What makes us human: comparisons between human and primate genomes.

What makes us human? The phenotypic changes in morphology, cognition and behavior translated from a small subset of functionally important gene differences, representing adaptive changes, between humans and our closest living evolutionary relative, the chimpanzee.

Uniquely human phenotypic changes have been identified. Since the human-chimpanzee split, humans became increasingly bipedal and the post-cranial anatomy changed accordingly. Head shape changed and brain size increased. Body hair was reduced, but the growth-cycle for scalp hair was extended. Human culture expanded to include complex language with syntax, elaborate tool production and manipulation, and sophisticated cognitive capabilities. Along with these advances has come an increase in the incidence and severity of many diseases.

Two complementary approaches, comparative genomics - identification of regions of the genomes that are evolutionarily important using different signatures of adaptive evolution (regions that differ between chimpanzees and humans and seem to have undergone positive selection along one lineage); and candidate gene approach - examination of genes that are likely to be involved in specific traits of interest, have identified functionally important genetic changes that are likely to underlie the phenotypic differences.

Genomic comparisons of human and chimpanzee with additional evidence from genetic studies of human disease, mutational analysis and gene expression profiles help to understand the genetic basis of the human phenotype.

The initial comparison of the human and chimpanzee genomes have generated a catalogue of the genetic differences that have accumulated since the human and chimpanzee species diverged from the common ancestor, constituting approximately 35 million single-nucleotide substitutions, 5 million insertion/deletion events, and various chromosomal rearrangements. The catalog was used to explore the magnitude and regional variation of mutational forces shaping the 2 genomes, and the strength of positive and negative selection acting on their genes. A few of the findings are listed below.

Genome evolution:

a) Nucleotide divergence:

. Single-nucleotide substitutions occur at a mean rate of 1.23% between copies of the human and chimpanzee genome, with 1.06% or less corresponding to fixed divergence between the species.

. Substitutions at CpG dinucleotides, which constitute 1/4 of all observed substitutions, occur at more similar rates in male and female germ lines than non-CpG substitutions.

b) Insertion and deletion (indel) events are fewer in number than single-nucleotide substitutions, but result in 1.5% of the euchromatic sequence in each species being lineage-specific.

c) Transposable element insertions: short interspersed elements (SINEs) - Alu elements , the most common type in humans - have been threefold more active in humans, whereas chimpanzees have acquired two new families of retroviral elements. SINEs likely altered functional genes and their regulatory elements.

1/3 of the gene duplications seen in the human genome seem to be human-specific (chimpanzees have fewer copies) and result in gene expression differences between the species (Cheng et al. 2005 ref.1,2)

Gene evolution:

a) Orthologous proteins in human and chimpanzee are extremely similar, with 29% being identical and the typical orthologue differing by only two amino acids, one per lineage.

b) The substitution rate at silent sites in exons is lower than the rate at nearby intronic sites, consistent with weak purifying selection on silent sites in mammals.

c) Patterns of evolution in human and chimpanzee protein-coding genes are highly correlated and dominated by fixation of neutral and slightly deleterious alleles.

d) Of 13545 orthologous protein-coding genes compared across the human and chimpanzee genomes, 4.4% show a potential signature for positive selection: genes that have undergone accelerated evolution have high ratios of non-synonymous to synonymous substitutions (ka/ks > 1) on the chimpanzee or human lineage.

The functionally important genetic changes underlying the phenotypic changes are only a subset of the 1.2% genetic difference between human and chimpanzee genotypes [humans and chimpanzees share almost all of the human genes and 98.8% (96% if insertions and deletions are included) of the human DNA] and are hypothesized to be of 3 not mutually-exclusive types: protein changes - changes in coding sequence resulting in important modifications to the encoded proteins, regulatory changes in gene expression, and gene losses - loss-of-function mutations at key loci and gene deletions.

Here are some examples of potentially functionally important genetic changes of all types along the human lineage: FOXP2 - speech (protein evolution and gene regulation), ASPM and MCPH1 - brain size (protein evolution), genes involved in the nervous system (protein evolution and gene regulation), PDNY and HAR1 - brain development (gene regulation), CMAH and CASP12 - pathogen resistance (gene loss), KRTHAP1 - hair protein (gene loss), MYH16 - jaw musculature (gene loss). These are just a few genes that drive the evolution of the traits that make us human.

Protein-coding genes showing signature for adaptive evolution include those functionally involved with immune response, cell signaling, amino-acid metabolism, and olfaction. Several protein-coding genes involved in hearing (e.g. TECTA) show signatures of accelerated change in the human lineage. These loci may play a role in understanding spoken language.

Specific candidate genes involved in controlling brain size have been shown to have undergone strong positive selection along the great ape lineage and especially along the human lineage. ASPM and MCPH1 (microcephalin) are two of six loci associated with autosomal recessive primary microcephaly, a developmental defect in which the overall architecture of the brain is preserved, but its volume is reduced 3-fold to the size of early hominid brain. Both genes contain a number of human-specific amino acid changes. It has been suggested that these genes may have played a role in human encephalization.

Among the most notable examples of human-specific protein evolution is the transcription factor FOXP2, implicated in cognitive processes underlying speech and language. Mutations of FOXP2 are associated with an inherited speech disorder. Evolutionary analysis of FOXP2 identified two human-specific amino acid changes, and showed that this gene has been subject to strong adaptive evolution in humans since the divergence between humans and chimpanzees. This has led to the hypothesis that the evolution of FOXP2 gene may have contributed to the emergence of human language.

The functional significance of evolutionary changes in regulatory elements in various tissues across various primates can be analyzed by studying gene expression profiles using microarray analysis. These studies suggest that gene expression in the brain has been significantly upregulated during human evolution; and that humans and chimpanzees might differ in gene expression profiles as much, if not more, in the liver than in the brain.

Phylogenetic shadowing can be used, alternatively, to identify human or primate specific regulatory elements. Human accelerating regions, corresponding to regulatory regions, were identified as positively selected along the human lineage when compared with highly conserved regions identified in mouse, rat and chimpazee genomes. The most dramatic acceleration was found in a novel RNA gene, HAR1, which is specifically expressed in neurons of the developing human neocortex, during cortical neuron specification and migration, thus playing an important role in the development of important neurological pathways.

Gene regulatory evolution seems to be of most importance in structuring the human brain and more evidence comes from studies of a gene involved in perception and memory, PDNY, a precursor for many neuropeptides. The PDNY promoter region shows human specific mutations and a signature of strong positive selection.

Gene expression differences among species can also be due to differences in gene copy number. In a study of gene expression profiles in human tissues using Western blotting and immunofluorescence followed by a genome-wide survey of gene duplication, the most striking human-specific amplification has been found to be in DUF1220 protein domains, highly expressed in regions of the brain associated with higher cognitive function and having a neuron-specific expression pattern in the brain. Humans carry 212 copies of the gene encoding DUF1220 while chimpanzees have only 37 copies of the gene.

Gene losses are dramatic genetic changes having a great effect on phenotypes and fitness. Genes lost in the human lineage, since the chimpanzee-human split, include genes involved in taste perception and sense of smell. Human-specific loss-of-function mutations are also at genes involved in pathogen resistance. The gene CMAH, which produces a certain type of sialic acid - cell surface molecule involved in cell-cell interactions and pathogen binding, seems to have been inactivated (Alu-mediated inactivation) in humans. The gene CASP12, whose null version might confer resistance to infection in certain environments, has also become inactive in humans.

The inactivation of the gene MYH16, most prominently expressed in masticatory muscles in mammals, may have been resulted in smaller masticatory muscles and changes to cranio-facial morphology and expansion of the human brain case but contrary evidence exists.

The hair keratin protein gene KRTHAP1 is functional in chimpanzees and gorillas but was inactivated in humans. The degree to which this gene influences the hair phenotype is speculative.

Though the focus is on human specific gene losses, many known human genes are fully, or partially missing from the chimpanzee genome. This may be due to incomplete coverage of the chimpanzee genome or these might simply missing in the one chimpanzee whose genome was sequenced.

There is strong evidence that several genes have been inactivated or are degenerate on the chimpanzee Y chromosome (e.g. USP9Y, TNSB4Y) but remain conserved and functional on the human Y. Some of these genes play a role in spermatogenesis and the chimp-specific gene decay is perhaps driven by sperm competition.

Both chimpanzees and humans show intra-specific variation in the copy numbers of many genes. More variation among chimpanzees is needed. This caveat applies to all conclusions based on genomic comparisons, which generally fail to consider possible intra-specific variation.

Completion of more primate genomes (for use as a better outgroup in comparing the human and chimpanzee genomes), new insights on gene function, and improved

means of detecting signatures of selection are the prospects in molecular analyses of human evolution, ultimately providing the answer to: What makes us human?

References:

1. Bradley BJ 2008. Reconstructing phylogenies and phenotypes: a molecular view of human evolution. J Anat. 2008 Apr;212(4):337-53

2. Mekel-Bobrov N. and Lahn B. T. 2006. What makes us human: revisiting an age-old question in the genomic era. Journal of Biomedical Discovery and Collaboration 2006, 1:18

3. Mikkelsen TS, Hillier LW, Eichler EE, et al. (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437, 69ñ87.