

Instant Update: Considering the Molecular Mechanisms of Mutation & Natural Selection

TINA HUBLER, PATTI ADAMS,
JONATHAN SCAMMELL

ABSTRACT

The molecular basis of evolution is an important concept to understand but one that students and teachers often find challenging. This article provides training and guidance for teachers on how to present molecular evolution concepts so that students will associate molecular changes with the evolution of form and function in organisms. Included are examples that illustrate how mutation followed by selection causes populations to evolve. Next month we will share lab activities that illustrate the concepts and reinforce the complementary roles of mutation and selection in the overall process of evolution.

Key Words: Molecular evolution; mutation; natural selection.

○ Introduction

Molecular evolution is a change in the chemical composition of molecules such as DNA, RNA, and proteins over time as a result of DNA mutation and selective pressure. While DNA mutations provide the raw material for evolution, conditions or pressures from the environment influence which organisms survive. With survival and reproduction comes a change in the frequencies of specific molecules that are observed within a population of organisms. The study of molecular evolution has blossomed in recent years with the development of technologies, such as DNA sequencing and bioinformatics, that allow us to determine and compare organisms' DNA sequences. This information has become so important for understanding evolution that numerous state and national science standards now include objectives related to the molecular basis of evolution. As an example, the *Next Generation Science Standards* (<http://www.nextgenscience.org>) contain two standards that address empirical evidence for biological evolution based on similarities in DNA sequences among organisms and genetic variation of individuals due to mutation. Despite this current emphasis, understanding the mechanisms that underlie molecular evolution

While DNA mutations provide the raw material for evolution, conditions or pressures from the environment influence which organisms survive.

is difficult for students. This article is designed to prepare teachers with the necessary background information so that, in turn, they can design lessons to help students understand important concepts in molecular evolution.

○ Molecular Evolution: A Review of Key Concepts

Genome Organization

Data from the Human Genome Project and similar sequencing projects revealed that a genome (all of the genetic information in an organism) is organized into genes (International Human Genome Sequencing Consortium [IHGSC], 2004; Alberts et al., 2007). The definition of a “gene” has been debated and revised, as scientists acquire a greater appreciation for the complexity of a genome (Gerstein et al., 2007). Genes may be described as DNA sequences used to produce functional products, such as protein or RNA. Eukaryotic genes include coding sequences called “exons” that provide the genetic information for the cell to assemble amino acids into specific proteins. These genes also include important non-coding sequences such as promoters, introns, 5′ and 3′ untranslated regions (UTRs), and RNA genes for transfer RNA (tRNA), ribosomal RNA (rRNA), small nuclear RNA (snRNA), and microRNA (miRNA). Promoters do not code for protein but are regulatory sequences associated with genes and control how genes are expressed. Introns are sequences that are removed from most RNAs after transcription. As introns are removed, exons are spliced together to form a “mature” messenger RNA that may be exported from the nucleus to the cytoplasm for translation into protein. The 5′ and 3′ UTRs are segments of messenger RNAs that are not translated into the amino acid subunits of proteins but that sometimes have regulatory roles. RNA genes are DNA sequences that are used to produce RNAs that

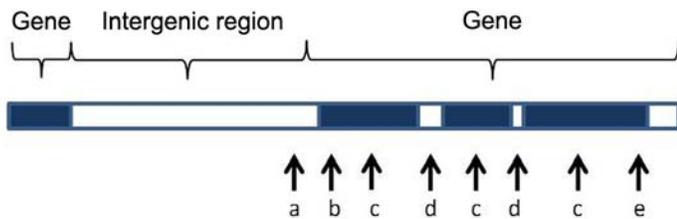


Figure 1. Schematic of the organization of DNA sequences composing a genome: a = promoter, b = 5' UTR, c = exon, d = intron, and e = 3' UTR.

Table 1. DNA sequence classification.

Examples of Functional DNA Sequences	Examples of Nonfunctional DNA Sequences
Protein coding sequences RNA genes Noncoding sequences: Promoters Other regulatory regions ^a	Noncoding sequences: Some 5' UTR, 3' UTR, intronic, and intergenic regions Pseudogenes

^aRegulatory regions may be located in noncoding sequences such as introns, 5' and 3' UTRs, and intergenic regions.

have a variety of functions: translation (tRNA and rRNA), removal of introns (snRNA), and regulation of gene expression (miRNA). Figure 1 is a schematic of genome organization that may help students visualize the relative locations of some coding and noncoding components in a genome. It is included as a PowerPoint slide (see online Supplemental Materials).

A recent analysis of >3 billion nucleotides of the human genome surprisingly estimated that it included only 20,000 to 25,000 genes (IHGSC, 2004). However, in addition to the coding sequences found in genes, the human genome contains an abundance of noncoding DNA (Lander et al., 2001; IHGSC, 2004) that constitutes about 95% of the human genome (Elgar & Vavouri, 2008). Noncoding DNA may be classified as functional or nonfunctional. It is considered functional if it is associated with a role such as regulation of gene expression. Functional noncoding DNA includes promoters and other regulatory sequences that may occur in introns, 5' and 3' UTRs, and intergenic (between genes) regions. Nonfunctional noncoding sequences are also found in introns, 5' and 3' UTRs, and intergenic regions. “Nonfunctional” is a conventional term used to describe noncoding DNA sequences that have no currently known function. The status of a nonfunctional sequence may be revised to “functional” if a role is identified. For example, a number of sequences in introns and UTRs have been assigned regulatory roles (Barrett et al., 2012). Nonfunctional noncoding sequences also include pseudogenes that are presumed to be mutated remnants of functional genes. Table 1 may be used to help students classify genomic DNA into functional and nonfunctional sequences.

The Dynamic Nature of the Genome

Data from sequencing projects also revealed that genomes are dynamic (Platzer, 2006). Mutations (changes in DNA sequences) occur spontaneously as a result of DNA replication error and recombination events. They may also be induced by physical agents

such as ultraviolet radiation, chemical agents such as polyaromatic hydrocarbons, or drugs such as doxorubicin. Although there are some regions of the genome that are frequently mutated (“hotspots”), mutations occur randomly in all regions of the genome. They appear throughout the genome as substitutions, insertions, deletions, and rearrangements. Therefore, as DNA is passed from one generation to the next, the genome accumulates variability.

The degree of variability observed after many generations is not the same in nonfunctional and functional regions. Nonfunctional regions are characterized by highly variable sequences. By contrast, functional regions such as those that encode important proteins are remarkably similar or conserved. The tendency for functional sequences to be conserved, while nonfunctional sequences accumulate variability, can be understood by considering the effect of natural selection.

The Effect of Selection on the Genome

Because mutations that occur in gametes are inherited by offspring, they provide a continuous source of variation for future generations. Mutations may be classified as neutral, harmful, or beneficial, based on the effect they have on an organism’s ability to survive and reproduce.

Mutations can have at least four possible types of effect on genome variability over many generations. First, any mutations in nonfunctional regions can be classified as “neutral” because they have no effect on the organism’s phenotype. These mutations will be perpetuated in offspring because there is no disadvantage to the organism that carries them. In successive generations, mutations will accumulate in these regions because they are free from the constraint of selection. Therefore, these genomic regions exhibit high variability compared with other regions.

Second, mutations that occur in functional regions may also be classified as neutral if the changes do not affect the organism’s phenotype. Examples of neutral mutations in functional regions are DNA substitutions that do not change the amino acid incorporated into a protein (a silent mutation) or DNA substitutions that change an amino acid to another with similar biochemical properties (a neutral mutation). As is the case with mutations in nonfunctional regions, neutral mutations in functional regions also accumulate over generations because there is no disadvantage to the organism that has these mutations. However, functional sequences will exhibit moderate to low variability compared with nonfunctional sequences because mutations in functional regions also have the potential to be harmful.

Third, as expected, harmful mutations in functional regions negatively influence the survivability of organisms. Fewer or no offspring may be produced, and these mutations will be found rarely, or not at all, in subsequent generations. So, harmful mutations may be “selected out” of the population and not observed in present-day genomes, even though the mutations occurred at some point in the past.

Fourth, beneficial mutations in functional regions confer an advantage to an organism. They positively influence survivability, which increases the production of offspring. These mutations will be inherited by offspring over multiple generations because it is unlikely that another random mutation will occur in the exact same location. Therefore, beneficial mutations tend to persist because survivability and reproduction are enhanced. In short, selection acts to remove

Table 2. The effect of a mutation's location and type on genome variability as a consequence of selection.

Mutation Location	Mutation Type	Effect on Organism	Effect on Variability Observed in Genome
Nonfunctional DNA	Neutral	None	DNA sequences are highly variable
Functional DNA	Neutral	None	DNA sequences are moderately variable
	Harmful	Deleterious	Rarely observed in genomes of surviving organisms
	Beneficial	Advantageous	DNA sequences are conserved

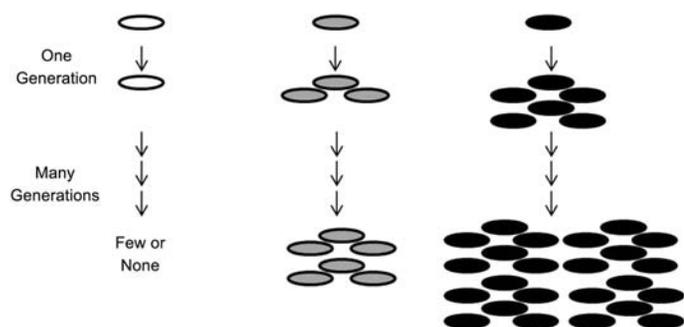


Figure 2. The effect of selection on the number of surviving organisms (dots) carrying harmful (clear), neutral (gray), or beneficial (black) mutations.

harmful DNA sequences and retain neutral and beneficial sequences. Table 2 summarizes the effect that the mutation's location and type have on the organism itself and on genome variability as a consequence of selection.

As discussed, the effect of selection on the nature of the genome is ultimately linked to survivability and reproduction. It may be helpful to present this schematically for students. Figure 2 can be used to illustrate the effect of selection on the number of surviving organisms (dots) carrying harmful (clear dots), neutral (gray dots), and beneficial (black dots) mutations. The schematic illustrates that, in comparison to organisms with neutral mutations, organisms with harmful mutations tend to produce fewer or no offspring. They may even become extinct. By contrast, organisms that acquire beneficial mutations tend to produce more offspring. After many generations, the surviving population consists mostly of organisms whose genomes house neutral and beneficial mutations. Two important points should be made to students: (1) Harmful mutations, particularly lethal ones, are removed from populations as a result of natural selection; and (2) beneficial mutations typically contribute to the production of larger numbers of offspring that carry the acquired mutation (as well as other mutations). This diagram is included as a PowerPoint slide (see online Supplemental Materials) to aid in the discussion of these concepts.

Mutations as a Historical Record of Evolution

Researchers gain insight into past molecular changes by employing comparative genomics. This method compares the DNA or amino acid

sequences of multiple organisms using sequence alignment tools, such as the BLAST tool provided by the National Center for Biotechnology (<http://www.ncbi.nlm.nih.gov>). For example, the amino acid sequence of cytochrome *c* has been compared in organisms. Cytochrome *c* is an essential component of the electron transport chain in mitochondria, where it functions to transfer electrons from donor to acceptor molecules during cellular respiration. Because of its role in cellular respiration, it is an essential protein. As expected, the DNA sequence of the gene and the amino acid sequence of the protein are highly conserved among organisms. The degree of cytochrome *c* amino-acid-sequence

similarity between two organisms is related to the length of time since the organisms diverged from a common ancestor. In distantly related organisms (e.g., humans and fish), the sequences are less similar than in closely related organisms (e.g., humans and chimpanzees). This is true because, in distantly related organisms, more time has passed since their divergence from a common ancestor. More mutations have accumulated, making the genomes more different. Because of the relationship between sequence similarity and the length of time since organisms diverged from a common ancestor, it was once proposed that sequence variation in conserved proteins such as cytochrome *c* and hemoglobin could be used as “molecular clocks” (Kumar, 2005). It is now appreciated that factors other than the length of time since divergence affect the degree of sequence variation. These include the particular proteins being compared, changes in population size, and variable generation times for different organisms. Still, a generalization can be made that closely related organisms share greater sequence similarity than distantly related organisms.

By comparing the sequences of organisms, researchers have also uncovered a number of specific DNA mutations that have been implicated in evolution. For example, DNA mutations are associated with changes in the coat color of rock pocket mice (Nachman et al., 2003) and the production of antifreeze proteins in Antarctic fish (Chen et al., 1997). In each case, a random change in DNA sequence conferred a selective advantage that was perpetuated in offspring. It is useful to realize that an organism's genome may “record” the presence of ancestral genes that are found in other organisms, as well as new genes that have arisen through mutation. These examples of currently living organisms are valuable in helping students envision the process whereby populations of organisms change over time.

Teaching the Concepts

We have found that teaching the concepts of mutation and selection as the molecular basis of evolution may best be divided into four stages. First is presentation of the concepts of sequence classification (noncoding, coding, nonfunctional, functional), mutations, selection, nonfunctional sequence variability, and functional sequence conservation. Second, students are prepared for laboratory activities by discussing polymerase chain reaction and electrophoresis. Third, laboratory activities are performed that expose students to basic principles of variable and conserved sequences. Finally, a reading assignment and video about icefish are employed to reinforce the evolutionary consequence of mutation and selection. Next month we will share a number of activities that concretely illustrate the concepts

and reinforce the complementary roles of mutation and selection in the overall process of evolution.

○ Supplemental Materials

Supplemental Materials are available at <http://www.buildingthepride.com/faculty/trhubler/>.

○ Acknowledgments

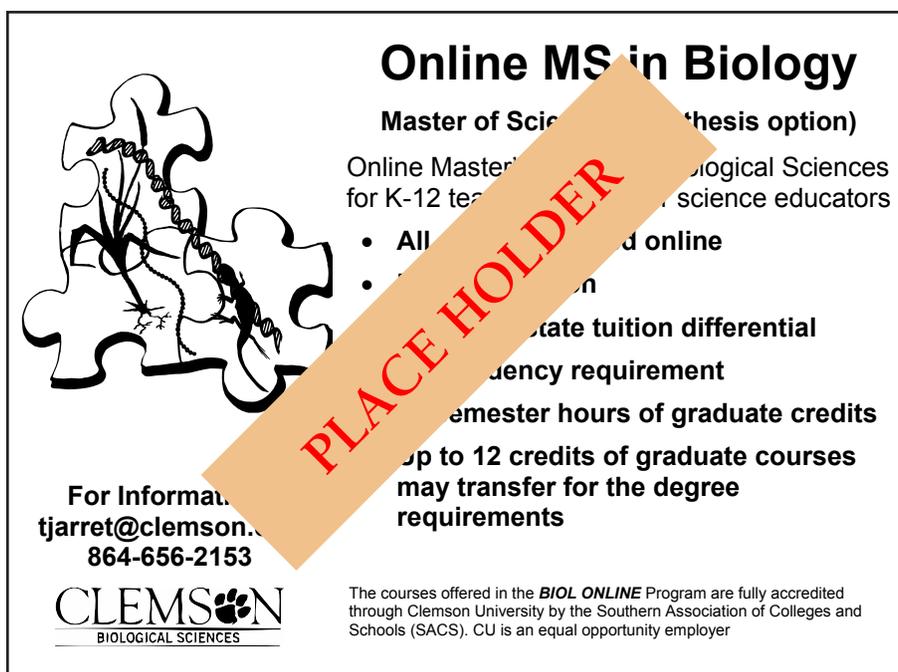
This work was supported by intramural grants from the University of North Alabama (UNA) and by the UNA Department of Biology and South University School of Pharmacy.

References

- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. (2007). *Molecular Biology of the Cell, 5th Ed.* New York, NY: Garland Science.
- Barrett, L.W., Fletcher, S. & Wilton, S.D. (2012). Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cellular and Molecular Life Sciences, 69*, 3613–3634.
- Chen, L., DeVries, A.L. & Cheng, C.C. (1997). Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. *Proceedings of the National Academy of Sciences, 94*, 3811–3816.

- Elgar, G. & Vavouri, T. (2008). Tuning in to the signals: noncoding sequence conservation in vertebrate genomes. *Trends in Genetics, 24*, 344–352.
- Gerstein, M.B., Bruce, C., Rozowsky, J.S., Zheng, D., Du, J., Emanuelsson, O. et al. (2007). What is a gene, post-ENCODE? History and updated definition. *Genome Research, 17*, 669–681.
- International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. *Nature, 431*, 931–945.
- Kumar, S. (2005). Molecular clocks: four decades of evolution. *Nature Reviews Genetics, 6*, 654–662.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J. et al. (2001). Initial sequencing and analysis of the human genome. *Nature, 409*, 860–921.
- Nachman, M.W., Hoekstra, H.E. & D’Agostino, S.L. (2003). The genetic basis of adaptive melanism in pocket mice. *Proceedings of the National Academy of Sciences, 100*, 5268–5273.
- Platzer, M. (2006). The human genome and its upcoming dynamics. *Genome Dynamics, 2*, 1–16.

TINA HUBLER (corresponding author) is an Associate Professor of Biology at the University of North Alabama, Florence, AL 35630; e-mail: trhubler@una.edu. PATTI ADAMS is an Assistant Professor of Pharmaceutical Sciences at South University School of Pharmacy, Savannah, GA 31406; e-mail: padams@southuniversity.edu. JONATHAN SCAMMELL is Professor and Chair of the Department of Comparative Medicine at the University of South Alabama, Mobile, AL 36688; e-mail: jscammell@southalabama.edu.



Online MS in Biology
Master of Science (thesis option)

Online Master of Science in Biological Sciences
for K-12 teachers and science educators

- All courses are online
- State tuition differential
- No residency requirement
- Up to 12 credits of graduate courses may transfer for the degree requirements

For Information:
tjarret@clemson.edu
864-656-2153

CLEMSON
BIOLOGICAL SCIENCES

The courses offered in the **BIOL ONLINE** Program are fully accredited through Clemson University by the Southern Association of Colleges and Schools (SACS). CU is an equal opportunity employer