Information about Evolution and Medicine

1.0 Fundamentals of Evolution and Medicine

Biologists use the processes of scientific inquiry to try to understand two fundamental observations about living entities. The first observation centers on diversity—there are millions of species on Earth, and within each one there is diversity among individuals. The second observation seems paradoxical to the first—despite life's incredible diversity, organisms share a number of characteristics. Biologists have proposed an explanation for both observations—evolution (National Academy of Sciences and Institute of Medicine (NAS and IOM), 2008). In the mid-19th century, Charles Darwin's *On the Origin of Species* (1859) set the stage for the scientific studies that would provide increasingly more sophisticated and insightful evidence supporting evolution through "descent with modification" as the explanation for life's unity and diversity. The concept of biological evolution is among the most important ideas ever developed by applying scientific inquiry to the natural world (NAS and IOM, 2008), and it offers many benefits to the field of medicine (Nesse and Stearns, 2008).

Evolution simply refers to change. In the context of biological evolution, this change refers specifically to a change in **allele** frequencies in a **population**. This change is **heritable** and occurs over time among successive generations. More simply, a population of descendants differs from an ancestral population in some characteristics (Darwin, 1859). The concept of descent with modification has tremendous explanatory power and shapes the two major types of questions that evolutionary biologists ask (Futuyma, 1998):

1. What mechanism of evolution caused a change in a **lineage** or trait?
2. What is the evolutionary history, or phylogeny, of a living lineage or a specific trait?

Figure 1. *The science of evolutionary biology has important implications for medicine.*

© James Steidl / Dreamstime.com

1.1 Processes of Evolution

"Descent" in "descent with modification" emphasizes the history of lineages, whereas "modification" refers to the fact that lineages change over time. But what are the mechanisms that cause the change? This is one of the main questions that evolutionary biologists ask. Studies involving mathematical models and organisms in nature suggest that four different mechanisms can cause changes in the genetic makeup of a population: **mutation**, **gene flow** (migration), **genetic drift**, and **natural selection** (Futuyma, 1998).

All four processes of evolution depend on and affect genetic **variation** within populations. Ultimately, all genetic variation arises from **mutation**. Genetic recombination reshuffles existing variation. Because many students struggle to consistently identify the origin

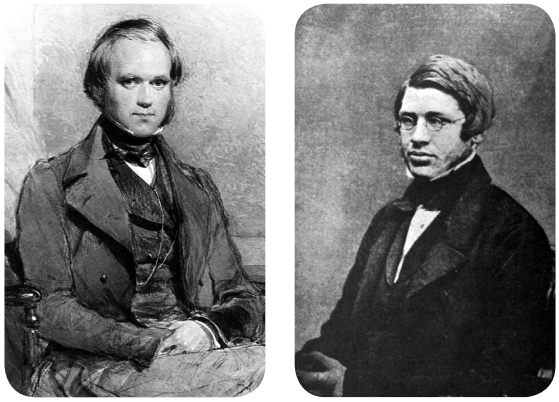
of genetic diversity when constructing explanations of natural selection, we ask them multiple times in the supplement to reflect on the role of mutations.

Though mutation is essential for generating genetic diversity, by itself it is not a major mechanism for changing the genetic makeup of a population from one generation to the next, because mutations happen in a single chromosome within an individual member of what is typically a large, often diploid population**. Gene flow**, also known as migration, is the movement of alleles from one population into another through immigration and emigration of individuals or through their gametes. Gene flow typically *reduces* the genetic differences among populations.

The remaining two mechanisms of evolution cause population divergence, that is, they *increase* the genetic differences among populations. **Genetic drift** refers to random changes in allele frequencies due to chance, or sampling error. Imagine a population of plants called monkeyflowers, some of which have red flowers and some have yellow ones. If a boulder rolls down a hill and flattens and kills some of the plants, this is not due to the plant's genotype for flower color, but rather to chance. One key feature of genetic drift is that it has a bigger effect on small populations than large ones. We can easily see that in a population where 75 percent of the plants have red flowers, rare events like falling boulders are more likely to drive the plants with yellow flowers to extinction if the total population size is four (that is, only one yellow flowering plant) than if the total population size is 400 (that is, 100 yellow flowering plants).

The final mechanism is **natural selection**. Alfred Russel Wallace and Darwin had jointly proposed this mechanism of evolution in 1858 through a paper delivered to the Linnean Society. Darwin described natural selection in detail in his book *On the Origin of Species* (Darwin, 1859). Darwin observed that within a species, characteristics among individuals vary. He was also aware that, for centuries, plant and animal breeders had bred organisms to emphasize or increase certain prized characteristics. Darwin reasoned that selection in nature could also bring about change in the characteristics of a population of organisms. Some organisms survive and reproduce better than others because of the characteristics they possess. Darwin called this process "natural selection." Natural selection provides a way to explain how new species could eventually appear from ancestral forms. Wallace developed a similar explanation around the same time.

Figure 2. *Charles Darwin (left) and Alfred Russel Wallace first described the process of evolution by natural selection.*



Natural selection is the only process of evolution that consistently yields adaptations. We can summarize the process of natural selection as three observations and one inescapable conclusion:

- *Observation 1*: Individuals within a population vary for many characteristics.
- *Observation 2*: Some of the differences in traits among individuals can be passed from parents to offspring. In other words, they are heritable.
- *Observation 3*: Individuals with certain variations have more offspring than others have.
- *Conclusion*: Individuals that possess heritable traits that enable them to better survive and reproduce will leave more

offspring, and these traits will increase in frequency over future generations, thus changing what the average member of the species is like.

Throughout the supplement, we ask students to explain certain observations by using natural selection. As they develop the arguments, they answer the following types of questions:

- In what ways does the population vary for an important trait?
- How did the variation arise?
- Can some of the differences in traits among individuals be passed from parents to offspring?
- Do individuals with certain traits survive and reproduce at relatively higher rates?
- How will the frequency of traits and the alleles affecting those traits change in the population over time?

The use of these types of questions across examples helps students frame the important pieces of an argument based on natural selection. A study by Bray Speth and colleagues (2009) suggests that these "concept frames" are a useful source of formative assessment data for instructors.

The activities in the supplement focus on natural selection. This is not meant to diminish the importance of the other mechanisms of evolution. In fact, the relative importance of genetic drift and natural selection is a long-standing debate among evolutionary biologists that continues to this day (Fisher, 1930; Wright, 1931; Provine, 1986; Coyne et al., 1997; Lynch, 2007). Many biologists accept that natural selection is the most powerful mechanism for phenotypic evolution, whereas genetic drift and mutation have played a pivotal role in shaping genomes and genetic architecture (Lynch, 2007).

The genome-wide approaches that researchers are using to detect positive natural selection in humans will vastly increase our understanding of the role natural selection plays in shaping the human genome (Sabeti et al., 2006). **Positive selection** occurs when variants of a gene, and the protein it produces, are continuously favored by natural selection and these "young" alleles spread rapidly in a population. One way that scientists detect positive selection in genetic sequences is by comparing the number of mutations that lead to no changes in amino acids (due to redundancy in the genetic code, so-called synonymous mutations) to the number of mutations that do lead to different amino acids (nonsynonymous mutations). In positive selection, more changes lead to different amino acids than would be expected by chance. Students get a very brief introduction to positive selection in influenza viruses in Lesson 4. Three of the examples we explore in the supplement show positive selection: MRSA (for methicillin-resistant *Streptococcus aureus*)(Harris et al., 2010), lactase in humans (Bersaglieri et al., 2004), and influenza viruses (Bush, 2001).

Understanding mechanisms of evolution, particularly adaptation by natural selection, provides many insights that enhance medical practice and understanding. A famous case involves the role of natural selection in helping researchers better understand sickle cell anemia. Sickle cell anemia affects millions of people and is a serious lifelong condition. With adequate healthcare, people with sickle cell anemia can live nearly normal lives with reasonably good health. Without adequate care, the disease can be debilitating and cause early death. This disease is caused by a recessive genetic disorder, a mutation in the *HBB* gene (which encodes β-globin). The allele that leads to sickle cell disease in homozygotes is called *HbS* (with the resulting protein hemoglobin S) and was one of the first specific genetic variants to be associated with a molecular defect (Pauling et al., 1949). *HbS* has four distinct forms, suggesting that it may have arisen independently multiple times in different locations (Kwiatkowski, 2005).

The frequency of the *HbS* allele in some regions of the world is high (about 10 percent). Biologists noticed that populations with a high frequency of this allele occurred in geographic areas with high rates of **malaria**. Malaria is thought to be the strongest selective agent known in recent human history (Kwiatkowski, 2005). Allison (1954) first hypothesized that the sickle cell allele is advantageous in certain environments because it protects carriers against malaria. Homozygotes for the typical hemoglobin allele do not have sickle cell anemia, but they are susceptible to malaria. Heterozygotes who carry one normal allele and one sickle cell allele have a 10-fold reduced risk of malaria and are only slightly anemic (Kwiatkowski, 2005). Natural selection favors the heterozygote in geographic regions with high rates of malaria and maintains both alleles in the population. The sickle cell–malaria scenario is a classic example of how selection explains why human populations vary for some genetically determined traits that affect health.

It is difficult to imagine how we would explain the high susceptibility to sickle cell anemia in some human populations without invoking evolution in the past. In other words, the frequency of the *HbS* allele is higher than we would expect if it did not influence survival in people with malaria. In fact, investigations of sickle cell were the first evidence of natural selection operating in humans (Allison, 1954). This scenario is featured in many high school biology curricula. It is important to keep in mind, however, that **heterozygote advantage** is probably relatively infrequent, and for good evolutionary reasons. The disadvantages that accrue to the homozygotes may provide a strong selection force for an alternative, superior solution. However, if the heterozygotes have a strong advantage, the polymorphisms can be maintained in populations for a very long time. Interestingly, it appears that heterozygote advantage is more common in populations exposed to a relatively recent environmental change.

Figure 3. *Sickle cell anemia was one of the first diseases that was better understood by considering evolution and the impact of malaria on human populations. Upper left: microscopic view of blood cells from a person with sickle cell anemia; lower right: female mosquito taking a blood meal.*

In this supplement, students explore the high prevalence of **thalassemia** in certain populations. Similar to sickle cell anemia, the high prevalence of thalassemia is partially explained by the fact that individuals with thalassemia have higher protection against severe malaria. Learning of a second disease that follows a pattern similar to sickle cell anemia's should help students generalize the main concepts of natural selection to a broader range of problems.

In general, Nesse and Williams (1994; see also Nesse, 2007) suggest several categories that help explain human vulnerability to disease, based on principles involving natural selection:

- *Mismatch to the Environment*: Modern environments in the industrialized world are radically different from those that predominated during most of human evolution. We are not yet well adapted to our current environment (Leonard, 2008), for the spread of adaptations in human populations is much slower than the rate of

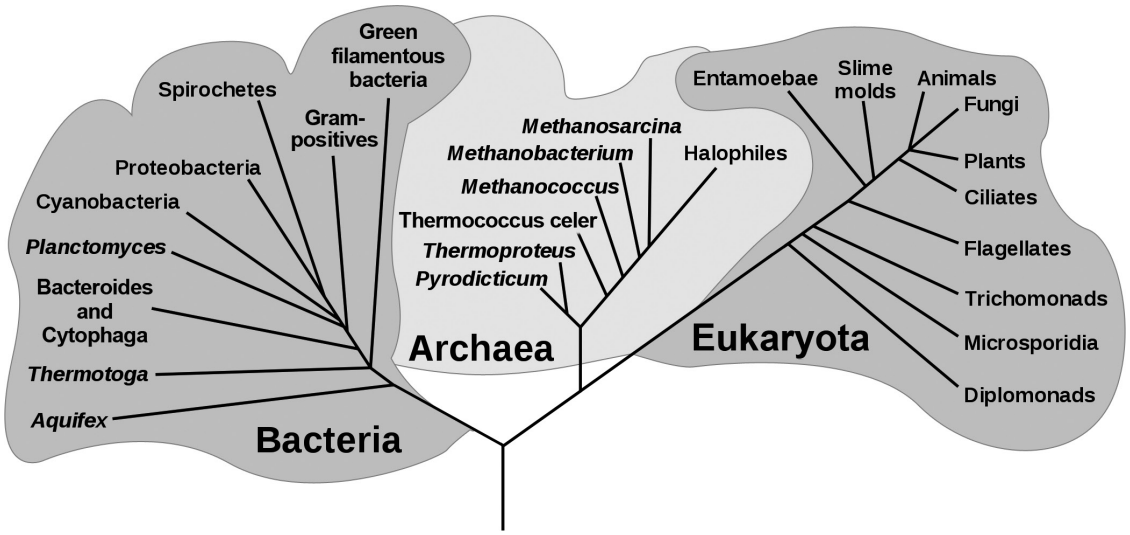
cultural change. In this supplement, students explore the evolution of **lactase persistence**, which has evolved multiple times in different populations of humans since the domestication of dairy animals. Different alleles of the lactase gene are associated with persistence in different populations (Ingram et al., 2007). Selection for this trait and on lactase persistence alleles has been very strong for the last 3,000–10,000 years (Bersaglieri et al., 2004; Tishkoff et al., 2007). Lactase persistence is described in more detail on page 34.

- *Rapid Pathogen Evolution*: Pathogens usually have much shorter generation times, higher mutation rates, and vastly larger numbers of offspring than their hosts. They also face higher selection coefficients than humans and any other organism with a generation time longer than a few weeks. This results in pathogens displaying more rapid adaptation (Hillis, 2004). Human evolution occurs over longer time periods, making protection from infection a persistent challenge.
- *Constraints*: One example of a constraint relates to genetic variation. Selection can only act on the variation present within a population. Exceedingly complex structures do not evolve de novo; instead, they evolve stepwise from preexisting structures that often have a different function (for example, the bacterial flagellum; Liu and Ochman, 2007). The variation present limits what selection can shape.
- *Tradeoffs*: A tradeoff occurs when an evolutionary change in one trait that increases fitness is linked to a change in another trait that decreases fitness. All organisms (including humans) must make tradeoffs, and this means that it is natural that some traits are not in an ideal state but are byproducts of selection acting on other traits. One important tradeoff found in many organisms is between reproduction and survival. Forms of genes that increase reproductive success will increase in frequency even if they negatively affect health and longevity. For example, men with high testosterone levels may compete more successfully for a mate, but they may suffer from decreased resistance to pathogens (Muehlenbein and Bribiescas, 2005). Selection for increased reproduction may result in decreased survival.

1.2 Common Ancestry

Early naturalists noticed that species can be clustered naturally into a hierarchical pattern of groups within groups—that is, species into

Figure 4. *All living organisms are related in one great phylogenetic "Tree of Life."*



genera, genera into families, families into orders, and so on. But it was not obvious why this nesting pattern occurs. Darwin (1859) realized that "this natural subordination of organic beings into groups under groups" could be explained by descent with modification from common (shared) ancestors. There is no logical reason to expect species to be arranged hierarchically if they arise separately.

Although scientists today may disagree about some of the natural groupings of organisms, they all agree with the idea of descent with modification from common ancestors. Descent with modification explains two features that are characteristic of organism groupings. First, the pattern is *hierarchical*, or made of groups within groups. Second, it is *branching*, or treelike.

A branching pattern of groups results whenever an ancestral group splits into related subgroups that come to differ in some way. This pattern allows us to trace the ancestry of the subgroups back to their common ancestor. We can trace the ancestry of this ancestor back to another shared ancestor, and so on. It is like working backward along the branches of a tree from the twigs to the trunk. We can trace the growth of all twigs back through a series of branch points to the trunk. The twigs represent existing species, nearby branch points represent recent shared ancestors, and the trunk represents a distant ancestor that is common to many branches.

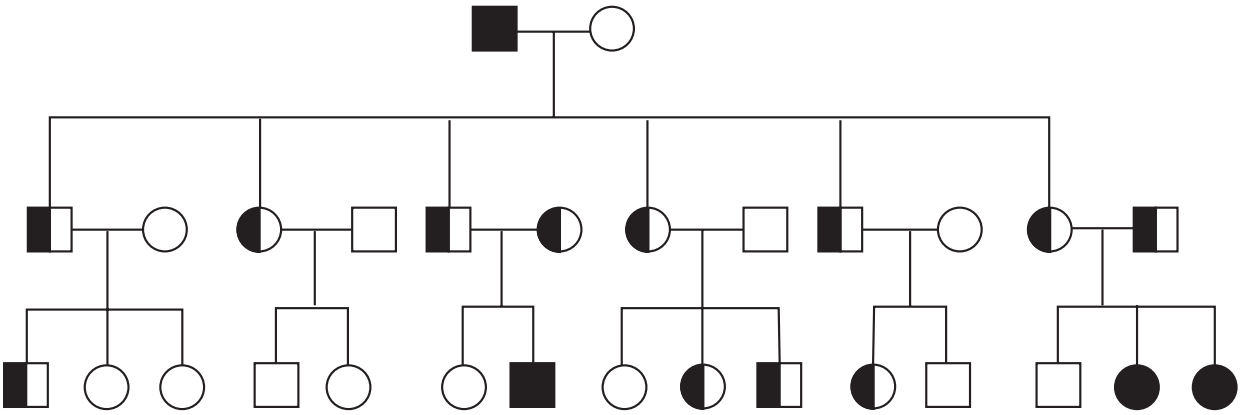
The similarities among organisms are evidence of their descent from a common ancestor. Scientists gather data from observable characteristics in organisms to estimate relationships. These characteristics include structural similarities (for example, skeletal features or cellular structures), patterns of embryological development, and, increasingly, molecular data. Since the development of molecular techniques in the 1980s, the use of DNA, RNA, and amino acid sequences as well as careful analyses of when and where genes are used in the development and maintenance of living organisms has sharpened scientists' ability to ask and answer fine-scale evolutionary questions. Many hypotheses based on morphological characteristics have gained further support. However, hypotheses of relationships can change when new data are acquired, and, in some cases, research has overturned previous ideas of relationships.

Diagrams that summarize the evolutionary history of the relationships among organisms are called "**evolutionary trees**" or "phylogenies." The characteristics of living organisms have been shaped by their long evolutionary history. Evolutionary biologists seek to answer questions about the relationships among living and extinct species, the history of specific populations within a species, the timing and geography of diversification events, the reconstruction of **ancestral states**, and the timing and origin of specific characteristics or processes in organisms. Phylogenetic hypotheses are being rapidly developed (Hillis, 2004), and tools and models for reconstructing relationships are becoming increasingly sophisticated. Students gain experience with some of the important NIH-sponsored tools and databases through the lessons in the supplement.

The explosion of phylogenetic information afforded by the sequencing of genomes from diverse organisms across the "Tree of Life" offers many insights that may inform medicine (Nesse and Stearns, 2008). The following are just a few examples of how phylogenetics has informed medicine:

- Alleles associated with specific phenotypes are more frequent in certain human populations of different geographical origin. For example, persons of Ashkenazi Jewish ancestry living in the United States have a higher frequency of *BRCA1* and *BRCA2* mutations (Ewald, 2008; Narod and Offit, 2005), as do Icelandic, Dutch, and Polish populations (Narod and Offit, 2005).

Figure 5. Pedigrees, like the one shown below, and genetic screening help researchers understand patterns of genotypes and phenotypes.



Thus, knowing the ancestry of individuals can provide some insight into probabilities of specific genetic conditions, which may influence the genetic screening and counseling these individuals receive. Large-scale representative sampling from populations across Earth for high-risk alleles is currently under way (Crews and Gerber, 2008). Students explore allele frequencies in different human populations in two lessons in the supplement.

- The evolutionary origin of pathogens is now routinely investigated by using phylogenetic methods. Phylogenetic analysis of the human immunodeficiency virus (HIV) and simian immunodeficiency virus (SIV) clearly shows that there are two major types of HIV that moved into humans from two separate hosts. Each type has become established in humans more than once (Rambaut et al., 2004; Sharp et al., 2001). The same types of analyses led to the identification of the coronavirus that causes severe acute respiratory syndrome (SARS; Ksiazek et al., 2003; Peiris et al., 2003). Phylogenetic analysis also led to the identification of bats as the reservoir for the coronavirus (Li et al., 2005). In the supplement, students use genetic sequences to explore the history of **influenza**. Scientists use these same tools and skills to identify the origin of emerging pathogens.

2.0 The Value of an Evolutionary Perspective for Medicine

The questions evolutionary biologists ask illuminate many matters that affect human health. The field of evolutionary medicine uses the models and theory of evolutionary biology to inform problems encountered in medicine and public health (Nesse, 2008). Many applications of evolutionary biology are already well established in medicine and are very useful. These include population genetics, phylogenetic analysis, and studies of antibiotic resistance. Even in these well-established areas, new evolutionary insights are leading to rapid advances.

Insights from evolution can also provide a theoretical framework for understanding why organisms are vulnerable to disease. Advocates suggest that evolutionary biology should be put on par with other fundamental basic sciences, such as biochemistry, and that teaching it will make medical education more coherent (Nesse and Stearns, 2008).

3.0 Specific Applications of Evolution in Medicine

3.1 The Relationship of Genetic Variation to Health

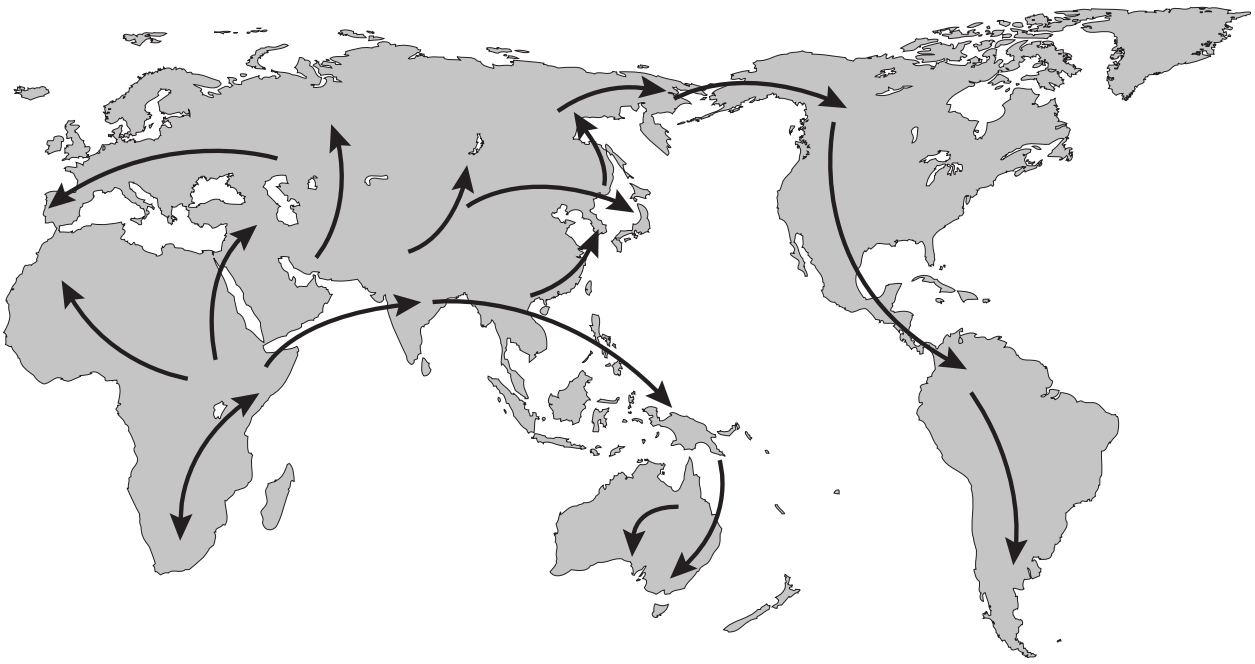
Variation is the raw material on which evolutionary processes operate (Futuyma, 1998). Though individuals in a population may show variation in a phenotype, only the proportion

of that variation that is heritable will respond to natural selection. Evolutionary biologists try to assess the proportion of phenotypic variance attributable to genetic variance, environmental variance, and genotype × environment interactions. Variation is often quantified within and among populations.

Humans vary across the world. Every independently conceived individual is genetically unique. This seems paradoxical in light of the fact that all humans have a high degree of genetic similarity. It is often reported that two humans are 99.9 percent similar in their DNA. However, the human genome is immense, providing multiple opportunities for genetic variation to arise; the 0.1 percent by which we differ amounts to 3.3 million nucleotides (Kidd and Kidd, 2008). Findings from the International HapMap Project confirm previous studies and show a relatively low amount of differentiation among human groups defined by ethnicity and geography (Govindaraju and Jorde, 2008). There is much more genetic variation within (about 90 percent) than among (about 10 percent) human groups. This means that the similarities among different groups of humans far outweigh the differences. To learn more about the HapMap Project, visit http://hapmap.ncbi.nlm.nih.gov.

As the ability to decipher the genotypes of individuals improves and becomes more widely available, medical practitioners will be better able to give patients specific information about their health. Individual genetic profiles provide useful information about disease susceptibility and predispositions. Crews and Gerber (2008) suggest three possible medical-clinical applications of individual genetic profiling: improved screening, more-informed counseling, and individualized drug formularies. Until more individualized data are available, however, researchers continue to try to determine whether disease susceptibility is linked to specific genetic

Figure 6. *Genetic data support hypotheses that humans migrated out of Africa.*



(Image created from data from the National Geographic Society's Genographic Project, https://genographic.nationalgeographic.com/genographic/lan/en/atlas.html.)

factors and, if so, whether the genetic factors are distributed differentially among geographic groups (Kittles and Weiss, 2003). Patterns of variation among humans have been shaped by migration, genetic drift, mutation, and natural selection. These evolutionary mechanisms lead to a correlation between geographic distribution and genetic variation (Ramachandran et al., 2005; Soo-Jin Lee et al., 2008) that may be medically relevant (Kidd and Kidd, 2008).

Genetic variation in health-related traits may be simple (one gene) or complex (multiple genes). Complex traits are often described as being multifactorial, meaning that the genes interact with each other and with the environment. Decades ago, much of the focus was on human diseases that have a relatively simple genetic basis (for example, Tay-Sachs disease or Huntington's disease). Modern genetic insights show that many of the supposedly simple traits are more complex than people thought. For example, phenylketonuria (PKU) is caused by mutations to a gene that affects the production of phenylalanine hydroxylase. However, scientists have now identified multiple mutations to the gene, each of which can cause different symptoms and outcomes for people with PKU (Kidd and Kidd, 2008). NIH maintains a site that contains large amounts of up-to-date information on human genes and genetic phenotypes, Online Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/omim).

Though diseases with a relatively simple genetic basis are important to study, a much larger fraction of genetic variation is likely to influence an individual's interaction with parts of the environment that influence health (Kidd and Kidd, 2008). Complex diseases that have significant genetic and environmental influences have a large impact on public health and are likely to command the attention of the biomedical community in the near future. Unraveling the causes of complex diseases has been advanced by genome-wide association studies, which look across the entire genome for genetic influences on disease risk. To date, such studies have found many new genes that influence disease risk. But the overall amount of variation among individuals in disease risk that can be explained by genetics has remained small: on the order of 3 to 5 percent. These results suggest that much of disease risk is tied up in complex gene × environment interactions, which means that both the particular genes a person inherits and the particular environments the person is exposed to are important. For example, the degree to which smoking and cholesterol increase the risk of heart disease depends on the particular versions of multiple genes that the person has inherited.

3.2 The Role of Evolution in Drug Response

The relationship between genetic variation and health is specifically manifested in the study of the genetic basis for variation in individuals' drug responses. The role of evolution in shaping this variation is also often relevant to these studies. We illustrate two classic cases, though numerous examples exist.

1. Various oxidant drugs can cause the destruction of red blood cells in certain people. Half a century ago, researchers showed that this response was caused by an X-linked, single-gene, recessive trait resulting in a deficiency of glucose-6-phosphate dehydrogenase (Carson et al., 1956). Alleles for this trait were common in people of African and Mediterranean ancestry. Similar to the case of sickle cell disease, researchers suspected that this allele is retained in these populations due to the protection conferred against malaria for heterozygote females (Ruwende et al., 1995), a hypothesis that has received recent support from genetic analyses (Sabeti et al., 2002; Saunders et al., 2005; Tishkoff et al., 2001). Response to selection caused

by malaria provides an explanation of the evolutionary origin of this drug response.

2. Many other drug responses are based on complex interactions rather than the action of a single gene. For example, patient response to warfarin, an anticoagulant drug, varies widely. A significant portion of this variation is attributable to genetic causes. Careful dosing and monitoring of the drug are required both to ensure an adequate patient response and to avoid excessive bleeding. Concerns about the variability in patient response and potential complications have discouraged physicians from prescribing warfarin, and an estimated 50 percent of patients who could benefit from anticoagulant therapy do not receive it (Friberg et al., 2006). Researchers have identified two genes that play significant roles in patient response to warfarin. Taken together, alleles for these two genes account for about 40 percent of the variability in the response. Selection, drift, and gene flow have shaped the patterns of genetic variation for these two genes in different groups of people, which partially explains patterns in the responses of different groups. Genetic testing of individuals aims to improve the safety of warfarin therapy. While the effect of these two genes is significant, other genetic and environmental interactions (such as diet and interactions with other drugs) also affect warfarin responses.

3.3 The Role of Evolution in Infectious Diseases

One of the most direct applications of evolution to medicine is in the well-documented evolution of antibiotic resistance in bacteria. Antibiotics exert strong selective pressures on populations of bacteria, which quickly increase the proportion of individuals that can resist the antibiotic (Bergstrom and Feldgarden, 2008). Bacterial resistance to a new antibiotic almost always becomes prevalent just a few years after the antibiotic is introduced. The cost of antibiotic resistance in the United States is estimated to be $80 billion annually (Bergstrom and Feldgarden, 2008). Similarly, viruses—especially retroviruses with RNA genomes—quickly evolve resistance to antiviral drugs, as in the case of HIV (Rambaut et al., 2004). This point is particularly compelling in light of the fact that the majority of emerging pathogens causing new infectious diseases are viruses, especially RNA viruses (Woolhouse and Antia, 2008).

The evolution of resistance has important implications for medicine. For example, some hospitals initiated antibiotic-cycling regimes (antibiotics are rotated on a schedule) to help reduce the evolution of resistance. However, mathematical models that incorporate principles of ecology and evolution show that this approach may be ineffective, a result supported by metaanalyses of clinical trials (Bergstrom and Feldgarden, 2008). Additionally, models of evolution help researchers formulate the vaccinations against influenza each year and design drug regimes for the treatment of AIDS (for example, highly active antiretroviral therapy; Rambaut et al., 2004).

Figure 7. *The evolution of resistance in HIV requires patients to use many medications.*

© Norberto Mario Lauria / Dreamstime.com

Studying the evolution of virulence provides powerful insights into infectious diseases. Before the rise of genetic engineering, researchers used the evolution of virulence to develop attenuated live viruses. Pathogens were grown in culture, and adaptations to culture conditions involved tradeoffs so that this evolved strain grew more poorly in the host (Ebert and Bull, 2008). More recently, some researchers have proposed using virulence evolution to help manage parasites. Others argue that we need better models that incorporate more details of specific biological processes before we start measures that entail risks (Ebert and Bull, 2008).

An additional role for evolution in helping us understand infectious diseases is through the use of phylogenetic methods to identify and trace the evolutionary origin of pathogens and the reservoirs of pathogens. (The phylogenetic approach for understanding HIV and SARS is described on pages 26 and 27.)

4.0 Students' Prior Conceptions about Evolution

In addition to including vivid examples of evolution and medicine, the supplement takes into account research on student preconceptions. Educational research on evolution shows that students hold several naïve preconceptions about evolution that interfere with their learning (Sinatra et al., 2008), in particular, naïve ideas about natural selection and the interpretation of evolutionary history. We list below some of the most common preconceptions that we address in the supplement. We suggest that you do not use these as a list of lecture topics for your students, but rather use them to inform your teaching as the preconceptions emerge.

4.1 Natural Selection and Processes of Change

- *Evolution is Only a Theory*: Many students do not understand the meaning of the term theory in science and equate it with meaning a hunch or a guess. This misunderstanding stems from a large problem in evolution education; namely, a poor understanding of the nature of science. (American Association for the Advancement of Science (AAAS), 2001).
- *Acquired Characteristics Can Be Inherited*: The inheritance of acquired characteristics suggests that changes that parents make in their lifetime are passed on to the next generation, whether or not they have a genetic basis. These incorrect ideas are remarkably persistent (Crow, 2004). In contrast, biologists recognize that evolution only occurs when changes are heritable and that change occurs at the population level, not the individual level (AAAS, 2001). For students to truly grasp the idea that populations, not individuals, evolve, they must have an understanding of variation within a population. They must also recognize that a shift in an average characteristic of a population represents change. We designed the activities in this supplement to help students recognize and confront these ideas.
- *Species Have an Underlying Nature That Cannot Change*: The concept of essentialism, which is incorrect, makes it difficult for students to recognize both variation within species and that species can change over time. Many students struggle to recognize this implicit and incorrect assumption.
- *Natural Selection Leads to Perfection*: Phrases such as "more highly evolved" can be misleading. Species adapt to conditions in the present, but these conditions can and do change. Often, an adaptation that helps a species survive in one environment is a disadvantage when the environment changes. Additionally, many traits are involved in tradeoffs. The idea of a workable compromise is a much better descriptor of the state of an organism than the idea of perfect design.
- *Fitness Means Individuals Are Stronger or More Athletic*: Many student associate fitness with overall strength and the ability to fight.

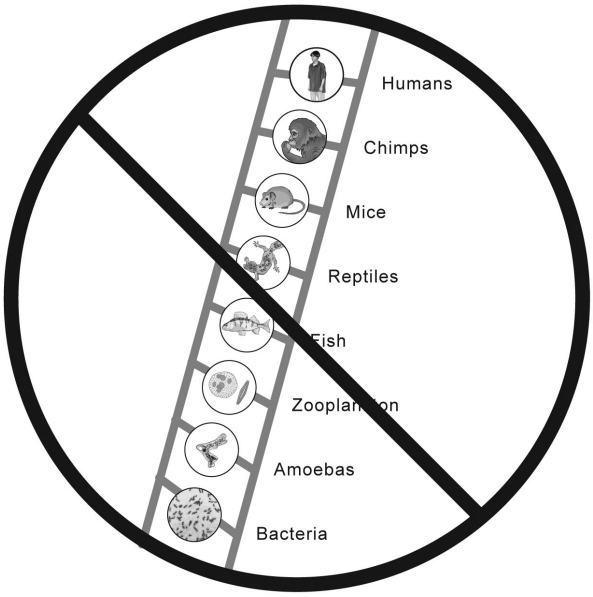
However, cooperation is an essential aspect for survival in many organisms. Offering students multiple and varied examples of selection documented in real time will help them construct a more accurate view of selection. It is important that students understand that natural selection does not operate to improve survival but rather to improve reproductive success. Of course, an individual must survive long enough to reproduce. But after maturation, evolutionary changes that increase reproductive success will spread in the population if the effects on reproduction are large enough and the effects on survival are small enough, even though they decrease survival.

- *Natural Selection and Evolution Mean the Same Thing*: There are multiple mechanisms for evolution, including genetic drift and gene flow. However, natural selection is one of the most powerful and well-documented mechanisms of evolution.
- *Greatly Different Time Spans are Equivalent*: Although many people understand that evolution has taken place over large expanses of time, they seem to place events in broad categories such as "extremely ancient," "moderately ancient," and "least ancient" (Catley and Novick, 2009; Libarkin et al., 2005; Trend, 2001).

4.2 "Tree-Thinking"

- *Evolutionary Trees Are a "Ladder of Progress"*: Some people equate a progression from simple to complex as "better." This line of reasoning may cause people to try to read evolutionary trees from left to right, with the organisms on the left evolving into the organisms on the right. Instead, evolution mostly proceeds by having one group bud off from an existing group (Meir et al., 2007). A true understanding of biodiversity shows that all species existing today are descended from ancestors that have survived since the origin of life more than 3.5 billion years ago. Plants, worms, bacteria, and humans all descend from ancestors that survived multiple mass extinctions and many major environmental challenges. To claim that one existing group is "more highly evolved" than another ignores this basic fact.
- *Modern Species Are Ancestors*: One way this misconception is manifest is the saying "humans evolved from chimpanzees." Instead, humans and chimpanzees evolved from a common ancestor. This common ancestor may have shared characteristics with modern chimpanzees, but both the modern chimpanzee lineage and the lineage that led to humans have been evolving for the same amount of time since splitting from the common ancestor (Baum et al., 2005).
- *The Ordering of Species at the Tips of an Evolutionary Tree Is Always Meaningful:* Many students ignore the pattern of branching in an evolutionary tree when trying to determine relatedness. They do not recognize that the most closely related species are those that share the most recent common ancestor (Baum et al., 2005).

Figure 8. *Many students mistakenly think of evolutionary trees as "ladders of progress."*



To help students determine the time since common ancestry and better understand the time involved in evolution, we included a time arrow on many of the evolutionary trees in this supplement.

In addition to these prior conceptions, students have difficulty interpreting evolutionary trees that are drawn in different ways. For example, many students have difficulty interpreting evolutionary trees drawn with the nodes as a V shape. When shown such trees, students often assume that the group at the end of the longest line is the ancestor of the other groups (Novick and Catley, 2007).

The activities in this supplement give students a chance to explore the widely accepted scientific explanation for the diversity of life on Earth: evolution. As a science teacher, it is your responsibility to your role is to present the scientific evidence for evolution, while respecting students' individual beliefs. Consider reviewing the differences in ways of knowing with students. Evolution is a scientific way of explaining biological change across time. Creationism (including creation science and intelligent design) is not scientific because it invokes supernatural causes. Evolution is not inherently at odds with religion. Religion offers other ways of knowing the world.

5.0 Featured Examples of Evolution and Medicine

5.1 Methicillin-Resistant Staphylococcus aureus *(MRSA)*

The evolution of antibiotic-resistant bacteria is a classic example of evolution by natural selection that has large impacts on humans. Today, controlling the spread of multiply resistant bacterial strains is a major global health concern. MRSA stands for **methicillin-resistant *Staphylococcus aureus***. Methicillin is in a group of antibiotics that includes penicillin and amoxicillin. Unfortunately, some populations of *S. aureus* are resistant to the entire group of antibiotics, which includes methicillin.

Health specialists differentiate between hospital-acquired MRSA and community-acquired MRSA. Hospital-acquired MRSA was recognized before community-acquired MRSA. It is not yet clear if community-acquired MRSA evolved from hospital-acquired MRSA or if it evolved from different strains of *Staphylococcus aureus*. Community-acquired MRSA typically causes the outbreaks found in schools and differs from hospital-acquired MRSA in that it is not multidrug-resistant and can be treated with other antibiotics (Dominguez, 2004). *Staphylococcus aureus* affects animals other than humans; it is one of the major causes of a disease in cows called contagious bovine mastitis, for example.

Nearly all MRSA skin infections respond to treatment. Serious infections, such as pneumonia and blood and bone infections, rarely occur in healthy people.

MRSA can be spread through skin-to-skin contact or sharing items such as towels used by infected individuals. The environments that are more conducive to spreading MRSA are called the "5 C's" by the Centers for Disease Control and Prevention (CDC): crowding, frequent skin-to-skin contact, compromised skin (cuts, for

Figure 9. *Infectious diseases affect our lives in many ways. Antiseptic dispensers like this are now commonplace in many work and school environments.*

© dblight / iStockphoto.com

example), contaminated items and surfaces, and lack of cleanliness. The CDC hosts a useful Web site with information about MRSA infections: http://www.cdc.gov/Features/MRSAinfections/.

In the first lesson, students use their prior knowledge to develop an explanation of how populations of MRSA evolve antibiotic resistance. If students are familiar with natural selection, they may suggest that the genetic variation required for the evolution of resistance came about through mutation. Indeed, mutation is the source of the requisite genetic variation in many cases. However, genes for resistance, virulence, or both can be acquired through horizontal gene transfer between closely or distantly related bacteria. Researchers at NIH used genetic sequences and an analysis tool called an "**evoprint**" to demonstrate that a strain of hospital-acquired MRSA shares multiple DNA sequence blocks with *Listeria monocytogenes* and *Staphylococcus saprophyticus* (Brody et al., 2008). Both of these bacteria can also cause bovine mastitis. These data suggest that bacterial co-infections may be an important factor in the evolution of resistance. Students gain experience analyzing an evoprint in Lesson 3.

5.2 Pax6

Pax6 is a gene that plays a crucial role in regulating the growth of eyes in a wide range of organisms, including humans, fruit flies, zebrafish, mice, squid, planarians, and even ribbon worms. In humans, the gene is involved in the early development of the eyes, brain, spinal cord, and pancreas. It is found on chromosome 11 and spans over 20,000 nucleotides. *Pax6* codes for a transcription factor in the paired box gene family (which is the origin of the name *Pax6*). By convention, *PAX6* (all capital letters) refers to the human gene, *Pax6* (just the first letter capitalized) refers to the gene in mice and rats, and *pax6* (all lowercase letters) refers to the gene in zebrafish. For simplicity, we use *Pax6* throughout the supplement.

Studies of *Pax6* caused scientists to radically reconsider the evolution of eyes in different lineages. Previously, many scientists inferred that eyes evolved independently many times, due to the large differences in anatomical structure among eyes in different groups of animals. For instance, arthropods have compound eyes, whereas vertebrates and squid have camera-type eyes (Carroll, 2006). However, the ubiquitous presence of *Pax6* as a regulator of eye development suggests that the gene was present in the common ancestor of these animals. Scientists infer that the common ancestor likely had a simple eyespot and may not have been able to form an image (Halder, 1997). Studies on *Pax6* support the claim that new, complex structures are rarely built anew throughout evolution; rather, they are often assembled from preexisting structures.

Some mutations to the *Pax6* gene cause the human disease **aniridia**. Scientists have identified over 250 such mutations in humans. Most of the mutations that cause aniridia are premature stop codons.

For more information on aniridia, visit http://ghr.nlm.nih.gov/condition=aniridia.

For more information on *Pax6*, visit http://ghr.nlm.nih.gov/gene=pax6.

5.3 Lactase Persistence

Like all mammals, infant humans have the ability to digest the sugar lactose in milk. However, in a large number of humans, the gene that codes for the enzyme lactase is "turned off" after weaning, causing these individuals to lose the ability to digest lactose. Individuals with this condition are called "**lactase nonpersistent**." The gene for lactase remains "turned on" for individuals who are lactase persistent (**lactose tolerant**). Strictly speaking, **lactose intolerance** is not equivalent to lactase nonpersistence because intolerance may also be caused by lactose malabsorption. An NIH Consensus Development Conference

on Lactose Intolerance and Health in 2010 defined lactose intolerance as "the onset of gastrointestinal symptoms following a blinded, single-dose challenge of ingested lactose by an individual with lactose malabsorption, which are not observed when the person ingests an indistinguishable placebo." Students use the terms "lactase persistence" and "lactase nonpersistence" in the activity. The majority of humans are lactase nonpersistent, and lactase persistence is a derived trait (Holden and Mace, 1997). This means that the common ancestor of all humans was lactase nonpersistent.

Lactase persistence has a strong genetic component and has evolved independently at least three times in human history. Scientists have found three main alleles that are associated with lactase persistence. Each allele is identified by a specific mutation to one nucleotide (called a single nucleotide polymorphism, or a SNP (pronounced "snip")). The SNPs are located thousands of nucleotides away from the coding region for the lactase gene but are all relatively close to each other. The first identified allele associated with lactase persistence is a SNP 13,910 nucleotides away from the lactase gene. The two SNPs associated with the other two other alleles are 13,915 and 14,010 nucleotides away. These nucleotides are located in an intron of a gene called *MCM6* that neighbors the lactase gene. Changes to the nucleotides in this region can affect the transcription of the lactase gene. Results of genetic analyses on each allele are consistent with positive directional selection.

Scientists conduct bioinformatic studies on the entire human genome, looking for genes that show a strong signal of positive selection. The regulatory region for the lactase gene shows one of the strongest signals for positive selection. Selection coefficients estimated in populations that have a high frequency of lactase persistence are as high as 5 to10 percent. A selection coefficient of 5 percent is strong enough to cause an allele to increase in frequency from 1 percent to 95 percent in 300 generations, or in about 9,000 years for humans. The rapid rise in frequency of lactase persistence alleles in some populations are referred to as selective sweeps. Recent analyses of 200 human genomes suggest that selective sweeps in human evolution are the exception rather than the rule, however (Hernandez et al., 2011). Instead, human adaptation was mostly shaped by relatively small changes in many genes.

To learn more about recent findings on the evolution of lactase persistence, see Enattah et al., 2007; Enattah et al., 2008; Gerbault et al., 2009; Holden and Mace, 1997; Ingram et al., 2007; Ingram et al., 2009; Swallow, 2003; and Tishkoff et al., 2007.

A recent study examined different hypotheses for the evolution of lactase persistence in different populations (Gerbault et al., 2009). The culture-historical hypothesis was consistent with the results obtained for the evolution of lactase persistence in Africa, whereas the UV–vitamin D–calcium hypothesis was consistent with the evolution of lactase persistence in northern Europe. However, the data for Europe are complex. Alternative hypotheses for selection pressures and the effects of population history and demography are difficult to disentangle. See the masters associated with Lesson 2 for more information about these hypotheses and other recent research about them.

5.4 Thalassemia

As described in the masters associated with Lesson 3, thalassemia refers to several diseases characterized by reduced or no production of the globin proteins that form hemoglobin. Hemoglobin in adults is composed of two alpha-globin protein chains and two beta-globin protein chains. Alpha-thalassemia is a condition in which a person produces less or no alpha-globin protein. Beta-thalassemia occurs when there is a defect in the beta-globin gene, the gene for the beta-hemoglobin chain. A useful description of the history and future of research

on the thalassemias can be found in a 2004 paper by D. Weatherall.

Alpha-thalassemia is caused by mutations in the alpha-globin genes on chromosome 16. Normally, each person has four copies of the alpha-globin gene, two on each chromosome of the pair. In alpha-thalassemia, one to four of the alpha-globin genes are not functional. When people have three functional alleles of the gene, they have few if any symptoms. When two or three of the alpha-globin alleles are nonfunctional, people have more-serious symptoms. If all four alpha-globin alleles are nonfunctional, no alpha-globin protein is produced, and this state is fatal. Alpha-thalassemia is most commonly found in Africa, China, India, the Middle East, Southeast Asia, and occasionally the Mediterranean region.

The eminent scientist J.B.S. Haldane first proposed the "malaria hypothesis" for the existence of thalassemia in 1948. Haldane suggested that individuals who are heterozygous for thalassemia may be more resistant to malaria. He also noted the co-occurrence of malaria and thalassemia in certain parts of the world. Recent studies show that the high frequencies of milder forms of alpha-thalassemia are indeed related to protection from malaria (Dronamraju and Arese, 2006). A recent study by Fowkes et al. (2008) explored whether people with different numbers of nonfunctional alpha-thalassemia alleles have higher protection against severe malaria. Their results suggest that the increased number and small size of the red blood cells in children who have some nonfunctional alleles of the alpha-thalassemia gene contribute to their protection against malaria.

Beta-thalassemia is a serious health problem worldwide and causes the deaths of hundreds of thousands of children per year. Compared with alpha-thalassemia, beta-thalassemia is more common in the Mediterranean region, but it is still much more frequent in areas where malaria is endemic. Malaria used to be endemic in the Mediterranean and is currently coming back into southern Italy. In the early 19th century, tourists often got malaria in Rome when they visited the Coliseum at night. The prevalence of beta-thalassemia in the Mediterranean is a good example of a past signature of selection that has not yet disappeared and in some areas is experiencing selection again. Beta-thalassemia is caused by different molecular defects that reduce or abolish the synthesis of the beta-globin chains.

5.5 Van der Woude Syndrome and Irf6

Much of the relevant information about **Van der Woude syndrome** and the *Irf6* gene is included in the masters in Lesson 3. In brief, Van der Woude syndrome is an inherited developmental disorder. Syndrome means a group of symptoms or signs that are characteristic of a specific disease. Individuals with the disorder may have pits of the lower lip and a cleft lip, a cleft palate, or both. Cleft lip and palate may or may not be a part of a syndrome. Van der Woude syndrome is the most common form of cleft palate that is associated with a syndrome.

For more information about Van der Woude syndrome, see http://ghr.nlm.nih.gov/condition=vanderwoudesyndrome.

Mutations to the *Irf6* gene cause Van der Woude syndrome (Kondo et al., 2002). In Lesson 3, students examine exon 3 and part of the flanking introns for the *Irf6* gene from a range of organisms. For more information about this gene, see http://ghr.nlm.nih.gov/gene=irf6.

The following Web site has an interesting interview with two of the authors of the paper that identified the relationship between Van der Woude syndrome and the *Irf6* gene: http://www.nidcr.nih.gov/Research/ResearchResults/InterviewsOHR/TIS102002.htm.

The phylogenies used in the activity are based on data from Prasad et al., 2008.

5.6 Influenza

Influenza is a disease that affects millions of people each year. Understanding the evolution of influenza viruses is critical for learning how to better avoid future pandemics and for treating people with influenza. Scientists distinguish between two main types of changes in influenza, **antigenic drift** and **antigenic shift**. Antigenic drift causes relatively small changes that occur among flu strains from year to year. Infrequent and relatively large changes in the composition of the predominant influenza virus are called antigenic shift. In the supplement, students explore antigenic drift.

Antigenic drift is not the same as genetic drift. Each year, one strain of influenza circulates widely throughout the world. People infected with influenza develop antibodies to the **hemagglutinin** antigen present on the outside of the virus. When the antibodies bind to the hemagglutinin antigen, the virus is prevented from entering the cells of the body. As a result, individuals are protected against reinfection from the same strain. Influenza vaccinations also result in the formation of protective antibodies. Because a large number of people develop immunity to a strain of influenza virus, forms of the virus that have mutations that change the shape of the hemagglutinin region—and allow the virus to avoid detection by the immune system—are favored by natural selection. As a result, the influenza virus changes from year to year. This helps explain why new vaccines must be developed every few years. A fuller explanation of why the diversity in influenza differs from that in other viruses, such as measles, mumps, and rubella, is described in Lipsitch and O'Hagan (2007).

Figure 10. *Outbreaks of influenza influence our lives. Wearing a mask can protect us from inhaling flu viruses.*

© Mcpics / Dreamstime.com

Dramatic changes to type A influenza are an example of antigenic shift. Antigenic shift occurs when a new strain of influenza A develops in which the hemagglutinin gene differs substantially from strains that have circulated in humans in recent decades. In the examples known to date, this has resulted from "reassortment" of genetic material from two or more influenza viruses. In this reassortment, the gene for the novel hemagglutinin comes from one virus and the rest of the genetic material comes from a different virus or multiple viruses. Such reassortment can occur when a host cell (in humans, pigs, birds, or other host species) is infected simultaneously by more than one genetically distinct virus. If an antigenically shifted influenza A strain is capable of being transmitted from human to human, it may cause a new pandemic of influenza. Pandemics are defined by genetically novel viruses that spread widely and infect more humans than are infected in a normal influenza season because few individuals have immunity to the novel hemagglutinin. The novel strain may also cause more severe disease than usual due to the lack of immunity, the strain's novel genetic characteristics, or both. Pandemics have occurred most recently in 1889–1892, 1918, 1957, 1968, and 2009.

More information about antigenic shift can be found at http://www3.niaid.nih.gov/topics/Flu/Research/basic/pages/AntigenicShiftIllustration.aspx.

Students may be interested to learn that in 2005, scientists sequenced the entire genome

of an influenza virus from someone who died of the flu in 1918 (Taubenberger et al., 2005). The person was buried in permafrost and was remarkably well preserved.

5.7 Vitamin C Biosynthesis

All eukaryotic organisms contain nonfunctional sequences of DNA called "pseudogenes." These genetically silent sequences of DNA are vestiges of once-functional genes inherited from ancient ancestors. One such pseudogene explains why all mammals, except humans, guinea pigs, and bats, can synthesize vitamin C. Mammals inherited a functional version of the gene for synthesizing vitamin C from an early ancestor. But in independent events in humans, guinea pigs, and some bats, the gene, even though present, later mutated and became nonfunctional. As a result, humans, guinea pigs, and some bats have to obtain vitamin C from their food.

Scientists have found the pseudogene for an enzyme that's required for vitamin C biosynthesis, L-gulonolactone oxidase (GULO), in the human genome (Nishikimi et al., 1994). Comparisons with the functional gene in other species show that the human pseudogene has many **substitutions** that eliminate its function. Some of these substitutions have produced stop codons within the human pseudogene.

Additional information about the evolution of vitamin C metabolism in humans is available in Lesson 5.