Human Lineage-Specific Gene Inactivation

Wendy E Grus, *University of Michigan, Ann Arbor, Michigan, USA*
Jianzhi Zhang, *University of Michigan, Ann Arbor, Michigan, USA*

Advanced article

Article Contents

Online posting date: 20th September 2013

Pseudogenes are nonfunctional vestiges of genes. Investigating genes that were inactivated specifically on the human lineage or within humans can reveal the genetic basis of interspecies differences between humans and chimpanzees and interindividual differences within humans. It can also help understand the selective pressures that altered during human evolution or vary among human populations. Recent genome-wide surveys showed that human-specific pseudogenisation events occurred much more frequently to chemosensory and immune response genes than to other genes. Although non-deleterious pseudogenisation events were traditionally considered to be neutral, emerging evidence suggests that some of them are beneficial. One well-characterised case is the human *CASP12* locus, where the null allele decreases the incidence of sepsis and has apparently been selected for in human evolution. Exploration of roles of pseudogenisation in genome evolution and phenotype evolution has started.

Introduction

Pseudogenisation is an evolutionary phenomenon whereby a gene loses its function by disruption to its regulatory or coding sequence. Such loss of function is generally thought to be detrimental to an organism and selectively disadvantageous. Because pseudogenisation leads to immediate loss of gene function, it likely affects organisms to a greater extent than do most amino acid replacements. However, if a nonessential gene is pseudogenised, there may be little disadvantage and such pseudogenisation can be selectively neutral or nearly so. Olson (1999) proposed the 'less-is-more' hypothesis, suggesting that gene loss may serve as an engine of evolutionary change. This hypothesis is particularly intriguing for human evolution, as several human gene losses have been proposed to provide opportunities for adaptations and be responsible for human-specific phenotypes. Thus, studies identifying gene losses specific to the human lineage could help unravel the genetic basis of the biological differences between humans and their close primate relatives and determine how selection influenced these genetic changes. Furthermore, studying pseudogenisations that occurred within humans can reveal among-individual differences in the genetic makeup and their functional consequences. **See also**: Pseudogenes and Their Evolution

Currently, over 17 000 pseudogenes are estimated to be present in the human genome (http://www.pseudogene.org). Many of these are unprocessed pseudogenes, meaning that they contain remnants of the exon–intron structure of a functional gene but have insertion, deletion, or substitution mutations that eliminate their ability to produce a complete functional protein. These pseudogenes represent previous functional genes in the genome that have been lost during evolution. Humans share the majority of these unprocessed pseudogenes with our close primate relatives such as chimpanzee or gorilla, because many pseudogenisation events occurred in the common ancestor of these primates. However, pseudogenisation events that occurred along the human lineage since the human–chimpanzee split can divulge information regarding selection pressures and genetic changes unique to our species. **See also**: Genome Organization: Human

Pseudogenisation is an ongoing process in genome evolution, and many human genes have segregating pseudogene

eLS subject area: Evolution & Diversity of Life

How to cite:
Grus, Wendy E; and Zhang, Jianzhi (September 2013) Human Lineage-Specific Gene Inactivation. In: eLS. John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0020835.pub2

alleles, meaning that some human individuals contain functional alleles whereas others have nonfunctional alleles. The segregating pseudogenes allow for different individuals to have different sets of functional genes and thus may explain interindividual differences in phenotype. The emergence of population-level human variation databases such as 1000 Genomes (http://www.1000genomes.org) and NHLBI Exome Sequencing Project (http://evs.gs.washington.edu/EVS) allows for systematic identification of segregating pseudogenes within and between human populations. The presence of segregating pseudogenes in the human genome reveals that these gene inactivation events occurred relatively recently. In addition, they allow the use of population genetics to determine the evolutionary forces driving the spread of the pseudogenised allele in the population. **See also**: Genetic Variation: Human; Whole Genome Resequencing and 1000 Genomes Project

Human-specific pseudogenes, with putatively functional chimpanzee orthologues, have been identified with both systematic comparative genomic studies and in individual case studies. Both types of studies identified human-specific pseudogenes that explain some of our obvious phenotypic differences from our primate relatives. Systematic studies of fixed and segregating human-specific pseudogenes reveal an overwhelming overrepresentation of genes of chemosensory or immune response function that became pseudogenised in humans. **See also**: Chemosensory Systems; Divergence between the Human and Chimpanzee Genomes and its Impact on Protein and Transcriptome Evolution; Gene Deletions in Evolution; Primate Evolution: Gene Loss and Inactivation; Vertebrate Immune System: Evolution

Fixed Pseudogenes: Complete Gene Loss in the Human Lineage

To evaluate the 'less-is-more' hypothesis, several case studies and systematic studies identified putative human-specific pseudogenes, genes that are nonfunctional in humans but have a functional orthologue in chimpanzees (Rose and MacDonald, 1997; Chou *et al*., 1998; Tatnell *et al*., 1998; Angata *et al*., 2001; Winter *et al*., 2001; Hamann *et al*., 2003; Stedman *et al*., 2004; Hahn and Lee, 2005, 2006; Wang *et al*., 2006). With more primate genomes available for comparison and revised reference genome sequences for improved accuracy, Kim *et al*. (2010) investigated 121 pseudogenes reported as fixed, human-specific pseudogenes and found that only 38 were actually human-specific pseudogenes. The primary causes of misclassification were the presence of a closely related paralogue that led to miscalling, the presence of the pseudogenising variant in gorilla, orangutan or macaque or the removal of the pseudogene locus from the human genome assembly (Kim *et al*., 2010). The majority of the human-specific pseudogenes (25 of 38) are odourant receptors, and are discussed below.

Segregating Pseudogenes: Recent, Ongoing Gene Loss in Humans

Data from large-scale human sequencing projects such as 1000 Genomes and high-throughput genotyping technologies allow for investigations into the population distributions of segregating pseudogenes and the evolutionary forces behind such variations. In a large-scale analysis of nonsense single nucleotide polymorphisms (SNPs), Yngvadottir *et al*. (2009) identified 169 segregating pseudogenes and found that there was weak negative selection acting to remove the majority of nonsense SNPs. MacArthur *et al*. (2012) identified 1285 pseudogenising variants from 185 individuals. Overall, MacArthur *et al*. (2012) found that, compared with frequency-matched synonymous SNPs, pseudogenising variants were not significantly more common in regions exhibiting positive selection. However, both large-scale studies identified outlier pseudogenising SNPs that could be advantageous as suggested by high derived allele frequency, high population differentiation and/or inclusion in an extended haplotype or being found in regions exhibiting positive selection (Yngvadottir *et al*., 2009; MacArthur *et al*., 2012). Among these SNPs are the nonsense mutation in *CASP12*, described below, which had both high derived allele frequency and high population differentiation. Further population genetic analysis of the locus indicates that the pseudogenised allele has been selectively favoured over the functional allele (Wang *et al*., 2006). However, unlike the case of *CASP12* where the pseudogenised allele decreases the risk for severe sepsis, there is no documented physiological benefit gained by the pseudogenised allele in most segregating pseudogenes. For instance, Yngvadottir *et al*. (2009) identified a pseudogene allele of *MAGEE2* as an outlier nonsense SNP by high population differentiation, as the pseudogene allele was common in Asian and South-American populations but virtually absent in European and African populations. Further population genetic analysis revealed reduced diversity in haplotypes carrying the nonsense variant (**Figure 1**). Neutrality tests were consistent with positive selection driving the increase in the frequency of haplotypes containing the nonsense SNP (Yngvadottir *et al*., 2009). *MAGEE2* encodes a member of the melanoma antigen family of unknown function, making it difficult to discern the physiological effects of the pseudogenised allele.

Similarly, the fast skeletal muscle fibre protein α-actinin-3 (ACTN3) has a pseudogenised allele that has a frequency of 10–50% in different human populations. Homozygotes for the pseudogenised allele are over-represented in endurance runners, but not in sprinters or power athletes. Interestingly, mice that are deficient for Actn3 can run 33% further than wildtype mice (MacArthur *et al*., 2007). A population survey in humans indicates that selective sweeps for the pseudogenised allele occurred 15 000 and 33 000 years ago in European and Asian populations, respectively. Although MacArthur *et al*. (2007) suggested

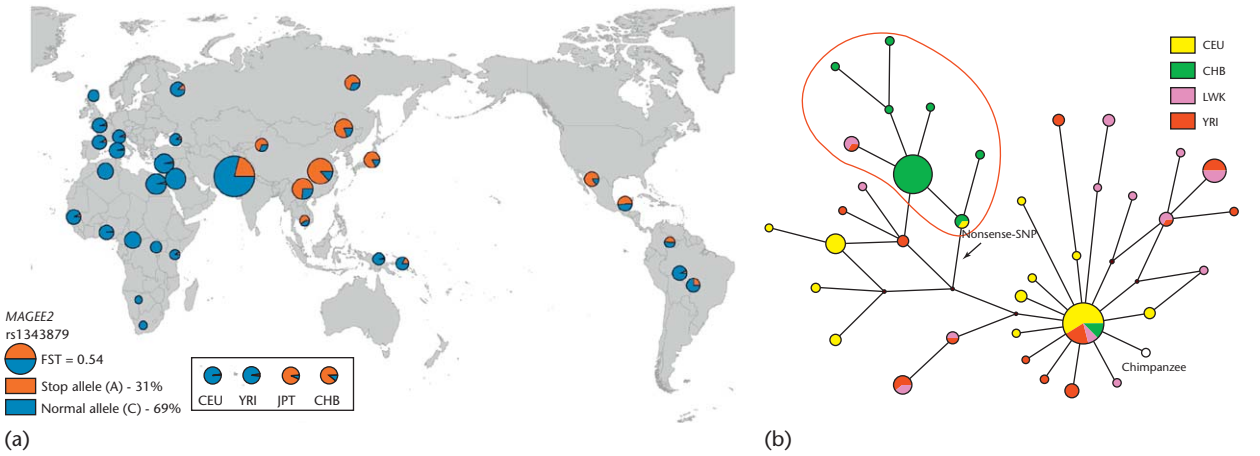

Figure 1 (a) Geographical distribution of pseudogene (orange) and intact (blue) *MAGEE2* alleles in different populations. Pie areas are proportional to sample sizes. (b) Median-joining network of inferred *MAGEE2* haplotypes reveal lower diversity levels in haplotypes containing the nonsense SNP rs1343879 (arrow). Circle areas are proportional to the haplotype frequency and are color coded according to population: CEU in yellow, CHB in green, LWK in pink, and YRI in red. Lines represent mutational steps between them (one or two steps, according to length). Used with permission from Yngvadottir *et al.* (2009). © Elsevier.

positive selection for more efficient aerobic muscle metabolism as the driver of these sweeps, the direct benefit of this physiological difference to humans remains unclear.

Another example of an increase in pseudogene allele frequency in human evolution is in the *SERPINA2* locus, which encodes a serpin peptidase inhibitor (Seixas *et al.*, 2007). A population genetic survey reveals that the frequency of a *SERPINA2* pseudogenised allele is 85% in Sao Tome, West Africa and 67.5% in Portugal. Linkage disequilibrium studies reveal decreased polymorphism in the pseudogenised allele haplotypes indicative of a selective sweep increasing their frequencies. However, the physiological benefit of this pseudogenised allele remains unclear. Based on the expression patterns of the functional alleles, this gene may be involved in reproduction. Consistent with this explanation, comparative genomic studies of proteases and protease inhibitors reveal many lineage-specific gene gains and losses in protein families including serine protease inhibitor (SERPIN) that could reflect differences in reproduction. Alternatively, the loss of *SERPINA2* might reflect differences in host–pathogen interactions, especially pathogens found in the reproductive tract. Interestingly, the *Serpina2* locus has been independently pseudogenised in the chimpanzee lineage, suggesting that similar selective pressures affect this locus in other primates. **See also**: Serpins: Evolution

Chemosensory Gene Loss in the Human Lineage

In a systematic identification of 67 human-specific pseudogenes, Wang *et al.* (2006) found that over half were odourant receptors and two were bitter-taste receptors. Among 169 segregating human-specific pseudogenes, olfactory receptors were significantly overrepresented (Yngvadottir *et al.*, 2009). Massive loss of odourant receptor genes appears to have also occurred in nonhuman Old World primates. The loss of these olfaction-related genes may reflect a shift from relying on olfactory cues as a means of communication to relying more on visual cues with the acquisition of a full trichromatic vision in the common ancestor of Old World primates (Zhang and Webb, 2003; Gilad *et al.*, 2004). **See also**: Olfaction; Olfactory Receptors

Segregating odourant receptor pseudogenes have been observed in humans, suggesting vast intrapopulation and interpopulation differences in human olfactory ability. For example, the human reference genome GRCh37/hg19 has 851 odourant receptor loci with 387 annotated as intact and 464 annotated as pseudogenes. Olender *et al.* (2012) investigated variation at these sites using 145 individuals from the 1000 Genomes data. Of the 387 'intact' genes, 218 loci had at least one pseudogenising variant due to frameshifting indel or nonsense mutation in at least one individual (Olender *et al.*, 2012). Of the 464 pseudogenes, 26 loci had an intact allele in at least one subject. Thus, in addition to 169 intact odourant receptors, the human odourant receptor repertoire consists of 438 fixed pseudogenes and 244 segregating pseudogenes (Olender *et al.*, 2012). These segregating pseudogenes resulted in a unique set of functional odourant receptor alleles in every subject with an average of 21 loci heterozygous for a pseudogene allele and 11 loci homozygous for pseudogene alleles in each individual (Olender *et al.*, 2012). On the one hand, among the 20 pseudogenising SNPs in regions exhibiting positive selection identified by MacArthur *et al.* (2012), 9 are in odourant receptor genes. On the other hand, they also determined that odourant receptors were enriched in a set

of tolerant genes that often contain homozygous pseudogenising variants (MacArthur *et al*., 2012). It remains unclear whether the interindividual and interpopulation differences in human functional odourant receptor repertoire are shaped entirely by random inactivation mutations and genetic drift or are in part due to differential natural selection in different geographic regions and environments.

Similar to odourant receptor gene loss, pseudogenisation of bitter taste receptors appears to have increased in the human lineage compared with other primate lineages. Since the last common ancestor of humans and chimpanzees, two bitter receptors have become fixed pseudogenes in the human lineage, but none in the chimpanzee lineage (Wang *et al*., 2004). Because bitter taste receptors are mainly used to detect toxic compounds in diet, loss of bitter receptors likely reflects a change in diet behaviour in human evolution such as the reduction in the consumption of plant tissues, which contain more toxins than animal tissues, and the use of fire for cooking and detoxification. However, some bitter receptors were also lost in other primate lineages and the causal relationship between diet and bitter receptor pseudogenisation is not always clear. In addition, similar to odourant receptors, not all gene loss in bitter receptors has been fixed. For example, two human segregating bitter receptor pseudogenes of low to moderate frequency of pseudogenised allele (2.3% and 25%) have been reported (Wang *et al*., 2004). These fixed and segregating pseudogenes reflect a relaxation from purifying selection on bitter receptors that occurred in human evolution. Although some studies have found positive selection acting at some human bitter receptor genes, the majority of the human bitter receptors appear to be evolving neutrally (Wang *et al*., 2004). Similarly, four pseudogenising SNPs were observed in components of the sour-taste receptor, including a frameshift variant in a region exhibiting positive selection (MacArthur *et al*., 2012). **See also**: Genetics of Taste Perception; Taste: Cellular Basis

Immune Response Gene Loss in the Human Lineage

Besides the overrepresented inactivation of chemosensory genes, human-specific pseudogenisation is frequent in genes that had a function in immune response and host defense. An additional 10 pseudogenes identified in Wang *et al*.'s (2006) systematic study had function in immune response or host defense. This overrepresentation is also observed in the individual case studies. For example, the human-specific inactivation of the gene encoding the enzyme cytidine monophosphate (CMP)-*N*-acetylneuraminic acid hydroxylase (CMAH) led to the deficiency of the mammalian common sialic acid *N*-glycolylneuraminic acid (Neu5Gc) on the human cell surface. This inactivation was due to an Alu-mediated sequence replacement that occurred approximately 2.7 Ma and may have had several important consequences to human biology and evolution (Chou *et al*., 1998, 2002) such as malaria susceptibility. Interestingly, parallel loss has occurred in the sialic acid binding immunoglobulin-like lectin (SIGLEC) gene family, and SIGLEC12 has been lost in the human lineage (Angata *et al*., 2001). Whereas a segregating nonsense SNP in *SIGLEC12* was identified as an outlier by Yngvadottir *et al*. (2009), a fixed loss-of-function missense variant is already known and common coding-disrupting variants in *SIGLEC12* are likely due to relaxation of selection (Angata *et al*., 2001; Mitra *et al*., 2011). In addition, one human-specific pseudogene is similar in sequence to the mannose-binding lectin genes, which mediate responses against invading microorganisms as part of the innate immune system (Wang *et al*., 2006). Another gene in this gene family, *MBL1*, is a pseudogene in all hominoid primates and a third gene, *MBL2*, is a segregating pseudogene in humans. Interestingly, mice deficient of *Mbl1* have increased survival rate for peritoneal sepsis, suggesting a possible selective agent for this gene loss in primates.

There are other immune function differences between humans and our primate relatives that are likely mediated through human-specific gene loss events. For example, the napsin B peptidase gene (*NAPSB*) lacks an in-frame stop codon and has a missense mutation at a crucial catalytic site, yielding no protein product despite abundant transcription in cell lines derived from immune cells (Tatnell *et al*., 1998). In addition, the haptoglobin gene cluster has undergone parallel gene deletion events during primate evolution that likely resulted in differences in immune response in humans and other primates (Puente *et al*., 2005). Humans and chimpanzees each have a unique two gene combination of an ancestral three gene locus. The gene lost in chimpanzees, *HPR*, could contribute to the chimpanzee's susceptibility to Trypanosome parasites, and the parallel loss of the *HPP* gene in humans likely has similar immune effects (Puente *et al*., 2005).

Yngvadottir *et al*. (2009) also highlighted a nonsense SNP in *CD36*, which encodes a thrombospondin receptor that acts as a receptor for *Plasmodium*-infected red blood cells (Love-Gregory *et al*., 2008). Individuals heterozygous for this variant have reduced susceptibility to malaria and metabolic syndrome that results in cardiovascular disease, although homozygous individuals suffering from complete *CD36* deficiency do not perceive these same physiological benefits (Love-Gregory *et al*., 2008). In addition to this common nonsense SNP in *CD36*, MacArthur *et al*. (2012) reported two novel splice-disrupting variants also found in African individuals, suggesting that selection may be driving the generation of multiple pseudogene alleles of the same gene.

Flavin-containing monooxygenase 2 (*FMO2*) has been almost completely pseudogenised, although the functional copy has a 4% frequency in African populations (Dolphin *et al*., 1998). This pseudogenising SNP of *FMO2* was identified as an outlier SNP because of high allele frequency and high population differentiation (Yngvadottir *et al*., 2009). This gene is the pulmonary member of the

FMO gene family that catalyses the oxidation of xenobiotics and drugs (Dolphin *et al.*, 1998). Similarly, two genes involved in drug metabolism, the putative *N*-acetyl transferase Camello-like 2 (*CML2*) and transporter 2, adenosine triphosphate (ATP)-binding cassette, subfamily B (*TAP2*), are segregating pseudogenes in humans (Hahn and Lee, 2006). Many other human-specific pseudogenes belong to gene families related to immune response or host defence such as T-cell receptors and immunoglobulins. This abundance of pseudogenes suggests a shift in pathogen infection and susceptibility in the human lineage since its divergence from the chimpanzee lineage, as reflected by the difference in susceptibility to human immunodeficiency virus (HIV)/simian immunodeficiency virus, malaria and other diseases between humans and their great ape relatives. In addition, interpopulation differences in the frequency of pseudogenised alleles for genes in these gene families suggest population differences in encounters with certain pathogens. **See also**: ATP-binding Cassette (ABC) Transporter Supergene Family: Genetics and Evolution; Immunology: Comparative Immunology of Mammals; Simian Retroviruses

As mentioned, the observation of a high-frequency pseudogene allele might indicate that the pseudogenisation is beneficial and the pseudogenised allele might be increasing in frequency by positive selection. The best example of this type is the pseudogenisation of *CASP12*. Outside of Africa, all human populations have a pseudogene allele of *CASP12*. However, in African and some Indian populations there is still a 10–20% incidence of the functional version of this gene. Interestingly, it was found that those with the functional alleles are more likely to suffer from severe sepsis than those with pseudogenised alleles. To investigate whether the pseudogene allele increased in frequency by positive selection for its ability to reduce severe sepsis or by chance, Wang *et al.* (2006) investigated the nucleotide sequence variation in functional and pseudogenised alleles. Based on decreased nucleotide variation in the pseudogenised allele near the pseudogenising mutation (**Figure 2**) and other population genetic evidence, the authors inferred that the spread of the pseudogenised allele was promoted by positive selection. They also estimated that *CASP12* became a pseudogene shortly before the out-of-Africa migration of modern humans. The functional human *CASP12* acts as a dominant-negative regulator of essential cellular responses including the nuclear factor-kappa B and interleukin-1 pathways; it attenuates the inflammatory and innate immune response to endotoxins. Because an appropriate level of immune response that is neither excessive nor

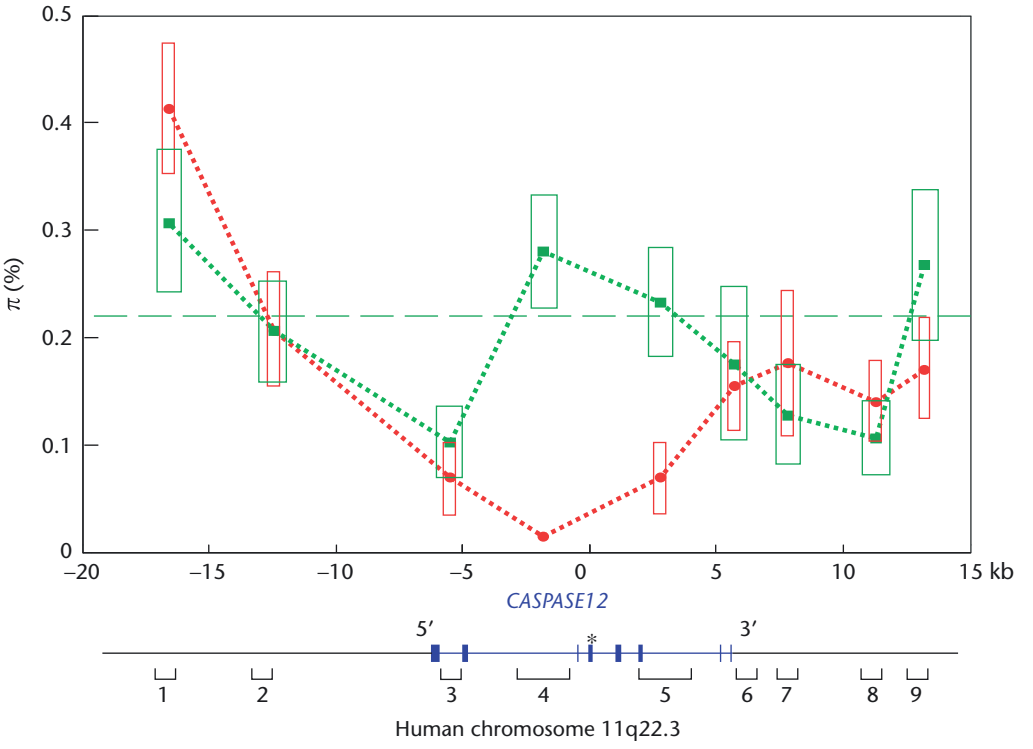

Figure 2 Nucleotide diversity in the *CASP12* gene in noncoding regions of the pseudogenised allele and functional allele. The exon structure for *CASP12* is shown in thick blue bars. Nucleotide diversity (π) is significantly lower near the disrupting mutation (*) in exon 4 of the *CASP12* pseudogenized allele (red) than is observed in the functional allele (green). Red circles and green dots show the percent nucleotide diversity in the pseudogenized and functional alleles, respectively, at each of nine noncoding regions within and around the gene (black bars). The rectangles surrounding the dots and squares show ± 1 standard error. The green dotted line is the average nucleotide diversity for the functional allele. The decreased diversity in the pseudogenised allele near the disrupting mutation is indicative of a selective sweep for this pseudogenized allele. Adapted from Wang *et al.* (2006).

insufficient is important to an organism, the immune suppression function of *CASP12* likely becomes harmful when the immune system cannot fully respond to a challenge. It is possible that during human evolution alterations in our genetic and/or environmental background resulted in a malfunction of the immune response to endotoxins, which rendered the previously necessary function of *CASP12* deleterious in humans and the pseudogenised allele advantageous over the functional one. **See also**: Cytokines as Mediators of Disease

Although the selective pressure for the spread of the *CASP12* pseudogenised allele has been discerned, not all such pseudogenes have an obvious selective advantage. For example, a pseudogenised allele for the chemokine receptor *CCR5* was thought to have increased its frequency by positive selection. The pseudogenised allele *CCR5Δ32* is present at relatively high frequencies in some European populations (up to 16%). Homozygotes for this pseudogenised allele are resistant to certain infectious diseases such as acquired immunodeficiency syndrome. Because HIV has been around for less than a century, its impact on the frequency of *CCR5Δ32* is minimal. Thus, it was hypothesised that resistance to diseases earlier in human history such as the bubonic plague or smallpox, led to selection of *CCR5Δ32*. However, further population genetic studies and allele frequencies for this gene from Bronze Age samples (long before *CCR5Δ32* was thought to have arisen) suggest that there was no positive selection promoting the spread of *CCR5Δ32* and that its relatively high frequency among Europeans was simply due to genetic drift (Sabeti *et al.*, 2005). **See also**: Human Immunodeficiency Virus (HIV) Infection: Genetics

Morphological Impacts of Gene Loss in the Human Lineage

Other human-specific pseudogenes help discern some of the physical differences between humans and chimpanzees. For example, the pseudogenisation of the sarcomeric myosin gene *MYH16* at the time of the emergence of the genus *Homo* is thought to be responsible for the marked size reduction in hominin masticatory muscles, which may have allowed the brain size expansion in *Homo* (Stedman *et al.*, 2004). However, a subsequent study gave a much earlier time for the pseudogenisation event, proving it unrelated to the brain size expansion (Perry *et al.*, 2005). Regardless of its effect on brain size, the pseudogenisation of *MYH16* probably resulted in distinct craniofacial muscle and skeletal structure in humans compared to other primates (**Figure 3**). In addition, human type I hair keratin is a pseudogene with functional orthologues in chimpanzee and gorilla (Winter *et al.*, 2001). Loss of this gene in the human may reflect the difference in body hair patterns between humans and their great ape relatives. **See also**: Keratins and Keratin Diseases; Myosin Superfamily

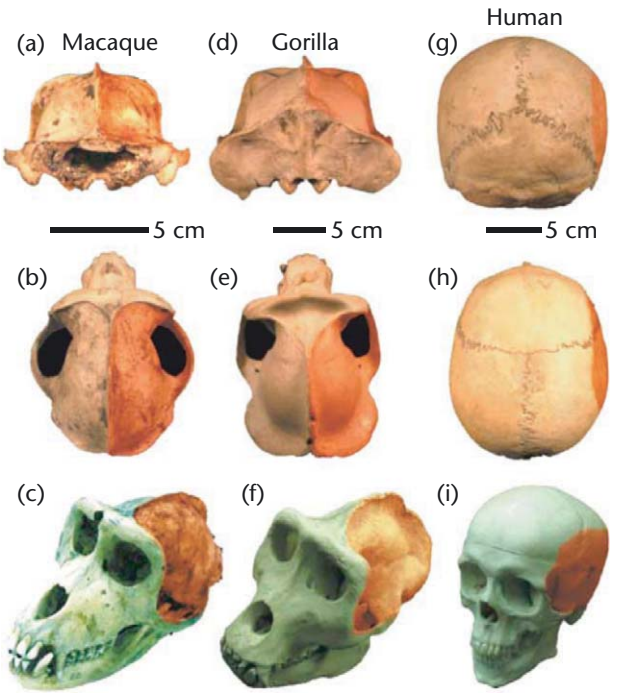

Figure 3 The relative size of chewing muscles differs greatly between humans and their primate relatives as a result of pseudogenization of human MYH16. The differences in musculature are reflected in the morphology of such craniofacial features as the temporal fossa and zygomatic arch (highlighted in red) in skulls from macaque (a–c), gorilla (d–f) and human (g–i). Adapted by permission from Macmillan Publishers Ltd., Stedman *et al.*, Copyright (2004).

Functional Implications of Gene Loss in the Human Lineage

In some cases, 'inactivating mutations' do not seem to result in loss of function. For example, the secretoglobin *SCGB1D4* has a frameshift mutation in the *C*-terminal end of the coding region that does not affect the functional domains of the protein, and the protein has been shown to be functional in experimental assays (Hahn and Lee, 2005). However, a mutation in the *C*-terminal region of a protein is not always benign. For example, a mutation in the stop codon of the natriuretic peptide (*NPPA*) gene increases the protein length by just two amino acids, but this lengthened form of NPPA is associated with increased risk of stroke (Rubattu *et al.*, 2004). Besides coding region changes, human-specific inactivating mutations have occurred in the regulatory regions. The promoter and enhancer regions of human pancreatic elastase 1 (*ELA1*) gene have acquired many mutations compared with other mammals and it is no longer expressed in the pancreas (Rose and MacDonald, 1997). However, the coding region of human *ELA1* remains intact, suggesting that this gene might be expressed elsewhere (Rose and MacDonald, 1997). In addition, human *ELA1* was found to be expressed in skin cells

(Talas *et al*., 2000). Because other ELA genes are expressed in the human pancreas, the functional consequences of this human-specific expression change are unknown (Talas *et al*., 2000).

Although the authors have highlighted selection for nonsense SNPs, most segregating pseudogene alleles are under negative selection as they contribute to rare Mendelian disorders or common complex diseases. Among the pseudogenising variants identified by MacArthur *et al*. (2012), 26 variants were known to be pathogenic in a homozygous state and an additional 21 were candidate disease-causing variants. In addition to being disease-causing, pseudogenising variants can affect disease treatment. This complication is exemplified by the inter-individual variation in cytochrome P450 (*CYP450*) genes. These genes are involved in the metabolism of xenotoxins in most vertebrates. Many human *CYP450* genes are segregating pseudogenes with variable frequencies of pseudogenised alleles in different populations. These population differences likely reflect the varied presence of different xenotoxins in different geographical regions. However, *CYP450* genes often aid in metabolism of modern drugs and the variation in CYP450 function results in a wide range of individual response to certain drugs (Lynch and Price, 2007; Zhou *et al*., 2009). **See also**: Cytochrome P450 (*CYP*) Gene Superfamily; Drug-Metabolising Enzymes: Genetic Polymorphisms

Conclusions

Gene loss that occurred specifically in the human lineage reveals a great deal about the evolutionary history of our lineage. The overrepresentation of human-specific pseudogenes of chemosensory and host-defense functions suggests that many of the differences between human and chimpanzee stem from our different environments. By contrast, segregating pseudogenes offer a more recent glimpse of human evolution. With the increased amount of polymorphism data generated from human population genomic studies, more functional variation will be identified in human populations. These differences may help identify the genetic mechanisms underlying the phenotypic variations among major ethnic groups.

Genetic differences between humans and their close primate relatives and those among human individuals have significant biomedical implications. First, these lineage-specific genetic changes present potential restrictions to using model organisms for understanding human biology. Second, the interindividual genetic differences resulting from segregating pseudogenes impact the efficacy of some medical treatments. This problem is beginning to be addressed by the fields of pharmacogenetics and personalised medicine and exemplifies the importance of understanding our past genes to maintain our future survival. **See also**: Pharmacogenetics and Pharmacogenomics

References

Angata T, Varki NM and Varki A (2001) A second uniquely human mutation affecting sialic acid biology. *Journal of Biological Chemistry* **276**: 40282–40287.

Chou HH, Hayakawa T, Diaz S *et al*. (2002) Inactivation of CMP-*N*-acetylneuraminic acid hydroxylase occurred prior to brain expansion during human evolution. *Proceedings of the National Academy of Sciences of the USA* **99**: 11736–11741.

Chou HH, Takematsu H, Diaz S *et al*. (1998) A mutation in human CMP-sialic acid hydroxylase occurred after the Homo–Pan divergence. *Proceedings of the National Academy of Sciences of the USA* **95**: 11751–11756.

Dolphin CT, Beckett DJ, Janmohamed A *et al*. (1998) The flavin-containing monooxygenase 2 gene (FMO2) of humans, but not of other primates, encodes a truncated, nonfunctional protein. *Journal of Biological Chemistry* **273**: 30599–30607.

Gilad Y, Przeworski M and Lancet D (2004) Loss of olfactory receptor genes coincides with the acquisition of full trichromatic vision in primates. *PLoS Biology* **2**: E5.

Hahn Y and Lee B (2005) Identification of nine human-specific frameshift mutations by comparative analysis of the human and the chimpanzee genome sequences. *Bioinformatics* **21**(Suppl. 1): i186–i194.

Hahn Y and Lee B (2006) Human-specific nonsense mutations identified by genome sequence comparisons. *Human Genetics* **119**: 169–178.

Hamann J, Kwakkenbos MJ, de Jong EC *et al*. (2003) Inactivation of the EGF-TM7 receptor EMR4 after the Pan–Homo divergence. *European Journal of Immunology* **33**: 1365–1371.

Kim H, Igawa T, Kawashima A *et al*. (2010) Divergence, demography and gene loss along the human lineage. *Philosophical Transactions of the Royal Society Series B Biological Sciences* **365**: 2451–2457.

Love-Gregory L, Sherva R, Sun L *et al*. (2008) Variants in the CD36 gene associate with the metabolic syndrome and high-density lipoprotein cholesterol. *Human Molecular Genetics* **17**: 1695–1704.

Lynch T and Price A (2007) The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. *American Family Physician* **76**: 391–396.

MacArthur DG, Balasubramanian S, Frankish A *et al*. (2012) A systematic survey of loss-of-function variants in human protein-coding genes. *Science* **335**: 823–828.

MacArthur DG, Seto JT, Raftery JM *et al*. (2007) Loss of ACTN3 gene function alters mouse muscle metabolism and shows evidence of positive selection in humans. *Nature Genetics* **39**: 1261–1265.

Mitra N, Banda K, Altheide TK *et al*. (2011) SIGLEC12, a human-specific segregating (pseudo)gene, encodes a signalling molecule expressed in prostate carcinomas. *Journal of Biological Chemistry* **286**: 23003–23011.

Olender T, Waszak SM, Viavant M *et al*. (2012) Personal receptor repertoires: olfaction as a model. *BMC Genomics* **13**: 414.

Olson MV (1999) When less is more: gene loss as an engine of evolutionary change. *American Journal of Human Genetics* **64**: 18–23.

Perry GH, Verrelli BC and Stone AC (2005) Comparative analyses reveal a complex history of molecular evolution for human MYH16. *Molecular Biology and Evolution* **22**: 379–382.

Puente XS, Gutierrez-Fernandez A, Ordonez GR, Hillier LW and Lopez-Otin C (2005) Comparative genomic analysis of human and chimpanzee proteases. *Genomics* **86**: 638–647.

Rose SD and MacDonald RJ (1997) Evolutionary silencing of the human elastase I gene (ELA1). *Human Molecular Genetics* **6**: 897–903.

Rubattu S, Stanzione R, Di Angelantonio E *et al*. (2004) Atrial natriuretic peptide gene polymorphisms and risk of ischemic stroke in humans. *Stroke* **35**: 814–818.

Sabeti PC, Walsh E, Schaffner SF *et al*. (2005) The case for selection at CCR5-Delta32. *PLoS Biology* **3**: e378.

Seixas S, Suriano G, Carvalho F *et al*. (2007) Sequence diversity at the proximal 14q32.1 SERPIN subcluster: evidence for natural selection favoring the pseudogenization of *SERPINA2*. *Molecular Biology and Evolution* **24**: 587–598.

Stedman HH, Kozyak BW, Nelson A *et al*. (2004) Myosin gene mutation correlates with anatomical changes in the human lineage. *Nature* **428**: 415–418.

Talas U, Dunlop J, Khalaf S, Leigh IM and Kelsell DP (2000) Human elastase 1: evidence for expression in the skin and the identification of a frequent frameshift polymorphism. *Journal of Investigative Dermatology* **114**: 165–170.

Tatnell PJ, Powell DJ, Hill J *et al*. (1998) Napsins: new human aspartic proteinases. Distinction between two closely related genes. *FEBS Letters* **441**: 43–48.

Wang X, Grus WE and Zhang J (2006) Gene losses during human origins. *PLoS Biology* **4**: e52.

Wang X, Thomas SD and Zhang J (2004) Relaxation of selective constraint and loss of function in the evolution of human bitter taste receptor genes. *Human Molecular Genetics* **13**: 2671–2678.

Winter H, Langbein L, Krawczak M *et al*. (2001) Human type I hair keratin pseudogene phihHaA has functional orthologs in the chimpanzee and gorilla: evidence for recent inactivation of the human gene after the Pan–Homo divergence. *Human Genetics* **108**: 37–42.

Yngvadottir B, Xue Y, Searle S *et al*. (2009) A genome-wide survey of the prevalence and evolutionary forces acting on human nonsense SNPs. *American Journal of Human Genetics* **84**: 224–234.

Zhang J and Webb DM (2003) Evolutionary deterioration of the vomeronasal pheromone transduction pathway in catarrhine primates. *Proceedings of the National Academy of Sciences of the USA* **100**: 8337–8341.

Zhou S, Liu J, Chowbay B (2009) Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metabolism Reviews* **41**: 89–295.

Further Reading

Balakirev ES and Ayala FJ (2003) Pseudogenes: are they 'junk' or functional DNA. *Annual Review of Genetics* **37**: 123–151.

Chimpanzee Sequencing and Analysis Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**: 69–87.

Clarke B (2006) Pharmacogenetics and cytochrome P450. *Biomedical Scientist* 38–40.

Gerstein M and Zheng D (2006) The real life of pseudogenes. *Scientific American* **295**: 48–55.

Menashe I, Man O, Lancet D and Gilad Y (2003) Different noses for different people. *Nature Genetics* **34**: 143–144.

Pink R, Wicks K, Caley D *et al*. (2011) Pseudogenes: pseudo-functional or key regulators in health and disease? *RNA* **17**: 792–798.

Poliseno L (2012). Pseudogenes: newly discovered players in human cancer. *Science Signaling* **5** re5.

Xue Y, Daly A, Yngvadottir B *et al*. (2006) Spread of an inactive form of caspase-12 in humans is due to recent positive selection. *American Journal of Human Genetics* **78**: 659–670.

Web Links

Comprehensive Pseudogene Database http://www.pseudogene.org accessed on 14 June 2013.

1000 Genomes Project http://www.1000genomes.org accessed on 14 June 2013.

NHLBI Exome Sequencing Project http://evs.gs.washington.edu/EVS accessed on 14 June 2013.