Disease consequences of human adaptation

Justin C. Fay

Washington University School of Medicine in St. Louis

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation
https://digitalcommons.wustl.edu/open_access_pubs/4773

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.
Disease consequences of human adaptation

Justin C. Fay

Department of Genetics, Washington University, St. Louis, MO, United States
Center for Genome Sciences and Systems Biology, Washington University, St. Louis, MO, United States

Abstract

Adaptive evolution has provided us with a unique set of characteristics that define us as humans, including morphological, physiological and cellular changes. Yet, natural selection provides no assurances that adaptation is without human health consequences; advantageous mutations will increase in frequency so long as there is a net gain in fitness. As such, the current incidence of human disease can depend on previous adaptations. Here, I review genome-wide and gene-specific studies in which adaptive evolution has played a role in shaping human genetic disease. In addition to the disease consequences of adaptive phenotypes, such as bipedal locomotion and resistance to certain pathogens, I review evidence that adaptive mutations have influenced the frequency of linked disease alleles through genetic hitchhiking. Taken together, the links between human adaptation and disease highlight the importance of their combined influence on functional variation within the human genome and offer opportunities to discover and characterize such variation.

© 2013 The Author. Published by Elsevier B.V.

Contents

1. Introduction ............................................................... 42
   1.1. Phenotypic divergence and disease consequences ........................................... 43
   1.2. Balancing selection and disease alleles ................................................ 43
   1.3. Hitchhiking and interference between advantageous and deleterious alleles ......................... 44
   1.4. Genes that have evolved under positive selection during human history . 45
2. Conclusions ................................................................. 45
Acknowledgments ............................................................... 45
References ............................................................... 45

1. Introduction

One of the initial motivations for sequencing the chimpanzee genome was to help define how humans are different from their primate relatives (Olson and Varki, 2003). Not only are human–chimpanzee genetic differences relevant to understanding human adaptations, but they are also relevant to understanding differences in the prevalence and susceptibility to disease (Varki, 2000; O’Brien et al., 2012). While documentation of disease in chimpanzees is scant in comparison to humans, especially for wild chimpanzees, a number of diseases are notably more common in humans (Varki, 2012). The incidence of such diseases cannot only be related to dietary or environmental differences, but can also be related to genetic changes that have occurred during evolution. For those genetic changes driven by positive “Darwinian” selection, there is no guarantee that adaptation is without health consequences, but only that there is a net gain in fitness, as measured by survival and reproduction.

With completion of the human and chimpanzee genomes and comprehensive surveys of human polymorphism, numerous genes and genomic regions exhibit evidence of positive selection along the human lineage (Sabeti et al., 2006; Aksey, 2009). Changes along the human lineage that bear evidence of positive selection include chromosomal rearrangements, duplications/deletions of genes, and point mutations in both coding and noncoding sequences (Kehrer-Sawatzki and Cooper, 2007; O’Brien et al., 2012). Evidence of adaptive evolution is generally based on a pattern of non-neutral molecular evolution indicative of
positive selection or a unique evolutionary history in humans (Fay and Wu, 2003; Harris and Meyer, 2006). In most instances, evolutionary changes within such genes have not been associated with any pheno-
typic consequence. Even so, many genes that have been influenced by positive selection are known to contribute to human disease (Bakewell et al., 2007), and a composite analysis of 27 studies of positively selected genes found an enrichment for genes expressed in the central nervous system and with synapse-related functions (Huang et al., 2013). Furthermore, positive selected genes are often relevant to understanding disease susceptibility and its genetic basis (Crespi, 2010; Quintana-Murci and Barreiro, 2010). Of particular relevance are those human diseases which are not well recapitulated in animal models, e.g. normal parturition and preterm birth where a fall in circulating levels of the pregnancy hormone progesterone is responsible for the initiation of labor in mice but not humans (Ratajczak et al., 2010).

In this review, I describe various mechanisms by which adaptation is known to influence human disease. First, I describe adaptations where-
by new or substantially altered phenotypes encumber us with new or an increased disease burden. Then I cover cases of balancing selection whereby changes in a gene can provide a benefit under some circumstances but impart a health cost under others. I also describe evidence for linkage effects whereby adaptive mutations can increase the fre-
quency of disease mutations through hitchhiking. Finally, I describe genes that have evolved under positive selection along the human line-
age and the impact of these changes on human health. Taken together, there are an increasing number of examples that point to a direct relationship between adaptation and disease, highlighting the value of incorporating adaptive evolution into our understanding of human disease.

1.1. Phenotypic divergence and disease consequences

Adaptive phenotypes can often be directly linked to common human health problems. One of the most dramatic differences between humans and other primates is bipedal locomotion. Bipedalism requires upright posture and results in an increased pressure on various structures that can give rise to a number of health problems, such as hernias, hemorrhoids, varicose veins, and back, hip and knee problems, e.g. osteoar-thritis (Jurmain, 2000; Varki, 2012). Another major difference lies in humans’ increased mental faculty, which may be linked to the preva-
ience of mental disorders, such as schizophrenia (Horrobin, 1998) and Alzheimer’s disease (Varki, 2000). While the incidence of major psycho-
ses, depression, phobias, obsessive–compulsive disorder, and mental retardation is difficult or impossible to compare with other species, the high incidence of mental disease compared to other diseases has generated considerable interest in various evolutionary explanations (Keller and Miller, 2006).

Perhaps one of the most clear health consequences of human adaptation is our cephalopelvic disproportion. The combination of our larger brain size with a more narrow pelvis, which facilitates biped-
al locomotion, has greatly complicated labor and delivery (Rosenberg and Trevathan, 2002). While the chimpanzee neonatal cranium can eas-
ily pass through the inlet, midplane and outlet of the pelvis (Fig. 1). In modern humans, unassisted childbirth is exceedingly rare, and with advent of modern medicine mortality rates associated with childbirth are currently 40–
50 times lower than even 70 years ago (Loudon, 2000). The dramatic cephalopelvic changes that have occurred during human evolution in comparison to changes in gestation length is one line of evidence supporting the hypothesis that humans have evolved to give birth earlier than other primates, at least on a developmental timescale (Montagu, 1961; Plunkett et al., 2011). The abundance of altricial characters in human neonates compared to neonates of other pri-
mates, such as skull development (Penin et al., 2002), emergence of teeth (Holly Smith et al., 1994) and vision (Booth et al., 1985), lends further support to this hypothesis. Being born early, and especially too early, might alleviate the cephalopelvic disproportion, but also raises a different set of health challenges faced by neonates, e.g. ear and sinus infections (Bluestone, 2005; Behrman and Butler, 2006).

1.2. Balancing selection and disease alleles

A number of human disease alleles have been associated with a fitness advantage under certain circumstances. In such cases, the disease allele can become more common than in the absence of the fitness advantage. Many examples are associated with resistance to infection. The classic example is mutations in HBB, which in heterozy-
gous form confer resistance to malaria but as homozygotes cause sickle-cell anemia. Resistance to malaria is also associated with muta-
tions that cause G6PD deficiency, thalassemia (HBA and HBB) and other erythrocyte defects (SLC4A1 and DARC) (Kwiatkowski, 2005). Because many other loci have been associated with resistance to malaria (Driss et al., 2011), the strong selective pressure for malaria resistance may have influenced other, as yet unknown, disease alleles.

Another compelling though complex example of balancing the pos-
tive and negative effects of human polymorphism can be found at the Major Histocompatibility Complex (MHC). The MHC locus encodes cell surface glycoproteins important to the immune system’s fight against infection and is one of the most variable regions within the human genome. Both population genetic and functional studies indicate that the high levels of diversity at the MHC locus are maintained by balancing selection (Hughes and Yeager, 1998). Coincident with one of the strongest signals of balancing in the human genome, the MHC locus also contains the largest number of disease associations within the human genome (de Bakker and Raychaudhuri, 2012). Many of the associates are related to inflammatory and autoimmune diseases, such as multiple sclerosis, rheumatoid arthritis, psoriasis, and celiac dis-
ease. Thus, there are potentially many alleles that increase resistance to certain pathogens but increased susceptibility to other pathogens or increased risk of disease unrelated to infection. Although linkage disequilibrium makes it difficult to tease apart genotype–phenotype associations as well as pinpoint the evolutionary signals of selection, the historical influence of selection at the MHC locus is compelling and likely includes both direct effects (directional or balancing selection) as well as
indirect effects (hitchhiking, as described in Section 1.3) on disease alleles (van Oosterhout, 2009; Shiina et al., 2006).

Although many other examples of balancing selection exist, e.g. Leffler et al. (2013), their association with disease alleles is not often known. However, a number of examples have accumulated. These examples include Celiac disease and bacterial infection at SH2B3 (Zhernakova et al., 2010), kidney disease and resistance to African sleeping sickness at APOL1 (Rosset et al., 2011), Crohn’s disease and selection at NOD2 and other loci (Nakagome et al., 2012; Cagliani et al., 2013), and either direct or indirect selection on a hemochromatosis allele at HFE (Distante et al., 2004). Even when phenotype associations are present, it can be difficult to know how balancing selection has led to an increased frequency of disease alleles. For example, a 32-bp deletion in CCR5 is associated with resistance to HIV infection (Dean et al., 1996), but susceptibility to West Nile Virus infection (Glass et al., 2006). While CCR5 also exhibits a strong signal of balancing selection and the CCR5 deletion allele has rapidly increased in frequency (Bamshad et al., 2002), the age of the deletion predates the emergence of HIV resistance as an important component of fitness (Novembre et al., 2005), suggesting a role for some other selective agent, such as smallpox (Galvani and Novembre, 2005).

1.3. Hitchhiking and interference between advantageous and deleterious alleles

As populations evolve, natural selection strives to increase the frequency of advantageous mutations and decrease the frequency of deleterious mutations (Fig. 2). Because of linkage, there are numerous opportunities for interferences between advantageous and deleterious mutations (Hill and Robertson, 1966). Thus, a strongly advantageous mutation has the potential to increase the frequency of linked deleterious mutations (Fig. 2). While the frequency of interference between advantageous and deleterious mutations is not known, recent work suggests that interference is common enough to have influenced disease alleles in humans.

Numerous regions in the human genome have been found to exhibit evidence of positive selection in the recent past. In a review of 18 genome scans for selection, 5110 regions covering 14% of the genome were found in one or more scans, 14% of which were found in more than one study (Akey, 2009). Furthermore, the composite hitchhiking signal based on all amino acid substitutions fixed along the human lineage suggests that many of these substitutions generated hitchhiking effects over small genomic regions, ~100 kb (Hernandez et al., 2011). The strength of positive selection is important because it determines the size of the region influenced by hitchhiking and whether it can overwhelm the influence of selection against any linked deleterious alleles (Hartfield and Otto, 2011). As such, strongly advantageous mutations have the potential to increase the frequency of many deleterious alleles.

While hitchhiking may only influence certain genomic regions, deleterious mutations are a pervasive feature of all functional elements within the human genome. Estimates of the number of deleterious mutations range from 200 to 800 per individual, based on evolutionary criteria (Fay et al., 2001; Chun and Fay, 2009; Abecasis et al., 2012).

The extent to which there is interference between positive and negative selection depends on the rate of recombination. When recombination is low or absent, the possibility of interference is increased. As expected under a model of interference, the abundance of deleterious mutations is enriched relative to neutral variation in regions of low recombination in humans (Chun and Fay, 2011). While background selection against deleterious mutations can also explain this observation, an enrichment of deleterious to neutral alleles was also found in regions of high recombination baring signatures of hitchhiking (Chun and Fay, 2011). In addition to alleles predicted to be deleterious, hitchhiking regions are also enriched for clinical variants influencing auto-immune, energy metabolism and mental, neurological or neurodevelopmental disorders. However, there is no association between GWAS hits and hitchhiking regions (Hindorff et al., 2009; Chun and Fay, 2011). The absence of an association with GWAS hits may be due to lower power to detect associations in regions with little variation remaining after hitchhiking, neutrality of many GWAS variants (Dudley et al., 2012), or some form of heterogeneity. For example, some associations with hitchhiking have been found for subsets of GWAS data. There is an association between positive selection and GWAS hits within conserved gene clusters (Preuss et al., 2012). Also, positive selection is associated with susceptibility alleles for type 1 diabetes, but protective alleles for Crohn’s disease (Corona et al., 2010). In another study, the risk alleles at inflammatory disease loci were found associated with positive selection and eQTL (Raj et al., 2013). However, some of these associations may not be related to linkage effects but rather direct positive effects of the disease alleles. For example, the increase in frequency of type 1 diabetes susceptibility alleles may be caused by historical periods of famine, as held by the thrifty genotype hypothesis (Neel, 1962).

In addition to genome-wide studies, there are a number of disease alleles at specific genes that are thought to have been influenced by
positive selection via hitchhiking. A 250 kb haplotype associated with inflammatory bowel disease is associated with an allele of OCTN1 that has increased in frequency due to hitchhiking and provides increased absorption of ergothioneine (Huff et al., 2012). A pleiotropic 1.6 Mb haplotype associated with a number of common diseases has increased in frequency in European populations consistent with positive selection (Soranzo et al., 2009). As a final example, a common inversion polymorphism has spread to high frequency in Europeans and has been associated with a number of diseases (Stefansson et al., 2005; Steinberg et al., 2012). The increased risk of microdeletions associated with the inverted haplotype illustrates another mechanism by which selection can influence disease incidence: changes in mutation rate. While not covered in this review, there are excellent examples of an increased rate of deletions or other types of mutations associated with prior segmental duplications or chromosomal changes (Mefford and Eichler, 2009; Stankiewicz and Lupski, 2010), some of which exhibit evidence for positive selection (Samonte and Eichler, 2002; Gokcumen et al., 2011). Further evidence for an increased mutational burden can be found in a number of human-specific genes which are known to cause disease when mutated (Cooper and Kehrer-Sawatzki, 2011).

1.4. Genes that have evolved under positive selection during human history

In contrast to the clear consequences of many adaptive phenotypes, the health consequences of genetic changes that have been fixed by positive selection are more difficult to come by. This can largely be attributed to the difficulty of linking signals of positive selection, of which there are many, to adaptive and disease phenotypes. Even when a phenotype is suspected, our ability to determine whether certain substitutions between species increase our susceptibility to disease is often limited. There are, however, a few striking examples that illustrate the potential importance of historical adaptations to current health problems along with a number of functional enrichments found present within rapidly evolving gene sets.

With the completion of the chimpanzee genome, numerous statistical tests of neutrality have been used to detect coding or noncoding regions that have evolved under positive selection along the human lineage. While the signal of selection can be difficult to distinguish from relaxed constraint and even neutral evolution, a number of gene ontology enrichments have been found. The most common classes of genes are those involved in sensory perception, immunity, and reproduction (Clark et al., 2003; Nielsen et al., 2005; Kosiol et al., 2008). However, positively selected genes have also been found to associate with schizophrenia and other psychiatric disorders (Crespi et al., 2007; Moalic et al., 2010), and apoptosis (da Fonseca et al., 2010). However, many of these same categories are enriched in genes rapidly evolving along the chimpanzee lineage and so may represent a general class of rapidly evolving proteins rather than classes specific to humans (Arbiza et al., 2006).

In the case of noncoding sequences, rapidly evolving regions have been found to be enriched nearby genes that function in cell adhesion (Prabhakar et al., 2006).

Some of the best characterized genes are those involved in viral attenuation, which exhibit strong signals of selection and for which functional studies have been completed. For example, adaptive evolution within TRIM5α can in part explain human-specific susceptibility to HIV infection (Sawyer et al., 2005). TRIM5α recognizes incoming retroviral capsids and targets them for destruction (Stremlau et al., 2006). However, the human allele of TRIM5α does not protect against HIV infection whereas that of macaques provides resistance to infection (Keckesova et al., 2004). Because the human allele of TRIM5α is able to restrict an endogenous retrovirus found in the chimpanzee and gorilla genomes but not the human genome, it is likely that TRIM5α evolved under positive selection to restrict the activity of this retrovirus in the ancestral human genome (Kaiser et al., 2007). The consequences of an arms race between viral evasion and host infection are not limited to TRIM5α and are also found in a handful of other genes that confer species-specific differences in viral protection (Daugherty and Malik, 2012).

Another compelling case involves genes that function in sialic acid biology (Varki, 2010). Sialic acids are a family of monosaccharides that function in self recognition and pathogen infection. The sialic acid N-glycolyneuraminic acid is absent in human blood, being replaced by N-acetylglycosaminic acid. While it is hard to know whether the gene responsible for this difference, CMAH (Irie et al., 1998), evolved under positive selection, this change is likely to have mediated escape from certain pathogens, such as malarial parasites, but susceptibility to others (Varki and Gagneux, 2009). Another likely consequence of loss of CMAH activity was a hyperimmune state that leads to multiple changes in other sialic acid binding proteins (Varki, 2010), some of which have evolved rapidly in humans (Angata et al., 2004).

More often than not, links between human adaptive changes and disease are as yet suggestive. For example, changes in genes responsible for the timing of labor are thought to have been under positive selection along the human lineage, making us particularly susceptible to preterm birth (Plunkett et al., 2011). While a number of genes associated with preterm birth have evolved rapidly in humans (Chen et al., 2008; Plunkett et al., 2010; Plunkett et al., 2011), the consequences of such changes are not easily measured. Intriguingly, one of the rapidly evolving genes is the progesterone receptor, which might contribute to humans' lack of response to progesterone therapy compared to mice (Muglia and Katz, 2010). Rapid evolution has also been found in both the coding and a nearby noncoding region of FSHR (Plunkett et al., 2011). However, FSHR primarily functions in the establishment of pregnancy, rendering its association with preterm birth not only intriguing but also enigmatic.

2. Conclusions

Adaptive evolution is rarely free from genetic, functional, development and selective constraints (Arnold, 1992). In the limited evolutionary trajectories to higher fitness, some traits may be compromised at the expense of an overall improvement in fitness. Inherent to this scenario is the idea that trade-offs are common to fitness associated traits (Stearns, 1989). However, linkage can also impose costs to adaptation whereby advantageous and deleterious mutations can only be disentangled by recombination. Regardless of how the cost arises, the negative consequences of many adaptations may be ameliorated over time, through subsequent evolutionary changes (Denver et al., 2010).

In this review, I highlighted recent progress in understanding how adaptive evolution during human history has incurred some cost to the health of modern humans. The examples used have various levels of support for the association between human adaptation and disease. More importantly, they illustrate the variety of mechanisms and types of supporting evidence used to connect current human health with our evolutionary past. Finally, it is important to recognize that only a subset of human diseases has been influenced by our adaptive history. In this regard, understanding the molecular basis of adaptation may provide insight into certain disease mechanisms.

Acknowledgments

This work was supported in part by a National Institutes of Health grant GM098896 to J.C.F.

References


