forkhead domain (FOX) and an acidic C-terminus (Acidic). s1–s3 are alternatively spliced untranslated 5′ exons

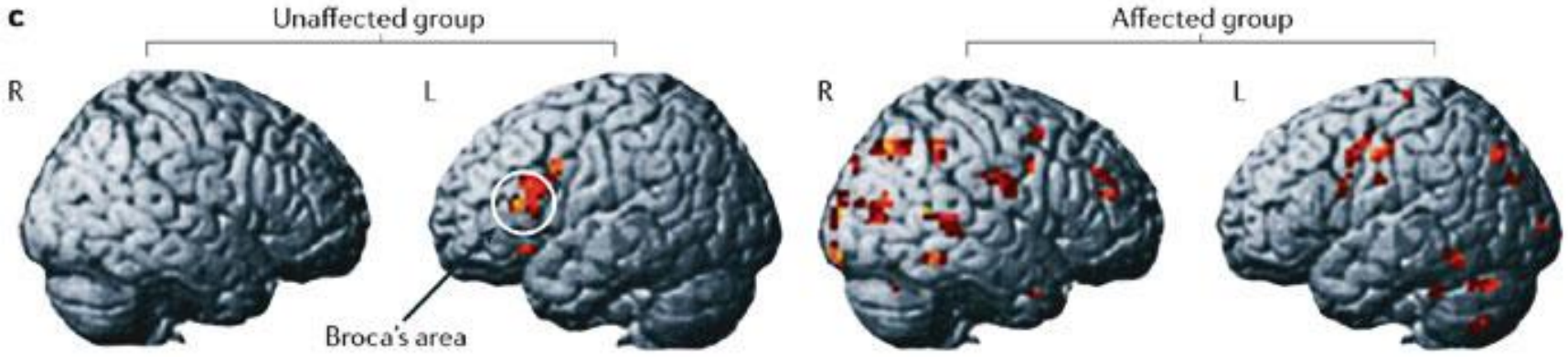
nature REVIEWS GENETICS



Evolution — nucleotide substitutions in the FoxP2 coding region for different lineages during primate evolution, shown as **non-synonymous over synonymous substitutions** (horizontal bars, nucleotide changes over time; shaded bars, amino-acid changes

nature REVIEWS GENETICS

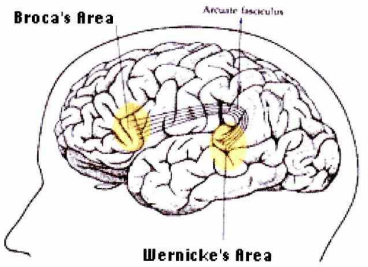
FOXP2 Gene disruption



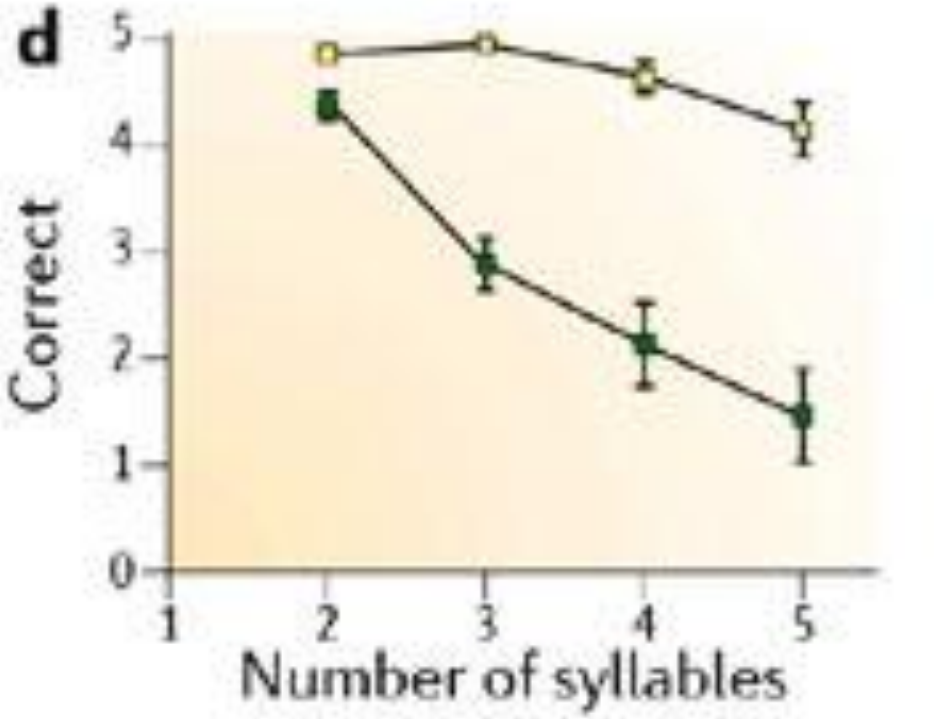
Neuroimaging

Humans carrying disrupted FOXP2 show functional abnormalities when carrying out a language task, even when producing verb forms mentally rather than aloud. **The anomalies involve underactivation of Broca's area** and bilateral activation in multiple cortical regions. The diagram shows the group average activation in the unaffected and affected members of the KE family, which is displayed at a threshold of $P<0.05$, corrected for multiple comparisons

- Area di Broca (III circomvoluzione frontale, c. motorio linguaggio)
- Area parieto-temporale di Wiernicke (comprensione del linguaggio)



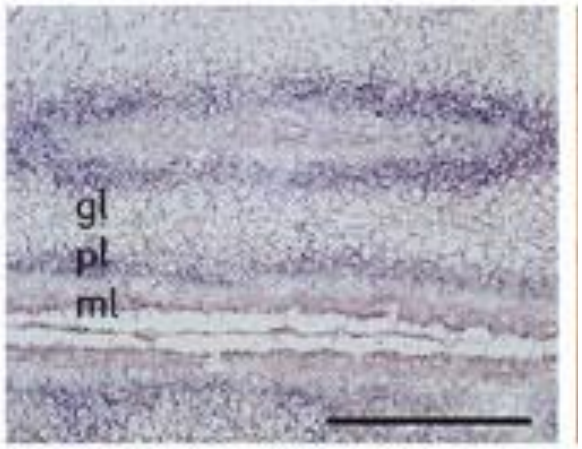
nature REVIEWS GENETICS



Neuropsychology — FOXP2 disruption leads to difficulties with coordinating speech.
Affected KE family members (green squares) perform worse than unaffected members (yellow squares) on **word-repetition tests that involve simple articulation patterns** (error bars, standard error of the mean). Impairment increases with syllable length. **Similar results are seen when repeating nonsense words, with greatest deficits on multisyllabic words that have complex articulation patterns**

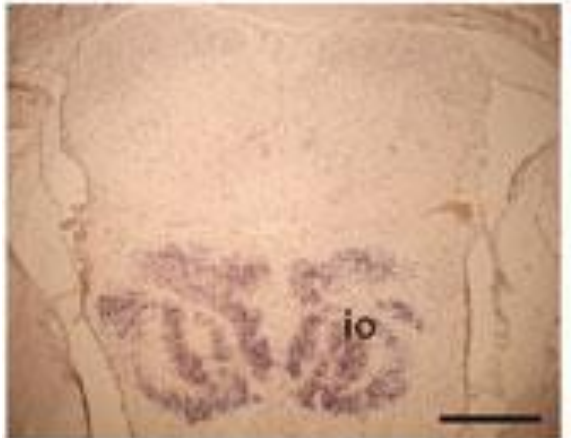
nature REVIEWS GENETICS

Molecular neuroscience — example sites of high Foxp2 mRNA expression in transverse sections from a newborn mouse brain (scale bars represent 0.5 mm).

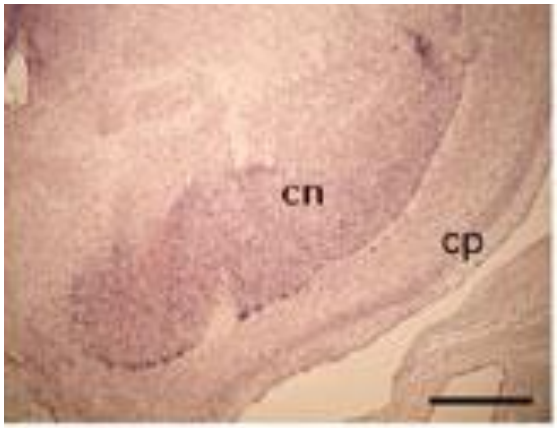


In the **cerebellum** Foxp2 expression is limited to Purkinje cells (pl), and absent from molecular (ml) and granular (gl) layers.

Neural expression patterns for this gene are highly conserved in all vertebrate species that have been studied, which range from humans to zebrafish.



In the **medulla** (middle panel) Foxp2 is expressed in the inferior olivary nucleus (io).



In the **forebrain** (right panel) there is strong expression in the caudate nucleus (cn) and the deepest layers of the cortical plate (cp).

A Humanized Version of Foxp2 Affects Cortico-Basal Ganglia Circuits in Mice

Cell 137, 961–971, May 29, 2009 ©2009 Elsevier Inc. 961

The “basal ganglia” refers to a group of subcortical nuclei responsible primarily for **motor** control, as well as other roles such as **motor** learning, executive functions and behaviors, and emotions.

Introduction of the evolutionary substitutions into the endogenous Foxp2 gene of mice.

- Mices generally healthy,but show qualitatively different ultrasonic vocalizations, decreased exploratory behavior and decreased dopamine concentrations
- This suggests that the humanized Foxp2 allele affects basal ganglia.
- In the basal ganglia affected in humans with a speech deficit due to a nonfunctional FOXP2 allele, medium spiny neurons have increased dendrite lengths and increased synaptic plasticity.
- Since mice carrying one nonfunctional Foxp2 allele show opposite effects, this suggests that alterations in cortico-basal ganglia circuits might have been important for the evolution of speech and language in humans.

2

E. Accellerated evolution

There are a number of genomic regions that have undergone substantial alteration of sequence or rearrangement in the human lineage. Accelerated evolution refers to situations in which sequence changes occur at a rate greater than the neutral mutation rate. Accelerated evolution implies that the changes have been selected because of their advantageous nature and thus have undergone rapid fixation. Identification of these regions relies on multiple methods and differs depending on whether the change is at the coding sequence, non-coding sequence, copy number or other structural level.

At the protein-coding sequence level, a comparison of K_a/K_s values between sequences is often used, where K_a is the number of nonsynonymous substitutions and K_s is the number of synonymous substitutions. Most gene-coding regions will have K_a/K_s ratios well below 1.0 owing to the effects of purifying selection. By contrast, coding regions under positive selection will exhibit a higher frequency of nonsynonymous changes and, as a result, a higher K_a/K_s ratio. Through this method, studies have identified accelerated evolution in the human lineage of a number of genes, one example being genes involved in nervous system function[111]. However, these estimates can be confounded by gene conversion events that erase evidence of selection by creating stretches of identical nucleotide sequences between homologous genes[112].

Evaluation of non-coding sequences is not as straight forward because of the difficulty in interpreting the importance of a change. Thus, studies identifying regions of accelerated evolution in non-coding regions have relied on looking for human-specific mutations in sequences that are highly conserved across mammals. An example of this is the identification of highly accelerated region 1 forward (*HAR1F*)[113]. *HAR1F* is a non-coding RNA expressed in the fetal brain that colocalizes with reelin, a protein that is important for cortical development.

2 E. Accellerated evolution

Table 3. Some candidate genes and gene families that may contribute to phenotypic differences between humans and apes[a]

Gene(s)	Gene product(s)	Unusual hominid or human-specific features	Potential relevance to the human condition
ASPM	Modulator of mitotic spindle in neural progenitors?	Accelerated evolution in ape and human lineages	Deletions in ASPM lead to microcephaly. Presumed to be related to increased brain size and/or other features of human brain (Zhang 2003; Dorus et al. 2004; Evans et al. 2004; Kouprina et al. 2004b; Mekel-Bobrov et al. 2005)
MCPH1	Microcephalin	As above	As above (Dorus et al. 2004; Evans et al. 2004, 2005)

The **Abnormal spindle-like microcephaly-associated protein** (ASPM) gene is the human equivalent of the Drosophila melanogaster 'abnormal spindle' gene (asp), which is known to be essential for normal mitotic spindle function in embryonic neuroblasts. **MCPH-1** is a DNA damage response protein. Inactivation of the MCPH1 gene in humans results in a severe condition called microcephaly, i.e. a small brain, associated with mental retardation.

Inactivation of these genes in humans results in a severe condition called microcephaly, i.e. a small brain, associated with mental retardation.
it had undergone significantly higher rates of protein sequence evolution in primates than in rodents and even more rapid evolution in humans, including evidence of a selective sweep in human populations.

2 F. Changes in regulatory sequences

Human-specific loss of regulatory DNA and the evolution of human-specific traits
McLean et al, Nature 2011

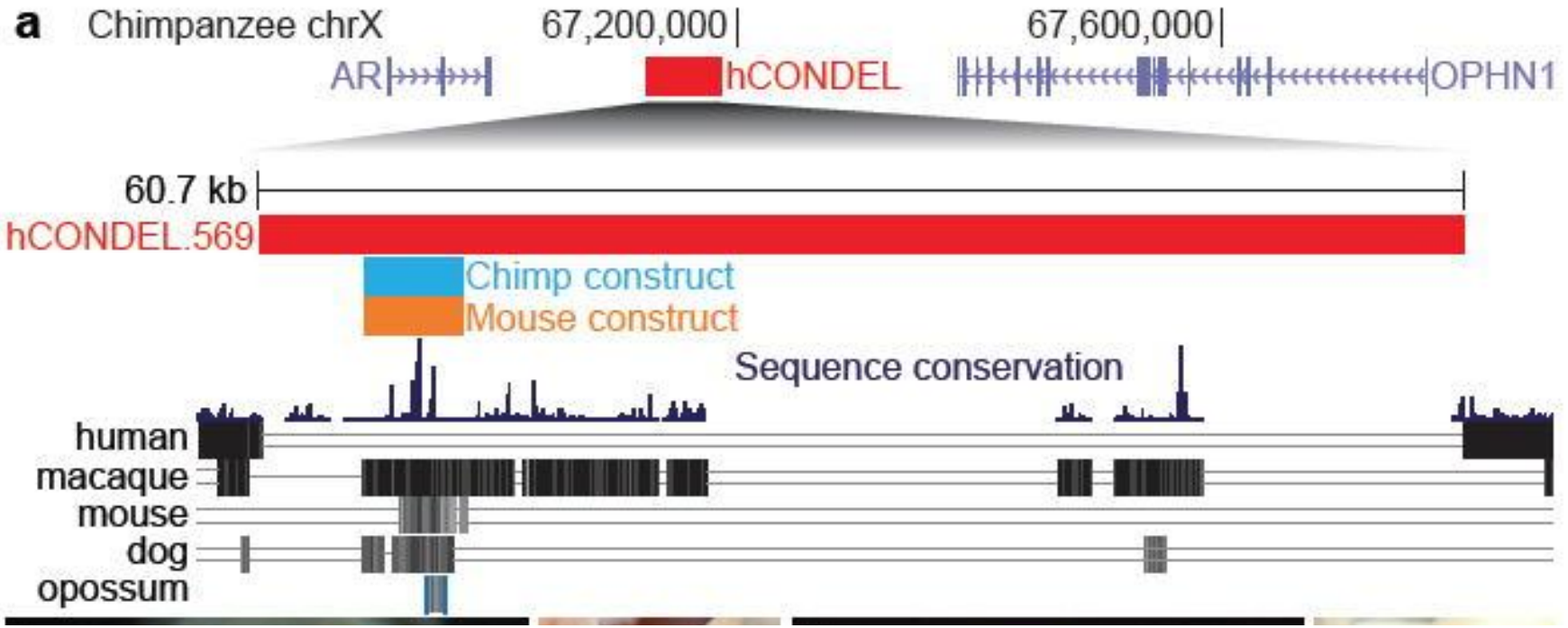
Here we identify molecular events particularly likely to produce significant regulatory changes in humans**: complete deletion of sequences otherwise highly conserved between chimpanzees and other mammals**.

510 such deletions in humans, which fall almost exclusively **in non-coding regions** and are enriched near genes involved in **steroid hormone** signaling and **neural function**.

One deletion removes a **sensory vibrissae** and **penile spine enhancer** from the human ***ANDROGEN RECEPTOR* (AR) gene**, a molecular change correlated with anatomical loss of androgen-dependent sensory vibrissae and penile spines in the human lineage[9,10].

Another deletion removes a forebrain subventricular zone enhancer near the tumor suppressor gene *GROWTH ARREST AND DNA-DAMAGE-INDUCIBLE, GAMMA* (*GADD45g*)[11,12], a loss correlated with expansion of specific brain regions in humans.

Deletions of tissue-specific enhancers may thus accompany both loss and gain traits in the human lineage, and provide specific examples of the kinds of regulatory alterations[6-8] and inactivation events[13] long proposed to play an important role in human evolutionary divergence.

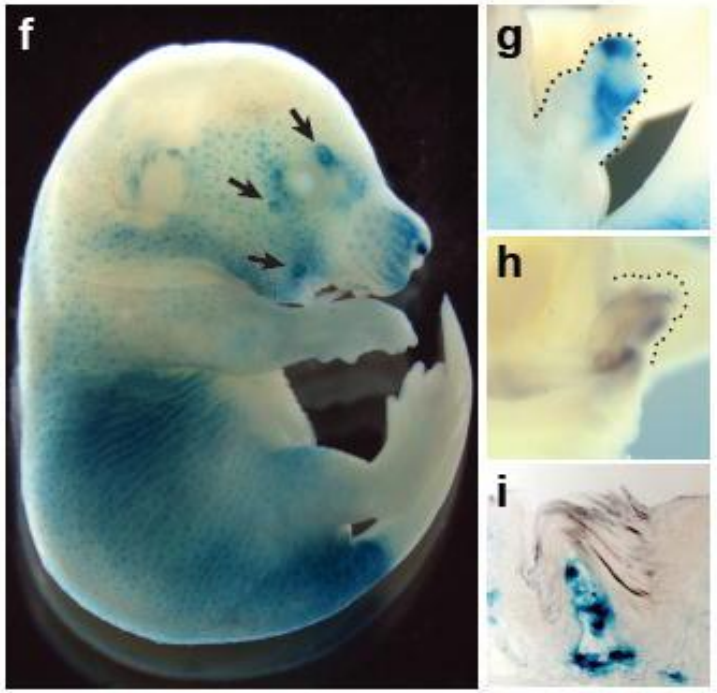


Transgenic analysis of a chimpanzee and mouse *AR* enhancer region missing in humans

a, *Upper panel*: 1.1 Mb region of the chimpanzee X chromosome. The red bar shows the position of a **60.7 kb human deletion removing a well-conserved chimpanzee enhancer between the *AR* and *OPHN1* genes**. *Lower panel*: Multiple species comparison of the deleted region, showing sequences alignable between chimpanzee and other mammals. Blue and orange bars represent chimpanzee and mouse sequences tested for enhancer activity in transgenic mice.

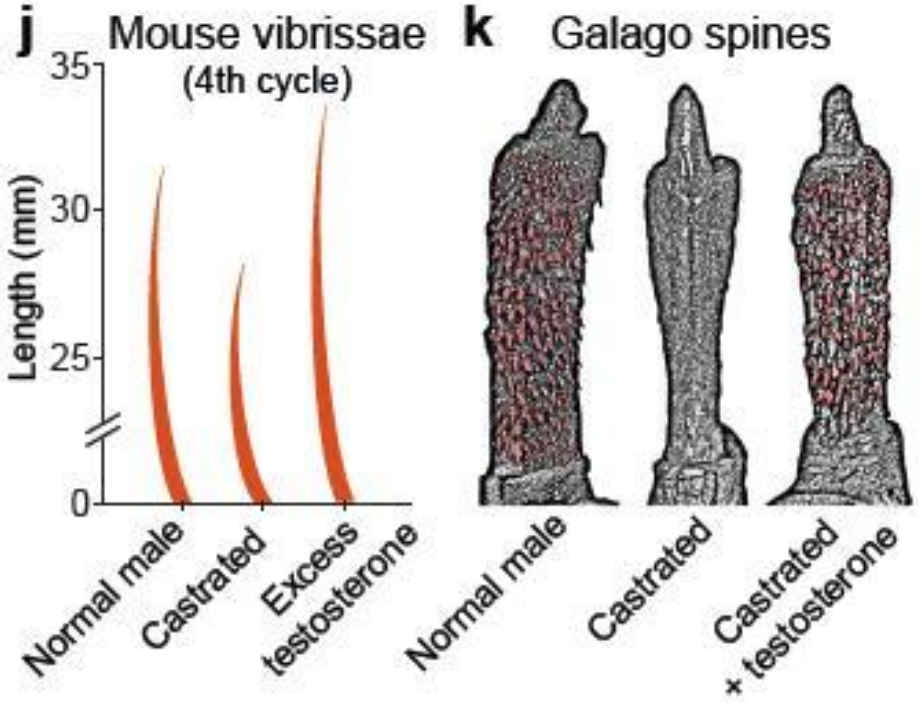
Transgenic analysis of a chimpanzee and mouse forebrain enhancer missing from a tumor suppressor gene in humans

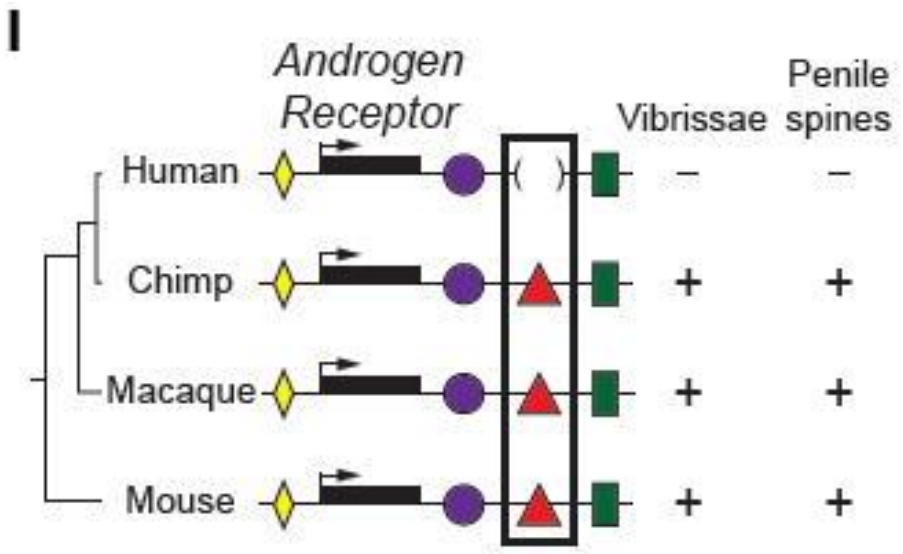
The chimpanzee sequence drives *lacZ* expression in, **b**, facial vibrissae (arrows); and **c**, genital tubercle (dotted line) of E16.5 **mouse embryos**. Histological sections reveal strongest staining in superficial mesenchyme of **d**, the prospective glans of the genital tubercle; and **e**, dermis surrounding the base of sensory vibrissae.



The mouse enhancer also drives consistent expression in **f**, facial vibrissae; **g**, genital tubercle; and hair follicles of E16.5 embryos. **h**, Endogenous *AR* is expressed in the genital tubercle (dotted line) as demonstrated by *in situ* hybridization. **i**, Histological section of a 60-day-old transgenic mouse penis showing postnatal *lacZ* expression in dermis of penile spines.

Vibrissae and penile spines are androgen-dependent, as shown by, **j**, changes in vibrissae length in **castrated and testosterone-treated mice** and, **k, loss and recovery of penile spines of a castrated and testosterone-treated primate** (*Galago crassicaudatus*) (modified from refs. [23] and [24]).





I, Model depicting multiple conserved **tissue-specific enhancers** (colored shapes) surrounding *AR* coding sequences (black bars) of different species. **Loss of an ancestral vibrissae/penile spine enhancer** in humans is correlated with corresponding loss of sensory vibrissae and penile spines.

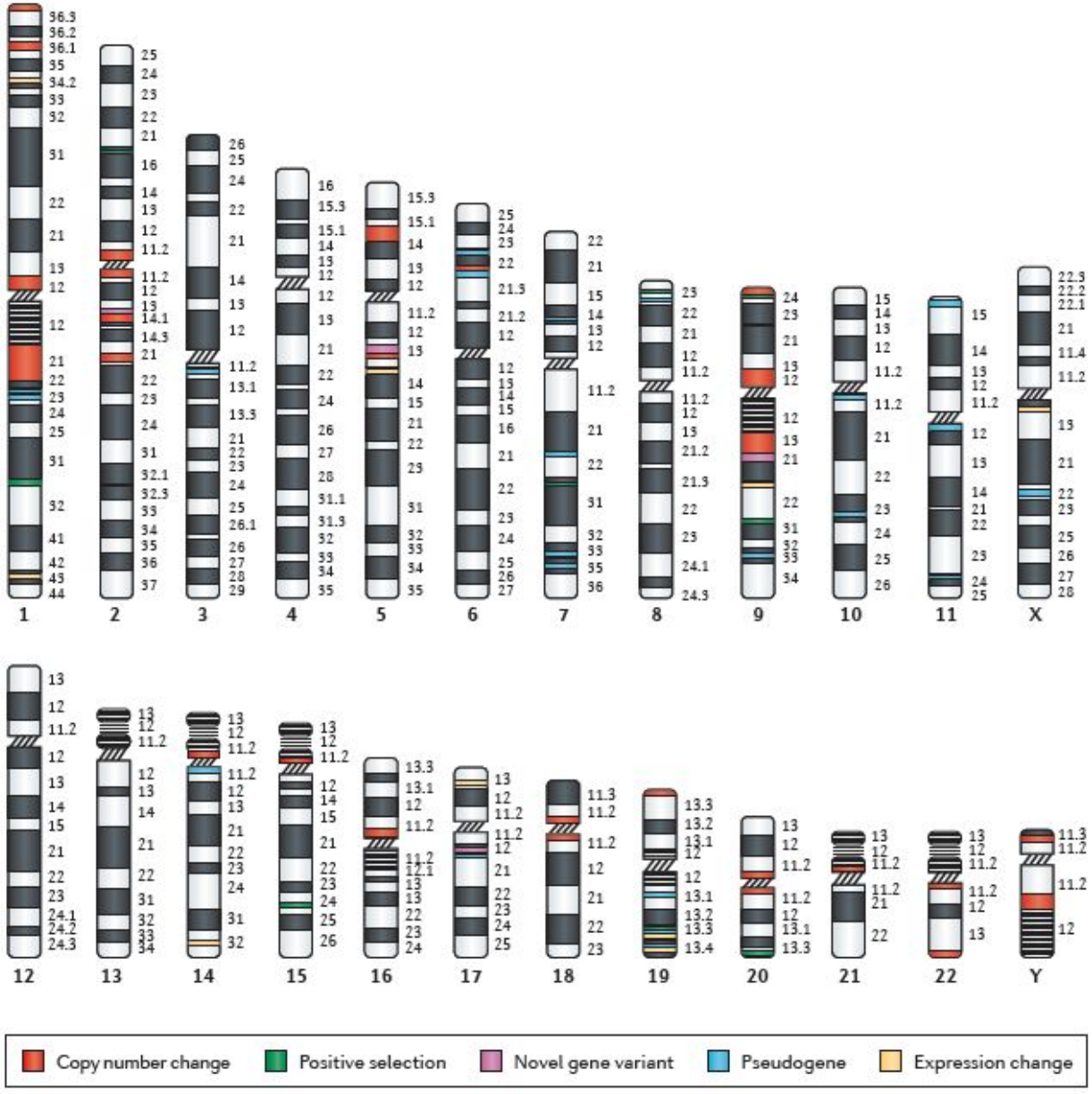


Figure 1 | **Genome positions of human-lineage-specific gene changes.** Human-lineage-specific (HLS) gene changes discussed in this paper are displayed in their corresponding genomic position across the human karyotype. The changes are divided into five categories that correspond to those listed in TABLE 1, and each type is colour coded. It should be noted that many genes have undergone multiple types of HLS changes, and in this case only one type is shown. For visualization purposes, the size of the coloured bands is not drawn to scale.