

***Forces causing gene frequency change***

- Random mating does not cause allele frequencies to change, but other forces do. Mutation creates new alleles but mutation rates are so low that that mutation has little effect on the frequencies of alleles already present in a population. Alleles frequencies change because of the combined effects of natural selection, genetic drift, gene flow and recombination.

***Mutation***

- *A mutation is the result of an error in DNA replication:* A change in a single nucleotide is called a point mutation. In higher plants and animals, the probability of a point mutation (i. e. the mutation rate) is very low, close to 1 in a billion ( $10^{-9}$ ). Rates of point mutations are higher in bacteria and higher still in viruses.
- *Other types of mutations occur:* small or large pieces of chromosomes can be deleted or duplicated. A duplication can create a second copy of a gene.
- *Whole chromosomes can be duplicated:* In humans, individuals born with 3 copies of chromosome 21 (trisomy 21) have Down's syndrome. The risk of trisomy 21 in the US population is between 1/650 and 1/1000; the risk increases with maternal age<sup>1</sup>.
- *Whole genomes can be duplicated:* If there is no reduction division during meiosis, diploid gametes are produced. A diploid gamete combined with a haploid gamete creates a triploid zygote. Bananas and many other domesticated plants are triploid.

***Natural selection***

- *Fitness:* If individuals with different genotypes differ in their chances of survival and reproduction, then there are differences in fitness that cause allele frequencies to change.

Average fitnesses of different genotypes may depend on the environment. For example, individuals with defective alleles of the PAH gene who eat a normal diet die at an early age because of the accumulation of phenylalanine. They suffer from phenylketonuria (PKU). If phenylalanine is removed from the diet, there is almost no reduction in survival rate.

---

<sup>1</sup> Information about trisomy 21 and other inherited diseases and conditions in humans can be found from OMIM (Online Mendelian Inheritance in Man), <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>.

- *Directional selection:* Directional selection occurs when one allele results in higher rates of survival and reproduction. For example, if there are two alleles A and a, and AA individuals have a higher fitness than Aa individuals who have a higher fitness than aa individuals, then A is the advantageous allele and a is the deleterious allele.

If a population initially contains only aa individuals and an advantageous allele A is created by mutation, then the frequency of A will increase every generation because of natural selection. Eventually A will be substituted for a.

In the 1920s, population geneticists showed that natural selection can cause the substitution of an advantageous allele in a relatively few generations even if fitness differences are small. They concluded that even very weak natural selection acting on Mendelian alleles could cause changes in allele frequencies in a few hundred or a few thousand generations. Although that may be a long time from a human perspective, it is a short time compared to the long times available for evolution to occur.

An example of directional selection in humans is on alleles conferring lactose tolerance in adults. Of other mammals, only cats can digest lactose as adults. Most adults of European ancestry and some adults of east African ancestry can digest lactose because the gene coding for the enzyme lactase-phlorizin hydrolase (LPH) is expressed in adults. One mutation that enhances production of LPH in adults is responsible in Europeans and three different mutations are responsible in east Africans. All four mutations have reached high frequency in these populations because of directional selection in roughly the last 7000 years.<sup>2</sup>

Past episodes of directional selection in humans may contribute to current health problems. Native Americans and some other groups are at a relatively high risk for type II diabetes (non-insulin-dependent diabetes) as adults. One explanation for this high risk is the “thrifty gene” hypothesis: ancestors of Native Americans experienced repeated famines that selected for efficient metabolism which causes type II diabetes in individuals who eat a modern diet. Some kinds of hypertension (high blood pressure) are caused by a tendency to retain too much salt, which possibly resulted from selection for salt retention during periods of low salt abundance. Although the thrifty-gene and salt-retention theories are plausible, they are not proved.

- *Purifying selection:* If a population contains only AA individuals and then a deleterious allele a is created by mutation, natural selection will tend to eliminate it. This is purifying selection.

Although mutation rates are very low, there are so many genes (more than 30,000 in humans) that deleterious alleles are created by mutation every generation.

A balance between new deleterious mutations and purifying selection acting to eliminate them will be reached.

---

<sup>2</sup> See <http://www.nature.com/ng/journal/v39/n1/full/ng1946.html> for details.

For many genetic diseases in humans, including PKU, allele frequencies are determined by mutation opposed by purifying selection, called mutation-selection balance.

- *Balancing selection*: If heterozygous individuals have a higher fitness than either homozygote, selection will maintain both alleles in the population. The high frequency of the S allele of the  $\beta$ -globin gene in populations with a high incidence of malaria is explained by balancing selection. AA individuals are more likely to die of malaria as children than are AS individuals. As a consequence, S is maintained in frequencies as high as 12% in malarial regions even though SS individuals suffer from a severe disease, sickle cell anemia which often causes early death.

In non-malarial regions, S is deleterious and A is advantageous. S alleles are very rare in non-malarial regions except as the result of immigration from malarial regions.

Malaria has been a particularly important selective force in human populations. In addition to the S allele of  $\beta$ -globin, alleles of the  $\alpha$ -globin gene which cause thalassemia provide partial protection against malaria. A defective allele (called A-) of the G6PD gene has a frequency of between 4 and 20% in malarial regions. Individuals homozygous for A- suffer from favism, a type of anemia induced when fava beans are eaten.

- *Rapid evolution occurs in response to changes imposed by humans*: Resistance to antibiotics, insecticides, herbicides, and poisons intended to control parasites, pests and pathogens, evolves quickly in virtually all species. Rapid evolution of resistance tells us that resistance alleles are already present.

### ***Genetic drift***

- *Genetic drift occurs because populations are not infinitely large*: On average, allele frequencies do not change after random mating. But in any one population, allele frequencies will change slightly each generation because the population size is finite.

The rate of change in allele frequency because of genetic drift is small in large populations and larger in small populations.

- *Bottlenecks and founder events*: A bottleneck in population size and a founder event both affect allele frequencies. A bottleneck is a reduction in size of an existing population while a founder event is the establishment of a new population of smaller size. Both cause more rapid changes in allele frequencies than would occur in a large stable population.

Extreme bottlenecks and founder events lead to the loss of many alleles along with the increase in frequency of a few alleles.

Cheetahs probably experienced an extreme bottleneck in size. There are so few genetic differences between individuals that skin can be successfully

grafted between unrelated individuals, something that is impossible in other mammals.

- *Founder events are important for the study of human genetic diseases:* Isolated populations founded by a few individuals may have some genetic diseases in high frequency that are rare elsewhere.

Huntington's disease (HD) is a late-onset dominant lethal condition. In the US population, the frequency of HD is about 1/10,000. The gene causing HD was identified by studying an extended family in San Luis, Venezuela, where almost 25% of the residents develop HD. One of the individuals who founded San Luis carried an HD allele. Selection against the HD allele is weak because its effects usually appear relatively late in life.

A founder effect could also explain the high prevalence of type II diabetes in Native Americans. Native Americans are probably descended from relatively small populations that came from eastern Asia 15,000 to 20,000 years ago. Other evidence that Native Americans are descended from a small number of ancestors is that they lack the B allele of the ABO blood type gene and the APOE\*2 allele.

- *Eventually, drift will cause the loss of one of two neutral alleles initially present:* Neutral mutations are those that have no effect on fitness. An example of a neutral mutation is one that changes the DNA sequence of a gene but not the amino acid sequence.

How long it takes for a neutral allele to be lost depends on how large the population is, but it will eventually happen even in very large populations.

Many new neutral alleles are created each generation by mutation. Most will be lost because of genetic drift but a few will be substituted. The substitution of neutral alleles is called non-Darwinian evolution or neutral evolution.

- *The rate of substitution of neutral alleles is approximately constant:* By comparing DNA sequences in species that diverged at a known time in the past, the substitution rate in animals has been estimated to be roughly  $2 \times 10^{-9}$  per nucleotide per year.

Once the rate is known, it can be used to estimate the time a pair of species diverged.

The constant rate of substitution is called a molecular clock.

### ***Gene flow***

- *Gene flow reduces differences between populations:* Gene flow results from the movement of individuals or gametes.

In the absence of gene flow, isolated populations will tend to become more different because of the combined effects of genetic drift, mutation and natural selection.

- *Gene flow is important for genetically engineered plants and animals:* Gene flow spreads engineered mutations from one variety of a species to another of the same species. A variety of rice called Liberty Link Rice (LL601) was

engineered to be resistant to the Liberty herbicide<sup>3</sup>. This variety was not approved for human consumption but was grown in a few test plots in several states from 1998 to 2001. In July 2006, traces of LL601 were found in other varieties of rice in Arkansas and Missouri. The news of this discovery resulted in a 10% drop in rice prices. In September 2006, 33 of 162 samples of rice shipped to Europe contained traces of LL601. Subsequently, Japan declared a month-long ban on the importation of rice from the US, and the European Union required testing of all imported US rice.

Gene flow also spreads genes from domesticated species to closely related natural species. Fish raised in hatcheries can interbreed with wild fish in the same streams. Farmed Atlantic salmon are known to interbreed with other salmon species on both the east and west coasts of the US. In this case, the concern is with making the wild species less well adapted. Farmed salmon grow faster than wild species but have higher death rates.

Gene flow of engineered genes to wild relatives might result in the evolution of weeds that are resistant to some herbicides or herbivores. Varieties *Brassica napus*, the species that produces canola oil, have been engineered to carry genes conferring resistance to several herbicides used for weed control. Some of those genes have been found in *Brassica rapa*, a wild relative. At present, no “superweeds” have because of gene flow from genetically modified species. It is hotly debated as to whether there is a significant risk.

### ***Example test questions***

Q1. Assume you determine the genotypes of 1,000 adults at the locus that codes for  $\beta$ -globin. In which one of the following data sets is there evidence for selection in favor of the heterozygous individuals?

- A. AA: 200, AS: 600, SS: 200
- B. AA: 640, AS: 320, SS: 40
- C. AA: 600, AS: 300, SS: 100
- D. AA: 1000, AS: 0, SS: 0
- E. None of the above

Answer: A

Q2. Assume you are studying the fitnesses of individuals with different genotypes at the  $\beta$ -globin locus in a population. The following are lists of the fractions of newborn individuals with each genotype that survive to adulthood. Which one of the following indicates that there is directional selection in favor of A in this population?

- A. AA: 0.9, AS: 0.95, SS: 0.2
- B. AA: 0.9, AS: 0.95, SS: 0.2

---

<sup>3</sup> See <http://www.sciencemag.org/cgi/content/full/313/5794/1714a>

- C. AA: 0.9, AS: 0.85, SS: 0.2
- D. AA: 0.9, AS: 0.85, SS: 0.9
- E. AA: 0.9, AS: 0.95, SS: 0.95

Answer: C

Q3. Which pair of words best fills the blanks in the following sentence? In areas with a high incidence of malaria, the S allele of the  $\beta$ -globin gene is maintained in high frequency by \_\_\_\_\_ selection; in regions where malaria is not present, S will be eliminated by \_\_\_\_\_ selection.

- A. adaptive, maladaptive
- B. directional, purifying
- C. purifying, directional
- D. balancing, purifying
- E. directional, balancing

Answer: D

Q4. Which one of the following is **not** a mutation?

- A. A change in a single nucleotide
- B. The loss of a chromosome
- C. The duplication of a chromosome
- D. A change in allele frequency
- E. The duplication of a gene

Answer: D

Q5. Which word best fills the blank in the following sentence? The discovery of an allele from LL601 rice in other varieties of rice illustrates the importance of \_\_\_\_\_.

- A. natural selection
- B. genetic drift
- C. recombination
- D. gene flow
- E. mutation

Answer: D

Q6. Which pair of words best fills the blanks in the following sentence? The substitution of neutral alleles is called \_\_\_\_\_ evolution because it is not caused by \_\_\_\_\_.

- A. Darwinian, genetic drift

- B. non-Darwinian, genetic drift
- C. Darwinian, natural selection
- D. non-Darwinian, natural selection
- E. None of the above

Answer: D

Q7. What is the approximate rate of substitution of neutral mutations in animals?

- A.  $2 \times 10^{-11}$  per nucleotide per year
- B.  $2 \times 10^{-9}$  per nucleotide per year
- C.  $2 \times 10^{-7}$  per nucleotide per year
- D. 1/1000 per nucleotide per year
- E. 2% per year

Answer: B

Q8. Which phrase best fills the blank in the following sentence? The high frequency of an allele causing Huntington's disease in San Luis, Venezuela is the result of \_\_\_\_\_.

- A. gene flow
- B. a founder event
- C. directional selection
- D. balancing selection
- E. Hardy-Weinberg

Answer: B