43-44 Selection in humans

- Natural selection has and continues to affect human evolution. There are several ways to study selection in humans.
- One was is to direct estimate differences in survival rates. In lethal genetic diseases, the relative fitness of the affected genotype is 0. Modern medicine can dramatically increase the survival rate. Until the development of antibiotics, the average age of death of a CF child was 3 years. Until a special diet for PKU children was developed, they rarely survived to adulthood and almost never had children.
- It takes very large studies to detect less drastic selection. The S allele of the beta globin locus is in high frequency in areas with a high prevalence of malaria in Africa and Asia, suggesting in the 1940s that it provided some protection against malaria. An early study found that in a group of 30,923 adults in West Africa typed for the A and S alleles of the β-globin gene, 25,374 AA, 5482 AS, and 67 SS adults were found. The genotype frequencies are 0.821, 0.177 and 0.002 and allele frequencies are 0.909 and 0.091. The HW frequencies are 0.826, 0.165, 0.008. If the newborns were in HW frequencies and all the difference between HW frequencies and the observed frequencies is caused by natural selection, then we can estimate the selection coefficients. You do this by assuming that the allele frequencies in the newborns are the same as in the adults and that there is no mortality of the AS individuals. You do not know the number n of newborns before some died of malaria, but you can estimate that number. For the heterozygotes, $2pqn$=5482 so n=5482/(2x0.909x0.091)≈33,163. Of those newborns, p^2=0.909^2=0.8263 of them of had genotype AA. Therefore there were 0.8263x33,163≈27,402 AA newborns, of whom only 25,374 survived. The survival rate is 25,374/27,402=0.926. The selection coefficient, s, against AA individuals is 1–0.926=0.074. For the SS individuals, you should convince yourself that the survival rate is 0.244 and the selection coefficient, t, against SS individuals is 0.756. With these estimates of s and t, you predict that the equilibrium frequency of S is $s/(s+t)$≈0.089. The actual frequency in that area was somewhat higher. This calculation does not take reproductive compensation into account.
- Several other loci are affected by selection caused by malaria. Several mutations at the alpha- and beta-globin loci reduce the abundance of hemoglobin and cause thalassemia, a form of anemia in regions that have or had a high incidence of malaria, particularly in Mediterranean countries and southern Arabia and India. The carrier frequency is as high as 20%.
- G6PD is another gene (X-linked) with mutants in high frequency in malarial regions. The mechanism seems to be a reduced concentration of nonprotein glutathione (GSH) which is required by the malarial parasite. Some heterozygous females and hemizygous males suffer from favism, which is anemia and jaundice caused by the consumption of fava beans and some other foods.
- Another kind of evidence of selection comes from the observation the an allele in high frequency is associated with a relatively long haplotype, much longer than is associated with other alleles at the same locus. The idea is that if an allele has increased from a single copy in a short time, the ancestral haplotype would not have had much time to break down because of recombination. A rapid increase in allele frequency was probably caused by selection. One of many examples is A– allele of

G6PD, which probably increased from a single mutant in the past 3000-5000 years with a selective advantage of 10%-20%.

- There are other examples of the same type. Most adult humans cannot digest lactose because the gene LCT (also called LPH) is no longer transcribed. In a few groups, Europeans and some east Africans, continue to transcribe LCT and can digest milk as adults. The trait is called **lactase persistence** (LP). LP in Europeans is caused by one two mutations in the LCT gene, T-13910 A-22018, which are 14 kb and 22 kb upstream from LCT. There is LD associated over a region of 1 mb surrounding T-13910, which indicates that it arose by mutation ~5000-1000 years ago, roughly at the time the cattle, sheep and goats were domesticated. These two alleles are not present in African populations with high prevalence of LP. Instead, there are three other SNPs upstream of LCT, that associated with LP in three different African group. All three show signs of recent strong selection in the past 10,000 years. The fact that these mutations are different from the ones in Europe supports the idea that domestication of cattle occurred independently in Africa and Eurasia.
- Another example is SLC24A5, a gene associated with pigment differences both in zebra fish and humans. The allele A111T is found in high frequency in Europeans and is absent from other populations. It is associated with lighter hair and skin. The extreme difference in allele frequency among populations suggested that it had been favored by selection in Europeans.
- In addition to gene specific studies, there have been large scale surveys of human genomic data has that show evidence of selection in regions not previously known to be selected. The idea is to scan for unusually long haplotypes and then see whether there are large differences among populations in SNPs found in those haplotypes. One such study identified 6 regions that showed strong evidence of selection. Two of them, LCT and SLC24A5 were already known. Two others were genes in the Yoruban population that are known to play a role in the response to Lassa fever. Two others were in the Asian populations at loci known to play a role in hair development.
- SNPs are not the only variants that have been subject to selection in humans. Amylase 1 (AMY1) codes for amylase in saliva. Amylase is important for the digestion of starch. Different human populations differ in the amount of starch in the diet. Differences among people in the level of amylase in saliva are due to differences in the number of copies of AMY1. People from populations that have high starch diets have more copies of AMY1 than people from populations with low starch diets.

Saunders MA et al. (2005) The span of linkage disequilibrium caused by selection on G6PD in humans. Genetics 171:1219-1229
http://www.genetics.org/cgi/content/abstract/171/3/1219

Tishkoff SA et al. (2007) Convergent adaptation of human lactase persistence in Africa and Europe. Nature Genetics 39:31-40
http://www.nature.com/ng/journal/v39/n1/abs/ng1946.html

Lamason RL et al. (2005) SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science 310:1782-1786
http://www.sciencemag.org/cgi/content/abstract/310/5755/1782

Sabeti PC et al. (2007) Genome-wide detection and characterization of positive selection in human populations. Nature 449:913-918
http://www.nature.com/nature/journal/v449/n7164/full/nature06250.html

Perry GH et al. (2007) Diet and the evolution of human amylase gene copy number variation. Nature Genetics 39:1256-1260
http://www.nature.com/ng/journal/v39/n10/abs/ng2123.html

Problems

43.1 You sample a group of 900 adults at the beta-globin gene and find 560 AA, 320 AS and 20 SS individuals. Assume the frequency of A is at equilibrium under selection and that the newborns are in their HW frequencies, and estimate the survival rate of AA and SS individuals relative to AS individuals.

Ans. The allele frequencies are p=0.8 and q=0.2. You can assume no mortality, so you would expect to see 320 AS individuals in a sample of 1000 at HW. Therefore, there were 640 AA newborns, 560 of whom survived. The survival rate is 560/640=7/8 and s=1/8. There were 40 newborns and 20 survived, so the survival rate is ½ and so is t. Note that $s/(s+t)$=0.2.

43.2 a. What evidence would you look for to show that lactase persistence is favored by natural selection in group in which adults regularly consume milk?

Ans. You would test for differential survival of individuals possessing and lacking lactase persistence.

b. Do you think you would actually be able to detect survival differences?

Ans. It would depend on how important a part milk is to the adult diet. It could easily be that there is only a slight difference in survival rate or none at all, except during a famine.